KUNGL
TEKNISKA
HÖGSKOLAN

CID-185 ● ISSN 1403-0721 ● Department of Numerical Analysis and Computer Science ● KTH

Hybrid physical-digital artefacts

SHAPE IST 2000-26069 Workpackage1 Deliverable D1.1

Steve Benford, Mike Fraser, Boriana Koleva, John Bowers, Paul Chandler, Luigina Ciolfi, Chris Greenhalgh, Tony Hall, Sten-Olof Hellström, Shahram Izadi, Ian Taylor, Liam Bannon, Holger Schnädelbach, Martin Flintham, Malcolm Foster, Gustav Taxén

CID, CENTRE FOR USER ORIENTED IT DESIGN

Steve Benford, Mike Fraser, Boriana Koleva, John Bowers, Paul Chandler, Luigina Ciolfi, Chris Greenhalgh, Tony Hall, Sten-Olof Hellström, Shahram Izadi, Ian Taylor, Liam Bannon, Holger Schnädelbach, Martin Flintham, Malcolm Foster, Gustav Taxén

Hybrid physical-digital artefacts

Report number: CID-185, SHAPE IST 2000-26069 Workpackage1 Deliverable D1.1

ISSN number: ISSN 1403 - 0721 (print) 1403 - 073 X (Web/PDF)

Publication date: September 2001

E-mail of author: bowers@nada.kth.se

Reports can be ordered from:

CID, Centre for User Oriented IT Design
NADA, Deptartment of Numerical Analysis and Computer Science
KTH (Royal Institute of Technology)
SE- 100 44 Stockhom, Sweden
Telephone: + 46 (0)8 790 91 00
Fax: + 46 (0)8 790 90 99
E-mail: cid@nada.kth.se
URL: http://cid.nada.kth.se

Situating Hybrid Assemblies in Public Environments

26069

Deliverable D1.1

"Hybrid physical-digital artefacts"

Workpackage 1, Lead partner: University of Nottingham

19/12/2001

Table of Contents

Chapter 1: Introduction to the SHAPE typology

Steve Benford, Mike Fraser, Boriana Koleva
Mixed Reality Laboratory, University of Nottingham, UK.
{sdb, mcf, bnk}@cs.nott.ac.uk

John Bowers
KTH, Stockholm, Sweden
bowers@nada.kth.se

This chapter introduces Deliverable 1.1 of the SHAPE project. This deliverable is the product of Workpackage 1 (Technology Exploration), Task 1.1 (Hybrid Digital-Physical Artefacts). As required in the Workplan of the SHAPE project, this deliverable offers a typology of hybrid mixed reality artefacts, presents demonstrations of some SHAPE-characteristic mixed reality applications, while exploring various strategies for inter-media interaction (particularly relating graphical with auditory interaction and display). This chapter presents a typology of mixed reality systems and applications by adopting the notion of a 'mixed reality boundary' as a primary concept. Systems are seen as instantiating one or more boundaries of this sort. Boundaries are further analysed into component concepts, and different classes of system can be situated in terms of how these component concepts are negotiated. A notion of hybrid digital-physical object is offered in relationship to this mixed reality boundary concept. This chapter shows that the demonstrators which Workpackage 1 has developed within SHAPE can all be accommodated and illuminated in terms of these concepts, the particular emphases of SHAPE being shown to be addressing deficiencies in the research literature with respect to the coverage of the typology. Along the way, each of the seven chapters which follow this one are introduced.

The Concept of Boundaries

Boundaries pervade the fabric of public spaces. Whether physically demarcating exhibition areas or collecting artefacts with related properties, museums in particular create boundaries for us to make sense of and experience. In this chapter, we suggest that a potentially useful typology for designing for public places is to build on the idea of boundaries. In this approach, a device establishes a relationship between two otherwise distinct spaces. We hope to show that this perspective enables us to have a characteristic approach to the combination of media in so-called mixed reality systems and to the status of informational objects within such systems. This chapter develops this framework, uses it to make sense of existing mixed reality systems in the literature and situates our own work in the first year of SHAPE in comparable terms.

We hold that there are three general types of boundaries:

- **Physical-physical boundaries** connect two physical spaces. These devices can be composed from traditional building materials (e.g., walls, windows, curtains, doors etc.) or from telepresence materials (e.g., networked video and audio connections that effectively join two remote physical spaces).
- **Virtual-virtual boundaries** connect two virtual spaces. A range of such boundaries has been proposed for 3D virtual environments, affording different possibilities for seeing (and hearing) a remote virtual space or for crossing into it (Barrus et al. 1996). Hyperlinks might also be considered to be a form of virtual-virtual boundary where the spaces are defined by documents or parts of documents.
- **Mixed reality boundaries** connect physical to virtual spaces (Benford et al. 1998). A simple example of a mixed reality boundary is to project a view of a virtual environment into a physical space while texture mapping a view of the physical environment into the virtual, and aligning the two so that they appear to form two sides of a common window. The occupants of the physical space may look into the virtual space to see avatars looking back out at them.

Mixed reality boundaries (and possibly physical-physical and virtual-virtual boundaries too) can be designed according to set of generic properties (Koleva et al. 1999). These include:

- **Permeability** – how does a boundary affect information that flows across it? Possible effects include attenuation, amplification or transformation of information and these may apply to visual information (leading to the sub-property of visibility), sound (audibility) and other media. One interesting sub-property is solidity, the extent to which a boundary allows participants and objects to physically traverse it so that they become embodied in, and able to move through, the remote space. If it does, it is what we call a traversable interface (Koleva et al. 2000).
- **Situation** – the spatial properties of the boundary including its position in each of the connected spaces, its orientation to the various participants, whether it is segmented (comes as different sections, each with different properties) and mobile (whether it moves through the connected spaces).
- **Dynamics** – the temporal properties of the boundary including for how long it exists (lifetime), when it is scheduled to appear and whether its properties can be dynamically changed.
- **Symmetry** – are these properties the same for each side of the boundary (physical to virtual and virtual to physical)? They need not be – it is quite possible to have asymmetric boundaries.

Classifying Mixed Reality

In this chapter, we review different existing mixed reality approaches, showing that the boundary property framework might serve as a means for classifying any mixed reality in general. It is suggested that fundamentally we can view mixed reality approaches as an attempt to link physical and virtual spaces and that whenever we try to connect two remote spaces we establish some form of boundary between them. It is further proposed that we can describe the different mixes of real and virtual elements, created by the various mixed reality approaches, in terms of boundaries with different configurations of the above property set.

The following sections will examine how effective the boundary framework is at classifying different mixed reality approaches and example systems thereof.

Augmented reality

Augmented reality systems can be seen as linking a physical space to a virtual information space. As the latter is usually uninhabited we only need to consider the properties at the physical side of the boundary. Table 1.1 summarises the systems Ultrasound Imagery, AGROS, KARMA, Shared Space and Digital Desk in terms of boundary properties.

Augmented reality systems create boundaries that offer visibility of virtual objects from a physical space. The field of view is limited when a HMD is utilised as a display. The resolution of the virtual objects differs between systems. Usually the viewing perspective is updated when the viewpoint is changed and some systems (like AGROS and Shared Space) also provide a stereoscopic view of the virtual objects. Most augmented reality boundaries are non-solid at the physical side, allowing participants to interact with the virtual objects.

The boundaries created by augmented reality systems are the most difficult to conceptualise in terms of the property of location. Here it is suggested that the most consistent interpretation is to view these boundaries as being located at the system’s display surface. The means that the boundaries created by Ultrasound Imagery, KARMA and Shared Space are located at the HMD. On the other hand, the boundary created by AGROS is vertical, being located on a computer monitor and the boundary created by Digital Desk is horizontal, being located on a desk surface.

Augmented reality boundaries are usually mobile. In the case when a HMD is used as a display, the user’s movements steer the boundary through physical and virtual space (i.e. as the user’s head moves around different parts of the virtual information space become visible). In the case of the Digital Desk boundary mobility is realised in a different way – the user navigates the virtual space by opening different documents.

Usually augmented reality boundaries are not segmented. In terms of configurability, the Shared Space system allows users to manipulate not only the location of the boundary but also the visibility through it (a user can specify whether a new web page is to remain private or become visible to all).

Systems that involve wearing equipment, such as a HMD, are normally aimed at accomplishing a specific task, after which the boundary is removed. The lifetime of these boundaries is in the order of minutes and hours. On the other hand, systems like Digital Desk are suitable for linking a physical and a virtual space on a more permanent basis.

Finally, augmented reality systems usually aim to offer a view of the present state of a virtual space. However, sometimes due to processing delays the user may be connected to virtual space, as it was a few seconds in the past.

Property	Ultrasound Imagery	KARMA	AGROS	Shared Space	Digital Desk
Visibility	Image of 512 x 480 max res., viewpoint perspective	Graphics within 22º horizontal field of view, display res. of 720 x 280	Stereoscopic graphics	30º horizontal field of view, stereoscopic graphics, viewpoint perspective	1120 x 780 res. (for first prototype)
Audibility	None	None	None	None	None
Solidity	Indirect manipulation of virtual objects	Indirect manipulation of virtual objects	Manipulation of virtual objects	Manipulation of virtual objects	Manipulation of virtual objects
Location	Head-mounted display Variable location	Head-mounted display Variable location	Vertical (monitor)	Head-mounted display Variable location	Horizontal (physical desk surface)
Mobility	Limited mobility through virtual space	Mobile through virtual space	None	Mobile through virtual web space	Mobile (navigation of virtual data space by opening different digital documents)
Segmentation	None	None	None	None	None
Lifetime	Minutes/Hours	Minutes/Hours	Minutes/Hours	Minutes/Hours	Persistent
Configurability	Dynamically configurable location	Dynamically configurable location	None	Configurable visibility and location	None
Temporal location	Present (a few second delay)	Present	Present	Present	Present
Symmetry	Asymmetrical	Asymmetrical	Asymmetrical	Asymmetrical	Asymmetrical
Representation	Implicit	Implicit	Implicit	Explicit representation of visibility and awareness	Implicit

Table 1.1 Augmented reality systems in terms of boundary properties

Property	FreeWalk		Virtual Office		Media Space	Invisible Shape
	Physical side	Virtual side	Physical side	Virtual side	Virtual side	Virtual side
Visibility	Graphics, monitor res.	Video, 320x180 res., 10 frames/sec	Graphics, monitor res., 180º FOV, FOV extent depends on concentration level	Video	Video, variable frame rate	Video
Audibility	Virtual world sounds, volume depends on distance	Participant's voice	Virtual world sounds, field of hearing depends on concentration level	Participant's voice	None	None
Solidity	Interaction with virtual objects	Solid	Interaction with virtual objects	Solid	Solid	Solid
Location	Vertical, monitor	Avatar	Vertical, monitor	Avatar	Vertical wall or horizontal ceiling of a virtual office	Spatially and temporally organised
Mobility	Static	Mobile	Static	Mobile	Static	Static but mobile through physical space
Segmentation	None	None	None	None	None	Some
Lifetime	Minutes/Hours	Minutes/Hours	Minutes/Hours	Minutes/Hours	Minutes/Hours	Minutes/Hours
Configurability	None	None	Indirectly level of visibility and audibility	None	Indirectly video frame rate	Configurable temporal location
Temporal location	Present	Present	Present	Present	Present	Linked to the past
Symmetry	Asymmetrical		Asymmetrical		Asymmetrical	Asymmetrical
Representation	Implicit		Implicit		None	Through shape of video object

Table 1.2 Augmented virtuality systems in terms of boundary properties

Augmented virtuality

The systems FreeWalk, Virtual Office, Media Space and The Invisible Shape of Things Past are described in terms of boundaries with specific properties in Table 1.2 (the system InterSpace is not included because it creates boundaries with very similar properties to FreeWalk and Virtual Office).

We can view all augmented virtuality systems as creating asymmetrical boundaries between physical and virtual space. Based on their properties, it is proposed that we can divide these boundaries into two categories. The first type is the boundary between a participant's local physical space and a virtual world that is created by systems that use video to augment avatars. FreeWalk and Virtual Office establish such boundaries and they are described in Table 1.2. This category of boundary generally offers good visibility and audibility of the virtual space and limited visibility and audibility of the physical space (the field of view is restricted to the participant's face and the field of hearing to the participant's voice). However, specific FOV and resolution values differ between systems (see Table 1.2). Further the boundary is generally not solid from the physical side, allowing participants to manipulate virtual objects and it is mobile through the virtual space, allowing participants to navigate. On the other hand, this type of boundary is solid from the virtual side and it is static in the physical environment. In terms of location the boundary is usually vertical, it is integrated with an avatar geometry in virtual space and represented by a display device (a monitor in FreeWalk and Virtual Office) in the physical environment. The lifetime of this kind of boundary is variable. It depends on how long a participant is connected to a virtual world, typically lasts from a few minutes to hours. Finally, these boundaries are generally not configurable but there are exceptions that allow participants to choose levels of audibility and visibility.

The second category of boundary is created by systems that use video to introduce real world views into a virtual environment. Invisible Shape and Media Space are such systems (actually Media Space creates both types of boundaries but only the second category is illustrated in Table 1.2). This type of boundary is not represented in physical space; this is why only the virtual side is described in Table 1.2. Usually good visibility and possibly audibility are provided of the physical space that is linked to the virtual world. The boundary is generally solid and it is often static in both spaces (although like Invisible Shape some systems have mobile cameras in physical space). The lifetime of this kind of boundary is usually the same as the lifetime of the virtual space.

Most often the intended temporal location is the present (although as with augmented reality systems there could be processing and transmission delays). The boundaries created by Invisible Shape are unusual in that they link to the past of physical spaces (this temporal location is dynamically configurable as the user is able to jump to any point in the video sequence). Finally, the boundaries of augmented virtuality systems may be fragmented. Invisible Shape may create segmentation if a number of video film objects exit for the same physical space at the same point in time.

Ambient displays

Ambient display systems link an inhabited physical space to a virtual space through a boundary that is only represented in the physical space. Table 1.3 describes the physical side properties of the boundaries created by the systems AmbientROOM, Water Lamp, Pinwheels and Table Fountain.

The most distinguishing feature of the boundaries created by ambient media displays is that they perform transformations on the information passing from the virtual to the physical space. It is very common for virtual events to be expressed in terms of abstract visual patterns. Alternatively some systems transform information from cyberspace into audio effects.

Ambient display boundaries are solid; participants cannot manipulate the virtual space through them. In terms of location some boundaries are represented as physical object/fixture whereas others, like AmbientROOM, completely surround the user. In fact AmbientROOM surrounds the user with three separate boundaries that convey information from three different spaces.

Property	AmbientROOM	Water Lamp	Pinwheels	Table Fountain
Visibility	Transformation (abstraction) - Activity in an atrium expressed as pattern of illuminated patches	Medium transformation - Virtual events mapped to light and shadow patterns	Medium transformation - Virtual events mapped to speed of spinning pinwheels	Medium transformation - Virtual events mapped to water fountain patterns
Audibility	Medium transformation - Virtual events mapped to volume and density of a soundtrack - Hamster activity mapped to vibration	None	None	Medium transformation - Virtual events mapped to water fountain sound
Solidity	Solid	Solid	Solid	Solid
Location	An office	Horizontal (ceiling)	Fixture	Fixture
Mobility	Static	Static	Static	Static
Segmentation	None	None	None	None
Lifetime	Persistent	Persistent	Persistent	Persistent
Configurability	None	None	None	Choice of transformations
Temporal location	Present	Present	Present	Present and past
Symmetry	Asymmetrical	Asymmetrical	Asymmetrical	Asymmetrical
Representation	Implicit	Implicit	Implicit	Implicit

Table 1.3 Ambient Display systems in terms of boundary properties

None of the example systems are segmented, although it is possible for multiple ambient displays to be used to convey information about the same virtual space. In terms of lifetime, ambient displays are very suitable for persistent use. From the examples Table Fountain is the only system that creates a configurable boundary. It allows the users to select what effect(s) are to represent a specific information source. Finally, because some of the effects of Table Fountain remain visible for some time (due to state change) it can represent current and past events.

Graspable interfaces

Like augmented reality systems, graspable interfaces usually link an uninhabited virtual space to a physical environment. This means that we only need to consider the properties at the physical side of these boundaries. Table 1.4 summarises the systems MetaDesk, Illuminating Light and Passage.

The boundaries created by graspable interfaces usually offer good visibility of the virtual space and most characteristically they are not solid, allowing manipulation of virtual objects. Graspable interface boundaries are often horizontal (as such a surface conveniently supports physical objects) but there are also vertical ones. Some systems allow the user(s) to navigate the virtual space, e.g. MetaDesk and Passage. It is arguable whether the interaction with physical objects in Illuminating Light results purely in manipulation of the virtual objects or in a combination of manipulation and navigation).

The MetaDesk system creates a segmented boundary because it offers three different views onto the same virtual space. The Passage boundary is also segmented. We can view each bridge/screen pair as a segment of a single boundary that provides access to a virtual data space. The systems Illuminating Light and Passage link to the present state of a virtual space. MetaDesk, on the other hand, gives access to a model that was created sometime in the past (and which is not dynamically updated as the original changes). Finally, the boundaries created by graspable interface systems have the potential to exist persistently because they are realised through dedicated technology set-ups.

Property	MetaDesk	Illuminating Light	Passage
Visibility	1280x1024 res. display of 2D map 25cm, 640x480 res. LCD TFT panel shows 3D view of space 12cm diameter circle of fiber-optic cluster material, displays aerial photograph of the space	Graphics (simulated path of laser light and ancillary numerical and graphical information)	Monitor res. (view of virtual data space)
Audibility	None	None	None
Solidity	Non-solid - manipulation of virtual space	Non-solid – manipulation of virtual space	Non-solid – manipulation of virtual space
Location	2 segments nearly horizontal, 1 segment vertical	Horizontal (desk surface)	Vertical (monitor)
Mobility	Navigation of virtual space	None	Navigation of virtual space
Segmentation	3 segments (desk surface, passive lens, active lens)	Single segment	Multiple segments (each bridge/screen pair is a boundary to the virtual data space)
Lifetime	Potentially persistent	Potentially persistent	Potentially persistent
Configurability	None	None	None
Temporal location	Past	Present	Present
Symmetry	Asymmetrical	Asymmetrical	Asymmetrical
Representation	Implicit	Implicit	Implicit

Table 1.4 Graspable Interface systems in terms of boundary properties

Roomware

Table 1.5 summarises the physical side properties of the boundaries created by the systems HoloWall, DynaWall, InteracTable, CommChairs, Smart Floor and Door Badges.

Most roomware components are designed to provide visual information from a digital space into a physical environment. Systems such as DynaWall and Interact Table provide a view onto large information spaces with high resolution, whereas others like Door Badges utilise small displays in order to provide key contextual information. The boundary created by the Smart Floor component is different in that the information flows from the physical to the virtual space (the boundary senses the weight of physical participants/objects). However the boundary only has a representation in the physical space (in the form of a floor).

The physical location of the roomware boundaries is their most distinguishing feature. In the physical space these boundaries are always integrated with the architecture or furniture elements. The most common orientations are vertical and horizontal. It is common for roomware boundaries to be non-solid at the physical side, allowing the manipulation of virtual objects and to be mobile through the virtual space, allowing participants to navigate.

None of the example systems are segmented, although it is possible to use a combination of roomware components (such as CommChairs and DynaWall) to access the same digital space. In this case a single segmented boundary is established. In terms of lifetime, roomware components are very suitable for providing a persistent connection to a virtual space.

Finally, two of the example systems create dynamically configurable boundaries. The CommChairs were designed specifically to be mobile, which means that their boundaries can be easily relocated. The Door Badges system on the other hand offers very simple visibility configurability. Users can switch between static information on the room's usual occupants and current location data for the usual occupants.

Property	HoloWall	DynaWall	InteracTable	CommChairs	Smart Floor	Door Badges
Visibility	Rear-projected view Projector res.	Rear-projected 3072x768 res.	Rear-projected 1024x768 res.	Laptop res.	None	Small LCD display
Audibility	None	None	None	Possible	None	None
Solidity	Manipulation of virtual objects	Manipulation of virtual objects	Manipulation of virtual objects	Manipulation of virtual objects	Non solid from virtual side (senses *physical* space)	None
Location	Vertical (wall)	Vertical (wall) 4.5m by 1.1m	Horizontal (table)	Tilted (laptop display)	Horizontal location in physical space (floor)	Vertical (door sign)
Mobility	Navigation through virtual space	Navigation through virtual space	Navigation through virtual space	Navigation through virtual space	None	None
Segmentation	None	None	None	None	None	None
Lifetime	Potentially persistent	Potentially persistent	Potentially persistent	Variable	Potentially persistent	Potentially persistent
Configurability	None	None	None	Configurable location	None	Configurable visibility
Temporal location	Present	Present	Present	Present	Present	Present
Symmetry	Asymmetrical	Asymmetrical	Asymmetrical	Asymmetrical	Asymmetrical	Asymmetrical
Representation	Implicit	Implicit	Implicit	Implicit	Implicit	Implicit

Table 1.5 Roomware components in terms of boundary properties

Property	Augmented Reality	Augmented Virtuality	Ambient Media	Graspable Interfaces	Roomware components	Mixed Reality Boundaries
Visibility	At physical side only Often stereoscopic views and viewing perspective	Always at virtual side Sometimes also at physical side	At physical side only Medium transformations	At physical side only	Usually at physical side Often high res.	At physical side and at virtual side
Audibility	Usually none	Two-way for avatar boundaries Usually none for real-world views	At physical side only Medium transformations	Usually none	Usually none	At physical side and at virtual side
Solidity	Manipulation of virtual objects	Solid virtual side Sometimes non-solid physical side	Solid	Manipulation of virtual objects	Manipulation of virtual objects	Can be non-solid from both sides
Location	Display surface at physical side	Usually vertical (both sides)	Physical fixture or architecture	Usually horizontal at physical side	Integrated with physical architecture/furniture	Various, usually in form of vertical screen
Mobility	Usually mobile through virtual space	Mobile virtual side for avatar boundaries	Static	Often mobile through virtual space	Usually navigation though virtual space	Usually static
Segmentation	None	Usually none	Usually none	Sometimes segmented	Usually none	Usually none
Lifetime	Usually minutes/hours	Usually minutes/hours	Persistent	Potentially persistent	Persistent	Variable, can be persistent
Configurability	Often physical location	Sometimes visibility	Often none	Usually none	Sometimes (various)	Sometimes (various)
Temporal location	Present	Usually present	Usually present	Present or past	Usually present	Usually present
Symmetry	Asymmetrical	Asymmetrical	Asymmetrical	Asymmetrical	Asymmetrical	Fairly symmetrical
Representation	Usually implicit	Usually implicit	Usually implicit	Usually implicit	Usually implicit	Some explicit

Table 1.6 Mixed Reality approaches in terms of boundary properties

Discussion

Now that we have discussed each of the mixed reality approaches in terms of the property framework, we can summarise the general properties of the boundaries created by these approaches. Table 1.6 provides an overview of augmented reality, augmented virtuality, ambient media, graspable interfaces, roomware components and mixed reality boundaries.

As Table 1.6 shows the property set allows us to capture the main features of the different mixed reality technologies. For example the transformations performed on the visibility and audibility of a virtual space by ambient display boundaries or the location of roomware components within the physical architecture and furniture. Furthermore, as illustrated by Tables 1.1-1.5, the property set allows us to systematically describe the details of specific mixed reality systems. For example, with what resolution and field of view is a virtual space displayed and are there any perceptible delays due to network communication and processing.

It is argued that this section has shown that the boundary property framework can adequately classify different mixed reality approaches. Specifically it achieves the following goals:

- It gives a common foundation to the field by treating all mixed reality approaches in terms of boundaries.
- It allows the different mixed reality approaches to be compared and contrasted in terms of the property configuration of their boundaries.
- It allows the systems of a particular approach to be compared and contrasted in terms of the property configuration of their boundaries.

However, the framework is not perfect. It is not always immediately clear how to describe a particular mixed reality system in terms of a specific configuration of the property set and sometimes more than one interpretation may be possible. It is suggested, though, that overall the property framework is valuable. It encourages us to think systematically about the functionality provided by different mixed reality technologies and it is a first step towards the conceptual unification of the various approaches in the field of mixed reality.

Working with SHAPE Boundaries

As we have shown, mixed reality boundary properties can be used to classify systems involving the combination of physical and virtual spaces. The first mixed reality boundaries were simple windows between the two so that the occupants of each could see into the other (those in the physical would see a projection of the virtual and those in the virtual would see an embedded live video image of the physical). Later boundaries refined this through the idea of boundaries having the different properties – permeability, situation, dynamics, symmetry and representation (Benford et al. 1998) – leading to new kinds of boundaries such as traversable interfaces that create the illusion of participants crossing between virtual and physical (Koleva et al. 2000).

In a previous SHAPE deliverable (Deliverable 4.1), we outlined the concept of fragmented boundaries. Rather than solely concatenating single physical and virtual spaces, several different fragments of a mixed reality boundary could be used to connect a physical space to a virtual space at several locations, to create the illusion that the two spaces are overlaid on one another other. Demonstrators of these boundaries were intended to differ conceptually from our initial boundary designs along dimensions of boundary properties.

Notably, our we have designed several of our SHAPE demonstrators this year to concentrate on what we have termed ‘situational’ properties of boundaries. Our review shows that these properties are less explored in earlier work. Fragmented boundaries directly address situational properties of boundaries, including the ‘mobility’ and the ‘segmentation’ of boundaries (Koleva et al. 1999). The following section describes the motivation for investigating these ‘situational’ properties of boundaries in SHAPE. It then proceeds to highlight the prototype SHAPE demonstrators reported in this deliverable which investigate these properties.

Investigating situational properties of mixed reality boundaries

Situational properties of boundaries include “the location of the boundary, whether this location is fixed, and whether the boundary is segmented” (Koleva et al. 1999). Our realisations of mixed reality boundaries prior to the SHAPE project have concentrated on fixed, monolithic boundaries. However, there are interesting reasons why we might wish to begin varying these properties within the context of SHAPE technologies, specifically for public spaces such as museums and galleries. ‘Situational’ elements of the boundary topology have very strong relationships with the design sensitivities that SHAPE studies of interaction begin to highlight (SHAPE Deliverable 2.1). In particular, the properties of *mobility* and *segmentation* have clear conceptual parallels with the kinds of interaction work that visitors accomplish around and with exhibited artefacts. For each of these properties, a motivating relationship between studies and design is outlined below.

Mobility

The concept of static boundaries in some ways mirrors traditional notions of the design of public spaces. Exhibition designers tend to demarcate and order spaces with the intention of demarcating and ordering visitor behaviour. However, studies reported as part of SHAPE Deliverable 2.1 have begun to delineate the extent to which this ordering directs actual instances of visitor behaviour. For example, consider the discussion of co-participation in Chapter 2 of that deliverable. What we see is visitors themselves demarcating exhibit encounters in situated interaction with one other. The order of museum space is massively and intricately a *situational* order, informed and affected by the movements of visitors. We should consider the demarcation of exhibits as practices to be supported and designed for. This leads us to believe that our technical designs need to take seriously the situational property of mobility. We therefore have decided that our technical designs should explore the presentation of non-static boundaries to address this core concern. Our aim in producing many of the technology demonstrators that are reported in this deliverable has been to mobilise boundaries in public places.

Segmentation

Mixed reality boundaries in early forms presented singular views of a concatenated virtual environment. Our boundaries have tended to allow monolithic hybrid environments; monolithic in the sense that the meeting of the virtual and real is specific and sedate.

However, contrary to a commonly held view of museum interaction, our studies reported in Deliverable 2.1 show it to consist of far more than simple, passive, monolithic observation. Visitors often co-ordinate their actions amongst themselves and with others in the locality. For example, consider the experience of the *Ghost Ship* piece reported in companion Deliverable 2.1 (Chapter 4). Co-ordinating the discovery of exhibits and exhibitions as assemblies-in-interaction is key to an experience of the exhibit. Thus, our designs should reflect the potential of discovering multiple aspects and perspectives of exhibits commonly found in the 'work' of visitors. This design sensitivity has led us to consider the segmentation of boundaries as a way of providing different views and perspectives on exhibits and public information. In particular, what is the benefit of designing for the 'work' of discovery? Can we, indeed how should we, design to provide multiple perspectives on public technologies? This deliverable reports on our first attempts to design for the situational property of boundary segmentation. In this way, we hope to provide a resource for visitors to, both alone and together, encounter and make sense of the multiple perspectives - the 'ecologies of participation' (Deliverable 2.1, Chapter 4).

Mixed reality boundary situational properties in SHAPE demonstrators

Over the course of the first year of SHAPE, we have constructed a number of technologies and demonstrators which have contributed to investigations of the properties of boundary mobility and boundary segmentation in interaction. The following brief summaries relate the subsequent chapters of this deliverable to this enterprise. This will highlight the complementary ways in which we have approached the perspectives of mobility and segmentation of hybrid boundaries. The next section then outlines some limitations of the boundary approach described thus far with respect to physical-digital objects, and mentions directions in which our typological enterprise might be taken to remedy these limitations.

Chapter 2 - Unearthing Virtual History

This chapter describes an application in which museum visitors hunt for virtual history outdoors, capture it, and bring it back indoors for detailed inspection. This application provides visitors with access to a parallel virtual world as they move through an extended physical space. A mobile, wireless boundary device is used for locating 'fragments' of virtual activity outdoors. Radically different experiences of the virtual are provided, depending upon location, task, and available equipment. The experience of fragments is segmented therefore across spaces and times of discovery.

Chapter 3 – Developing Hyperphysical Links

This chapter explores the notion of 'hyperphysical links' connecting real world physical objects with their informational counterparts. In particular, a wearable display technology (Glasstrons) is investigated to scope the practical and technical problems involved in using such devices in a museum setting. The Glasstrons have the mobility of a personal device and address boundary segmentation using semi-transparent display technology to simultaneously reveal physical and digital perspectives on an object. Various problems for individual-wearable solutions for mobile, segmented boundaries are raised as a result of this work.

Chapter 4 - Augurscope

The Augurscope is a mobile mixed reality boundary for outdoors. A tripod-mounted display is wheeled to different locations and rotated and tilted to view a virtual environment that is aligned with the physical background. Video from an onboard camera is embedded into this virtual environment, and shown on a secondary display, segmented from the mobile device to provide the logical converse view of the boundary. Analysis of use reveals problems with the movement of the device and relating virtual and physical viewpoints, and shows how environmental factors and physical form affect interaction with and around the boundary.

Enhancing interaction and permeability at mixed reality boundaries

Just as the situational properties of mixed reality boundaries have been under-investigated, thereby setting SHAPE a prioritised research agenda, so have only a restricted set of interaction styles been investigated in mixed reality research. An examination of Tables 1.1-1.5 reveals that very rarely do systems symmetrically investigate the visibility and audibility properties of boundaries. Concomitantly, graphical interaction and display on the one hand, and auditory interaction and display on the other, tend to be treated in a separate, non-integrated fashion. Another major theme of our work in SHAPE has been to enhance interaction at mixed reality boundaries to bring together graphical and sonic interaction. In particular, we have examined various uses for sound in making information available about the real world embodied activity of participants on the one hand, and sonifying the properties of informational (virtual) objects on the other. These are reciprocal investigations of how to enhance, in the terms of the framework we have been developing, the audible permeability of a boundary.

Chapter 6 - Publicly Deploying ToneTable as a Multi-Participatory, Mixed Media Installation

This chapter describes ToneTable, an installation we have developed in SHAPE to explore different strategies for associating sound and graphical elements. While ToneTable itself can be thought of as a kind of ‘roomware component’, the strategies for sound-image-interaction linkage can have application in mixed reality beyond just this particular kind of system. For example, we describe a number of strategies for sonifying the activity of participants at the table. These strategies could be employed whenever participant-activity needs to be made audible across a mixed reality boundary. We also discuss how virtual objects in a graphical display can be sonified with their movement being rendered on a multi-speaker array. Again, these strategies could be employed whenever features of objects need to be spatialised in a real world physical environment. In addition, Chapter 6 explores a method for combining the interactional influence of a number of participants in a combined collaborative outcome, thereby offering some specific mechanisms for the support of collaboration in response to emerging work in Workpackage 2. (Further strategies for incorporating sound in mixed reality settings are to be found in Chapters 2 and 7.)

Integrating Objects with Boundary Technology and Typology

Our discussion thus far on boundaries has only considered how *participants* experience a boundary. However, some of our investigations in SHAPE suggest new possibilities and extensions to this approach. Specifically, devices such as The Invention Observatory (Chapter 7), the Storytent (Chapter 8) and the Unearthing Virtual History demonstrator (Chapter 2) involve *objects* being used with boundaries. For example, the Storytent can recognise RFID tagged objects that enter and leave it, and can trigger effects in the projected virtual world as a result. Similarly, The Invention Observatory involves the use of combinations of tagged objects in order to trigger digital effects (a virtual invention can be created). In the Unearthing demonstrator, fragments of virtual objects are collected outdoors and brought back to an indoor space for more inspection on a periscope. In these cases, an object is triggering effects or revealing information as a consequence of its relationship to a boundary. The objects could even be thought of as traversing the boundaries – disappearing from the physical space of (at least some of) the participants and entering a remote virtual space.

We can generalise this idea further. An object may be composed of many attributes. Some of these can be perceived and manipulated in the physical world. They have physical attributes. Some can be perceived and manipulated in the digital world. They have digital attributes. For example, an object from The Invention Observatory has physical attributes (the properties of the physical representation of the invention) and digital attributes (the properties of the invention itself).

In this context, a boundary is a device that establishes a connection between the physical and digital worlds. As a result it can reveal the digital attributes of an object to a person who is in the physical world or conversely can reveal the physical attributes of an object to a person who is in the digital world. This might naturally be achieved by pointing the boundary device at the object or placing the object near to on or in the boundary. Placing physical inventions on a plinth reveals sound and graphics effects. Bringing an object into the Storytent triggers a digital effect. Pointing the Augurscope at an object might reveal its properties in a connected space.

A further possibility is that on coming into contact with the boundary, the physical object's physical attributes become hidden and its virtual ones revealed. In this case the object appears to traverse the boundary – apparently leaving one space to enter the other. This is the effect implied by the Storytent, where objects entering the tent become part of the digital story.

Some points are worth noting about this proposal:

- It separates our typology into two kinds of device: objects and boundaries. Independent boundary devices are used to reveal the hidden attributes of object devices. However, there are other ways in which this could be done. The object itself can be extended to reveal its own attributes. For example, inventions could play their own sounds without the need for any external device such as the plinth. Both approaches are valid and can co-exist. In part, choosing between them will depend upon the practical feasibility of extending objects. Is it technically possible and economically feasible to add displays to many (possibly small objects) or does it make sense to incorporate the displays into (presumably fewer) independent devices that can then be related to potentially many different objects?
- The concept of mixed reality boundaries makes an explicit and deliberate assumption that there are digital spaces that are somehow separate from physical spaces and that consequently, the two need connecting is some way. This, of course, is largely a matter of metaphor – the same metaphor that underpins the whole of virtual reality – that

there is a separate virtual world that you can somehow enter, experience and share with others. There are other views. Whether it is useful to think of there being separate digital spaces and properties that are connected to physical spaces and properties by mixed reality boundaries may depend upon a number of factors. How important is it for an application to create the illusion of a separate virtual world? For some, such as historical simulations, it may be important. To what extent does the application involve many participants at different locations, using different interfaces navigating and sharing a common information space? If it does, such as in the case of assemblies of devices, then it may be useful to think of them as sharing a common digital space. However, there may be many cases where the metaphor is not appropriate.

Conclusions

We believe that the SHAPE approach to the use and application of a boundary typology has broader implications than simply providing yet another metric against which to judge quality or correctness. The design sensitivities provided within the project (reported in both this deliverable and in companion Deliverable 2.1) allow us to develop a taxonomic approach which can be applied to our technical design. Here, our use of the word 'applied' is meant to indicate the sense that the typology is used to inform design, but also the sense that our use of typology is sensitive to the pragmatics of technology deployment.

There are interesting technical design paths suggested by our experiences with the typology we have employed. In Unearthing Virtual History (Chapter 2), we have extended experiences across both space and time. Participants are expected to discover artefacts across a relatively large area, and experience events related to these artefacts in the 'same' virtual place at different times and in different physical places (outside when searching; and then inside when using the periscope). We have explicitly designed according to features of our typology, to make these boundaries as coherent and seamless as possible for participants. Our strategies have included:

- Creating a virtual space which visually imitates a physical space on which it is overlaid.
- Using location-aware technologies such as GPS to match the virtual and physical boundaries of tracked objects as closely as accuracy allows.
- Using the same virtual locations at different times and in different physical places to indicate the correspondence of parts of an experience.

Despite our design strategies, however, it is still noticeable that participants' interaction makes the sense of the experience. Our search parties, and indeed ourselves as designers, interact to weave a sense of the experience. In some cases, participants succeeded in making coherent sense of the world, in others things make less sense.

For example, of particular concern for our Augurscope demonstrator (Chapter 4) has been the problematic nature of designing a coherent experience. The Augurscope design has benefited from consideration of many of the features of our boundary typology - for example, the mobility of the device itself, and the segmentation of a secondary display. Nonetheless, it is still noticeable in analysing participants' use of the Augurscope that these design features are not sets of constraints. Rather, they provide a set of resources from which individuals, groups of visitors, strangers, friends and families can begin to discuss, make sense of, and weave an experience for one another. Coherent hybrid technologies in public settings seems to benefit and be guided by our use of a boundary approach, without being overly constrained by it.

Nonetheless, we also keep in mind that participants' sense-making efforts in these experiences are not necessarily something to be suppressed. Indeed, we may actually wish participants to have difficulties in making sense of situations in order to provide a slowly unfolding narrative. For example, we might like our demonstrators to allow participants to experience the mystery and fascination that an archaeologist experiences when they move from discovering an object to making sense of its origin and use.

Therefore, within the context of research over the course of the next year, it seems that we should further explore the relationships between boundaries and objects. In particular, our aim is to explore the situational properties of mobility and segmentation for objects and participants *in particular situations*. This will allow us to further explore the properties of our typology as they occur in public interaction.

Further Contributions

Two further chapters of this deliverable require explicit note, as they complete the reporting of our work so far in Workpackage 1.

Chapter 5 - Wasa: Towards a Set of Technologies for Producing Public Mixed-Reality Learning Environments

Prior to the start of the SHAPE project, consortium partners had tended to use standardly available computer graphics libraries and techniques in their work, much of which had a history in research on virtual reality and collaborative virtual environments (CVEs). These graphical techniques tend to focus on giving a first person view (associated with a particular avatar) with conventional perspective rendering techniques (e.g. a finite, small number of vanishing points, or orthographic projection). In addition, in the name of efficiency, several of the standard renderers available in the CVEs the consortium had used or developed, employed quite simple rendering techniques which did not always make for a rich visual experience.

It is questionable whether this approach is appropriate for investigating mixed reality settings for public deployment. For example, one might want to experiment with projection onto unusual surfaces or indeed surfaces which might change their character at run-time. Perspective, level of detail, and rendering model may all need to change dynamically. We are beginning, therefore, to identify the requirements for computer graphical techniques for mixed reality settings. As an experimental vehicle for this, Chapter 5 of this deliverable reports on Wasa, a set of computer graphical techniques for mixed reality applications. Amongst other things, Wasa enables exploration of recent rendering techniques not commonly seen in CVE systems, for example, particle systems, environment maps, variable lighting models. Recently, we have begun trials integrating Wasa with other infrastructural systems in use in SHAPE to enable our varied display and input devices to show and interact with high quality graphics.

Chapter 7 – The Invention Observatory

We have already mentioned how this demonstrator has enabled us to explore objects which, when brought to a suitable device (e.g. a plinth), make their associated digital properties available in an informational or virtual environment. As such, The Invention Observatory gives an elementary demonstration of what we mean by hybrid objects in relationship to boundary devices such as plinths. The Invention Observatory also gave us the opportunity to combine a number of devices into a simple technical ‘assembly’. In this demonstrator, the plinth, the Augurscope and a large projected display systematically interwork. Furthermore, we extend our techniques for participant gesture sonification in giving an auditory rendering of activity at the plinth. As such, Chapter 7 reports on what is for us a ‘summative’ demonstrator of the work in this deliverable while giving a forwards glance to the concern with ‘assemblies’ of mixed reality artefacts that will occupy us later in Workpackage 1.

References

Barrus, J.W., Waters, R.C. and Anderson, D.B., “Locales: Supporting Large Multiuser Virtual Environments”, *IEEE Computer Graphics and Applications*, 16 (6), November 1996, pp. 50-57.

Benford, S., Greenhalgh, C., Reynard, G., Brown, C., and Koleva, B., Understanding and Constructing Shared Spaces with Mixed Reality Boundaries, *ACM Transactions on CHI (ToCHI)*, 5 (3), pp. 185-223, ACM Press, 1998.

Koleva, B., Benford, S. and Greenhalgh, C., The Properties of Mixed Reality Boundaries, in *Proc. ECSCW’99*, pp. 119-137, Copenhagen, Denmark, 1999.

Koleva, B., Schnädelbach, H., Benford, S. and Greenhalgh, C., Traversable Interfaces Between Real and Virtual Worlds, in *Proc. CHI 2000*, The Hague, April 2000, ACM Press.

Chapter 2:
Unearthing Virtual History: Using diverse interfaces to reveal hidden virtual worlds

Steve Benford[1], John Bowers[2], Paul Chandler[1], Luigina Ciolfi[3], Martin Flintham[1], Mike Fraser[1], Chris Greenhalgh[1], Tony Hall[3], Sten-Olof Hellström[2], Shahram Izadi[1], Holger Schn delbach[1], Ian Taylor[1]

[1] The Mixed Reality Laboratory, University of Nottingham, UK
{sdb, pdc, mdf, mcf, cmg, sxi, tar, hms, imt} @cs.nott.ac.uk

[2] Centre for User-Oriented IT-Design (CID), Royal Institute of Technology (KTH), Stockholm, Sweden
{bowers, soh} @nada.kth.se

[3] Interaction Design Centre, University of Limerick, Ireland
{Luigina.Ciolfi, Tony.Hall} @ul.ie

We describe an application in which museum visitors hunt for virtual history outdoors, capture it, and bring it back indoors for detailed inspection. This application provides visitors with ubiquitous access to a parallel virtual world as they move through an extended physical space. Diverse devices, including mobile wireless interfaces for locating hotspots of virtual activity outdoors, provide radically different experiences of the virtual depending upon location, task, and available equipment. Initial reflections suggest that the physical design of such devices needs careful attention so as to encourage an appropriate style of use. We also consider the extension of our experience to support enacted scenes. Finally, we discuss potential benefits of using diverse devices to make a shared underlying virtual world ubiquitously available throughout physical space.

Introduction

Museums, galleries, cultural heritage and tourism are promising application domains for ubiquitous technologies. Personal and handheld devices coupled with embedded and projected displays can enrich experience inside a traditional museum or gallery (Aoki and Woodruff

2000, Benelli et al. 1999, Oppermann and Specht 1998). Mobile technologies can enhance cultural experiences when exploring a surrounding city (Cheverst et al. 2000). We are interested in how a combination of the two might provide visitors with rich and engaging cultural experiences that connect a conventional museum or gallery to a surrounding city or site of special interest. Our approach involves providing participants with diverse interfaces for detecting, revealing and experiencing events that are taking place in a parallel 3D virtual world; that is a virtual world that is hidden behind, but potentially ubiquitously available from, everyday physical space in both indoors and outdoors locations.

A First Demonstration: Unearthing virtual artifacts on the Nottingham campus

We have created a museum experience where participants explore an outdoors location, hunting for buried virtual artifacts that they then bring back to a museum for more detailed study. Inspired by the results of previous museum projects, our intention is that the process of actively searching for history will be engaging for visitors and will also lead them to critically reflect on the information that they discover (Rayward and Twidale 1999). At the beginning of the experience participants are told the following (fictional) back-story:

During the construction of our campus in 1999 builders unearthed four ancient artifacts: a samurai sword, a maiolica dish, an ivory box of dominoes, and a bell (objects from the collection at the nearby Nottingham Castle museum). Scientists have since discovered that physical artifacts radiate traces of their history. When they are buried for long periods of time these traces can leak into the surrounding earth, and can subsequently be detected and captured using specialized sensing instruments. Unfortunately, our builders failed to note the locations where the objects were unearthed.

Part 1 (outdoors): locating the target objects and capturing their history

Groups of participants head outside and search an island on our campus using a “virtual history meter”, a device that informs them of their proximity to various virtual objects that actually exist in a 3D virtual model of the campus. This device consists of a laptop and a Compaq iPAQ that communicate with one another and also with computers in the nearby laboratory over a WaveLAN network. The global position of the laptop as given by an attached GPS device is transmitted back to computers in the laboratory, enabling them update the position of an avatar in the virtual world that represents the search party. In turn, the computers running the virtual world update the mobile laptop with measures of this avatars’ proximity to different fragments of the virtual objects. For brevity, our initial demonstration is limited to three fragments from one object, the maiolica dish.

It can be difficult to create a satisfactory and reliable visual overlay of a virtual world on an outdoors scene due to a combination of limited tracking accuracy and variable lighting conditions (Azuma 1999). We have therefore opted for an alternative approach that is primarily based upon audio information. The proximities of the search party to the three fragments are sent to a computer in the laboratory that is running an application that sonifies the party’s location in relation to the fragments. Each fragment has a different pulsing synthesised tone associated with it that increases in amplitude and pulse rate as the search party get closer to it. The mix of the three tones is transmitted to the wireless laptop. When our participants are within a (configurable) distance of a fragment, they are deemed to have acquired it. They now hear a different tone and the iPAQ device displays an appropriate image and some accompanying text.

Figure 2.1 shows two participants searching our campus next to the corresponding image of their avatar in the virtual world as it encounters one of the fragments.

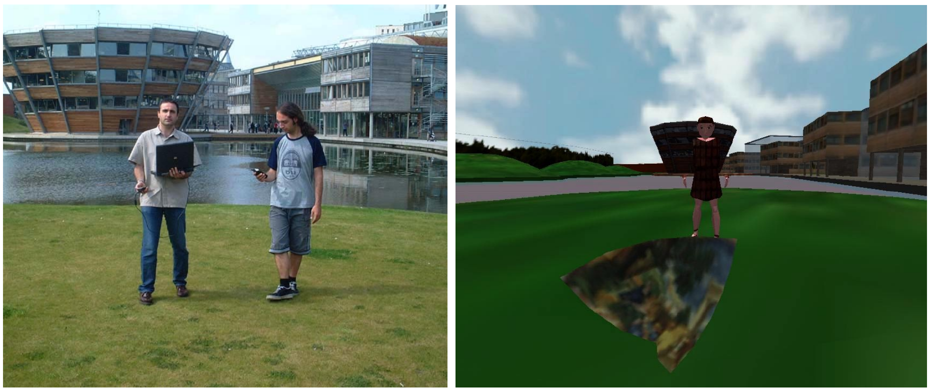

Figure 2.1. Hunting for fragments of virtual objects outdoors

Part 2 (inside the museum): Viewing the captured history

The search party brings the captured virtual history back to the museum in order to view it in detail. Each captured fragment is loaded onto a periscope, a rotating ceiling-mounted screen that allows a user to view and hear a virtual world by turning about a fixed virtual location. Grasping and rotating the periscope rotates one's viewpoint in the virtual world and controls the mix and spatialisation of associated audio (both ambient sounds and commentary) that is heard through wireless headphones. A small projector mounted on the base of the periscope projects an additional view onto surrounding screens that is supplemented with four external audio speakers. This public display is intended to allow the experience to be shared by groups of visitors (e.g., families) and to attract other visitors to the exhibit.

The periscope user finds that they have been transported back to the island, but this time as a 3D model, and that a historical scene from the object's past has now appeared at the location where they found the fragment. By rotating the periscope, they can explore the scene, trigger spoken information and mix related sounds. Inspired by the presentation of the actual dish at the Castle Museum, one scene depicts the event that is painted onto the dish, a second tells of how and where the dish was made, and a third tells so of how it came to be in Nottingham. Figure 2.2 shows the periscope and an example scene.

Implementation

Our demonstration has been implemented through the coordinated use of the MASSIVE-3 and EQUIP software platforms and applications authored in the MAX/msp audio programming environment (www.cycling74.com). MASSIVE-3 is a platform for distributed virtual worlds. It can support between ten and twenty mutually aware avatars in a shared virtual world communicating using real-time audio and has previously been used to create a variety of on-line storytelling events (Craven et al. 2001). EQUIP is a dynamically extensible framework for integrating C++/Java based applications with a variety of interfaces and devices, ranging from wireless portable devices through to fully immersive systems. EQUIP provides applications and devices with one or more shared tuple-like data spaces through which they can publish and subscribe to each others' data. For example, our virtual history meter publishes its GPS updates to an EQUIP data space so that they can then be read into a MASSIVE-3 virtual world. In return, MASSIVE-3 publishes the positions of fragments in the world that can then be read by the meter. An EQUIP module was also developed which (in part 1) published the GPS data as MIDI continuous controllers to communicate with the sonification and (in part 2) sent periscope angle to a panoramic mixing application which triggered commentary sound files, generated ambient textures algorithmically, and controlled the overall mix and spatialisation. The operation of EQUIP, its integration with MASSIVE-3 and its role in supporting this demonstration are not covered here (but see Greenhalgh et al., forthcoming). Equip is described in further detail at www.mrl.nott.ac.uk/~cmg/Equator/.

Figure 2.2. Viewing the captured fragments indoors

Immediate reflections

Our initial demonstration was given to a small audience of invited participants, including staff from Nottingham's Castle Museum. A round-table discussion suggested a number of possible refinements. For this paper we focus on those comments that relate to the design of the virtual history meter.

- It was possible, but quite slow and sometimes difficult to locate the target fragments outdoors. The island presents a featureless landscape and the virtual fragments were not sited at obvious locations. Users probably needed to be more systematic and painstaking about searching than they were.
- Several participants commented that they expected the handheld devices – the iPAQ and GPS – to be sensitive to orientation. In other words, that the sonification would change according to the direction in which these devices were being pointed.

Of course, we might directly address these comments by changing the locations, spacing and target sizes of the fragments and adding a directional compass to the wireless device. However, it is also interesting to speculate whether a different physical design for the wireless device would have encouraged a more appropriate style of use to suit the original set-up. What if the GPS sensor had been embedded into something resembling a metal meter? Would this have encouraged more systematic searching over the ground? Would participants expect a larger ground-hugging device such as a metal meter to be as sensitive to orientation as a hand held pointing device? In future experiences we should pay greater attention to the physical design of devices in order to ensure that they communicate their intended style of use rather than just relying on off-the-shelf devices.

Extending with Enacted Scenes

In order to be able to create richer historical experiences we have extended our system to support enacted scenes. Users, represented as avatars enter the virtual world, move around, manipulate objects and talk to one another. MASSIVE-3 allows such scenes to be enacted live or to be saved as 3D recordings that can subsequently be replayed in a live virtual world (Greenhalgh et al. 2000). In this way, actors can stage scenes from the past and tour guides, curators and teachers can quickly create customized virtual tours that to be played out into physical space.

Two further technical innovations support these ideas. First, we have created a version of the virtual history meter that tracks avatars as they move around the virtual world (showing their positions on a radar style display on the iPAQ) and that allows users to listen in to their dialogues (via the laptop). Second, we have experimented with a technique for projecting “shadow avatars” into physical environments so as to give fleeting impressions of ghostlike figures from a parallel world. When an avatar passes by a specific location in the virtual world, its shadowy image (with associated sound) is projected onto the wall or floor of the equivalent physical location. This shadow technique demonstrates a further class of device for revealing the virtual world, one in which users do not have to carry any specialised equipment at all or even have any intention to experience the world. Such techniques could be used to attract the attention of bystanders so as to draw them into virtual events.

Diverse Interfaces onto a Ubiquitous Virtual World

We finish with some general reflections on the approach of using diverse devices to access a shared virtual world. Conventional augmented reality employs physical or video see-through displays to overlay a virtual world on the physical (Azuma 1997). Recent projects have begun to move augmented reality outdoors, for example exploiting handheld displays and wearable computers with see-through head-mounted displays (H llerer et al. 1999, Azuma 1999). Other researchers have also explored the use of hand-held devices to interact with immersive virtual environments (Krebs et al. 2000, Watsen et al. 1999).

Our approach focuses on how very diverse devices can provide radically different experiences of a virtual world at different times and in different places. Some devices will offer high fidelity and accurately registered views of the virtual along the lines described above. However, others will offer more impressionistic views of the virtual, for example audio sonifications as demonstrated by our virtual history meter or projected shadow avatars. We propose that our approach offers a number of benefits:

- **Variable engagement** – heterogenous interfaces can allow participants to vary their engagement with the virtual world. An unfolding story may gradually introduce participants to a virtual world. Bystanders have only a fleeting awareness of virtual events, whereas committed players may be fully involved. Participants in a long-term event may vary their level of engagement over time.

- **Variable tracking –** the display of the virtual can be configured to match the accuracy of tracking in different locations. Where accurate tracking is available the user may be offered a fully 3D view of the virtual. Where it is not, they may be offered more impressionistic views.
- **Variable physical environments –** sound-based representations of the virtual may be able to accommodate bright and variable lighting conditions where it might be problematic to project detailed graphical views.
- **Orchestration** – staff in the virtual world can monitor and dynamically shape participants' experiences from behind-the-scenes, for example moving virtual objects to make them easier or harder to find. In fact, our demonstration supported an additional interface, a table-top projection of the virtual world as an interactive map, for this purpose.
- **Systems –** Finally, connecting multiple wireless physical devices to a common underlying virtual world brings advantages from a systems perspective. VR research has developed a repertoire of techniques that use virtual space to manage information flows among large numbers of communicating users (Singhal and Zyda 1999). These techniques can be directly applied to mobile devices that are tracked and represented in a virtual world. For example, a group of PDAs that are proximate in the virtual world (i.e., whose virtual "auras" have collided or who are in a common virtual "locale") would automatically join the same server or multicast group and so communicate with one another.

References

Aoki, M. and Woodruff, A., Improving Electronic Guidebook Interfaces Using a Task-Oriented Design Approach, Proc. DIS 2000, Aug 2000, 319-325.

Azuma, R. T., A Survey of Augmented Reality, Presence: Teleoperators and Virtual Environments, 6(4): 355-385, Aug. 1997.

Azuma, R., The Challenge of Making Augmented Reality Work Outdoors, In Mixed Reality: Merging Real and Virtual Worlds (Yuichi Ohta and Hideyuki Tamura, eds), 1999, Springer-Varlag.

Benelli, G., Bianchi, A., Marti, P., Not, E., Sennati, D. (1999a). HIPS: Hyper-Interaction within Physical Space, in Proceedings of IEEE ICMCS99, Florence, June 1999.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences, Proc. CHI'2000, 17-24, The Hague, Netherlands, ACM Press.

Craven, M., Taylor, I., Drozd, A., Purbrick, J., Greenhalgh, C., Benford, S., Fraser, M., Bowers, J., Jää-Åro, K., Lintermann, B, Hoch, M., Exploiting Interactivity, Influence, Space and Time to Explore Non-linear Drama in Virtual Worlds, *Proc. CHI'2001*, Seattle, US, April 2-6, 2001, ACM Press.

Greenhalgh, C., Purbrick, J., Benford, S., Craven, M., Drozd, A. and Taylor, I., Temporal links: recording and replaying virtual environments, *Proc. ACM Multimedia 2000*, L.A.

October 2000.

Greenhalgh, C., Izadi, S., Rodden, T. and Benford, S., The EQUIP platform: bringing together physical and virtual worlds (forthcoming).

H llerer, T., Feiner, S., Terauchi, T., Rashid, G., Hallaway, D., Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System , In: Computers and Graphics, 23(6), Elsevier Publishers, Dec. 1999, pp. 779-785

Krebs, A., Dorohonceanu, B. and Marsic, I., Collaboration using Heterogeneous Devices – from 3D Workstations to PDAs, In Proceedings of the IASTED International Conference on Internet and Multimedia Systems and Applications (IMSA'2000), pages 309-313, Las Vegas, NV, November 20-23, 2000.

Oppermann, R. and Specht, M, Adaptive Support for a Mobile Museum Guide, Proc. Workshop on Interactive Applications of Mobile Computing (IMC'98).

Rayward, W. B. and Twidale, M. B., From Docent to Cyberdocent: Education and Guidance in the Virtual Museum, Archives and Museum Informatics, 13, 23-53, 1999.

Singhal, S. and Zyda, M., Networked Virtual Environments – Design and Implementation, ch. 7, SIGGRAPH Series, ISBN 0-201-32557-8, July 1999, ACM Press.

Watsen, K., Darken, R., and Capps, M., A Handheld Computer as an Interaction Device to a Virtual Environment, 3rd International Immersive Projection Technology Workshop (IPTW'99), Stuttgart, Germany, 1999.

Chapter 3: Developing Hyperphysical Links: An initial museum-based technology exploration

Tony Hall, Luigina Ciolfi, Liam Bannon
Interaction Design Centre, University of Limerick, Ireland.
{Tony.Hall, Luigina.Ciolfi, Liam.Bannon}@ul.ie

One of the aims of SHAPE Research Challenge 1 is to explore new possibilities for linking physical artefacts and digital information, or hyperphysical links, which "extend the traditional idea of hypermedia links (that link different digital artefacts) to new kinds of link that link physical and digital artefacts" (SHAPE Annex 1, 2001). In this chapter we describe our experiences of creating and exploring hyperphysical links onsite in a museum and illustrate what hyperphysical links might be interesting for museums visitors. We used a particular technology, a Sony Glasstron PLM-S700E head mounted display (HMD), for our initial exploration. In this chapter we also document our findings from using this technology in the museum and discuss its possible utility in SHAPE.

Introduction

This chapter presents early SHAPE technical work, the purpose of which was to investigate new associations between physical and digital artefacts. 'Hyperphysical links' might be interesting or useful for visitors in the museum and thus we have investigated the possibility of using a Sony Glasstron PLM-S700E HMD as a means for visitors to access these new digital-physical links. Also, on the basis of our findings, we suggest a possible future scenario for visitor interaction with small assemblies of hybrid artefacts in the museum. However, before we document our experiences of creating hyperphysical links and users' experience of using the Glasstron HMD in the museum, we first briefly compare our work with other research projects, what they are trying to achieve and the approach they are taking in developing hyperphysical links.

Overview of Research in Hyperphysical Links

There are a number of other projects exploring hyperphysical links and their application in public settings. In this chapter, we focus briefly on two of these projects: *Whisper Space* and ARCHEOGUIDE. *Whisper Space* is exploring novel ways to link digital information and physical artefacts, specifically the use of voice recognition tools and wireless headphones as new means for associating museum artefacts and digital audio. The type of scenario that *Whisper Space* is trying to develop is: “The area around the painting [or artefact, exhibition, etc.] is equipped with speakers and the user has a wireless headset. As the whispers are heard, the user can participate by adding his own commentary [about the painting, etc.], or asking for more information about a comment from someone else.” (Hammond 2001) Researchers working on the ARCHEOGUIDE project have been exploring hyperphysical links at the Temple of Hera cultural heritage site at the Olympia, Greece. This project uses a Glasstron HMD, among other display devices like a pen-computer mobile unit. Wearing the HMD, and using head-movements, the user navigates a virtual reconstruction of the temple (now ruined), which they see overlaid on its ruins, containing virtual models of its original contents, statues, drapery etc. The aim of ARCHEOGUIDE is to enhance users’ experience of the ancient site with a superimposed virtual reconstruction, which enables users to appreciate somewhat what the site looked like in former times (Vlahakis 2001).

In the Hunt museum, we explored linking both aural and visual digital information with physical artefacts indoors, in the museum setting, using a Glasstron HMD to provide these links. We will presently describe the Glasstron technology in more detail but we will first outline our rationale for why we used this technology to prototype hyperphysical links in the museum. First, the Glasstron has a semi-transparent screen so one potential benefit was that the visitor would be able to see digital information beside or overlaid on actual artefacts, in the same visual field. We could therefore investigate how digital visuals (virtual models for example), might enhance visitors’ interaction with physical artefacts in the museum. The HMD is also equipped with headphones so there was also the possibility to supplement visual information about physical artefacts with digital audio. Another possible benefit of the Glasstron is its portability and because he would still see the actual museum and his environs, the visitor should also be able to walk around and look at different physical artefacts, while simultaneously receiving supplementary digital information. In summary, we intended (1) to explore hyperphysical links in the museum and (2) to investigate the possible usefulness of the Glasstron HMD in SHAPE. We presently describe our findings from the initial use of the HMD in the museum, but first we describe in more detail the Glasstron technology that we used in our initial technical exploratory work.

The Sony Glasstron PLM-S700E

The Glasstron PLM-S700E is an advanced visualisation technology including an LCD with over 1.5 million pixels per component – each display component is roughly the size of a small coin. When connected to a personal computer, the Glasstron enables users to see true SVGA (800_600) quality images from a PC or other video output device, e.g. S-VIDEO. The main features include: a 30-inch screen (up to SVGA 800x600) viewed as if from a distance of 1.2 meters when connected to a personal computer. In addition, the Glasstron contains built-in personal stereo headphones that deliver full stereo sound with Mega Bass and Automatic Volume Limiter System (AVLS). AVLS keeps the device volume constant, offsetting sudden peaks in sound. Also, a see-through mode allows users to view (or not) the surrounding environment by adjusting the transparency of the screen. The user can change the level of transparency to fade the physical environment or make the screen completely opaque so it just displays video. This see-through capability is perhaps the most interesting feature of the technology because one can potentially overlay physical artefacts and environs with digital information, animation, video et cetera. However, as we will presently describe, this see-through function can be very limited.

Figure 3.1a and 3.1b. The Sony Glasstron PLM-S700E (Sony Corporation) head mounted display and using the technology in the museum.

Interestingly, a number of other commercial interests are also pursuing see-through functionality. Microvision is currently building retinal scanning technology, the NOMAD device being an example, and is attempting to develop devices that will laser-scan to the retina to give the effect that the physical environs are overlaid with digital information. Microvision plans to test the application of their see-through technology for surgical work, where it might benefit doctors to see vital statistics synchronously and in the same visual field as their patients during surgery.

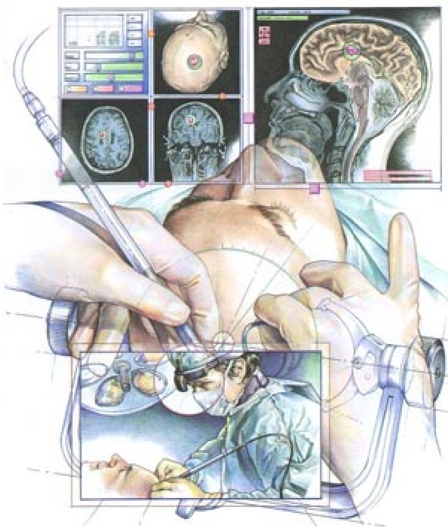

Figure 3.2. Using Microvision NOMAD retinal scanner for surgical work (Microvision Inc.)

In the museum, we used the Glasstron and specifically its variable see-through mode to create hyperphysical links and overlay and associate digital information (virtual models, digital audio) beside and with physical artefacts (antique and priceless museum pieces). We now present our findings from the initial use of the Glasstron HMD in the museum.

Using the Glasstron HMD in the Museum

At the expense of degradation in viewing the actual artefacts, visitors saw digital information, two virtual models, overlaid near the two museum artefacts – a horse statuette and ornamental dish. The first artefact visitors used the technology with is perhaps the centrepiece of the internationally renowned, eclectic collection in the Hunt Museum in Limerick. Housed in the beautiful Captain's Room of the museum, it is a bronze statuette of a rearing horse, purportedly after Leonardo Da Vinci. Although Da Vinci left no bronze sculptures after his death, it is argued that another artist cast the piece and other comparable pieces according to a model found at Da Vinci's studio. There are bronze horse artefacts, similar to this horse statuette in a number of other museums in Europe and the US. They include: the Metropolitan Museum of Art, New York; the Szepmuvezeti Museum, Budapest; and, the Jeannerat Collection, London. The second piece users used the HMD with was a Maiolica dish in a Study Collection in the museum. This dish depicts a famous scene from Greek mythology: *Achilles Triumphant*; Achilles is riding victoriously in his chariot, parading the desecrated body of his defeated nemesis Hector under the walls of Troy in an attempt to demoralise its defending army. Figure 3.3 gives an impression of what visitors saw while wearing the HMD and looking at the Maiolica dish in the museum. Although the Glasstron enables the projection of digital information next to or overlaying physical artefacts, as can be seen from Figure 3.3, there is a significant drawback in that its semi-transparent visor can also considerably occlude the view of the physical artefact for the user. The view of the physical artefact can also degrade further with changes in ambient light levels. Figure 3.3 shows somewhat what it was like to look through the Glasstron in a high state of degradation or occlusion.

Figure 3.3. Approximation of highly degraded Glasstron HMD view of digital and physical dishes

We found that the degradation/occlusion of the physical environment view detracts significantly from the possible usefulness of the technology, especially indoors - more on this presently.

While viewing the digital and physical artefacts, the visitor also heard related audio content and questions about the artefacts, through the HMD's headphones. The audio was activated/de-activated by the assisting person on request from the visitor (Figure 3.4). The first audio narrative, linked to the virtual Maiolica dish, which was based on a dish housed physically at Nottingham Castle Museum, Nottingham, UK described the similarities between the dish represented virtually on the HMD visor screen and the actual dish in the Study Collection in the Hunt museum (Figure 3.3). The narrative described the cultural and physical correspondence of the dishes. Besides depicting mythological scenes, (the actual dish, i.e. the Hunt dish, depicts *Achilles Triumphant*; and the virtual dish, i.e. the Nottingham Castle dish, depicts *The Legend of the Bull of Perillus*), the dishes, in their physical form, are both Maiolica. They are both made from a special type of earthenware that is fired with lead and oxide of tin. The audio asked the visitor to notice the similar physical construction of the dishes and to consider that glazed earthenware, at the time of its introduction, represented a significant technological breakthrough; people were able to use these ceramics to store liquids, foodstuffs etc. without risk of contamination through seepage. In later times, however, these dishes were usually just decorative, used mainly by the wealthy to adorn a mantelpiece, cabinet or dresser.

Figure 3.4. Visitor, assisted by one of the authors, using the Glasstron HMD to compare digital and physical Da Vinci horse artefacts in the Captain's Room of the Hunt museum

The visitor also heard examples of what people use such ceramics for today – inexpensive, hard wearing false teeth, and extremely heat-resistant tiling for the space-shuttle, being some of the example applications given. The audio narrative linked to the virtual horse described the custodial history of the Leonardo horse: how and when it was made, and where it probably originated. The technology exploration with the Glasstron was completed over two days and the visitors who participated were all adults, educated and astute museum visitors. They included the education officer from the Hunt museum itself, a couple and an individual visitor. The feedback from visitors was mixed. When asked about the technology, the first visitor (Figure 3.4) said that it detracted from her experience of the physical artefacts – "I can't see how this adds value to the real horse; I came here to admire the actual horse but the screen interferes with that." She added that she found the narrative too pedantic (or teacher-like) and boring. She found it particularly difficult to see the actual dish wearing the HMD and the light in the Study Room and reflections on the glass casing exacerbated this. The visiting couple were very positive about the experience, especially about the prospect of using new technologies, like the Glasstron, to collocate historical content. For example, “I can see real use in such technology to compare . . . I mean I’d like to compare the Picasso piece in the Captain’s Room with a similar piece elsewhere – this headgear could be used to *take me* to the museums that keep Picasso’s other works.” The education officer was also interested in the potential of the technology to collocate historical content for visitors to compare and contrast, and to enhance physical artefacts in the museum with digital information, audio, animation etc. However, she was concerned that the quality of the visuals would disappoint younger users, who would be used to higher quality graphics. However, we are beginning to render high-quality graphics in SHAPE (see Chapter 5) and aim to redress this in future developments. The education officer also expressed concern about the robustness of the technology and the variable quality of its see-through mode. There were also a number of other problems with the Glasstron, though we anticipated that these problems would arise. First, visitors had to interrupt their tour of the museum, initially to put on the headset, and then to adjust and fix the visor, headphones and different straps etc., before the HMD could be worn comfortably and used effectively. Furthermore, novice users required assistance to operate the HMD with the laptop. It is also an individual and not a collaborative technology. In addition, the Glasstron is not robust enough for general or frequent handling and use. However, notwithstanding the usability issues and problems with the technology in use, it did provide us with a readymade tool to rapid-prototype hyperphysical links and explore new associations of physical artefacts and digital information in the museum. Deploying the HMD for use by visitors in the museum, we attempted to implement some of the ideas suggested in the Cyberdocent research of Rayward and Twidale (1999), and explore how novel technology might be used to present museum visitors onsite with pointed and supplementary digital information about artefacts. We now describe briefly the hardware and software used for the demo, before we conclude and present our ideas for future work.

Technology Implementation

Initial prototyping, texturing etc. of the dish object was completed in Java 3D™. This object was subsequently imported into 3D Studio Max to replicate, as closely as possible, a sheen so the virtual representation of the Nottingham dish would be reasonably realistic and comparable with the *Achilles Triumphant* dish actually encased in the Hunt museum. Creating the sheen was important to give the effect of glazing. The horse on plinth object was created in 3D Studio Max and the virtual museum world, to house the virtual objects, was created in VRML. Full studio recordings of the audio narratives were recorded and subsequently edited using Sonic Foundry Sound Forge. To offset potential latency difficulties, the original WAV files were compressed to MP3 using a standard Windows Media CODEC, and then hyperlinked to the virtual objects in the VRML world. The Sony Glasstron PLM-S700E HMD was connected to a Dell Inspiron 5000e laptop with Pentium III 700/550 MHz SpeedStep and 256MB RAM. A ParallelGraphics Cortona VRML client was used to play the virtual world. The Glasstron require connection to a video output device with the screen resolution options up to 800_600.

Conclusions and Ideas for Future Work

However, notwithstanding the usability issues and problems with the Glasstron HMD, which perhaps limit its ultimate usefulness in SHAPE, it did provide us with a readymade tool to rapid-prototype hyperphysical links and explore new associations of physical artefacts and digital information in the museum. Also, it has prompted us to think about alternative technologies for creating hyperphysical links. For future work, we envisage scenarios where hybrid artefacts are used to render hyperphysical links in the museum. We include a first, tentative description of one of the scenarios we have in mind, which is based on our findings from field study and observational-analytical work we have completed in respect of the innovative educational programs in the Hunt museum (see also Deliverable 2.1). We envision these hybrid physical-digital objects or artefacts to be replicas of actual, priceless artefacts: spearheads, armlets, statuettes, for example, which contain smart technology, like RFID (Radio Frequency Identification tags). In Chapter 4 of Deliverable 2.1, we describe further application possibilities including ‘cabinet of curiosities’. For example, using an on-board potentiometer, we might be able to track the way in which a visitor turns a hybrid artefact, creating interesting sonic and visual effects for the visitor depending on how she orients the object. Our observations of handling sessions in the Hunt museum (see Deliverable 2.1) suggest that a ‘virtual archaeology’ scenario merits further research. In the educational wing of this museum there are a number of simulated archaeological pits, at which visitors, younger visitors usually, work together to unearth unusual replica antique artefacts, which are buried in sand in cordoned sections (see Figure 3.5).

Figure 3.5. Hunt museum sandbox for simulating archaeological digs

We are thinking of developing hybrid physical-digital artefacts, replica ancient or antique artefacts that contain RFID tags, which are used in combination by visitors to reveal digital information about historical content housed in the museum. Visitors unearth these artefacts in the sandboxes and combine them with other visitors’ artefacts at certain collections in the museum. Combining their respective hybrid artefacts, visitors create interesting sonic or visual effects or reveal information about the collections. The visitor-assembled artefacts would therefore act as a kind of collective activation device for revealing, according to variable levels of visitor collaboration and engagement, digital information about physical artefacts in the museum.

Acknowledgements

For their help in developing the content for the demonstrations, we thank Nora Hickey, Mikael Fernström, and Bob Armstrong. We would also like to thank the visitors who participated in the demos at the Hunt museum.

References

Hammond, K. "Whisper Space" Northwestern University, US (Online), available at http://dent.infolab.nwu.edu/infolab/projects/project.asp?ID=13, accessed on 12th November, 2001

Microvision Inc., US (Online), available at http://www.mvis.com/applications_health.htm, accessed on 8th November, 2001

Rayward, W.B. & Twidale, M.B. (1999) From Docent to Cyberdocent: Education and Guidance in the Virtual Museum, *Archives and Museum Informatics*, 13, 23-53.

Situating Hybrid Assemblies in Public Environments (SHAPE) Annex 1 (2001) "Description of Work"

Sony Corp., US (Online), available at http://www.ita.sel.sony.com/products/av/glasstron, accessed on 8th November, 2001

Vlahakis, V., Karigiannis, J., Almeida, L., Stricker, D., Gleue, T., Ioannis, T., Christou, R.C., & Ioannidis, N. (2001) ARCHEOGUIDE: First results of an Augmented Reality, Mobile Computing System in Cultural Heritage Sites, *Proceedings of Virtual Reality, Archaeology and Cultural Heritage (VAST'01)*, Glyfada, Greece, November 2001, in press.

Chapter 4: The Augurscope: A mixed reality interface for outdoors

Holger Schnädelbach[1], Boriana Koleva[1], Martin Flintham[1], John Bowers[2], Mike Fraser[1], Paul Chandler[1], Malcolm Foster[1], Steve Benford[1], Chris Greenhalgh[1], Shahram Izadi[1]
[1] The Mixed Reality Laboratory, The University of Nottingham, Nottingham NG7 2RD, UK
{hms, bnk, mdf, mcf, pmc, mxf, sdb, cmg, sxi, tar}@cs.nott.ac.uk
[2] Centre for User-Oriented IT-Design (CID), Royal Institute of Technology (KTH), Stockholm, Sweden
bowers@nada.kth.se

The Augurscope is a portable mixed reality interface for outdoors. A tripod-mounted display is wheeled to different locations and rotated and tilted to view a virtual environment that is aligned with the physical background. Video from an onboard camera is embedded into this virtual environment. Our design encompasses physical form, interaction and the combination of a GPS receiver, electronic compass, accelerometer and rotary encoder for tracking. An initial application involves the public exploring a medieval castle from the site of its modern replacement. Analysis of use reveals problems with lighting, movement and relating virtual and physical viewpoints, and shows how environmental factors and physical form affect interaction. We suggest that problems might be accommodated by carefully constructing virtual and physical content.

Introduction

The rapid spread of wireless communications, mobile computing devices and global tracking systems such as GPS has stimulated a growing interest in outdoors augmented reality in which the physical world is overlaid or enhanced with digital information (Azuma 1997, 1999). Wireless handheld tablets have been used to aid navigation or to deliver location-based information to tourists in a city (Cheverst et al. 2000). Wireless wearable computers, complete with see-through head-mounted displays have enabled digital information to be overlaid on and registered with an outdoors environment (Höllerer et al. 1999). These early examples hint at the potential to move beyond today's uses of mobile phones and PDAs to a new generation of more interactive and media rich mobile applications, providing that key challenges can be met concerning lighting, weather, power and tracking (Azuma 1999).

At the same time, recent advances in video processing and display technologies raise the possibility of new augmented virtuality experiences in which hitherto preprogrammed 3D virtual worlds can be enhanced with live information from the physical world or can even be constructed on the fly (Gong et al. 2000). These include advances in building 3D models from video, extracting the movements of people and objects, and displaying multiple video textures in a virtual world (Reynard et al. 1998).

Considered together, augmented reality and augmented virtuality represent two forms of mixed reality, a continuum of experiences in which virtual and physical are merged in different ways, stretching from the purely physical to the purely virtual (Milgram and Kishino 1994).

This paper explores the design of a mixed reality device for use outdoors. This device, called the Augurscope, supports both augmented reality and augmented virtuality. For augmented reality it allows a virtual environment to be viewed as if overlaid on an outdoors physical environment. For augmented virtuality it captures real-time video from this physical environment that can then be embedded into the virtual environment.

Two potential applications of the Augurscope are in cultural heritage and environmental planning. In the former, visitors to historical buildings and sites of special interest can experience scenes from the past as they explore an outdoors site. In the latter, they can explore scenes from the future as part of consultation over designs for new buildings, transport systems and other public facilities. Our paper describes a first public application of the Augurscope belonging to the former category, an open-air museum experience, and presents initial reflections arising from field observations.

Design Requirements

Our design for an outdoors mixed reality device has been driven by the following requirements:

- **Use by public groups** – our intended applications involve directly engaging the public. Our device should be open and inviting to the public. It should be immediately usable by non-expert first-time users with only minimal training. Deployment in public settings also requires consideration of aesthetic quality as well as robustness and maintainability. Our device should also be sharable by groups of users, a requirement that has emerged from previous studies of interactive museum exhibits (vom Lehn et al. 2001) and art installations (Büscher et al. 2001) that have shown how displays are frequently shared by small groups of family or friends. In such situations, several users will often view a display at once, even if only one is able to interact. Furthermore, users frequently learn by watching others, a practice that extends beyond the bounds of the local group to encompass more peripheral observers (vom Lehn et al. 2001). Our device might also be used by domain-experts such as tour guides and town planners to present to a small group.
- **Relocatable** – users will need to move our device to different viewing positions within an extended physical setting such as a town square, a building site or around the perimeter of a large building. Beyond this, it should be only a few minutes work to set-

up and configure the device to work in an entirely new setting. Typical operation might be for a tour guide or town planner to arrive at a location, set-up the device and then move between a number of different viewpoints.

- **Networkable –** there are several reasons why our device should be able to communicate with other devices. First, we are interested in its use with live (i.e., dynamically updated) virtual environments. We are especially interested in the device itself providing the live input to such environments, by capturing and transmitting information from its surrounding physical environment. Second, we are interested in our device as a potential communication tool between users out ‘in the field’ and those ‘back at base’, for example as part of remote guided tours and meetings. Third, we require the device to be able to link to and control other secondary displays such as projected interfaces in order to address more users and to support remote management.
- **Use outdoors –** as noted earlier, previous experiences with augmented reality have encountered a number of difficulties (Azuma 1999). Flat-screen displays can be difficult to read in bright sunlight. Protection is required against adverse weather conditions. Outdoors positioning systems such as GPS can suffer from variable accuracy and reliability and don’t work at all well under some conditions. Devices have to be self-powered, a particular problem where 3D graphics hardware is used as this is relatively power hungry and has only recently become available in laptop computers.

Design of the Augurscope

We considered several general designs that might meet this combination of requirements, including those based on head-mounted, wearable and handheld displays. We eventually opted for a design based on a tripod-mounted display that can be assembled in different outdoor locations and then carried or wheeled around the physical environment (see Figure 4.1). This display be moved to any accessible outdoors location and then rotated and tilted on its tripod in order to view a virtual environment as it would appear from that particular vantage point. At the same time, it captures and transmits a video view of the physical environment from this location.

We named our device an ‘Augurscope’ because it augments both reality and virtuality and also because one of its potential uses is to peer into the future (‘auguring’).

Of course, boom or stand mounted 3D displays are already familiar from virtual reality, where devices such as the Fakespace Boom (www.fakespacelabs.com/products/boom3c.html) are commercial products. Stand-mounted rotating displays have also been used in augmented reality, for example the Panoramic Navigator overlaid text and graphics on a video see-though view captured from an onboard camera, and also included hyperlinks that could be selected via a touch screen (Cook et al. 2000).

We based our design around a portable stand mounted display due to the core issue of physical scale:

- In contrast to wearables or PDAs, a stand-mounted display can be shared by a small groups.
- Users can engage and disengage by stepping up to and away from the display, an important issue when there is a regular turnover of users such as in a museum.
- The required combination of a laptop computer and various tracking, video and audio

peripherals is both bulky and weighty (especially as current laptops with 3D graphics hardware are relatively heavy). Early tests showed that users would quickly tire of carrying them, ruling out a handheld solution.

- The tripod provides a platform for mounting a variety of other devices such as GPS, cameras, speakers and other accessories as we shall see below.
- We were able to enclose the display in an outer casing, improving both its overall aesthetic and ruggedness and also making it more tamper-proof.

Our design process involved two major iterations of construction and testing. The first focused on general physical form factor and produced a standalone device. The second refined this initial design, extending it with video capture and networking. The remainder of this section describes the final design of the Augurscope.

Physical form

The Augurscope (Figures 4.1 and 4.2) is built around a laptop computer (a Dell Inspiron 8000 with a 15 inch display and NVIDEA Geforce2 Go 3d graphics). This is mounted on a rugged tripod using a camera mounting that allows indefinite horizontal rotation and vertical tilting between 25º degrees upwards and 90º downwards (when the display becomes completely horizontal and can potentially be used as an interactive table). The laptop and its mounting are boxed in a wooden casing that features:

- two handles for easy manipulation
- a counterweight for a well balanced and smooth rotating and tilting action so as not to tire users and to maximize accuracy of use.
- a button that when pressed zooms in the virtual viewpoint by a factor of six times and that when released returns to the normal setting.
- a removable cover that bears simple instructions and also conceals the keyboard from users but that can easily be removed if it is needed for administration or maintenance.
- surrounding wooden panels that provide shielding from bright light.

In designing the shielding we were aware of a tradeoff between shielding from sunlight and restricting peripheral viewing and hence inhibiting group use. Indeed, at one point we had considered incorporating a waterproof fabric hood (similar to that used with old fashioned cameras) but decided that this would compromise the open and inviting nature of the device and use by groups. Bearing in mind that current laptop screens offer a relatively narrow viewing angle, a sensible compromise is to allow shielding to restrict the viewing angle up to but no further than the viewing angle afforded the laptop screen. On-board shielding might also be supplemented with external shielding such as parasols.

Wheels were added to the base of the tripod to facilitate movement to new locations. During the course of development we experimented with two sets of wheels. The first was an off-the-shelf accessory wheel-base supplied by the manufacturers of the tripod (Manfrotto). These featured three small rotating wheels on a rugged base with a foot-pedal operated brake. These proved suitable for smooth surfaces, but generally unsuitable for rough surfaces and grass where they were difficult to move and resulted in a very rough ride for the on-board technology. As a result, we then built a second set of more outdoor wheels with inflatable tyres that were more suited to grass and rougher surfaces.

Position and movement tracking

The most basic interaction with the Augurscope is carry it to a new location, set it down and then rotate and tilt it in order to look around. This is made possible through a combination of three onboard tracking technologies.

An etrex GPS receiver with electronic compass attached to the display mount on the tripod gives the position and orientation of the Augurscope relative to the surrounding environment. Position data has a typical accuracy of between two and four meters, although this varies according to weather and proximity to buildings. The compass provides rotational data with a typical accuracy of 1°. However, there is a delay between moving the device and receiving an update of more than a second. Furthermore, position and orientation readings fluctuate by approximately two meters and one degree respectively, even when the device is held stationary.

A rotary encoder is attached to the tripod mounting in order to provide rapid and accurate measurement of the rotation of the display relative to the tripod. This consists of a cannibalized mechanical mouse, where the wheel that normally detects vertical mouse movement is fixed so that is presses against the tripod. The display can therefore be rotated indefinitely.

Figure 4.1: The Augurscope

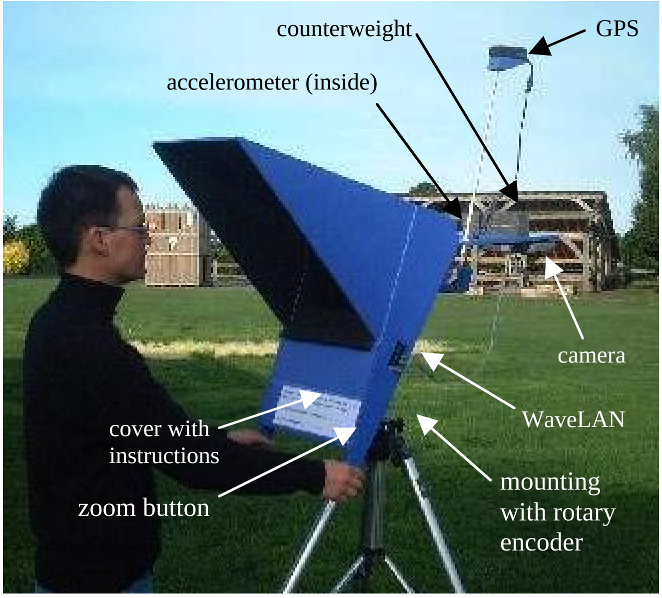

Figure 4.2: The Augurscope close up

A solid-state accelerometer mounted on the wooden frame measures the tilt of the display relative to the tripod. The delay and fluctuation associated with this and the rotary encoder are negligible compared to the GPS receiver and compass. The following table summarises the roles and characteristics of these three tracking technologies.

Technology	**Purpose**	**Characteristics**
etrex GPS compass	Global position of display mount Global orientation of display mount	Two second delay, fluctuates by 2 meters Two second delay, fluctuates by 1°
Rotary encoder	Rotation of display relative to tripod	Negligible delay and fluctuation
Accelerometer	Tilt of display relative to gravity	Negligible delay and fluctuation

The fluctuation and delay in the GPS data causes problems. Applying each update directly causes the virtual viewpoint to continually jump around. We therefore only apply position updates when they show a significant movement (sufficient to clearly indicate that the Augurscope has been moved). A threshold of two meters seems to give a satisfactory balance between stability and responsiveness. In a similar way, the electronic compass reading cannot be used directly while the display is being rotated because of the considerable latency. The absolute orientation of the device is therefore only updated from the electronic compass when it is stationary, in our case after five stable readings have been obtained.

Virtual world display and interaction

The Augurscope's display presents the user with a viewpoint into a 3D virtual world. This provides a first person perspective from the point of view of the device itself so that the virtual world appears to be overlaid on the physical world. Sound is played out through a pair of small battery powered speakers hidden inside the frame.

Our current prototype uses the MASSIVE-3 collaborative virtual environment software for the virtual environment. This supports multi-user/device access to a shared virtual world. An additional software platform called Equip supports the integration of the tracking system with standard MASSIVE-3 interface components.

Video capture

In order to support augmented virtuality applications, video capture is enabled through an onboard camera. This is positioned to look out from the Augurscope so as to capture the physical scene at which it is pointing. However, it could also be turned round to face the user to support conferencing applications.

Networking

Networking with other devices is supported through a WaveLAN card on the laptop. This enables the Augurscope to receive live updates from remote virtual world servers and to transmit position and orientation data as well as live video back to these servers. The Augurscope can therefore synchronise with other devices by publishing its position and orientation data via the shared virtual world. These other devices can then subscribe to this data in order to follow its viewpoint as it moves.

Given that WaveLAN is a local area technology and may not always be available, the Augurscope can also be configured to work in a stand alone mode in which the 3D environment is stored locally and there is no need to access a remote world server.

Recreating Nottingham’s Medieval Castle as a Public Demonstration

For our first public trial of the Augurscope we chose a historical application: recreating Nottingham’s medieval castle on the site of its modern castle.

The problem – Nottingham’s missing medieval castle

A fortified castle was first built at Nottingham on a large outcrop of sandstone in 1067. Over the next six hundred years the castle was extended by a succession of kings to become one of the most important and impressive medieval castles in England as well as the backdrop to the adventures of the mythical character Robin Hood.

Figure 4.3 is an artist’s impression of the medieval castle as it was in the late 14th century. It shows how it was divided into three main areas: the Upper Bailey, the smallest, highest and most protected; the larger Middle Bailey containing many buildings including the Great Hall; and the Outer Bailey, a large open space with no buildings.

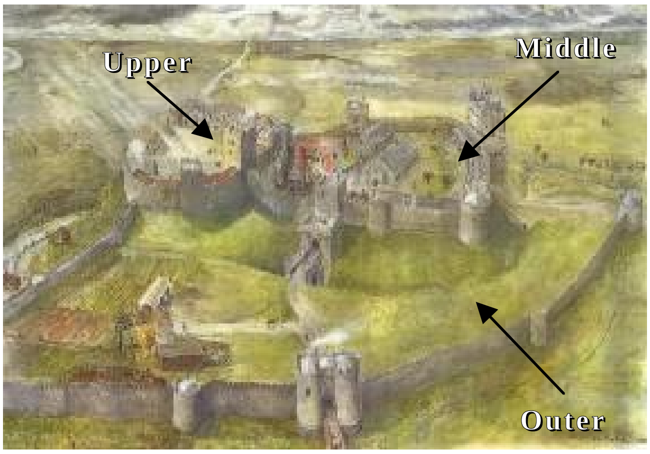

Figure 4.3: Artist's impression of the medieval castle

However, in 1651 following the Civil War, the victorious parliamentarians ordered that the castle be 'slighted', that is demolished to the point where it would be indefensible. The ruins were subsequently cleared in 1674-79 so that the Duke of Newcastle could construct the modern 'Ducal Palace' that occupies the site to this day, currently as a museum (see Figure 4.4).

Herein lies a major problem. Tourists expect to see a fine example of a medieval castle, but instead are presented with the 17th century Ducal Palace in its place. Not only is this disappointing, but it is also difficult understand how the more complex medieval castle was structured, where its parts would have been in relation to the current site, and how they would have appeared. For example, what was once the Middle Bailey (containing the Great Hall, the state apartments, the kitchens and surrounded by high stone walls) is now a now open grassy space called The Green (Figure 4.4). The problem of recreating the missing medieval castle on the current site seemed to us to provide an ideal test application for the Augurscope.

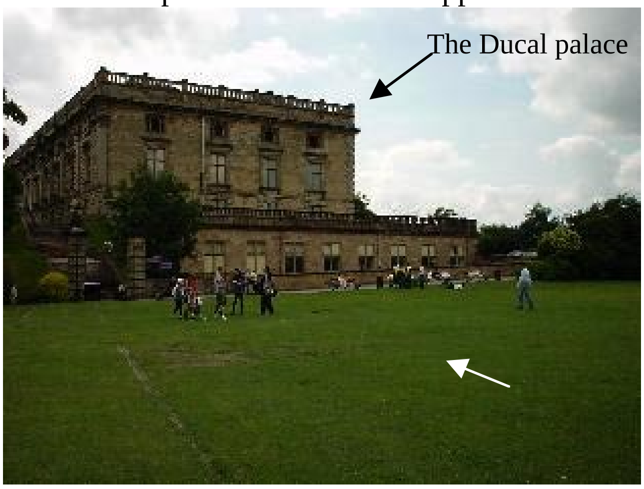

Figure 4.4: The Ducal Palace today seen from The Green

The castle museum already employs various mechanisms to give visitors some sense of the medieval castle: a physical model is on display inside the museum; a slideshow presents the history of site; guides, brochures and text books are available; and museum staff and local guides are highly knowledgeable and sometimes give guided tours. In addition, the locations of some of the original walls are marked out on the ground of the current site, and public displays with maps and diagrams have been placed at key viewpoints. Finally, costumed actors play medieval characters roaming the site (Figure 4.5), a technique occasionally extended to larger scale recreations of medieval markets and tournaments.

Figure 4.5: Actors in role at the castle

Deploying the Augurscope at Nottingham Castle

We were fortunate to obtain an existing 3D model of the medieval castle that could be readily adapted and imported into MASSIVE-3, calibrated with GPS readings from the current site, and then run on the Augurscope.

We also hired one of the castle’s actors who regularly plays the character of a medieval guard to come to our laboratory and pre-record several 3D scenes involving a guard avatar moving through the 3D model and talking as they went. This made use of MASSIVE-3’s record and replay mechanism to capture all of this avatar’s movements and speech in the virtual world so that this could be replayed as live at a later time (Craven et al. 2001). In a day-long session we recorded five separate scenes at different locations in the virtual model in which our guard described various features of the castle and medieval life in general. Figure 4.6 shows our actor making the recording As an aside, note that we attached a polhemus sensor to his spear so that he could use it to gesture during the recordings. In this way we avoided infeasible movements of the virtual spear – such as passing it through his own body – that occurred at first when he was empty handed. Figure 4.7 shows the avatar in the model.

Figure 4.6: recording the medieval guard character

Figure 4.7: replaying the character in the castle model

Our public deployment at the castle involved two networked displays. First was the Augurscope itself. Second was a further public display that was located under a portable gazebo on the castle Green (this can be seen in the background of Figure 4.1). This second display showed a view of the virtual model, with its viewpoint slaved to the Augurscope, but offset so that a graphical representation of the Augurscope was visible in the foreground. This representation included an embedded live video texture taken from the Augurscope’s onboard camera. Figure 4.8 is a screenshot from this secondary display, showing the graphical embodiment of the Augurscope in the foreground. We see that the user is currently looking at the medieval guard avatar.

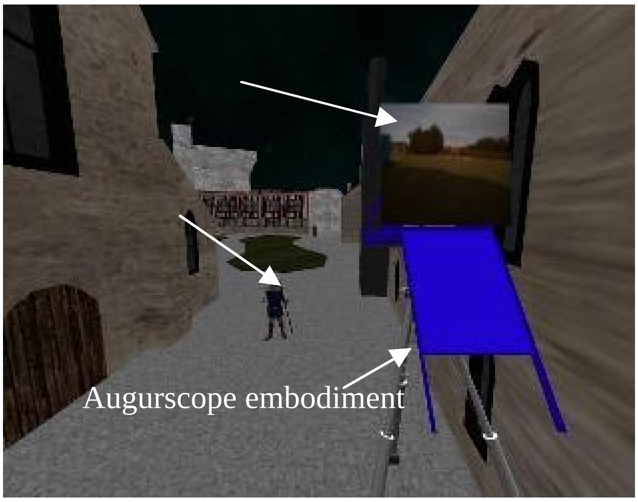

Figure 4.8: View from the secondary display

This slaved display demonstrated the logical reverse of the Augurscope by showing a view of today's castle inset into the 3D model of the medieval castle. Our configuration therefore spanned two points on the mixed reality continuum: augmented reality (the Augurscope) and augmented virtuality (the slaved display).

Finally, a third networked computer in the gazebo provided an interface for manually launching the different 3D recordings from behind the scenes.

Public Trials and Initial Reflections

The development of the Augurscope and the castle application involved a sequence of site visits and public tests. A first iteration of a stand-alone Augurscope was tested in a busy town centre, leading to several refinements including the addition of wheels and the zoom facility. Several visits to the castle were made to select key sites for deployment, obtain reference GPS coordinates, clarify the relationship between the medieval castle model and the physical site, measure WaveLAN signal strength at different locations, and calibrate and test the Augurscope. It is important to realise the sheer contingency that has to managed in deploying technologies in public environments, especially mobile technologies. A mobile device might enter into interaction of some sort with any other artefact at or feature of the location. For example, the various real-world things on The Green which could 'get in the way' of either the wireless network, or moving the Augurscope, or assembling a crowd, or maintaining a line of sight included: a flagpole, a mocked up wooden battlement, a sandpit, children's play equipment, flower beds, trees, steps, bumps in the terrain, et cetera et cetera. These and many other contingencies had to be negotiated in making the exploration of the medieval castle with the Augurscope practically viable.

Eventually, the Augurscope was ready for public trials. These were carried out over a day with the set-up described previously. The weather was mostly sunny, but with some overcast periods. A sign was placed near to The Green inviting visiting members of the public to try out the Augurscope. Approximately thirty members of the public used the device during the day. These ranged from individuals to groups of family and friends. They included tourists (with several overseas groups), local residents, museum designers and staff, the managers of a large public construction project, experts in planning and architecture, other virtual reality and augmented reality researchers, and the media. The pattern of the experience varied between visitors. On the whole, we tried to minimise the amount of training and other scaffolding that was given and instead encouraged visitors to use the Augurscope as independently as possible. The duration of use varied from approximately a minute up to fifteen minutes. Visitors were also encouraged to view the secondary display, in most cases after they had used the Augurscope.

Over the course of the day we collected video of these visitors using the Augurscope. A video camera was placed some distance away, with the zoom facility being used to capture visitors' movements. Audio data was captured via a wireless microphone that was mounted on the Augurscope. Subsequent analysis of this data revealed some interesting aspects of visitors' interaction.

In general, the public and professionals who tried the Augurscope appeared to comprehend its purpose and responded with enthusiasm. Most could operate the device with little training. Rotation, tilting and zooming were used frequently and movement of the device did occur, although infrequently as we discuss below. We saw examples of groups using the device. Often one person would grasp the two handles to rotate the display while others looked over their shoulder (indeed, we suspect that providing a single central handle might have encouraged more equally shared control). However, there were some problems with differences in height, especially for family groups where we saw instances of parents having to lift children (this was less of a problem for groups solely composed of children because the display could be raised and lowered via a handle on the tripod).

We therefore feel that the Augurscope was broadly successful as an outdoors public mixed reality interface for small groups. That said, the remainder of this paper now focuses on several key issues that were more problematic and that suggest possible directions for the future development of the Augurscope and similar devices.

Shedding Light

Despite our attempts to shield the laptop screen, it was noticeable that users sometimes had difficulty seeing the image, even when directly facing it. This became particularly obvious during sunny spells of weather. For example, one visitor adamantly maintained that there was nothing at all displayed on the screen, until directed to stand closer in order to block the sunlight with his head. When instructions continued that "you can rotate and tilt the device", his partner advised that his rotation should "probably not [be] towards the sun, [it] might be better coming that way". What is interesting here is not so much the (already reported in Azuma 1999) observation that bright sunlight is a problem for outdoors displays, but rather the ways in which users react when they are able to freely orientate a display. Turning the Augurscope so that the screen faced away from the sun was a common reaction, even though this sometimes meant that objects of potential interest were missed. Othertimes, there was no real alternative but to engage in extended movement across The Green without looking at the Augurscope screen simply to find a new location where the sun's direction would not be problematic. Another method we observed in one individual was to move the Augurscope with his head buried deep inside its light shield. Making a rather curious sight, this person explored, alone, a large amount of the site this way but, tellingly, had to give up this method when two colleagues approached him and he explained to them what was going on, gesturing towards the screen.

There are several potential approaches to coping with bright light. We might change the graphics, perhaps making them bolder and brighter (in fact, we had lightened the textures following initial testing and prior to the public trials). At the risk of compromising group use, we might extend the shielding, perhaps with a parasol or in the extreme case a blackout hood. However, the above observations suggest an interesting alternative. We might deliberately encourage users to adopt an orientation that shields the display from the sun. This includes taking advantage of shade at different times of the day. We might do this by marking different vantage points in the physical environment. We might also modify the locations of virtual material. For example, actors might be advised to position live avatars on the sunside of the Augurscope whenever possible. Another possibility is to deploy a set of public stationary secondary displays so that there might be the opportunity to consult these when the screen on the mobile device is unclear.

Moving Pictures

Visitors generally appeared reluctant to move the Augurscope, possibly due to the weight of the onboard equipment and frame combined with the rough grassy surface. With two notable exceptions (when the Augurscope was taken on extensive tours of The Green), visitors seemed to prefer viewing the virtual world at a single location, and movement of the device was limited to short distances or to times when the supporting technical team offered help in moving to other viewpoints. In particular, it seemed highly problematic to move the device while holding an object of interest in view. This made quite simple activities such as approaching an object or other ‘target’ on a continuously controlled trajectory hard to achieve. Rather, visitors would accomplish such activities in a ‘fragmented’ fashion: take a few steps, check the target is still in view, take a few more steps, check again, and so on. Equally, it was problematic to ‘back off’ from a location while holding the view orientation. This difficulty is especially notable as backing off in this fashion is a standard method for troubleshooting navigational difficulties in virtual environments where users have a restricted field of view. For example, if one gets caught ‘in a corner’ or ‘too close’ to an object, it would be often be convenient to back off until, say, the peripheries of the field of view are more informative of where one is or the object is revealed. Again, this familiar navigational method was often accomplished in a fragmented fashion: take a few steps, see if the view makes better sense, take a few more steps and so on.

While these difficulties of moving the Augurscope make some kinds of explorations hard to achieve, there are other activities which, interestingly, are facilitated. One particularly noticeable effect was that visitors tended to engage in detailed discussions of those phenomena that were easily available by simply panning and tilting the device. That is, the difficulty of laterally moving the Augurscope gave a motivation to exploring ‘panoramically’ and discussing what could found through so doing. Visitors also made extensive use of the zoom facility, perhaps as a way of compensating for physical movement. Exploring in this fashion often enabled ‘discoveries’ to be made. For example, a visitor might ‘look around a corner’ in the virtual model and discover something, which, if it were easier to sustain lateral movement, might have been missed. Indeed, there is a sense in which the occasionally unwieldy nature of the Augurscope, together with the fact that the virtual model allowed access to both internal (e.g. in the Great Hall) and external points of view (e.g. in the Middle Bailey), added to the fascination of exploring The Green.

Initially, groups and individuals would explore the encountered scene with great deliberation. Only subsequently might movement of the device be attempted. For example, one family explored the Great Hall and its finery by taking it in turns to use the zoom facility to examine all possible content (particularly, the tapestries on the floor and walls), even attempting to zoom through windows. When unable to distract a member of the support team, they continued to view the hall for a short time, and then moved on inside the castle to "go and see the tapestries for real".

It is interesting to note the relationship that different movement methods have to group and individual use. If the Augurscope is swung through a large angle by a user, co-group members are quite likely to lose visual contact with the screen. A co-group member who is shoulder-to-shoulder with whoever tilts or pans the device might be able to maintain the view but someone at the rear of the group might have to jump or quickly run to a new position! This is likely to be at least part of the reason why our examples of extensive movement around The Green seem to be confined to individual use, and groups seem to engage with the Augurscope more tentatively, especially initially. Maintaining views, while maintaining group engagement with what the Augurscope screen is displaying, is a non-trivial accomplishment.

Again, we might deal with these issues by redesigning the physical device and/or the experience. In the first case, an obvious step would be to reduce the weight of the Augurscope, through the use of a different frame made from lighter but similarly sturdy materials, or by using a more lightweight display. We might also change its shape, making the frame smaller (but trading off shielding from sunlight) and providing convenient handles for grasping and lifting. In the second case, the ground surface might be made smoother by the addition of pathways (the Augurscope is far more maneuverable on smooth surfaces using the alternative wheelbase – it can even be gently pushed along as it is tilted and panned). More visible physical markers might be used to indicate key viewpoints and these might also be marked on maps and guides. We might also adapt the content of the virtual environment, for example using avatars' movements and dialogue to encourage users to move to new locations (it was noticeable that some visitors became quite fixed on the guard avatar, although others tended to focus more on the background model).

Doing Legwork

An interesting feature of the use of a tripod was the way in which the three legs appeared to constrain rotation of the display. The legs of the tripod protrude at enough distance from the central axis to maintain the stability of the device. However, this also seems to have the effect of 'framing' the use of the Augurscope into three 120˚ segments, each defined by two of the three legs.

Users, whether individually, or in small or large groups, appeared to treat the legs as cut-off points for standing. We recorded instances of up to twelve people standing within the 120˚ segment containing the display. Users rarely stood across tripod legs, preferring to move quickly past the legs, before pausing to spend time in the next segment. Where visitors did stand outside this segment in proximity to the device, they often appeared uncomfortable or detached. Such visitors were observed to move to a better viewpoint; to encourage rotation of the device into 'their' segment; or even to disengage and try to engage others in related activities, such as side conversations.

A further key aspect of this segmentation, is that the Augurscope was often used for long periods of time within one segment. Movement of the display from one segment to another required the effort of traversing a leg, and was far less frequent than movement of the display within a segment. A typical pattern of use was to explore a location by thoroughly investigating each segment in turn, before traversing to the next.

Again, these observations provide interesting cues for designing and managing the overall experience. Dynamic content can be introduced into the current viewing segment if continuity of experience is required or alternatively, can deliberately be placed in one of the other segments in order to encourage the movement of the users and Augurscope across a tripod leg into another segment. Actors controlling avatars might choose their direction of approach to the Augurscope based on these criteria.

The tripod design also exacerbated the difficulties of lateral movement while maintaining a target in view which we have already mentioned. Our video contains an example of a trio of visitors finding themselves at a location where a wall texture fills the Augurscope's screen. They rotate the device to see themselves suddenly pass through the wall and then back again. Reckoning that their location is 'within a wall' (note, by the way, that collision detection was not implemented in the virtual worlds we used), they decide to fix the view angle and back off – something we have already noted as problematic. However, the path that they decide to take is not one that the tripod's legs and wheels are oriented for. Accordingly, the group actually lift the Augurscope and carry it with all legs off the ground (the Augurscope's legs that is). One group member bares the lower part of the device's weight, another steadies the top, while a third checks on the view unfolding on screen! While such examples naturally suggest the redesign of the device, they also testify to the group's motivation in continuing to engage with the activity of exploring The Green and the (virtual) castle. Indeed, it is arguable that some of these difficulties might add to the intrigue of exploring with the Augurscope and provoke concerted collaborative activity in so doing.

Relating Worlds

Our final issue focuses on resolving relationships between the physical and virtual worlds. Let us discuss a number of aspects of this.

First, being able to understand and anticipate the relationships between physical and virtual worlds was a particular problem for application developers. For example, making the 3D recordings raised some tricky issues. The guard avatar needed to be positioned so that he could make reference to key features of the model and also so that viewers could find good vantage points from which they could see him and the features being discussed. Indeed, acting when you don't know where the audience will be located is itself quite difficult. For example, you may often need to avoid spatial references (such as the names 'left' and 'right'). Care also had to be taken in designing gestures, especially those which, with a present audience, might be made with respect to them. However, these issues were further complicated by the constraints of the physical world: it was also necessary to understand whether there would be an appropriate vantage point on The Green from which to view the action. The most difficult case required the Augurscope to be quite precisely positioned on the edge of The Green so that it could see over the edge of a slope to a bridge below (a physical constraint) and yet was on the right side of a virtual wall that ran nearby (a virtual constraint).

Related issues became apparent when deciding when to replay different 3D recordings. A member of the technical team needed to judge when a visitor was in a good position to be able to see a prerecorded avatar. This relied on knowledge of both their physical and virtual location, as well as a sense of the right moment to intervene upon the activity the vistors were currently engaged in. Previous studies of mixed reality experiences indoors have shown how performers and crew rely on both physical and virtual monitoring in order to orchestrate an experience from behind the scenes (vom Lehn et al. 2001). This becomes more problematic when mobile devices are being used in outdoors locations. The secondary slaved display proved useful for the crew here as it showed the Augurscope's position and orientation in the virtual world. However, it still proved difficult to resolve the physical and virtual positions (as the two spaces were very different). This suggests the development of secondary displays that show positions overlaid on both the physical and the virtual.

While there were issues for the crew in resolving the relationship between the virtual and physical in supporting the visitor's experience, it must be remembered that the activity of exploring the virtual medieval castle required visitors to do this all the time! What our video materials reveal is that people were able to do this, that the activity made sense to them, that they could troubleshoot successfully when in difficulty because they had a sense of what should be going on, and that to do so was pleasurable enough to (at least in some cases) lead to significant periods of engagement. Perhaps the most striking examples we can offer concern the relationships between gestures towards the various displays and gestures towards the real world (a point we will analyse out at length in future work). For example, during the episode described above when the threesome manhandle the Augurscope to get a better view, one of the group members offers a gesturally rich account of where the group are and what they are seeing. "We are in a narrow gap…. between a big hall there and another there". As he speaks, each 'there' is accompanied by an expansive two handed gesture pointing away from the Augurscope first in one direction then in another, at…. nothing. Well…. Nothing that can be seen if one is expecting to see real 'big halls' 'there' and 'there' but, if in the light of having traversed The Green with the Augurscope, one is accustomed to see The Green as the site of parts of a medieval castle, then there is nothing senseless in how the man alternates his gestures between the screen and thin air. Indeed, the gazes of his co-group members follow his arms as they move away from pointing to the on-screen graphics to the imagined big halls. Through his talk and gesture, the man animates the on-screen scene in terms of what might have stood right 'there and there' in the past. Perhaps we can say that, through his conduct, the man is practically 'mixing realities' so as to assemble an account of the relationship between the imagined medieval castle and the contemporary physical environment. And the Augurscope is an adequate technical resource to help him accomplish this.

Summary

We have described the design of the Augurscope, a portable mixed reality interface for outdoors. We chose a design based on a tripod-mounted display so as to support use by public groups and to provide a stable platform for a 3D display, tracking technologies, a video camera and powered speakers. We explored design tradeoffs spanning physical form, user interaction and the integration of different tracking technologies. For the latter, we used a GPS receiver with electronic compass to locate the device within the surrounding environment and an onboard rotary encoder and accelerometer to support smooth local interaction. Wireless networking allowed communication with remote management tools and slaved displays.

Our first application involved exploring a destroyed medieval castle on the site of its more modern counterpart. User testing raised a number of issues for further exploration. In particular both environmental factors (such as the direction of the sun and the roughness of the ground surface) and physical form factors (weight, wheels and shape) affected interaction. We propose that apparent problems might be addressed by redesigning the device, but might also be dealt with through the careful design and management of both physical and virtual experience. There were also problems with resolving relationships between the physical and the virtual, particularly during application development and orchestration. These might be resolved through the careful design of secondary displays.

To conclude, we believe that the Augurscope provides an interesting contrast to other approaches to mixed reality outdoors and that an appropriately refined example might be suited to outdoors applications in public places such as museums. We hope that our experience provides useful insights for the design of other devices.

References

Azuma, R. T., A Survey of Augmented Reality, *Presence: Teleoperators and Virtual Environments*, 6(4): 355-385, Aug. 1997.

Azuma, R., The Challenge of Making Augmented Reality Work Outdoors, In *Mixed Reality: Merging Real and VirtualWorlds*, 1999, Springer-Varlag.

Benelli, G., Bianchi, A., Marti, P., Not, E., Sennati, D., HIPS: Hyper-Interaction within Physical Space, *Proc. IEEE ICMCS99*, Florence, June 1999.

Büscher, M., O'Brien, J., Rodden, T. and Trevor, J., 'He's Behind You': The experience of presence in shared virtual environments, *Collaborative Virtual Environments: Digital Places and Spaces for Interaction*, 77-98, Springer-Verlag, 2001.

Cheverst, K., Davies, N., Mitchell, K., Friday, A. and Efstratiou, Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences, *Proc. CHI'2000*, 17-24, The Hague, Netherlands, 2000.

Craven, M., Taylor, I., Drozd, A., *et al.*., Exploiting Interactivity, Influence, Space and Time to Explore Non-linear Drama in Virtual Worlds, *Proc. CHI'2001*, Seattle, US, April 2-6, 2001, ACM Press.

Gong, S., McKenna, S., Psarrou, A., Dynamic Vision, Imperial College Press, London, 2000.

Höllerer, T., Feiner, S., *et al.*, Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System, *Computers and Graphics*, 23(6), Elsevier, Dec. 1999, pp. 779-785.

Koleva, B., Taylor, I., Benford, S., *et al.*, Row-Farr, J., Adams, M, Orchestrating a Mixed Reality Performance, *Proc. CHI'2001*, Seattle, April 2001.

Milgram, P. and Kishino, F., A Taxonomy of Mixed Reality Visual Displays, *IEICE Transactions on Information Systems*, Vol E77-D (12), Dec. 1994.

Reynard, G., Benford, S., and Greenhalgh, C., Awareness Driven Video Quality of Service in Collaborative Virtual Environments, *Proc. CHI'98*, 464-471, LA, April 18-23, 1998.

vom Lehn, D., Heath, C. and Hindmarsh, J., Exhibiting Interaction: Conduct and Collaboration in Museums and Galleries, *Symbolic Interaction*, 24 (2), University of California Press, 2001.

Cook, J., Pettifer, S., Crabtree, A., Developing the PlaceWorld environment, *eSCAPE*

Deliverable 4.1 The Cityscape Demonstrator (eds. Mariani & Rodden) Lancaster University, ISBN 1-86220-079, 2000.

Chapter 5: Wasa: Towards a set of technologies for producing public mixed-reality learning environments

Gustav Taxén, John Bowers
Centre for User-Oriented IT-Design (CID), Royal Institute of Technology (KTH), Stockholm, Sweden
{gustavt, bowers}@nada.kth.se

This chapter introduces Cybermath, a shared virtual environment for the discovery and sharing of the mathematics of geometry. However, rendering its content into mixed reality for our first Living Exhibition requires a substantially different application platform. We have evaluated a selection of such platforms. A suitable candidate has been chosen and extended with mixed reality-specific components. The new platform will give us additional flexibility in building our Living Exhibitions.

Introduction

The first SHAPE Living Exhibition is concerned with hybrid physical digital and mixed reality artefacts, with a focus on mathematical content. Previous work at KTH has dealt with the creation of virtual environments for sharing and discovering the mathematics of geometry. This work has resulted in the Cybermath demonstrator (Taxén and Naeve 2001b) and we expect that at least a portion of its content will be present in the first Living Exhibition.

However, bringing virtual environments into a public environment such as a museum places place heavy requirements on the supporting technical platforms with respect to stability, efficiency and visual quality. Early on, it became obvious that the DIVE platform, which the original Cybermath demonstrator is built from, does not meet these requirements. In addition, being a traditional avatar-based virtual reality system, it is unclear whether DIVE is suitable for the special rendering requirements of mixed reality applications. Therefore, an effort was made to find an alternative application framework. Systems that were examined and rejected include ActiveWorlds, ALICE, MASSIVE, Open Inventor, OpenGL Performer, Java3D, and OpenSG. Instead, we have chosen to work with the Wasa platform that is under development at the Centre for User Oriented IT Design (CID) at KTH.

Wasa is a collection of utilities and components rather than a complete, monolithic system, which makes it flexible and extendable. It is also highly portable, efficient, and produces very high-quality visuals. Since Wasa currently lacks a stable network distribution model, KTH has begun the integration of Wasa with the EQUIP platform of the University of Nottingham. In addition, SHAPE are developing a number of new Wasa components that are suitable for mixed reality applications, including particle and fluid animation, collision detection, shading and projection onto non-linear surfaces.

The next section describes the background of the Cybermath environment and how it has informed the SHAPE project with respect to the design of our new virtual reality application platform.

Cybermath

Virtual Reality and Mixed Reality systems have the potential to allow students to discover and experience objects and phenomena in ways that they cannot do in real life. Since the early 90s, a large number of educational virtual reality applications have been developed. These include tools for teaching students about physics (Dede et al. 1996), algebra (Bricken 1992), colour science (Stone et al. 2000), cultural heritage objects (Terashima 1999) and the greenhouse effect (Jackson 1999).

There is convincing evidence that students can learn from educational VR systems (Winn 1997). However, issues relating to collaboration and learner motivation have largely been overlooked. Also, few authors have focused on the mathematics of geometry as content, even though geometry is particularly suited for graphical visualisation. KTH has built the Cybermath system to allow further studies of these issues (Taxén and Naeve 2001a).

Cybermath is a shared, avatar-based virtual environment built on top of DIVE (Carlsson and Hagsand 1993). DIVE has the ability to display shared interactive 3D graphics as well as distribute live audio. It can run on a number of hardware configurations, ranging from standard desktop PCs to head-mounted displays and CAVEs (Cruz-Neira et al. 1993). It is possible to allow different users to access the same virtual environment from workstations with different hardware configurations. CyberMath is built as an exploratorium that contains a number of exhibition areas (Figure 5.1). This allows teachers to guide learners through the exhibitions but learners can also visit CyberMath at their leisure, alone or together with others.

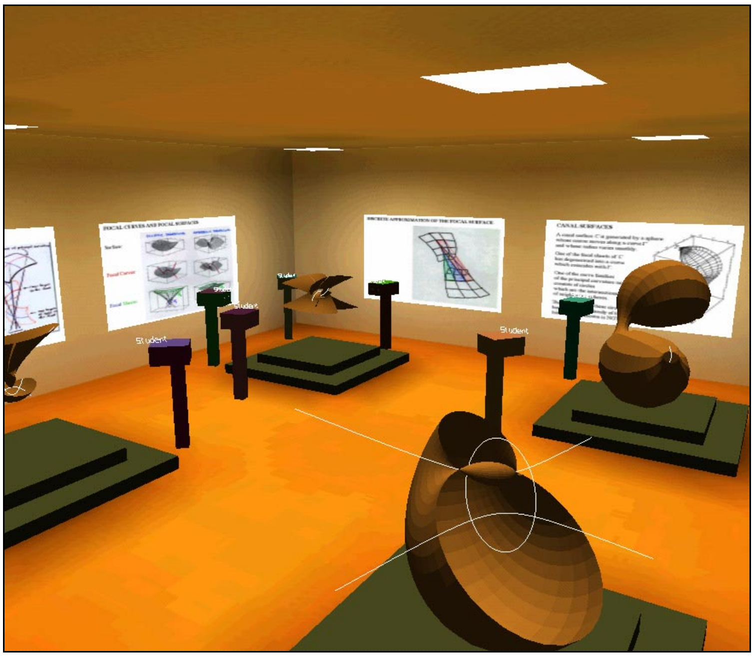

Figure 5.1. An exhibition in the DIVE version of Cybermath

DIVE supports rapid prototyping through Tcl/Tk scripts. We have complemented this support with a Mathematica-to-DIVE conversion utility that can be used to convert standard three-dimensional Mathematica objects and animations to the DIVE file format. It is then straightforward to add Tcl/Tk code to turn the converted Mathematica objects into interactive CyberMath exhibitions. This makes it possible to support rapid-turnaround teacher-driven development of new CyberMath exhibitions in the same fashion as in the QuickWorlds project (Johnson et al. 2000).

At the time of writing, three example exhibition areas in the exploratorium have been completed. Their respective content is:

- **Interactive transformations** (Figure 5.2). An $\mathbf{R}^3 \rightarrow \mathbf{R}^3$ transformation maps a three-dimensional point to another three-dimensional point. In this exhibit, users can investigate the effects of any such transformation on different mathematical entities, including points, lines, planes, boxes and spheres. The user can manipulate the entities and immediately see the results of the transformation, either in a separate coordinate frame or in the same coordinate frame as the untransformed surface.

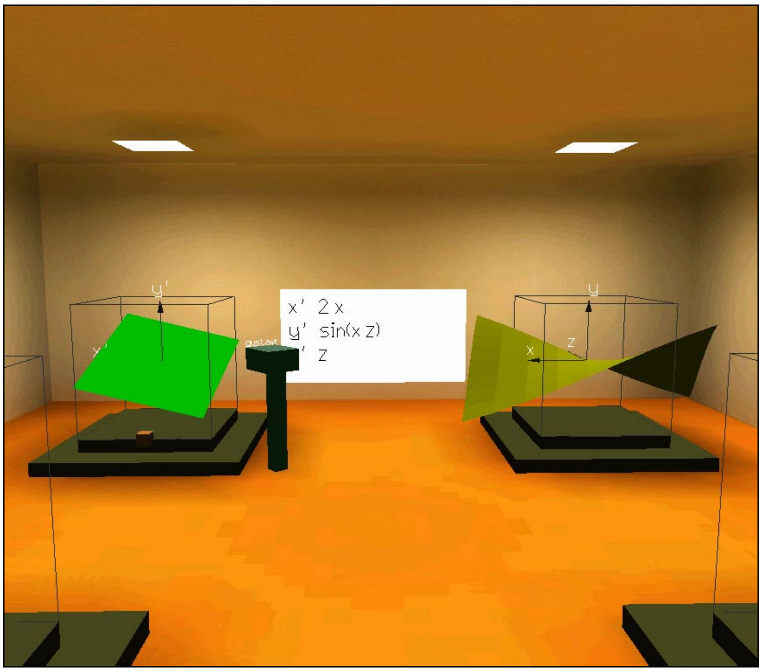

Figure 5.2. The interactive transformations exhibition

- **Differential geometry** (Figure 5.3). In this area, users can learn how to construct advanced three-dimensional surfaces using methods from differential geometry. The exhibition includes a number of three-dimensional animations and wall posters that illustrates these methods.

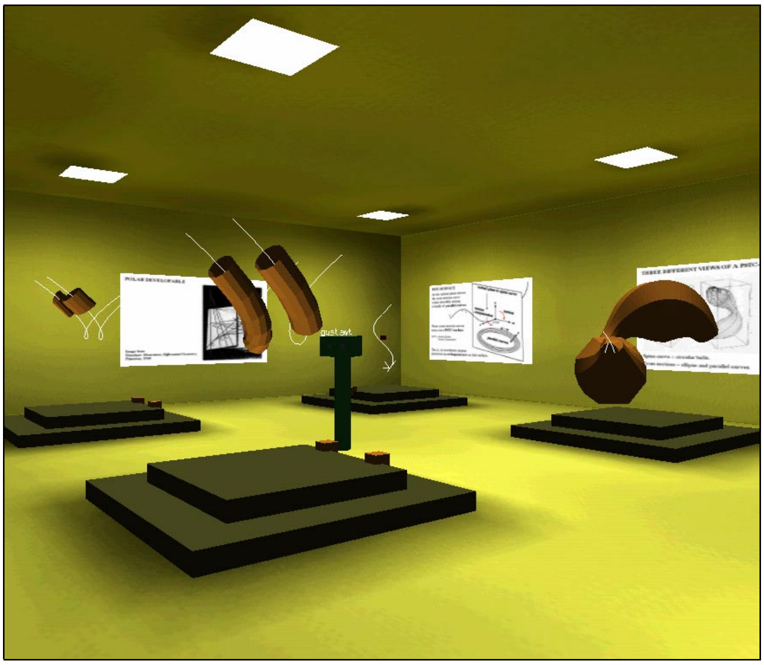

Figure 5.3. The generalised cylinders exhibition

- **Focal surfaces** (Figure 5.4). In this exhibit, we illustrate how two cylindrical mirror surfaces can, when used together, focus sunlight at a point in space in the same way as a paraboloid mirror can.

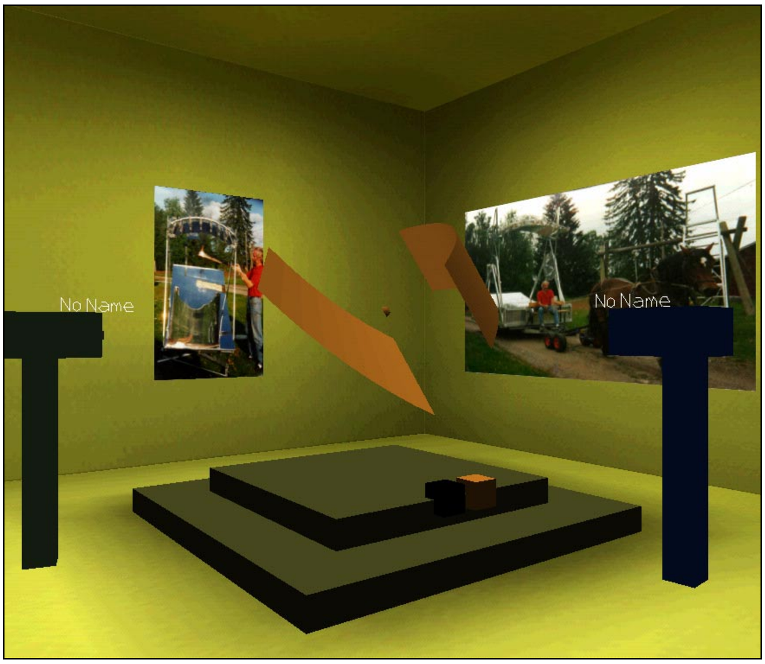

Figure 5.4. The focal surfaces exhibition

Bringing Cybermath content into public environments

We would like Cybermath, or at least a subset of its content, to be part of our first Living Exhibition. However, in order to bring technology into a public setting, it has to meet a number of requirements:

- *The software platform has to be extremely robust.* This is especially true for situations where young children are allowed to interact with it. For example, the KidStory project continuously updated its KidPad application with respect to robustness throughout the projects' three-year lifetime (Taxén et al. 2001). In museums, it is not uncommon that visitors go through huge efforts to break or vandalise the exhibits: computer hard drives are re-formatted, running software is broken into and hardware peripherals are broken or even stolen.
- *The software platform has to be secure and easy to manage and restore*. If our exhibits are to be maintained by museum staff, it must be easy to reset the application to its default configuration and to move it from one computer to another.
- *The interaction with the software has to be smooth and efficient.*
- *The software platform must be capable of producing visuals of high quality*. This is important for attracting the attention of visitors and keeping their interest in the exhibits.
- *The software platform must be flexible* in the sense that it can be used for mixed reality applications and not only standard avatar-based virtual environment applications.

After some internal evaluation, we have concluded that the DIVE platform is unsuitable for building software for public display. Our main concerns lie with its limited visual quality, fixed feature set, low execution speed and lack of security features. Also, DIVE is somewhat crash prone and hard to configure and set up. Thus, a search for a more suitable platform was initiated.

A very large number of systems for creating interactive graphics exist. Most of these can be categorised as either high-level application-building systems, where a number of fixed features and capabilities are combined to form a new application, or as low-level application platforms that provide a number of data structures and management systems that abstract the underlying hardware rendering pipeline. We believe that very few of these meet all our requirements for mixed reality applications. High-level systems are typically aimed at the construction of standard 3D, first-person-view walkthrough applications and low-level systems are typically tied too closely to the underlying hardware to be useful for implementing current state-of-the-art rendering algorithms and projections (or cannot be extended to do so). The following section highlights a few of the systems we have examined.

Review of software applications

The ActiveWorlds system (http://www.activeworlds.com) is attractive because of its robustness, security and ease of use, both in terms of system management and of producing new content. Also, KTH has used it for previous museum exhibits and for the public construction of virtual exhibitions (Walldius 2001). However, it is too limited in its ability to produce interactive components and high-quality graphics. Also, it is unsuitable for other applications than avatar-based virtual environment walkthroughs.

The ALICE system (Conway 1997) is similar to ActiveWorlds in the sense that it allows easy and straightforward construction of new virtual exhibits. In addition, quite sophisticated interactivity can be built into the exhibits. However, the resulting virtual environments cannot readily be distributed, they can only run on Windows-based PCs, and their graphical quality is somewhat limited.

The MASSIVE system of the University of Nottingham has been proven robust enough to be used in public settings (Benford et al. 1999)(Shaw et al. 2000). It has also been used within SHAPE as the technical platform for our Unearthing Virtual History and Augurscope demonstrators. MASSIVE entities can also be distributable by the EQUIP networking system developed under the Equator project (Greenhalgh et al., forthcoming). Since MASSIVE is an established platform within SHAPE, we also expect it to be used for several forthcoming demonstrators. However, it produces visuals of limited quality and does not take advantage of modern rendering methods.

We have also evaluated a number of scenegraph-based low-level software architectures that might be used as a foundation for a mixed reality rendering platform. Open Inventor (Wernecke 1994) is a well-known extendable scenegraph architecture. Unfortunately, it is somewhat awkward to implement modern rendering methods in Open Inventor because of its tight conceptual coupling with the OpenGL rendering pipeline. Also, SGI provides no implementation of Open Inventor for Windows.

The OpenGL Performer system (http://www.sgi.com/software/performer/) is somewhat less extendable than Open Inventor and also has a tight conceptual coupling with OpenGL. In addition, there is no non-commercial license available for the Windows platform.

OpenSG (http://www.opensg.org/) is an Open Source alternative to Open Inventor. Unfortunately, it is still undergoing substantial architectural redesign and is, at the time of writing, not extendable by application writers.

Java3D (http://java.sun.com/products/java-media/3D/) is similar to OpenGL Performer and runs on Java platforms. However, java applications are typically much slower than their C or C++ counterparts and the rendering capabilities of Java3D are limited.

The Centre for User Oriented IT Design at KTH is developing the Wasa platform, which is a collection of portable and extendable programming libraries for developing high quality graphical applications (Taxén et al. forthcoming). Wasa contains a lightweight scenegraph management component that is similar in design to Open Inventor together with a number of additional components for lighting, 3D file format conversion and view frustum culling. In addition, Wasa makes no direct assumptions of the configuration of the underlying pipeline that is used to render its graphics, which makes it straightforward to add support for new rendering algorithms to Wasa. It is also easy to add support for alternative surface representations, such as NURBS or Catmull-Clark subdivision surfaces (Catmull and Clark 1978) to the Wasa scenegraph management component. Wasa can be used together with VRJuggler (Bierbaum et al. 2001) to produce visuals for non-standard hardware configurations such as head mounted displays or VR CAVEs.

As a proof-of-concept, we have re-implemented selected Cybermath exhibits in Wasa (Figures 5.5 and 5.6). The resulting application is responsive, stable and visually attractive. In addition, it is easy to extract isolated parts of the application to include in the Living Exhibition, or to develop it into a fully featured public demonstrator suitable for exhibiting in a museum.

Figure 5.5. The Wasa version of Cybermath

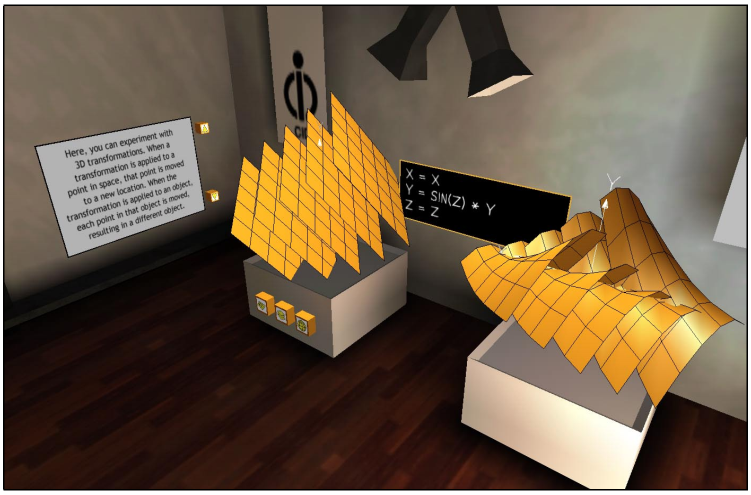

Figure 5.6. The transformation exhibit in the Wasa version of Cybermath

Extending Wasa

In anticipation of the Living Exhibition, SHAPE has extended Wasa with a generic particle animation component, a collision detection component and a new shading component. We are also developing components for projection onto non-linear surfaces and for fluid dynamics.

Particle systems can be used to visualise a wide range of phenomena (Reeves 1983). In particular, they are suited for renderings of dynamic flows of matter or interacting discrete entities, both of which we anticipate will play a role in the Living Exhibition. Our collision detection component, based on (Melax 2000), allows particles to interact with a virtual environment.

Our new shading component allows virtual objects to be lit by captured real light. Recent work by Debevec has produced algorithms where a so-called light probe is used to represent all light that arrives at a point (Debevec 1998). Light probes are typically obtained through photographs of real environments, although our shading component includes the capability of constructing light probes from environments that have been lit by the Wasa lighting utilities. When a light probe has been obtained, it can be used to illuminate virtual objects in real-time, as illustrated in Figure 5.7. Our shading component implements a recent algorithm described by (Ramamoorthi and Hanrahan 2001) that allows the irradiance reflected by Lambertian surface properties to be approximated by just nine values. An advantage of this algorithm is that these nine values are easy enough to compute to be useful in real-time applications. By generating light probes from a set of connected web-cams, we intend to light virtual objects with real, dynamic light. We believe that this is a new and unique approach for mixed reality applications.

Figure 5.7. A virtual object in Wasa, lit by light captured from photographs

We would like to be able to project images onto surfaces with unusual curvatures, or even onto surfaces that change dynamically. Such projection surfaces will distort the output from a traditional renderer, which may be an intentional effect. However, curved projection surfaces can be used to give the viewer an increased sense of depth in an image without the need for stereo viewing. Unfortunately, such a projection requires the geometry to be pre-distorted before it is sent to the rendering component. We are currently developing a utility for defining and executing such distortions.

The ToneTable demonstrator (Chapter 7) will very likely constitute an important part of the Living Exhibition. However, its current implementation is rather restricted in terms of animation of the water surface. Therefore, we are investigating alternative, physically based animation methods. Possibilities include the solving the Navier-Stokes equations using the real-time methods of (Stam 1999) or the solution of simplified equations from fluid mechanics.

At this time, the Wasa includes a simple network distribution library that can be used for small-scale applications. However, in order to build more advanced networked application, a more sophisticated distribution system, such as EQUIP, is needed. EQUIP also allows user interface components and peripherals to be dynamically added to and removed from applications in a controlled manner at run-time. This feature has the potential to greatly simplify user interface management. Therefore, we are adding the capability of distributing Wasa entities with EQUIP. Thus far, we have built a proof-of-concept prototype that allows mouse interaction events to be distributed across a network by EQUIP to a Wasa application, where they modify a virtual object. We are currently investigating different approaches for extending this work.

Conclusion

Throughout the course of the first year of SHAPE, we have begun to identify a number of requirements for a mixed-reality rendering platform. We believe that these requirements are likely to include the ability to render high quality graphics, fast responsiveness to user input, and a high level of robustness, security and flexibility. None of the currently available development platforms and graphical application builders we have examined meet these requirements. Therefore, we have begun to extend the Wasa rendering platform of the Center for User-Oriented IT Design at KTH with a number of components suitable for building mixed reality installations. We are also pursuing an integration between these components with the EQUIP network distribution system of the University of Nottingham.

References

Benford, S., Greenhalgh, C., Craven, M., Walker, G., Regan, T., Morphett, J. and Bowers, J. (1999) Broadcasting On-Line Social Interaction as Inhabited Television. In Proceedings of ECSCW, 1999.

Bierbaum, A., Just, C., Hartling, P., Meinert, K, Baker, A. and Cruz-Neira, C. (2001) VR Juggler: A Virtual Platform for Virtual Reality Application Development. In *IEEE VR 2001*, Yokohama, March 2001.

Bricken, W. Spatial Representation of Elementary Algebra. In *Proceedings of the 1992 IEEE Workshop on Visual Languages*, 55-62.

Catmull, E. and Clark, J. (1978) Recursively Generated B-Spline Surfaces on Arbitrary Topological Surfaces. In *Computer-Aided Design* 10(6):350-355, November 1978.

Carlsson, C., Hagsand, O. (1993) DIVE - A Multi User Virtual Reality System, In *Proceedings of IEEE VRAIS '93*, 394-400.

Conway, M. (1997) Alice: Easy-to-Learn 3D Scripting for Novices. Ph.D. Thesis, The University of Virginia.

Cruz-Neira, C., Sandin, D. J., DeFanti T. A. (1993) Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE. In *Proceedings of ACM SIGGRAPH '93*, 135-142.

Debevec, P. E. (1998) Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography. In *Proceedings of ACM SIGGRAPH 1998*.

Dede, C., Salzman, M. C., Loftin, R. B. ScienceSpace: Virtual Realities for Learning Complex and Abstract Scientific Concepts. In *Proceedings of IEEE VRAIS '96*, 246-252.

Greenhalgh, C., Izadi, S., Rodden, T. and Benford, S.. The EQUIP Platform: Bringing Together Physical and Virtual Worlds, forthcoming.

Jackson, R. L. Peer Collaboration and Virtual Environments: A Preliminary Investigation of Multi-Participant Virtual Reality Applied in Science Education. In *Proceedings of the ACM 1999 Symposium on Applied Computing*, 121-125.

Johnson, A., Moher, T., Leigh, J., Lin, Y-J. QuickWorlds: Teacher-Driven VR Worlds in an Elementary School Curriculum. In *Proceedings of ACM SIGGRAPH '00 Educators Program*, 60-63.

Melax, S. (2001) Dynamic Plane Shifting BSP Traversal. In *Proceedings of Graphics Interface*, 2001.

Ramamoorthi, R. and Hanrahan, P. (2001) An Efficient Representation for Irradiance Environment Maps. In *Proceedings of SIGGRAPH 2001*.

Reeves, W. T. (1983) Particle Systems - A Technique for Modeling a Class of Fuzzy Objects, *Computer Graphics*, vol. 17, no. 3, pp 359-376, 1983.

Shaw, J., Staff, H., Farr, J. R., Adams, M., vom Lehm, D., Heath C., Rinman, M-L., Taylor, I.

and Benford S. *Staged Mixed Reality Performance "Desert Rain" by Blast Theory*, eRENA Project Deliverable 7b.3, August 2000.

Stam, J. (1999) Stable Fluids. In *Proceedings of ACM SIGGRAPH 1999* August 1999, 121-128.

Stone, P. A., Meier, B. J., Miller, T. S., Simpson, R. M. Interaction in an IVR Museum of Color. In *Proceedings of ACM SIGGRAPH 2000 Educators Program*, 42-44.

Terashima, N. Experiment of Virtual Space Distance Education System Using the Objects of Cultural Heritage. In *Proceedings of the 1999 IEEE International Conference on Multimedia Computing and Systems*, vol. 2, 153-157.

Taxén, G. and Naeve, A. (2001a) CyberMath: Exploring Open Issues in VR-Based Learning. In *Proceedings of ACM SIGGRAPH 2001 Educators Program*, 49-51.

Taxén, G. and Naeve, A. (2001b) CyberMath: A Shared Virtual Environment for Mathematics Exploration. In *Proceedings of ICDE 20th World Conference on Open Learning and Distance Education*, Düsseldorf, April 1-5, 2001.

Taxén, G., Druin, A., Fast, C. and Kjellin, M. (2001) KidStory: a technology design partnership with children. *Behaviour & Information Technology*, Vol. 20, No. 2, April-March 2001, 119-125.

Taxén, G., Bäckström, P., Lenman, S. and Sunblad, O. Wasa: A Toolkit for the Production of High-Quality Graphics Applications. (Forthcoming.)

Walldius, Å. (2001) Shared 3-D Workplace Exhibitions as Sites for Community Meetings. *Behaviour & Information Technology*, vol. 20, no. 2, pp. 91-99, March-April 2001.

Wernecke, J. (1994) *The Inventor Mentor: Programming Object-Oriented 3D Graphics With Open Inventor*. Addison-Wesley, March 1994.

Winn, W. *The Impact of Three-Dimensional Immersive Virtual Environments on Modern Pedagogy*. University of Washington, HITL, Report No. R-97-15, 1997.

Chapter 6: Publicly Deploying ToneTable as a Multi-Participatory, Mixed Media Installation

John Bowers, Sten Olof Hellström and Gustav Taxén
Centre for User-Oriented IT-Design, Royal Institute of Technology (KTH), Stockholm.

This chapter describes the development and deployment of ToneTable as a research vehicle for exploring a variety of design principles for mixed media, multi-participatory artefacts. We situate ToneTable in relation to other 'mixed reality artefacts' as a kind of 'roomware' but emphasise its relative novelty in focussing on providing users with a rich set of interactive, inter-media (graphics and sound) relationships. We describe how ToneTable has been designed according to principles which are proposed in the light of emerging work studying people's interaction with mixed media artefacts in public places. These principles highlight the design of 'emergent collaborative value', 'layers of noticeability' and 'structures of motivation' in an installation which can be regarded as an 'ecology of participation'. Two public exhibitions of ToneTable are described, along with the modifications to design we made in the light of initial experience. The paper closes with some general remarks about the challenges there are for the design of collaborative installations and the extent to which we have met them.

Introduction

The current chapter describes the further work which has been undertaken in SHAPE, principally at KTH, developing and publicly deploying ToneTable, a multi-participatory, mixed media installation. ToneTable was developed in its first version within the second SHAPE Constructional Workshop at KTH, Stockholm during February 2001. This was described as part of Deliverable 4.1. Subsequent developments have taken into account our experience showing ToneTable to people at that workshop, on numerous formal and informal occasions in our laboratory, and, most notably, at a public exhibition. We have continually refined ToneTable and the strategies for mixing media that it explores so as to attempt to develop a smoothly working installation which is practically viable as a potential component for a public exhibit. In addition, we have undertaken some video analysis of people's use of the table and, in direct influence of the work emerging in Workpackage 2, refined ToneTable mindfully of the phenomena which are being documented there. The current chapter documents our latest work and its connection to the design sensitivities in Workpackage 2.

In the SHAPE project, ToneTable is a major vehicle for exploring a variety of strategies for combining graphical/visual with sonic/musical materials, within non-trivial multi-participant interactive formats so as support exhibits and installations which are sensorially rich and engaging. Primarily, ToneTable has been intended as a 'research vehicle' enabling us to explore interaction-graphics-sound relations without being committed to particular ambitions for its ultimate exhibition (e.g. as part of a Living Exhibition). Indeed, Deliverable 4.1 emphasises how it was developed to embody 'abstract, yet suggestive content'. Nevertheless, ToneTable has attracted some attention as a potential vehicle for the demonstration of various mathematical concepts relationships (e.g. wave dynamics), so it may yet form part of the exhibitions the project participates in. For the time being, though, it is best for the reader to think of ToneTable as a demonstration of certain interaction and inter-media design principles, albeit one with a greater attention to aesthetic detail than is common in 'demonstrations'.

Research Background: Mixed reality interaction surfaces

In the ESPRIT I^3 eRENA project, KTH developed a number of interaction surfaces based around a table on which visualisations could be projected. Let us give a short review of this research and its background literature as it provides the relevant research context for ToneTable.

A number of researchers have worked on interaction surfaces which present computer graphical information embedded within a table-like overall design. For example, the InteracTable (http://www.darmstadt.gmd.de/ambiente/activities/interactable.html) developed at GMD uses a large projection onto a table top with information manipulation being supported by pen and finger-touch based interaction at a touch sensitive surface. Local infra-red networking allows other devices to be brought to the table for interaction purposes (see also the discussion of InteracTable in Chapter 1).

A further development of this concept is to combine the manipulation of specially designed physical objects on the surface with a projection of a computer graphical world onto the surface. DigitalDesk (Wellner, 1991), Bricks (Fitzmaurice, Ishii and Buxton, 1995) and phicons (Ishii and Ulmer, 1997) are all concerned with the combination of computational media with a physical device or display surface. Several applications have been shown to successfully integrate physical interaction handlers and virtual environments or tasks, as in the system BUILD-IT (Rauterberg et al., 1998), where engineers are supported in designing assembly lines and building plants, or in URP (Underkoffler and Ishii, 1999) where a physical interface is used for urban planning, or the concept of 'Embodied User Interfaces' (Fishkin, Moran and Harrison, 1998) where the user physically manipulates a computational device.

In the table environment of Rauterberg et al. a menu area is proposed for object selection that, thereafter, can be placed on the virtual floor plan by moving the interaction handler. This approach uses the physical object as a general interaction device. The physical objects that are used in Underkoffler and Ishii for the urban planning example are mostly used in a less generic but more specific way which lowers the chances of errors due to user input, e.g. a building phicon would less likely be used as something else than a generic brick object.

Another approach is reported in Ullmer, Ishii and Glas (1998) where physical objects, the so called ‘mediaBlocks’, are used as digital containers that allow for physical manipulation outside of the original interaction area.

Work involving KTH in the eRENA project extended these approaches in a number of ways. First, we introduced a context sensitive functionality to the physical objects a user interacts with. That is, the exact significance of an action on a physical object can change in relation to the context in which the action is performed. This enabled us to support several different kinds of user action without proliferating the number of phicons which needed to be used and identified. Second, we propose a setup that combined physical interaction with abstract visualisation in an application that is not concerned with the off-line design of an environment, but real-time intervention in an environment. Finally, we emphasised the overall working ecology in which the physical interface we prototyped was designed to fit. We imagined a room-sized cooperative environment where physical interfaces might enhance and add to traditional interfaces and work activity. This concern for realistic cooperative working environments is rarely emphasised in the design-led demonstrations of physical interfaces and tangible bits which are commonly reported.

Hoch, Jää-Aro and Bowers (1999) describe The RoundTable in which a visualisation is projected up onto a table surface. On the table surface, a small number of phicons can be placed, which can have a variety of effects on the visualisation. The phicon positions, orientations and identities are extracted from video which is captured by a camera positioned above the table. Hoch et al. (1999) describe an application in which movements of the phicons control, amongst other things, the deployment and movements of virtual cameras in an on-line collaborative virtual environment, the table top visualisation providing a map-view of the overall environment. In an extension of this work, Bowers, Jää-Aro, Hellström, Hoch and Witfield (2000) describe an application of The RoundTable in which the positioning of objects on the table surface mixes sound sources, a kind of mixed reality mixer desk. The position, orientation and identity of objects in the visualisation denote sound sources, while the position et cetera of phicons placed on the surface denote virtual microphones with the mix at a selected virtual microphone being computed and rendered on a stereo loudspeaker system.

This work was taken as a starting point for the Second SHAPE Workshop in which ToneTable was developed. KTH initially sketched the idea of building a physical environment with a multi-speaker array around its perimeter and an interactive table centrally placed. The proposal was for activities at the table to influence both computer graphical projections onto the table surface and the mixing and spatialisation of sound to the enveloping multi-speaker sound system. In a preliminary way, such an environment would instantiate a number of the features of interest to the SHAPE project. By combining interactive computer graphics with sound control, we would be examining a combination of media and sensory modalities. By enabling physical interaction in relationship to graphical displays, we would be ‘mixing realities’. We envisaged this construction as at a room-sized level of scale. We were proposing that a display surface and sound environment would be the main ways in which participants or users would encounter our artefacts: supporting conventional computing technology and interfaces would be hidden. In this way, a participant’s encounter with the environment would not be one based around a computer screen and its conventional peripherals. In all these respects, we regarded the proposal to build a graphics/sound environment of this sort to be grounded in the interests of the SHAPE project and on-topic for the Disappearing Computer programme.

To date, we have preferred a development strategy which has concentrated on carefully designed relationships between computer graphics and digital sound. ToneTable has not yet incorporated any direct physical-tangible means for interaction with the display surface such as the phicons of The RoundTable or the touch sensitive surface of InteracTable. Rather, as we shall see, conventional trackballs located around the display surface have been preferred (it is worth noting parenthetically that this, at first sight, limited design choice has a number of interactional benefits). It should be easy to see though that ToneTable could be extended in the direction of more tangible interfaces and, indeed, the 'plinth' described in the next chapter, and which does support the use of phicons, took some of its motivation from the idea of a 'cut down ToneTable'. This brief anticipatory comparative discussion enables us to place ToneTable within the typology of mixed reality artefacts essayed in Chapter 1. ToneTable falls under the category of 'roomware' in that it is an (informational) artefact which embodies a sense of (real-world) furniture and room-sized scale. The particular novelty of ToneTable lies in the attention paid to inter-media relationships in a roomware artefact and how these relationships have been designed with a social scientific understanding of how people engage with artefacts in public places in mind.

Social Interaction in Public Places

On the basis of a series of social scientific studies, work in SHAPE's Workpackage 2 is developing a characteristic perspective on how people interact with and around artefacts in public places. Deliverable 2.1 ends with a series of 'design sensitivities' which mark a first step in reflecting this work back into design. A number of these sensitivities are worth pulling out as they have been kept in mind in the initial design of ToneTable and, even more explicitly, in its subsequent development and study. We itemise the following two areas of concern in particular.

- **Multiple forms of participation.** People manifest a variety of different orientations towards artefacts, installations and exhibitions. There is a range of forms of participation – central/peripheral, active/passive, overhearer/overseer et cetera – which need to be taken account of. Visitors who are alone, and those who come with others, need equally to be accounted for. If possible, one should design so as to support the simultaneous coexistence of these multiple forms of participation in an 'ecology of participation'.
- **Interaction and co-participation.** Interaction should not refer to just the interaction of a single 'user' with an exhibit but should address the multiple ways in which people engage with each other in, around and through the artefact. This may involve providing "enhanced or variable functionality when participants interact with each other in and through the exhibit" (Deliverable 2.1, Chapter 6). It also is important to recognise that participants commonly instruct each other in what they are or should be seeing, hearing or doing.

In developing ToneTable, we have worked with some specific design concepts to try to respond to these design sensitivities. We introduce these very shortly. However, for clarity of exposition, it is best if we next describe the initial design of ToneTable in more depth.

ToneTable: Interactive Graphics

ToneTable is a sound and computer graphics installation which enables up to four people to collaborate on exploring varied dynamical relationships between media. Physically the installation consists of a table as the focus of a room-sized environment which also contains a multi-speaker sound system. Top-projected onto the table is a visualisation of a real-time updated physical model of a fluid surface. The 'virtual fluid' has its own autonomous flowing behaviour, as well as being influenced by the activity of participants. Floating on the surface are a small number of virtual objects (initially, five). These move around the display in response to the dynamics of the modelled fluid surface. Through the use of trackballs, participants are able to move sources of virtual 'wavefronts' around the display. These wavefronts further perturb the virtual surface and enable participants to 'push' the floating objects. If the local force upon a floating object exceeds a certain threshold, the object suddenly orbits around the display before gently coming to rest and resuming the more gentle meandering behaviour characteristic of the objects moving as a result of the flowing surface alone. This sudden interruption in object-behaviour is intended to add interest to the graphics as well as being an outcome that is easier to achieve through concerted collaborative activity between participants. Thus, the threshold for the occurrence of orbiting behaviour is set so that it will tend to be exceeded by a local force produced by two or more proximal wavefronts. That is, two or more participants need to align their perturbations of the surface to produce the orbiting effect.

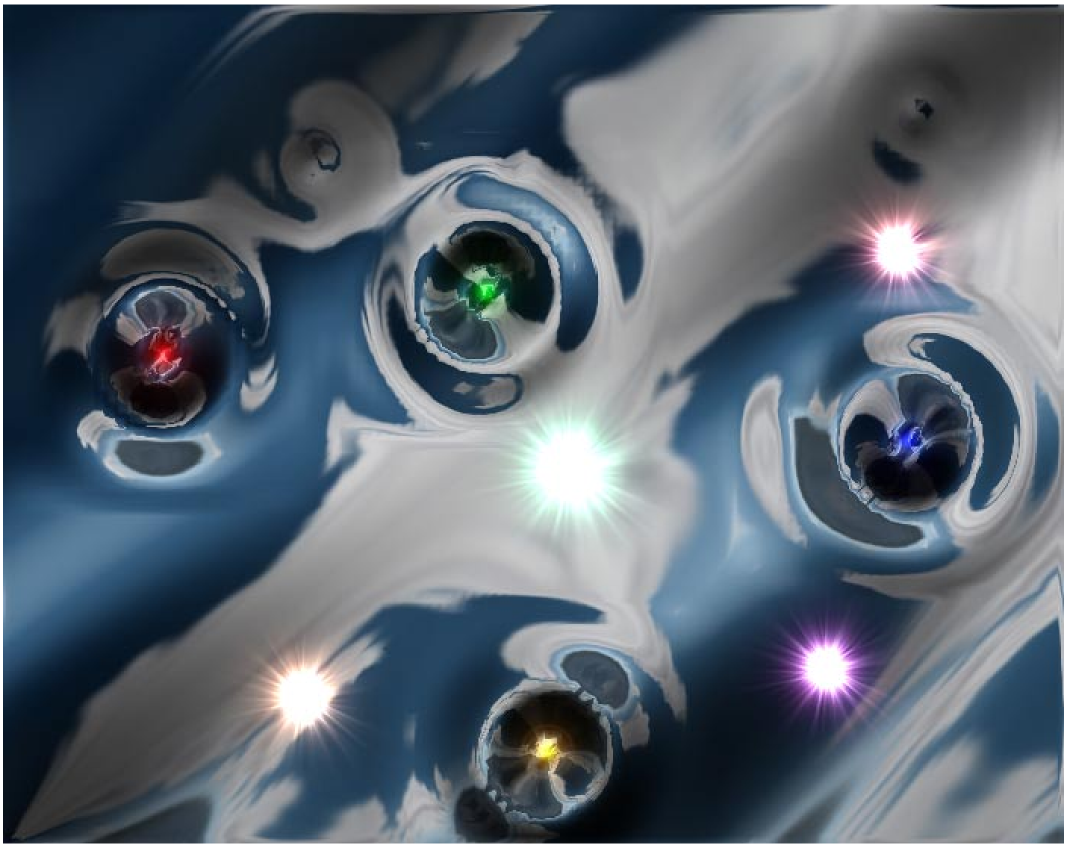

Figure 6.1. The graphical environment of ToneTable.

ToneTable: Sound Environment and Sonification

To achieve a mixed media installation, several notable features of the interactive computer graphics have sonic correlates. The floating objects each have a sound texture associated with them. A set of four speakers placed distally from the table creates a soundfield (approximately 3m by 3m) within which these sounds are heard. The sounds are spatialised so that their position on the table is spatially consistent with their heard-location in the soundfield. If an object gently meanders in the graphical environment, so will its location in the soundfield slowly change. If the object orbits the display, so will its sound orbit around the outer four speakers.

Beneath the table is a set of four further speakers, and a sub-woofer. These are principally used to carry sonifications of participants' activity and its effects on the virtual fluid surface. Associated with each trackball is a tone. The greater the movement in unit time of the trackball (and hence the greater the change in position of the wavefront associated with it), the greater the amplitude and high-partial content of the associated tone. The collective activity of participants is also sonified. A measure of the sum of individual trackball movements in unit time is taken, along with a measure of the separation of the four wavefronts in the display. These, when normalised, give two parameter values to a sound synthesis algorithm which generates various species of 'splashing' sounds. Great and little collective activity, close together and far apart wavefronts in the display all produce different splashing effects.

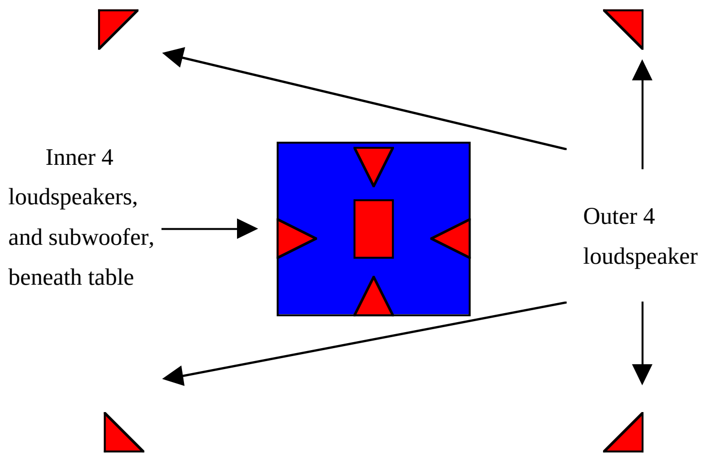

Figure 6.2. Loudspeakers around and under the table.

ToneTable: Implementation

ToneTable has been realised using a variety of inter-working machines, devices, systems and application development environments. MAX/msp (see http://www.cycling74.com) applications were authored to manage the mixing and diffusion of sounds and to calculate appropriate measures of participant-activity and surface perturbation for sonification purposes. Pulkki's (1997) VBAP algorithm was employed to spatially locate sounds. Activity sonifications involved synthesis models implemented on Clavia Nord Modular synthesisers (see http://www.clavia.se). The data from the trackballs is managed using the Multiple Input Device (MID) package developed at the University of Maryland package (Hourcade and Bederson 1999). An OpenGL application was authored to render the graphical surface in terms of the behaviour of its underlying virtual physical model. A local MIDI network linked machines and synthesisers. Where needed, Java/NoSuchMIDI applications provided the glue to attach some machines/applications to this network. (More thoroughgoing implementation details are given in Deliverable 4.1 and not repeated here.)

ToneTable: Critical design features

Let us bring out a number of critical features from the foregoing description of ToneTable, as these express some of our early lines of exploration of principles for the design of interaction for such mixed media artefacts in ways which are responsive to the design sensitivities emerging from social scientific work in Workpackage 2. These include the following.

Layers of noticeability, varieties of behaviour, and structures of motivation

ToneTable manifests a variety of sonic and graphical behaviours which can be progressively revealed through engagement (both individually and collectively) with it. This can give a 'structure of motivation' to its use. That is, we intended to provide an 'in-built' incentive to explore the table and its varied behaviours and image-sound relations. Indeed, in detail, the dynamical behaviours of ToneTable were defined and calibrated with various non-linearities. Our intention here was to make the exploration of ToneTable an open-ended affair with, indeed, some of the behaviours it is capable of being 'emergent' and not necessarily known to the designers in advance. As such we were hoping that ToneTable would make for a contrast with interactive installations where there is a 'key' or hidden, underlying principle that needs discovery and, once discovered, exhausts the interest of the piece. Finally, by 'layering noticeability and interaction' in the manner we have described, allowing the behaviours of ToneTable to be progressively revealed and explored, we wanted to create an artefact which could be explored over various timescales. There is immediate responsivity to use. There are further behaviours revealed with more extended engagement. In this way, ToneTable is intended to give value no matter how long participants have available to engage with it.

Interaction through a shared virtual medium and emergent collaborative value

ToneTable supports interaction between participants through them sharing a virtual medium. By coordinating their activity in that medium, they can engender 'added values': behaviours of ToneTable which parties acting individually do not so readily obtain. However, ToneTable does have a variety of behaviours available when just one person is engaging with it. Its resting state is also not without interest and variety. The intention here is to design an artefact which permits variable forms of engagement, both individual and collaborative, both 'hands-on' and spectating. What is more, by coordinating activity through a common virtual medium, we hoped that participants could gracefully move between one form of engagement and another. They could work individually or in close coordination with others through the use of the same devices and repertoire of gestures. As such, collaboration does not require a switch of ‘interface mode’ over individual activity (cf. the proposals for ‘encouraging collaboration’ in Benford et al. 2000).

Variable image-sound-activity associations

ToneTable relates image, sound and participant-activity in a variety of ways. Sound is associated with individual graphic objects. Sound is also associated with individual device-usage (the trackball tones). And so forth. This variety of strategies was intended to enable an approach to the mixing of media which is rich and more satisfying for participants than if just one technique had been employed. It has the consequence that a single gesture may well produce multiple sonic effects, each associated with a different aspect of it. This gives participants a rich set of resources in terms of which to compare their perceptions of ToneTable’s dynamical behaviour.

Figure 6.3. Publicly exhibiting ToneTable.

ToneTable: Public exhibition and evaluation (I)

ToneTable has been presented to the public on a number of occasions. We will concentrate on two such public presentations as these have been most closely studied by us. ToneTable’s first exhibition was as part of a workshop of the SHAPE project. About thirty people were in attendance. One of us started by giving a welcome and a brief account of the SHAPE project and the Workshop, as well as the broader Disappearing Computer research context. The in-development status of ToneTable as a demonstrator was emphasised, as was the intention to create something ‘abstract yet suggestive’. Suggestions at any level were welcomed. Following this, people were invited into a separate room to see and explore ToneTable. (A detailed discussion of public reaction to this presentation as it initially appeared to us appears in Deliverable 4.1. What follows here is a condensed and, with respect to some details, revised account.)First, from a technical perspective, it can be noted that the setup for the ToneTable worked very well. Using Java for managing multiple input devices and conversion between positional data and MIDI was very straightforward. In addition, the MID package made it unnecessary to use one computer for each input device. The graphical visualisation was effective in the sense that participants perceived it as a watery surface (as intended). The projection onto the table seemed to reinforce this illusion. However, the relatively low resolution of the water surface lattice caused some aliasing artifacts. These did not appear to interfere with the total experience, though. Feedback from participants at the first public demonstration was generally very positive. Indeed, some of it was extremely complimentary. A commonly praised point was that people experienced the ToneTable as having several different behaviour types and relationships between activity, sound and graphics and that these unfolded over time with increasing engagement and prolonged periods of observation.

The sonification of activity at the table was also well received and clearly several participants took some delight in making loud noises with vigorous trackball movements. The fact that a sound could be heard in an immediately responsive way to one’s individual activity through the presence of a tone emanating from under the table gave a clear indication that one was having an effect. The synthesised splashing sound was also appreciated. Good feedback was received about the high quality of the computer graphics and the sound, a quality far exceeding that ever experienced before in a computer-related installation by some of the attendees. Our public demonstration raised a number of interesting critical points and these are worth discussion.

Crowding the space

The room in which we first demonstrated ToneTable could not ‘carry’ a large number of people. While space existed between the table and the outer set of speakers, this could only be comfortably occupied by the four principal participants and a small number of on-lookers. When the environment became crowded, people could find themselves right next to a single loudspeaker and very far from any audio ‘sweet spot’. Indeed, from such a position, they would absorb some of the sound themselves! Generally, we had not allowed for large enough viewing and listening positions, except to support a small number of users and onlookers. Furthermore, we hadn’t specifically designed ToneTable to give a listening position for onlookers. While they might be within the outer set of speakers, their impression of both stationary and moving sounds would have been compromised.

Object-sound associations

While it was clear to participants that their activity was being sonified and that objects while orbiting moved around the sound space, it was not clear exactly which object related to which sound or whether, indeed, there was a fixed ‘standing-for’ relationship. It is possible that five sound objects is too many to individuate in such a setting.

Emergent collaborative value gained too cheaply

While we designed in a mechanism to allow new behaviours (specifically the orbiting animation) to emerge as a result of combined activity from participants, this outcome could be gained rather too cheaply. If two participants just thrashed around with their trackballs, there would be a good chance that sooner or later their ripples would coincide in such a way as to push an object into orbit. Accordingly, we observed few examples of the careful concerted coordination to move objects and yield new behaviours that we were hoping to provide for. Ironically, the sonifications of gestural activity might have been excessively rewarding, as thrashing around would have very notable sonic effects (a louder and more complex trackball-tone, a louder and more complex watery-splashing sound). This might have relatively reduced the incentive to concerted collaborative activity between participants. Finally, the crowding of the space already noted created a situation where participants did not want to overstay their time at the table. Again, this might have not allowed enough time for concerted coordinated activity with a co-participant to be explored.

Gestural legibility

A feature of trackballs (and mice) as devices is that they disassociate the locus of gestural engagement from the locus of display effects. This occasionally made it hard for participants to see which trackball was associated with which set of ripples. In turn, this made it hard to concertedly coordinate trackball activity with another as it would not be clear which other person was producing which effects on the surface. Trackball gestures then were not readily legible to other parties at this first presentation.

ToneTable: Public exhibition and evaluation (II)

In the light of these experiences we made a number of modifications to ToneTable for its second exhibition. This took place as part of the Connect Expo in Stockholm in April 2001, a major Swedish technology fair, where ToneTable was encountered by (we estimate) 600 visitors over a three day period. Our modifications were of four sorts.

Configuring the architectural space

To address some of the over-crowding problems, we gave careful consideration to the environment in which ToneTable would be embedded. Most notably, ToneTable was placed within a plexiglass room-within-a-room. This gave a 5 meter square space which could be occupied by people interacting at or around the table while giving those outside sight of it. The enclosure also contained the sound somewhat so that adjoining exhibits were less disrupted. The dimensions of this space, and the visibility of those already in it, helped to regulate the flow of people in and out, and prevent over-crowding problems. A small notice was placed at the entrance to the plexiglass enclosure describing ToneTable as manifesting varied methods for interrelating collaborative interactive computer graphics and digital sound.

Figure 6.4. ToneTable within a plexiglass room at Connect Expo, Stockholm, 2001.

A more integrated soundfield

In our first demonstration of ToneTable, the sounds corresponding to the floating objects were exclusively mixed to the outer four loudspeakers. When the graphical objects orbited and the sounds moved rapidly around the four loudspeakers, this gave effective results for those close to the table who would be within the ‘sweet spot’. However, stationary sounds tended to ‘collapse’ into the nearest loudspeaker to the listener which was carrying the sound, and a poor impression of location or movement would be given to listeners positioned away from the table. To address this, we mixed a portion of the signal going to the outer four to the speakers under the table. We boosted this portion in the 2.5kHz region. This added notably to the overall liveliness of the sounds, especially when orbiting. It also did something to ameliorate the problem of sources collapsing into the nearest speaker, as listeners both at the table and standing near it would hear sounds from speakers all around them. Finally, distributing the mix of the sounds associated with the floating graphical objects between the outer and inner speakers had the effect of heightening the perceptibility of the associations between the floating graphical objects and their sounds.

Smaller number of object-sound associations

Initially, we placed five graphical objects on the watery surface and associated a sound with each. We found it hard for participants to individuate five and notice the relationships. In our second exhibition of ToneTable, we reduced this number to four, which, together with the other changes we implemented, was intended to enable participants to more readily map particular behaviours with particular graphical objects.

Simplified sonification of gesture

In our first demonstration, a trackball movement would have two sonic consequences in addition to any effects it had on the sounds associated with the floating objects: the trackball-tone and the splashing sound. We simplified this by removing the trackball-tone and just sonifying the overall ‘perturbation’ to the virtual fluid surface through splashing sounds. In this way, we did not ‘over-reward’ large individual gestures, while making the sonification of participant-gesture more coherent. Though simplified, ToneTable still manifested a variety of image-sound-activity relationships and sonification strategies.

Observations

The changes we implemented, though not dramatic, enabled a more satisfactory exhibition of ToneTable than our first workshop presentation. Once again, visitors endorsed the points that were already strong in our first exhibition: the quality of sound and graphics, the existence of different behaviours which could be progressively uncovered. Some visitors were less tolerant than our workshop audience of something ‘abstract, yet suggestive’ and found the exhibition lacking in real content. However, amongst those who were willing to enter in a more playful spirit, and with our changes in place, we were able to see more examples of careful collaborative interaction between participants at the table as, on a number of occasions, people coordinated their gestures to jointly elicit the orbiting behaviour and other effects. The environment did not become over-crowded and the more careful design of the soundfield enabled participants at the table and those nearby to equally benefit from an ‘immersive’ sound experience. Interestingly, although we did not replace the trackballs, the difficulties participants had in the earlier demonstration with working out which trackball corresponded to which co-participant were not noticeably reported. The circumstances of the exhibition as well as our simplified gesture sonification scheme enabled participants to take a little more time to work such details out with a clearer sonification of activity to assist them. Finally, we noted numerous examples of visitors returning to ToneTable, bringing new people with them and encouraging them to explore the table and its behaviours. Again, the exhibition setting, along with our design of a special environment for ToneTable helped people to point the table out and instruct others in its features, even if they did not have hands on at the time.

In addition to making these gross observations, we are able to give a little more analytic detail to some aspects of people’s behaviour in relation to ToneTable. Video recordings were made at various times during the exhibition and, although these materials were not collected in a manner optimal for detailed interaction analysis (sound quality was poor, for example), we are able to illustrate a number of points in more depth. We discuss in turn:

- How ToneTable supports different gesture types with respect to the virtual medium.
- How participants coordinate these different gesture types and how the table enables a variety of participant-behaviours to coexist.
- How participants gesture in physical space.
- How participants manage their comings and goings from the table.

- How different sonic-graphical behaviours emerge through collaborative activity.

Figure 6.5. A variety of gesture types in use simultaneously at the table.

Gestural variety

Participants are supported in their ‘hands-on’ interactions with ToneTable by means of trackballs. Although we have used conventional devices, it should not be thought that there is necessarily anything lacking in them with respect to their usefulness in this setting. Indeed, we have observed a great variety of different gesture types being performed on the trackballs, with correspondingly a variety of different behaviours being achievable in the virtual environment projected on the table and in the soundfield.

As a suggestion of a ‘catalogue’ of some of the gesture types we have noted, let us describe the following.

- **Tickles** – by gently and in turn moving the fingers over the trackball a slow, continual, yet interruptible, trajectory of the wavefront across the table can be sustained.
- **Tremors** – by quickly moving a finger or the palm backwards and forwards or from side to side, the wavefront can ‘shudder’ on the display.
- **Rubbings** – by rolling the palm across the trackball, a large displacement of the wavefront on the table can be achieved. Such gestures have a characteristic acceleration and deceleration and a start-move-stop ‘envelope’. They are often followed by a rubbing in the reverse direction as large oscillations across the display are accomplished.
- **Circular rubbings** – by rolling the palm around the trackball, a large continuous circular path can be inscribed on the display.
- **Single finger rub** – a single finger, commonly the index, might be used to accurately

and delicately position the wavefront at a particular locus in the display.

- **Flickings** – a single finger, again commonly the index, is withdrawn under the base of the thumb and out of contact with the trackball, it is then suddenly released, hitting the ball which turns freely and then decelerates while the flicking finger follows through. This produces a trajectory on the table with sudden onset and rapid movement.

These are just a few of the characteristically different gesture types we have observed at ToneTable. It should be immediately apparent that the trackballs can flexibly support a variety of motions of the participant's wavefront. Indeed, each of the above can be used practically to achieve determinate ends. For example, we have commonly observed tremors and small circular rubbings to be employed in the earliest moments of a participant's exploration of the table. We imagine it is through these gestures that a participant discovers which wavefront corresponds to their trackball and can do so without great disruption to whatever else may be going on at the table.

For their part, flickings elicit sudden onsets and loud responses in the sonification of participant gesture. In contrast, single finger rubs may be used to precisely coordinate wavefront movement with respect to graphical objects on the table and other wavefronts. Finally, with tickles, a continuous movement is produced, as it were, through discontinuous means. A tickle can therefore be swiftly aborted if it seen that a wavefront might interfere with another or otherwise disrupt what someone else might be doing at the table. A tickle can be accelerated or decelerated in response to conditions on the table. It can also be easily interpolated into a flick or a single finger rub. Indeed, we commonly see sequences of gestural activity in which one gesture type is blended into another to achieve a complex compound of responsive behaviour within the graphical and sonic environment of ToneTable. That trackballs are able to support this is good testimony to this simple device.

Coordinating gestures

Our video recordings reveal a number of examples of co-participants closely coordinating the kinds of gestures they perform and their temporal patterning. For example, at one moment, Y initiates a rubbing gesture to perturb one 'corner' of the graphical display. Immediately following this, M moves his wavefront to the same corner and performs the same gesture type. After a couple of seconds of this joint activity, they both simultaneously 'expand' the rubbing behaviour so as to take in more of the display in their wavefront movements with a highly noticeable increase in intensity of the activity sonification accompanying their gestural expansion.

Figure 6.5 earlier shows three people at ToneTable. The two to the right of the picture are both jointly engaged in rubbing gestures, one with the middle and ring fingers in contact with the ball, one with the thumb. They are jointly achieving an extensive perturbation of the virtual surface at the corner between them. For her part, H with her back to the camera and to the left of the picture is rubbing the trackball vigorously with the palm of her hand, producing large movements of her wavefront over the rest of the display. At this moment, then, a pair of participants are coordinating their gestures with each other in close coordination, while a third person employs a gestural type which will enable her to make a big effect but without disturbing them. Importantly, then, the table is able to support the coexistence of a variety of gestural types and activities. It does not enforce all participants to collaborate with one another and is tolerant of variable groupings and foci for activity.

Most notably, we also have examples of co-participants coordinating their activity so as to produce the emergent orbiting behaviour. These we discuss in a separate section.

Figure 6.6. Gesturing towards the table.

Figure 6.7. Gesturing towards the table while co-participants make small rubbing gestures.

Gestures in physical space

So far we have discussed some of the different gestures which we have observed being made with respect to the trackballs and the different effects in the graphical and sonic environment they produce. We have also begun to discuss how participants coordinate their different gestures with each other. In this section, we shall discuss some other kinds of gestures, in particular, those not made on or with the trackball. For example, in Figure 6.6, K points to a region of the display just adjacent to where L and M are making their wavefront movements. In fact, he is using the shadow of his hand in the projection to precisely pick out a graphical object he would like his co-participants to try to perturb.

Gestures of this sort are often precisely timed so as to accomplish a kind of ‘commentary’ or ‘suggestion’ with respect to what is going on within the display, without disrupting it. Equally, activity on the table often accommodates such gestural commentaries and suggestions as they are being offered. In Figure 6.7, G is describing the sonic effects of large gestures and moves his hand across the display, and within its light, to illustrate. While he is talking, D and E initiate gentle rubbing movements which slowly move their wavefronts away from G’s gesture. Following G’s description, D and E then experiment with large, more noisy movements at the table.

In Figure 6.8, we see F pointing at the display while still maintaining a pattern of small movements with her left hand at the trackball. F is drawing attention to a graphical object which has just orbited the display having been perturbed by the coordinated movements of their wavefronts in the region to the left of the display.

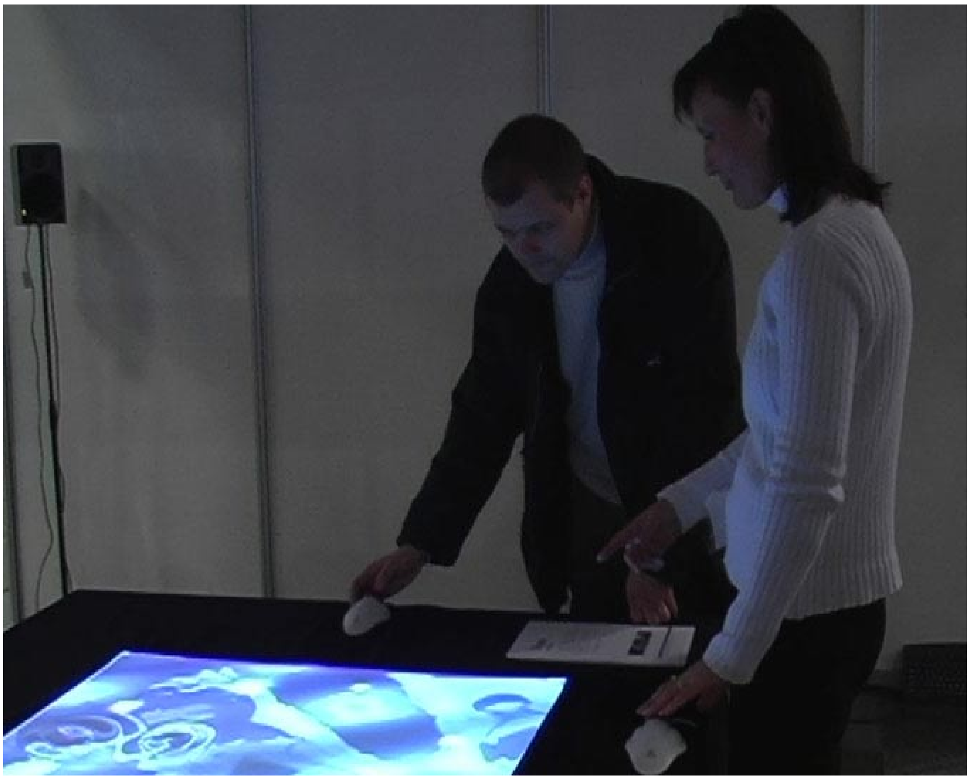

Figure 6.8. Pointing while manipulating the trackball.

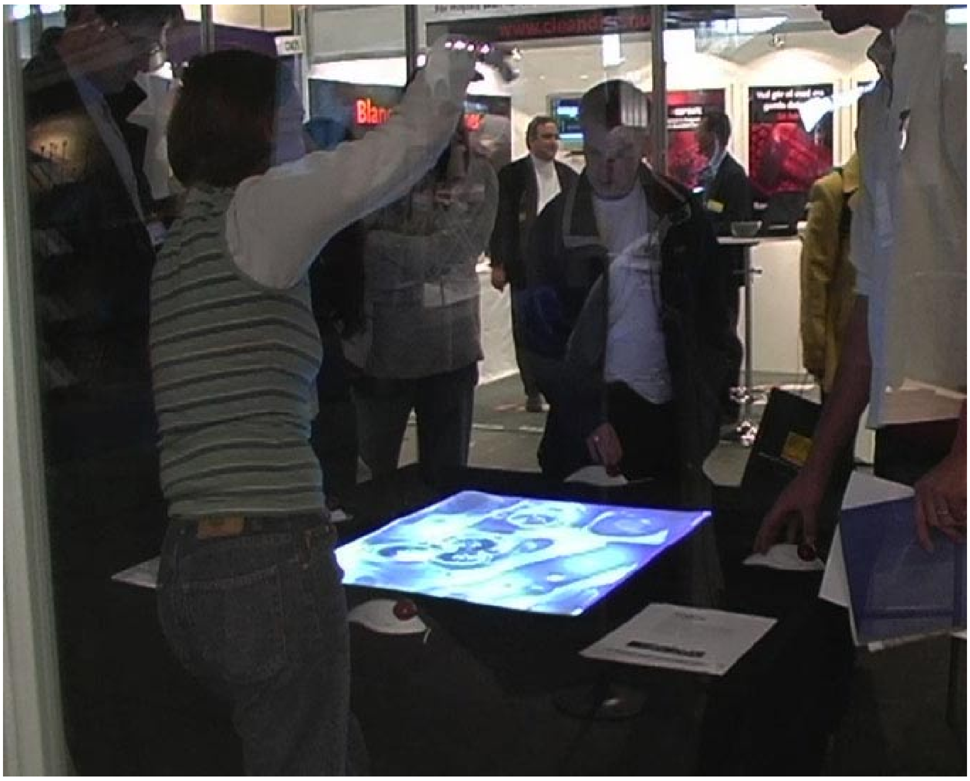

Figure 6.9. Gesturally ‘animating’ the moving sounds.

Similarly, in Figure 6.9, H is making a large circular gesture with her right hand to draw attention to the orbiting of a sound around the room’s soundfield. In this way, she picks out a particular consequence of her activity at the table and draws attention to the relationship between sound and graphics. This occurs just after the moment depicted in Figure 6.5 where H was dramatising the effect of large gestures. The table and her gestural activity with respect to it is enabling H to ‘instruct’ visitors to the installation in the graphical-sonic relationships it contains for her. Throughout all this, two other participants continue to explore the table with smaller gestures.

Coming and going

We have already noted how the plexiglass construction was designed to enable passers-by to have an awareness of activity within while offering an enclosed space appropriate to a sound installation. In addition, we were able to manage the overcrowding which was a problem when ToneTable was first exhibited. We have also described how we designed a more integrated soundfield to give a greater variety of satisfactory listening positions. In all these respects, we have been designing not just for hands-on use of the devices at the table but for a participant’s *trajectory* through the installation. Our design is very flexible in how it allows for ‘comings and goings’. A single person can explore the table, as can a pair both working together or separately. While up to four people can be accommodated hands-on, they can pattern their activity very flexibly. Equally, there is space allowed for others to peripherally participate, perhaps waiting their turn while watching, or allowing a friend to have their turn.

The simplicity of the trackball as an interaction device and the fact that it requires no special ‘tooling up’ or instruction (in contrast, say, to the Glasstrons discussed in Chapter 3) allows comings and goings at the table to be elegantly managed. A visitor can peripherally monitor the action at the table standing close to one of the participants. When that participant gives way, the new person can take over probably having already worked out the associations of particular trackballs to particular wavefronts and having observed a variety of behaviours and gestural types. Our design makes it easy for a newcomer to ‘pick things up’ where an earlier participant ‘left off’ and either extend the earlier person’s explorations or try something new.

Figure 6.10. A precisely timed ‘takeover’.

The frames in Figure 6.10 show S (on the left) claiming a turn at the table as his friend M turns away. Prior to this, M had extensively explored a variety of types of gesture at the table, spending a notable period engaged with it. When S becomes visible in shot he crouches down at the corner of the table reading the leaflet which can be seen resting there. Meanwhile, M is touching the ball lightly with his finger tips. M then removes his hand from the ball and gently curls his fingers into a fist. He then returns to small light touchings of the trackball as S rises and then crouches again. M’s gestures are increasingly ‘sparse’ involving ever lighter, ‘glancing’ contact upon the trackball. M then looks away to his left (Figure 6.10 left) and pulls back from the table. S *immediately* moves in to touch the trackball (Figure 6.10 right), his rising from the crouching position being exactly timed to mesh with M turning away. Of course, throughout his time near ToneTable and particularly while crouching down, S is able to monitor M’s conduct at the table. M for his part designs an extended series of gestures which indicate a trajectory of gradual disengagement. From continuous light fingering, through manually disengaging and forming the notable fist gesture, through intermittent light touches, M is giving his conduct at the trackball a ‘winding down’ shape. This enables S to very swiftly move in once his friend turns his body away.

Emergent collaborative value gained less cheaply

Throughout our work on ToneTable we have been concerned to provide behaviours (both graphical and sonic) which emerge when multiple participants coordinate their activity. Specifically, the orbiting behaviour is more likely to emerge through the careful superimposition of two wavefronts on the display. In our first exhibition of ToneTable, though, we were concerned that the orbiting outcome was gained too cheaply. If a group of participants thrashed around without close coordination of their gestures, they would produce (as if by chance) the orbiting behaviour fairly frequently. The circumstances of Connect Expo were much more favourable in allowing participants the time to try out various patterns of coordinated activity, and we have already noted that coordinations of gesture type, distribution and onset/offset do occur.

In several groups of participants we were able to observe a repeatable pattern of coordination which tended to elicit the orbiting behaviour of the graphical objects and their associated sounds. If two or more participants approach one of the floating objects together following approximately the same trajectory with their wavefronts passing over the object at approximately the same time, then the object is highly likely to start orbiting. By jointly pursuing the orbiting object, the participants are likely to get the object to orbit again once it stops. This strategy of 'co-chasing' one or more objects is likely to systematically elicit the orbiting behaviour and maintain it, if not continuously, then at least prominently. A number of groups of participants realised this and organised themselves to achieve this outcome. In particular, one pair of participants returned to ToneTable on a further occasion with an extra friend so as to more effectively chase the computer graphical objects around the projected display, and make the sounds move around the room.

Conclusions: Designing for mixed realities

In this chapter, we have presented ToneTable, an installation developed within the SHAPE project. ToneTable combines, in a number of different ways, high quality computer graphical and sonic materials in a room-sized environment. We have exhibited ToneTable on two major occasions, as well as demonstrating it on numerous others, and have adopted a design strategy of incremental improvement in the light of experience, while being guided by some substantive design principles and concepts, which we have proposed as responses to sets of social scientific sensitivities emerging from studies of interaction with and around artefacts within public places. Overall, we believe that we have developed an artefact which supports collaboration in and around it, which is tolerant of multiple forms of participation coexisting, which enables people to explore a variety of gestures and concomitant behaviours of graphical and sonic objects, and which has been exhibited with systematic regard for the trajectories people follow as they participate in relation to the artefact at different times and in varied relationship to other people. Furthermore, we believe that we have produced an engaging mixed media installation which is sensorially rich without being overwhelming, and which repays repeated visits.

This is not to say ToneTable is beyond criticism. Far from it. A number of difficulties endure. Let us highlight just three.

- **Content.** Throughout, we have avoided given ToneTable content which is as substantive as, say, the content of the Unearthing Virtual History or Augurscope demonstrators reported in Chapters 2 and 4 of this deliverable. We have preferred something 'abstract, yet suggestive'. While there are good reasons for this, it does yield an artefact which is somewhat wanting both aesthetically (it doesn't have the aesthetic 'pointedness' of an art installation) and practically (it doesn't directly convey content of relevance to our practical collaborations with museums). The future exhibition of ToneTable would have to attend to this, of course. For example, we might consider a different virtual wave dynamics if we were to provide an exhibit designed to teach wave mechanics in a technical museum or similar setting.

- **Object-sound associations.** Some of the sounds in play in ToneTable stood in a one-to-one relationship with particular graphical objects. However, even with a reduced number of sound-object pairings (just four), we do not have evidence of participants commonly 'decoding' the relationships so that they can, say, 'map' the grinding sound to a particular floating object. It has to be admitted that participants were not set this as any kind of 'task' to perform but neither did these *particular* object-sound relationships form part of their spontaneous discourse at the table. Other sound-image-interaction relationships were clear as intended, however. For example, the sonification of activity at the table was clearly notable and, even, 'performable/playable'. A number of visitors compared the installation to, or could imagine an extension of it, as a collaborative musical instrument.
- **'True' collaborative emergence.** While we have referred to 'emergent collaborative value' as a strategy for giving motivation to collaboration, it is questionable whether ToneTable truly manifests 'emergence' in the stricter senses one often encounters in the literature on complexity and non-linear dynamics. To obtain a greater likelihood of novel and unexpected behaviour as participants interrelate their conduct, we simply introduced a threshold in the underlying dynamics. This has the virtue of the dynamics being manually 'tuneable': the threshold can be set to taste with ease. A more thought-through non-linear dynamics could allow for a greater variety of behaviours emerging with different constellations of participants. In addition, a time-varying dynamics (e.g. possibly through the mutation of the underlying dynamical equations or a drift in their parameterisation) would allow for yet further behaviours to be encountered on re-visiting. Such dynamical systems would require a kind of 'in-line' calibration of their equations to user-input. This would be a difficult, yet fascinating challenge.

Let us finish this account of ToneTable by drawing out some lessons of more general interest from our design work and our evaluations of people's experience interacting with ToneTable. We do this under three headings.

Designing for variable participation

When interactive artefacts are deployed in public environments, it is noticeable that people take very varied orientations to interaction with them. They may be 'hands on', 'overseeing', 'passing by', 'in the distance, yet taking an interest', and so forth. They may encounter the artefact on their own or as part of a small group, in the presence of others and other groups, and so forth. An important challenge is to think how these multiple and varied participation formats can be designed for in an integrated fashion when developing an artefact (installation, exhibit or whatever) for a public environment. This is a much more complex question than those traditionally discussed in human-computer interaction research under the rubric of 'usability', and points beyond 'interface design' narrowly considered to the careful design of all environmental elements, both computational and architectural. In our development of ToneTable, we have tried a number of design strategies for addressing such settings. We have explored notions of 'collaboration through a virtual medium', 'emergent collaborative value', 'layers of noticeability', 'structures of motivation'. These are all concepts intended to suggest ways for orienting design for variable participation.

Multiple, coexisting inter-media strategies

We have also explored a number of strategies for relating media. We have sonified device gesture. We have sonified the effects such gestures have on a virtual medium. We have associated particular sounds with particular graphical objects. We have variably mixed sounds to different loudspeaker groupings, these different groupings having different relations with a graphical projection. Our experience is that a rich and varied set of strategies can be made to work together to create engaging environments, though it is important to ensure that one does not build excessive complexity.

Understanding practical contexts

It is important to understand the practical contexts in which artefacts are encountered. Specifics of particular settings may precipitate redesigns (e.g. the way in which we accommodated ToneTable within a larger exhibition) and observations of what participants actually make of an artefact should be taken into account (e.g. redesigning sound diffusion algorithms to minimise the ‘damage’ done by someone standing in front of a loudspeaker). In many ways, the concern to understand practical contexts of use and evaluating real participant-experience becomes more intense the more ambitious one’s design goals are, not less. If we are now seeking radical ways of embedding computation in everyday environments or producing perceptually rich inter-media installations, we need an equally radical understanding of what those environments and people’s activity in and perception of them is like.

Personnel and Acknowledgements

Initially, ToneTable was collectively developed by John Bowers, Sten Olof Hellström and Gustav Taxén (KTH), Steve Benford, Chris Greenhalgh and Claire O’Malley (Nottingham), Jon Hindmarsh, Christian Heath and Paul Luff (KCL), Luigina Ciolfi, Tony Hall and Mikael Fernström (Limerick). The support of Clavia DMI is acknowledged.

References

Benford, S. D., Bederson, B. B., Åkesson, K., Bayon, V., Druin, A., Hansson, P., Hourcade, J. P., Ingram, R., Neale, H., O'Malley, C., Simsarian, K. T., Stanton, D., Sundblad, Y. and Taxén, G. (2000). Designing Storyteling Technologies to Encourage Collaboration Between Young Children, in Proc. ACM Conference on Human Factors in Computing Systems (CHI 2000), Hague, Netherlands, April 2000, pp.556-563, ACM Press.

Bowers, J., Jää-Aro, K.-M., Hellström, S.-O., Hoch, M. and Witfield, G. (2000). Production Support Tools for Electronic Arenas: Using Tangible Interfaces for Media Editing. In Bowers, J. M. (editor) Deliverable D4.5 of the ESPRIT eRENA project, KTH, Stockholm.

Fitzmaurice G, Ishii H, Buxton W (1995), Bricks: Laying the Foundations for Graspable User Interfaces, ACM Proceedings CHI'95, Denver, Colorado, 1995.

Hoch, M., ., Jää-Aro, K.-M and Bowers, J. (1999). The Round Table: A Tangible Interface for Virtual Camera Control. In Bowers, J. M. et al. (editors) Deliverable D4.3/4.4 of the

ESPRIT eRENA project, KTH, Stockholm.

Hourcade, J. P. and Bederson, B. (1999) Architecture and Implementation of a Java Package for Multiple Input Devices, University of Maryland Technical report CS-TR-4018, UMIACS-TR-99-26.

Ishii H, Ullmer B (1997), Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms, ACM Proceedings of CHI'97, pp. 234-241. http://www.media.mit.edu/groups/tangible.

Pulkki, V. (1997). Virtual sound source positioning using vector base amplitude panning. Journal of the Audio Engineering Society, 45(6) pp. 456-466, June 1997.

Rauterberg M, Bichsel M, Fjeld M (1997), Computer vision based interaction with a planning tool for construction and design, Proc. of Gesture Workshop Bielefeld 1997, to appear in Lecture Notes in Artificial Intelligence, Springer 1998.

Wellner P (1993), Interacting with Paper on the DigitalDesk, Communications of the ACM, Vol. 36, No. 7, July 1993, pp 87-96.

Chapter 7: The Invention Observatory: A simple assembly of mixed reality artefacts

John Bowers[1], Sten-Olof Hellström[1], Ian Taylor[2], Martin Flintham[2], Mike Fraser[2], Chris Greenhalgh[2], Steve Benford[2] and Tony Hall[3]

[1] Centre for User-Oriented IT-Design (CID), Royal Institute of Technology (KTH), Stockholm, Sweden
{bowers, soh} @nada.kth.se

[2] The Mixed Reality Laboratory, University of Nottingham, UK
{imt, mdf, mcf, cmg, sdb} @cs.nott.ac.uk

[3] Interaction Design Centre, University of Limerick, Ireland
Tony.Hall@ul.ie

This chapter describes The Invention Observatory. In this, we sought to combine a number of devices and displays, and include some clear, simple instantiations of hybrid, mixed reality objects. We embedded radio frequency identity tags (RFID tags) in a number of everyday objects. These objects (including a light bulb, a mobile phone and a hot water bottle) were selected because, at one time or another, they can be said to have been 'invented'. A 'plinth' was designed incorporating a small projection surface, an embedded tag-reader, and a small sound system of five loudspeakers. Above the plinth can be placed a projector (e.g. to project image material over the real, physical object) and a video camera (e.g. to capture gestures made by someone interacting with the plinth). We constructed an assembly containing a plinth, a large vertical projection screen, and the Augurscope. When an everyday object is placed on a plinth, information about it (e.g. a story of its invention) is projected to the large projection. When a different object is placed alongside it on the plinth, information about the *relationships* between the two objects appears in the Augurscope. These relationships might concern their co-invention by the same individual, their historical connections, their use together in similar contexts or as components to a larger invention. When two selected objects have no known relationship, we make available in the Augurscope a 'fictitious invention', combining features of both in an imaginary object, together with a generated text describing a fictional history. As content, we are proposing this assembly of devices and objects as an 'invention observatory', enabling users to juxtapose objects in an entertaining and informative way. Indeed, even our imaginary objects and generated texts are designed to encourage reflection on the innovation process and the 'received' stories often told about it. As a technical demonstrator, we have an assembly of devices (plinth, projection, the Augurscope) and hybrid objects (RFID-tagged objects with counterparts viewable through the Augurscope) as an environment and set of components to summatively demonstrate the work on Task 1.1, as well as provide further resources for use in the First Living Exhibition (Task 3.1) which will take place in a museum of technology.

Introduction: Workshop objectives

In early October 2001, a Fifth SHAPE Workshop was held in Nottingham. Cross-partner collaboration in this constructional workshop involved KTH, Nottingham and the University of Limerick. Our overall objective in this workshop was to construct an assembly of interworking mixed reality artefacts. As such this workshop was intended to give us some preliminary insight into what would be involved in Task 1.2 when the project's attention more fully turns to combining artefacts in an integrated way. Our aims were primarily technical. We were concerned to prove that the infrastructures we had been working with so far in the project (e.g. EQUIP, see Chapter 2) could support communication between devices in a manner appropriate for our future work in Task 1.2 and in the Living Exhibitions, albeit in a simplified way. While technical integration was our primary objective, we wished to implement content which was interesting, entertaining and potentially of use in the collaborations SHAPE has with science and technical museums.

We were also concerned in the workshop to introduce new artefacts to the mix available to the project, as well as find new uses for existing ones. As well as deploying the Augurscope (see Chapter 4) in a novel manner, we prototyped a 'plinth' for the display and examination of objects placed upon it. Our proposal for the plinth was to embed a Radio Frequency Identity (RFID) tag reader within it, so that it could detect the presence of tagged objects placed upon it. Additionally, we sought to enrich the experience of using the plinth by sonifying the movement of objects and hands in proximity to the plinth. This would also, in some contexts, extend the 'perceptual reach' of an assembly containing the plinth as characteristic sonic changes might accompany its usage. In this way, we wanted to explore the possibility of using sound to give a cue to passers-by of the activity within the assembly. This, like our latest development work for the ToneTable (Chapter 6), indicates our growing sensitivity in Workpackage 1 for the emerging design sensitivities emerging in Workpackage 2. Finally, we wished to explore the possibility of projecting graphical information down upon the plinth and any objects resting on it. In this way, a computer graphical overlay could superimposed over real objects.

This full conception (plinth with embedded RFID tag reader together with a local sound system and provision for overlaid projections) could not be prototyped in the time available (we were unable, for example, to implement a graphical projection onto the plinth) but it is indicative of the kind of mixed reality extensions we are seeking to make in SHAPE. In the terms of the typology of Chapter 1, such a plinth would extend a 'roomware' artefact with augmented reality and graspable components. In an application where the components to an assembly are interrelated in terms of an underlying virtual model, the local sound and graphics of such a plinth could give a 'fragmented' view onto an underlying virtual world.

As a final objective of the workshop, we were concerned to extend the strategies for sonifying gesture that we had employed in ToneTable to good effect. In that work, a measure of individual and collective activity was taken by scaling and normalising trackball displacement data in unit time. This was then used to parameterise sound synthesis algorithms to give a sonic rendering of participants' activity. In the current context, we wished to look at another technology for capturing activity data (video image analysis) to explore its suitability for giving an extended sense to participant's engagements with the plinth, thereby extending the perceptual reach of the plinth itself. Our interest in using video analysis as a sound interaction technology goes beyond this particular context as we wished to add in this interaction technique to the repertoire of those available to use for working with sound in mixed reality settings.

In contrast to our earlier constructional workshops in SHAPE, the scale of the current one was much reduced. We worked together for just 3 working days, rather than the 5-7 working days of earlier workshops. A smaller team assembled too, with advance agreement to a greater depth than in earlier workshops as to what would be constructed. In contrast to earlier workshops, we also preferred a local demonstration to lab members at Nottingham, rather than a public presentation. All of these changes seemed appropriate for the more focussed objectives we were working towards.

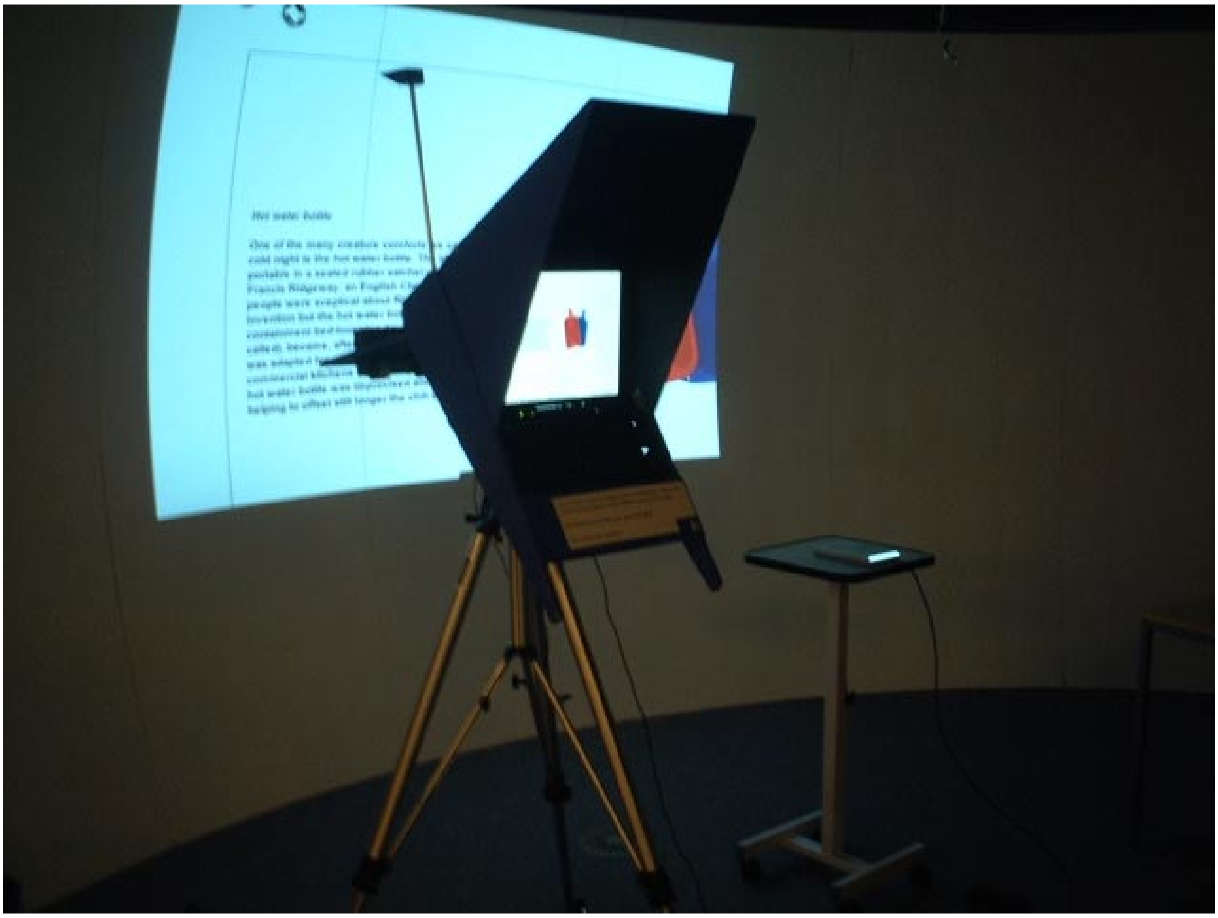

Figure 7.1. Augurscope, large projection and plinth (note: plinth is shown here without camera and loudspeakers)

The Invention Observatory: A simple technical assembly

Our demonstration consisted of three devices working together: the plinth, the Augurscope and a large projection (see Figure 7.1). To most clearly describe how these devices worked together, let us introduce the overall concept of The Invention Observatory which gave content to our demonstration.

Concept and content

The idea of The Invention Observatory is that everyday objects can be brought to a plinth for examination. Placing the real, physical object on the plinth enables further information about it to become visible to devices capable of displaying such information. In the terms of Chapter 1, the everyday object has a hybrid existence as a real, physical object and a collection of informational properties. In The Invention Observatory, the properties of the object which become visible to other devices in the assembly concern its history as an invention. By extension of the ideas in Chapters 2 and 4, it is *historical information* which becomes available to appropriate devices (cf. the use of the periscope and panoramic display in Chapter 2 and the Augurscope in Chapter 4). In The Invention Observatory, an historical account of the invention and subsequent refinement and use of the everyday object in question is displayed as illustrated text on the large display.

When two items are placed on the plinth, The Invention Observatory displays information concerning the *relationship* between them on the Augurscope. That is, the Augurscope is the means for displaying (possibly hidden or little known) relationships between two inventions. These relationships could be related to the circumstances of their invention (e.g. both objects were invented by the same person, or in response to the exigencies of war), their co-membership within a significant category of artefact (e.g. both could be electronic or made of rubber), their co-existence as joint components of some larger artefact, et cetera.

In any population of everyday objects, it is likely however that many pairs do not have a salient relationship. In such cases, The Invention Observatory generates a fictitious invention combining the two objects. For example, a fictitious invention might have the visual form of one object, yet the function of another. Alternatively, it may contain one object within the other as a component. The fictitious invention is displayed on the Augurscope, while the larger display shows historical information concerning the two component objects.

Our intention in exploring fictitious inventions was in equal parts playful and instructive. Naturally, great humour is to be had in thinking of a pseudo-invention which combines a mobile phone with a hot water bottle. However, through this humour, one can present received accounts of inventions in critical relief. Often inventions are a surprising combination of formerly unrelated components. Often inventions do project unusual and, at first sight, fantastic new forms of human behaviour which only subsequently become recognised as legitimate ‘uses’. Our intention with The Invention Observatory was to provoke reflection on the process of invention through the presentation of both historically accurate and fictitious material.

Naturally, for any population of objects of size N, there are N(N-1) pair-wise combinations of them. This mandates an algorithmic approach to the computation of relationships between selections or in the generation of fictions, at least for any large N. The artist Bill Seaman has proposed a ‘Hybrid Invention Generator’ which bears some similarity to our proposals though he does not concern himself with the devices which might be used to display inventions and information about them as we crucially do. Seaman’s project, which is in receipt of funding from Intel, proposes a ‘genetic code’ for inventions which supports the recombinative generation of new possibilities. http://www.cda.ucla.edu/faculty/seaman/texts.html contains a number of Seaman’s writings on what he calls ‘recombinative poetics’ as well as pieces containing some specific information about his Hybrid Invention Generator. Of course, there are a variety of methods for automatically generating material related to combinations of objects. The exploration of these was beyond the scope of our workshop except in a preliminary way (our invention narratives tend to have a recognisable template structure within which it would be possible to substitute elements).

In summary, then, The Invention Observatory enables everyday objects to be examined at a plinth. Information related to the objects is presented on a large display. If a pair of objects is placed on the plinth, information concerning their relationship or a fictitious combination of the two of them is displayed in the Augurscope, with the large display showing information about the objects individually. Throughout, the large display shows information about the objects while the Augurscope is a means for uncovering their relationships.

Figure 7.2. A mobile phone and a light bulb placed on the plinth (note: plinth covered in black cloth; note also: one of the plinth's five loudspeakers can be seen top-right)

Technical Implementation

For this preliminary demonstration, the main structure of the plinth was provided by a platform of the kind used for supporting overhead or slide projectors in lecture theatres and the like (see Figure 7.1). An RFID tag reader was placed on top of this surface. To support the sonification of gestures made with respect to the plinth and of the movement of object onto it, a camera was placed above the plinth's surface directly pointing down. The field of view and height of the camera was adjusted so that the square top of the plinth filled the view. A standard, inexpensive webcam was used with a USB connection to Apple Macintosh PowerBook. This machine was running software developed to use analyses of the video image as parameters for sound generation. The video image was divided into a 3 by 3 grid and the greyscale level in each of the nine regions was periodically reported (at about 15 fps). These values were combined and normalised to give a greyscale value for the front, back, left, right and centre of the image. Associated with each of these regions was a sound file. The level of grey for each region was used to compute a mix coefficient for the associated sound file, the whiter the region the louder its sound. To facilitate contrast with the (white) hands of workshop participants, the plinth was covered in black cloth. Sounds were played back using an algorithm which introduced pitch shifts and crossfades to avoid the sound file sounding looped. In this way, a variety of continually changing sound textures could be interacted with, and with varying responsivity depending upon the movements and gestures of a person at the plinth. Specifically, the plinth would generate a different sound mix depending upon whether an object was placed on it (in the central region) or not and would respond characteristically to a hand and arm being brought in from the side while giving yet another response if someone were to lean over the plinth from the front, and so forth. The video analysis and sound mixing application was developed using the MAX/msp environment (see http://www.cycling74.com) with Singer's Cyclops code-object providing the peripheral video analysis (see http://www.ericsinger.com). A local sound environment was created around the plinth using a set of inexpensive multimedia loudspeakers in a 4.1 configuration. That is, a pair were placed to the left and right of the plinth, a second pair to the rear, with a subwoofer beneath the plinth. Slight frequency dependent phase shifts were introduced to the rear pair so as to create a local, 'immersive' soundfield for anyone interacting with the plinth. The two pairs of speakers made approximately a 2m square. (See Figures 7.2 and 7.3.)

Just three everyday objects were selected for use in this demonstration: a mobile phone, a light bulb and a hot water bottle. Each was given a RF tag with an ID to individuate it. Texts and images were prepared for these objects describing their invention and history. Mobile phone and light bulb are both electronic items, so a text concerning electronics was created for display on the Augurscope when both of these items were brought to the plinth. Fictitious inventions were also described combining mobile phone and hot water bottle, and light bulb and hot water bottle. For demonstration purposes, each combination was hand crafted rather than algorithmically generated. Applications in Macromind Director were authored to manage the presentation of text and image content on the large display and within the Augurscope. The XML mark-up language was used to define simple state machines which would select the correct text and image material given the currently identified RFIDs. EQUIP (see Chapter 2) was used to manage the publication and subscription of data between these various applications and devices.

Figure 7.3. Handling tagged objects at the plinth within view of overhead camera.

The Invention Observatory: A pictorial walkthrough

In this section, we present a pictorial walkthrough of an interaction sequence with the plinth. For simplicity, the plinth is displayed without its attendant video camera and sound system. For clarity, we also duplicate the Augurscope display on the large screen.

The walkthough starts with a tagged mobile phone being placed on the plinth….

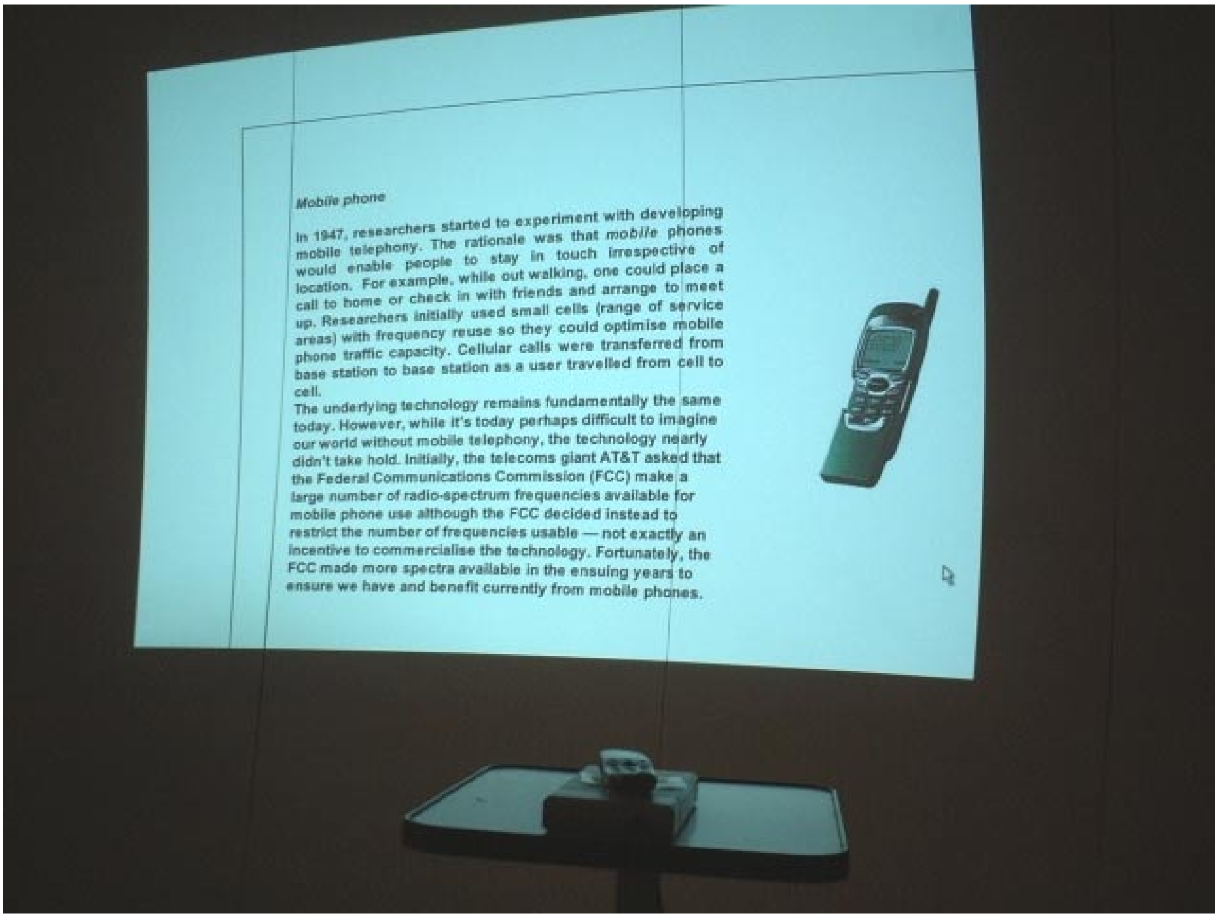

Figure 7.4. Plinth with mobile phone and associated text displayed on large projection.

The identification of the mobile phone causes text and image related to this invention to be displayed on the large screen….

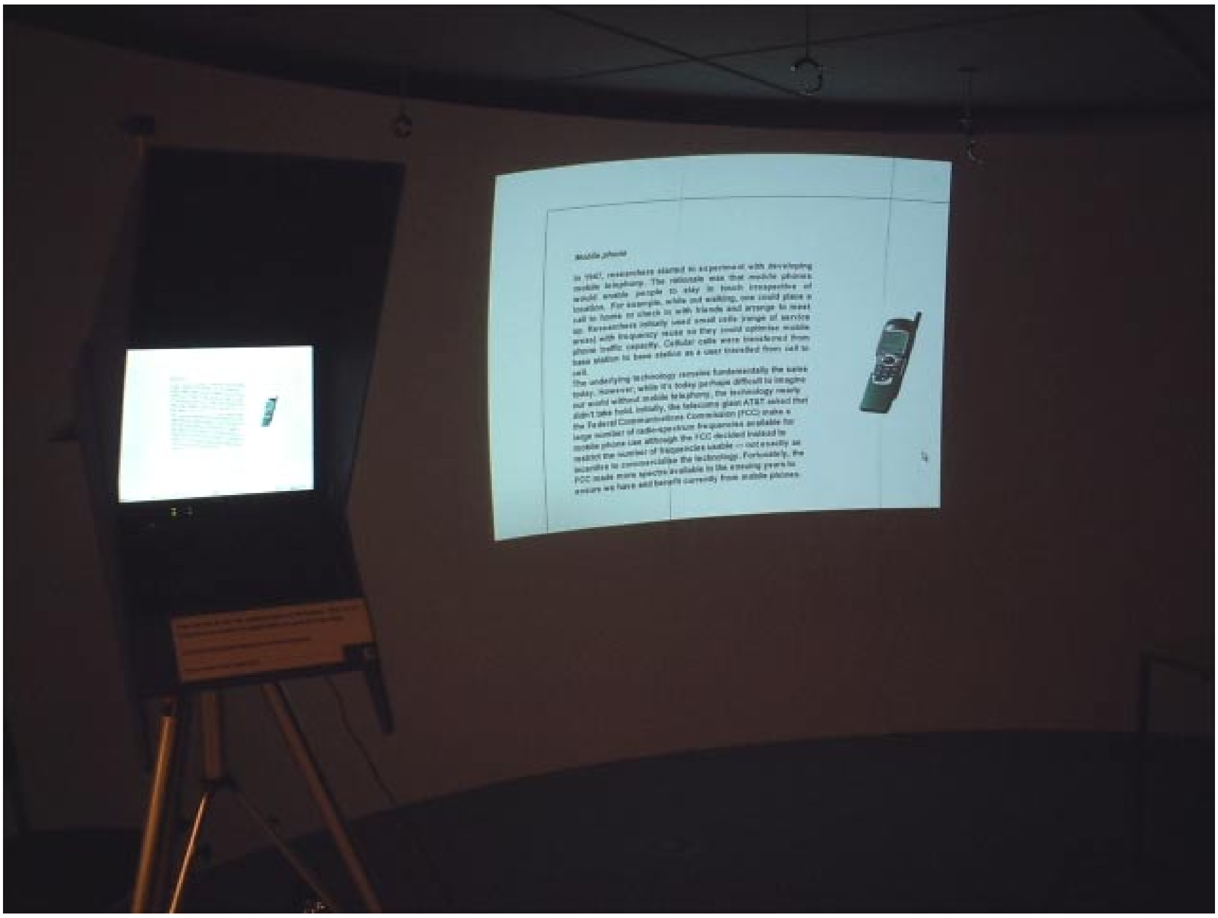

Figure 7.5. Mobile phone related information displayed on large screen and Augurscope.

…. and on the Augurscope….

Figure 7.6. Placing a light bulb alongside the mobile phone on the plinth.

While the mobile phone remains on the plinth, a tagged light bulb is placed alongside it.

Figure 7.7. The large projection shows the relationship between the two objects on the plinth (mobile phone and light bulb) – they are both electronic artefacts.

Information concerning the *relationship* between the two objects is now displayed. The user picks up a hot water bottle in readiness of combining it with one of the objects on the plinth.

Figure 7.8. With a mobile phone and a hot water bottle placed on the plinth, a fictitious invention is generated on the large display (a phone bottle).

The light bulb is removed and the hot water bottle is placed on the plinth. These (otherwise unrelated) objects yield a fictitious invention which combines features of the pairing. An amusing text and graphic describing the invention and use of a phone bottle appears.

Figure 7.9. A hot water bottle is placed on the plinth and information related to this invention appears on the large display and on the Augurscope; then a light bulb is added….

The mobile phone is removed, leaving the hot water bottle. Information related to the hot water bottle as an invention is displayed. While the hot water bottle remains on the plinth, the light bulb is added.

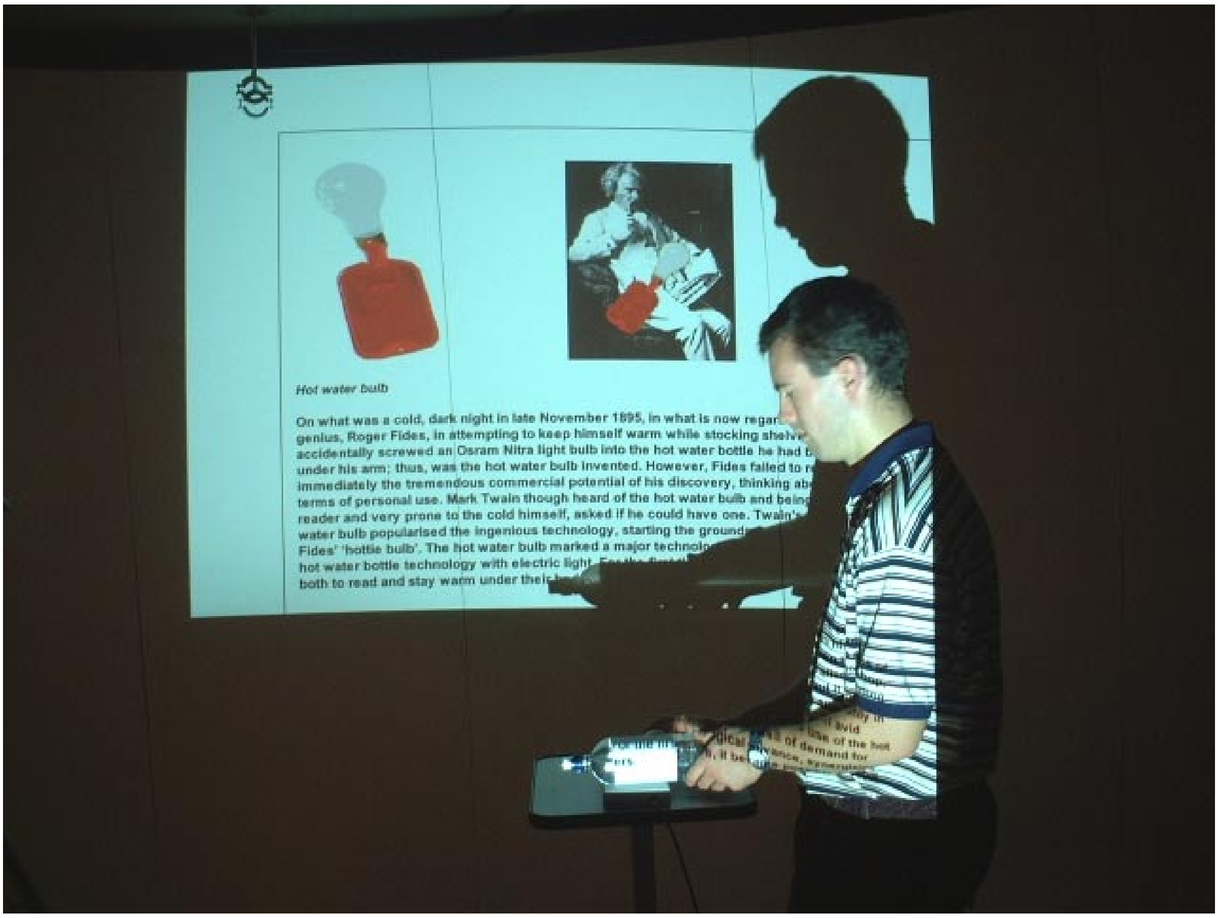

Figure 7.10…. another fictitious invention is generated (the hot water bulb).

The combination of hot water bottle and light bulb gives rise to another fictitious invention.

Early Experience and Future Directions

Although quite simple in conception, The Invention Observatory demonstrates a number of core research concerns in the SHAPE project. It shows how physical objects can have informational (or ‘virtual’) counterparts using the technology of RFID tagging. It shows how different devices can selectively display information derived from the tags which have been identified. It shows how a simple assembly of such devices can be composed each with a specific function. It shows how some of the interaction with sound techniques we have been developing can be deployed to give an extended sense of the gestural activity of participants. Although simple, The Invention Observatory summatively demonstrates much of the work of Task 1.1, as well as introducing a new piece of ‘roomware’ (a ‘bare-bones’ plinth in preliminary form) which can be extended and deployed alongside other devices we have worked on. Let us close with some further brief evaluative remarks before looking to the future of The Invention Observatory and its components.

- **RFID tagging.** This technology was highly effective. The identification of a tag is, for all practical purposes, instant and highly reliable. Any delays which occurred between bringing an object to the plinth and the display of associated text were attributable to other, known sources (e.g. network-related delays), not due to any infelicities in RFID technology. We strongly anticipate making more use of this technology as a technical means for realising ‘hybrid objects’ (see also the following chapter).
- **Sonification.** We feel that the sonification of gesture at the plinth worked very well. Distinctive sound mixes could be heard giving a cue to the status of the plinth and whether anyone was interacting with it. A set of sounds were selected which (together) created a watery-effect. This prompted one workshop attendee to refer to the plinth poetically as ‘the well of invention’! Certainly, when using the plinth within the ambit of the 4.1 sound system, such vivid impressions were encouraged.
- **Problems of video.** Using video to support the plinth sonification was problematic in a number of (known) respects. Video technologies commonly require the control of ambient lighting and, even with the use of a black background cloth, there were fluctuations in ambient lighting which occasionally mixed the sounds in unplanned for ways. This was not always objectionable however as it gave the plinth some apparently autonomous ability to change its sonic character. It has to be noted that a black cloth background is only maximally effective with hands of white pigment, at least if simple greyscale calculation is the analysis method used. While we feel we have demonstrated the in principle utility of video analysis to sonify activity, there is a strong argument for more sophisticated image processing algorithms to address some of these familiar problems.
- **EQUIP.** EQUIP (see Chapter 2) again proved its utility in managing the dissemination of information between heterogeneous devices.
- **XML.** We used XML for defining the state machines which underlay the interaction between IDs and the presentation of materials. This proved to be an elegant solution which enabled us to prototype and test different interaction sequences efficiently.
- **The Augurscope.** The full value of the Augurscope was rather underplayed in this demonstration. In Chapter 4, we see the Augurscope being used to navigate a terrain, viewing a superimposed virtual model along the way. In The Invention Observatory the

mobility of the Augurscope did not really feature. At one stage, we were hoping to develop two plinths and make something of the Augurscope's movement from being pointed at one plinth to the other, perhaps displaying a virtual model containing two virtual plinths on its screen together with models of the objects on the plinths and other information related to them. We realised early in the workshop that there was not time to develop this more extended realisation of The Invention Observatory.

- **The First Living Exhibition.** The Invention Observatory has enabled us to add to the repertoire of devices and interaction techniques which are available to the project to explore in the First Living Exhibition. In particular, the development of a prototype plinth is idiomatic for us in a museum context, and the idea of a museum visitor selecting what object or combination to place on such a traditional display device is an appealing metaphor for enhancing visitor interactivity in such a setting. However, it must be remembered that the workshop was principally concerned with demonstrating a technical assembly of devices and display technologies. Whether such a collection of components would adequately work together to support people assembling a sense of invention (or whatever else one were to examine as content) is a different question, one which we will be able to more fully address when we study our first assemblies of technologies in the hands of the public.

Chapter 8: Equator Related Activities in SHAPE

Steve Benford and Mike Fraser
Mixed Reality Laboratory, University of Nottingham, UK.
{sdb, mcf}@cs.nott.ac.uk

Over the course of the first year of SHAPE, we have been benefited from sharing a partner (Nottingham) in a separate project. 'Equator' is a 10-million pound 6-year programme of research, consisting of 8 UK partners and funded by the UK's Engineering and Physical Sciences Research Council (EPSRC). Equator is dedicated to investigating, uncovering and supporting the variety of possible relationships between physical and digital worlds. In this chapter, we describe developments in the Equator project in which SHAPE personnel have had a part to play. Previously in this deliverable, we have presented SHAPE project output in which Equator personnel have had an occasional part to play. In this chapter, we report the converse; work that SHAPE personnel have profited from by maintaining links with Equator. This includes research on traversable interfaces for public deployment in the form of a 'storytent'; and work involving city-wide augmented reality experiences. We also take this opportunity to describe more precisely the boundaries between SHAPE and Equator research.

Introduction

In its first year, SHAPE has taken part in several mutually beneficial collaborative activities with the Equator project. Equator is an Interdisciplinary Research Collaboration (IRC) funded by the UK's Engineering and Physical Sciences Research Council (EPSRC) to investigate the interweaving of physical and digital interaction. Equator is funded for six years (beginning in October 2000), involves eight UK partners and is led by The University of Nottingham. It has a broad scope, covering the development of new devices, adaptive software platforms, research methods and the application of these to different areas of everyday life. More details can be found at: www.equator.ac.uk.

Collaborations between SHAPE and Equator have taken the form of a series of workshops to develop and evaluate mixed reality applications at which both SHAPE and Equator personnel have been present. These can be divided into two kinds. First have been two SHAPE-led workshops to develop the Unearthing Virtual History and Augurscope demonstrators that were reported in previous chapters. These have focused directly on mixed reality interfaces for museum experiences and have been driven by the SHAPE project with support from Equator. SHAPE personnel have been primarily responsible for concept and content development and for evaluation. Equator has provided software infrastructure in the form of the Equip and MASSIVE platforms as well as input on use of GPS and wireless networking. Both SHAPE and Equator personnel have contributed to the development of devices.

Second have been two Equator-led workshops to prototype new mixed reality interfaces for storytelling and performance in mixed reality environments. The first of these was a workshop focused on early development of a storytent, a tent-like interface to a virtual world. The second was a workshop to brainstorm different ways of overlaying a virtual city on a physical city when outdoors. In both, Equator personnel were primarily responsible for concept and device development. SHAPE personnel were involved mainly in content development and in supporting public experiments.

The following sections report on the two Equator-led workshops. The final section discusses the relevance to SHAPE of the concepts and technologies that were developed.

The Storytent as a Traversable Interface

A team of researchers from the Equator and SHAPE projects along with several external collaborators met in Nottingham for the week of 2nd – 6th July 2001 for a workshop to explore the design of a storytent – a portable immersive environment for interacting with virtual environments. A storytent is a tent-like structure onto which images can be projected (our current design resembles an A-frame tent with a separate projector being used for each side). Speakers and microphones can be located outside and inside the storytent to support sound. Additional technologies might enable tracking of participants and objects moving outside, inside and into and out of a storytent as well as various kinds of interaction with the tent surface.

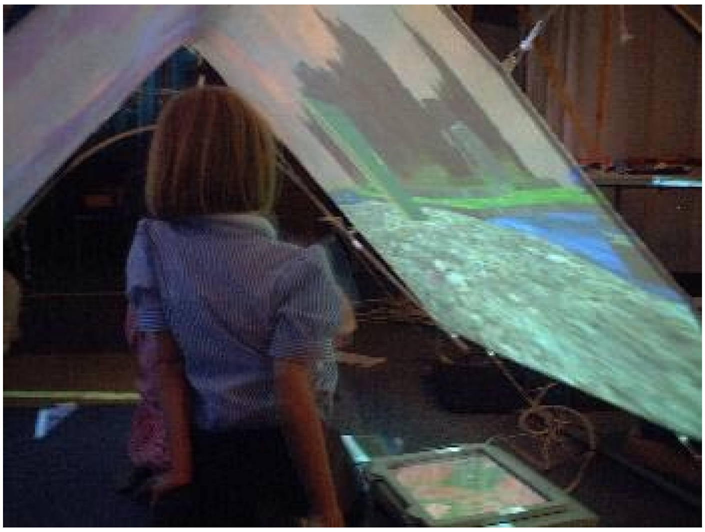

Figure 8.1. The Storytent

At a practical level, the workshop focused on the physical design of the storytent and explored the potential of various interaction techniques. At a more conceptual level, the workshop was intended to inform our understanding of the design of traversable interfaces – interfaces that support some notion of transitions between the physical and digital. More specifically, the storytent allows us to explore:

- traversable interfaces that support object traversals (the tent is such an interface – you step in side of it to pass into a digital space and can carry objects with you as you go)
- the affordances of different interaction technologies (RFID, ultrasonic pens, touch screens, tablet and video tracking) for use with traversable interfaces
- how to design projection surfaces that display interaction to third party observers in interesting ways as well as to direct participants (the tent is interesting because it can easily be viewed from both inside and outside). This is an important issue for use in public environments.

In addition to researchers from SHAPE and Equator, two external collaborators were involved: Hayley Newman, an artist and performer who had previously worked with the team at Nottingham on creang kids storytelling experiences as part of the KidStory project (an I^3 ESE project) and Mat Wand, a musician and sound artist working with Hayley.

The main activities carried out through the week were

- Monday and Tuesday: setting up two storytents that could be used to explore a common virtual environment in the MASSIVE-3 system. Between them, these were equipped with various interaction technologies including RFID for tracking people and objects passing into and out of the tents, touchscreen and tablet based interactive maps for steering a tent, and a mimeo ultrasonic pen system for interacting with the surface of a tent.
- Wednesday and Thursday: exploring the design of RFID aerials that might enable us to identify tagged kids and objects as they move in and out of the tents
- Thursday and Friday: working with Hayley and Matt to evaluate the potential of the different interaction technologies to support different kinds of interaction with a virtual environment. As well as the technologies that had been set up earlier in the week, we also explored (using Wizard of Oz techniques) the potential of video tracking, especially using a camera outside of the tent to track the light cast from flashlights inside the tent onto its surface.
- Thursday: involving two visiting 6-7 year olds for an hour of testing with kids.
- Friday: a final discussion session.

The following summarises the key design issues and options arising from the workshop.

Why the storytent?

The tent is an interesting interface technology because:

- It provides us with an example of and experience with a traversable interface
- It allows us to focus on how objects cross traversable interfaces
- Its size and portability means that we can deploy multiple tents in order to experiment with traversal between multiple spaces
- It offers a two-sided interface and so deliberately supports an experience that is aimed at participants outside as well as inside the tent
- Its two-sided nature potentially enables some interesting ways of improvising and orchestrating interaction (e.g., those outside monitoring and responding to the actions of those inside)
- It is an appropriate and exciting story environment
- It represents an inexpensive immersive environment with potential for future use in

museums, exploratoria, classrooms and maybe even the home.

Discussion of interaction techniques for the tent

We established the following experimental set-up in order to explore different interaction techniques or the storytent:

- Two tents in a single physical space interfacing to a single (i.e., shared) virtual world
- Tent 1: RFID reader, 2 x mimeo, radio mic, speakers inside and outside
- Tent 2: RFID, touch screen map, tablet map, radio mic, speakers inside and outside
- Occasional use of additional computers to control an avatar in the world or the MASSIVE-3 helper-monkey (a dedicated management interface for controlling objects in a virtual world).
- An additional audio system that could route different mixes of the audio from the mics in the tent and external audio sources to the speakers inside and outside the two tents.

Using RFID tags to track participants and objects as they enter and leave

We tested using RFID to track objects and people going into and out of various tents.

- We explored deploying aerials on the tent doors. We got to the point where we could track kids crawling into a constrained opening with tags attached to their backs most of the time. Objects were more of a problem. We tried aerials at various locations around the tent openings but the best we could manage was about a 12 inch sensing distance from an aerial. We have begun to subsequently explore other designs – e.g., loop shaped aerials.
- Tracking objects entering and leaving the tent is more of a priority than tracking people (although both would be ideal).
- We would be able to work with putting objects into a container inside the tent – box, bag, etc.

Our general feeling is that RFID technology might be suitable to time delayed interactions such as triggering quite general effects in the virtual world such as changes in weather or time of day.

Using a Mimeo ultrasonic pen system to interact with the tent surface

We carried out various tests with using a Mimeo ultrasonic pen system (a cheap commercially available package to turn a conventional whiteboard into an interactive whiteboard) as a mouse to interact directly with the surface of the tent from inside.

- We tried variants on both the side and floor of the tent.
- There were interference problems with two sides at once.
- There was interference from people's bodies when used on the floor
- There were calibration problems when trying for the whole tent surface and the resulting interaction was inaccurate with our current tent size.
- Tests with kids found that they found it difficult to use (e.g., not being sure how far away from the surface to hold the pen).
- Although this might be a useful approach, we haven't seen a usefully working test yet.
- We discussed the possibility for embedding in a flashlight if we could get it working.

This technology would probably be more suitable to direct interaction with objects in the world, although currently seems too unreliable.

Using (video tracked) flashlights to interact with the tent surface

We discussed the potential for interaction based on video tracking.

- Early discussions had focused on casting shadows onto the surface of the tent. This could be done with flashlights from the inside or directly by passing in front of the projectors from outside. We had discussed whether to work with kids gestures, but felt that this might be too difficult and had also discussed using shadow puppets to get more reliably shaped shadows.
- In the end we focused more on tracking the light (not the shadow) cast by a flashlight used inside the tent as seen by a video camera outside of the tent.
- We ran some improvisation sessions using a flashlight with a simple cardboard mask taped to its end so as to throw a star shaped light source into the tent. People inside the tent moved this around and those outside improvised various interactions in the world (Wizard of Oz style). An example was trying to catch a balloon that was floating in the virtual world. These sessions were very encouraging. This feels like a fun and appropriate way of interacting and it seems to be technically feasible to at least track the position of the center of the light source as it moves. It may also be possible to recognize its shape, introducing the possibility of multiple flashlights.
- This use of flashlight was aesthetically pleasing both in terms of interaction inside the tent and the effect for those outside.
- There will be issues with people moving outside the tent and accidentally casting shadows that might interfere with the video tracking.
- It will be important to think about when is the flashlight deliberately being used as an interaction tool and when is it just being waved about or being used for some other purpose. A flashlight with a spring switch may be a useful option here.
- We need to think about where to mount the video cameras. On the projectors is one possibility (saves on stands and mountings and gives a god shaped image). On the tent frame is another.

This seems like the most aesthetically interesting interaction techniques, although so far it is completely unimplemented. It might be best to aim for direct but fairly imprecise interactions. For example the flashlight could be used to attract and influence a creature such as a moth that flits around in the virtual world. This would leave lots of space for improvising responses if necessary.

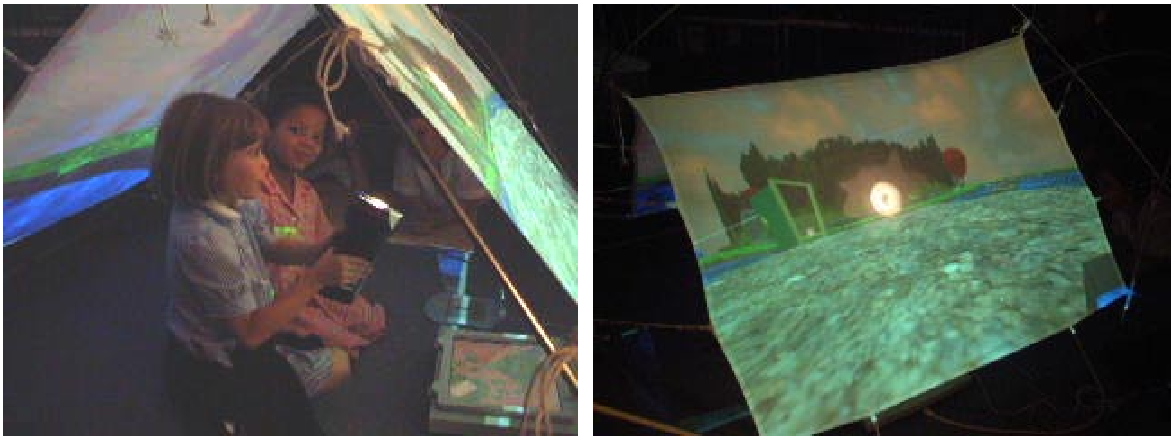

Figure 8.2. A flashlight as a potential interaction device

Interactive maps

We explored two different ways of providing an interactive map within the tent: a touch-screen on a flat panel display and a commercially available tablet/stylus system. The touch-screen option was based on a flat-panel display with a touch-screen displaying a birds-eye orthogonal projection of the world. Touching the map would zoom the tent view to the selected location. Dragging your finger across the map would let you steer the tent. Because we used some existing code for our interactive table-top this was configured so that the tent was always orientated to face the center of the world.

Our kids seemed to get some sense of the map this after a while, although finding specific objects (such as moving near a balloon) was quite hard, perhaps due to the reorientation of viewpoint and fairly coarse scale of the map. It might be interesting to use a simplified more abstract map with few selection points and larger selection areas that take the tent to a carefully planned pre-set position.

We saw some competition for the display at first – two kids touching the screen at once which doesn't work. This seemed to settle down a bit after a while.

We also experienced the problem of one of the kids trying to use the flat of her hand to interact instead of her finger.

An advantage of using the touch screen in this way is that the map can be active or live. It can show the dynamically changing positions of objects, characters and tents within the world. It could update its display according to where the tent had already been or according to which tagged objects were in the tent (physical objects in the tent might sow up as virtual objects on the map display).

The second option was a WACOM tablet and stylus. A paper map (screen shot of the birds-eye orthogonal projection) was placed upon it and the stylus was used to select location. Otherwise, it worked the same as for the touch screen.

We saw competition for the stylus (verbal requests such as 'it's my turn'), but because there was only one stylus there was only ever one kid interacting at a time.

The tablet did not display a live map. On the other hand, the use of a paper map brings some advantages:

- Kids can draw on it
- They can even draw their own map that they take between tents
- The use of a physical map provides a tangible quality that allows us to consider its properties for mixed reality applications.

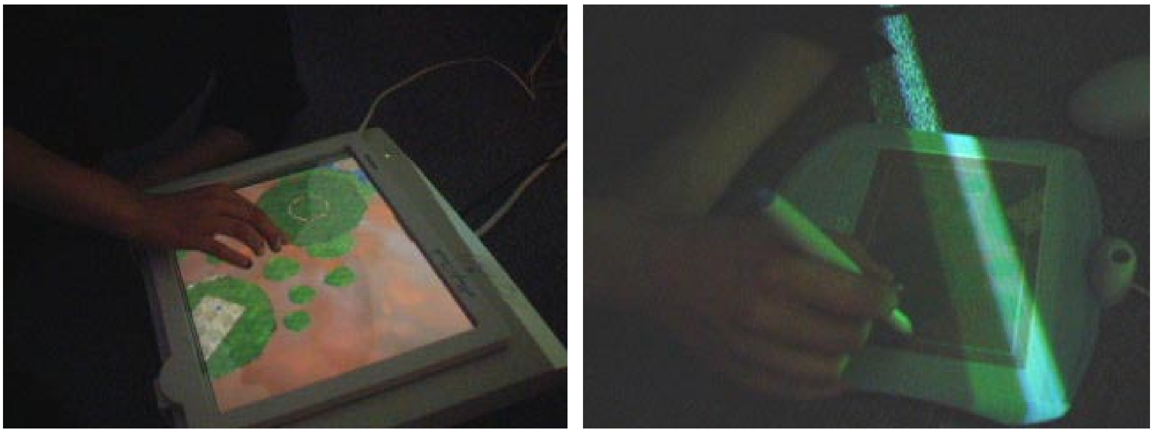

Figure 8.3. The touch screen (left) and tablet (right) based interactive maps

Overall, the maps seemed to be useful and potentially usable. Obviously, they are good for navigating viewpoint. Would it be possible to combine both to get a live display with some kind of transparent physical map overlaid?

Physical tent design

We briefly reviewed the physical design of the tents. The current design aims for:

- Portability – folding frame for easy positioning and can be disassembled and carried in a bag
- Openable – the tent can easily be opened up to varying degrees, right the way to being a plat screen that can be propped against a wall.

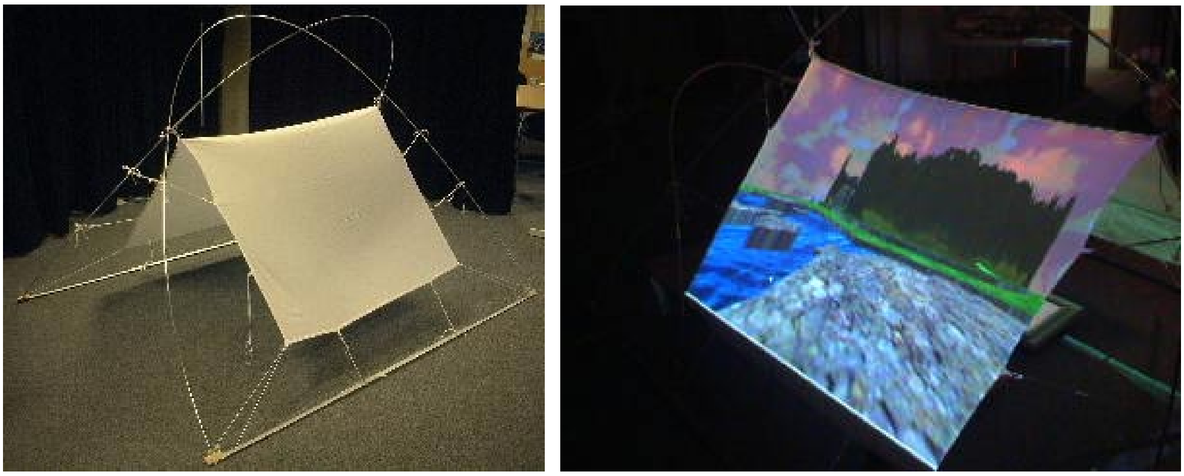

Figure 8.4. The physical tent design (left) and in use (right)

Taking Augmented Reality Outdoors

A team of researchers from the Equator project, with the additional involvement of researchers from SHAPE and three members of the performance group Blast Theory met in Nottingham from the 9th to the 13th July for a weeklong workshop to interfaces for explore outdoors augmented reality as part of Equator's Citywide Performance project.

The aim of the workshop was to carry out a series of practical tests of different technologies that might be used to experience a collaborative virtual environment when moving around a physical city. The technologies included portable devices with GPS and WaveLAN being used to track participants around a region of the city; and projecting graphical shadows of avatars and other objects in virtual worlds into public spaces at night. The workshop unfolded as follows.

Monday

Blast Theory arrived at Midday and we began planning the rest of the week. We also began gathering and reviewing supporting materials including maps of Nottingham and a 3D model of Nottingham's Market Square that we had acquired from some colleagues at Nottingham and imported into the MASSIVE-3 system.

Tuesday

In the afternoon we carried out our first test in the Market Square in the centre of Nottingham. The plan was for a participant to wander around the square carrying a laptop equipped with a GPS receiver and displaying the 3D model of the square. The viewpoint on the model would automatically be updated from the GPS position so that the participant moved through the virtual Market Square as they moved through the actual Market Square. The main aims were to test GPS accuracy and software and to gain some experience with using computers in outdoors public spaces. Technically the test failed as we could not get the software working. It also rained (which it did on three of the four days in which we worked outdoors). People seemed interested in what we were doing from a distance.

Between 9:30 PM and 12:30 we toured around Nottingham carrying out a series of nighttime projection tests. The aim was to broadly explore the kinds of effects that we might be able to create using different sizes of projectors, with different styles of graphics in a variety of locations, under varied lighting conditions, and on different surfaces. Our set up involved:

- A van containing a petrol driven power generator
- Three projectors - a very-large (two person lift) long throw projector, a medium size projector typical of the kind currently used for presentations in small rooms, and an A5 notebook size projector that was small enough to be lifted up and pointed around by one person.
- A laptop rendering a viewpoint from a virtual world. In addition to a standard MASSIVE renderer, we created a black-and-white renderer and also experimented with rendering as wire-frame.

We tried projecting in four locations:

- At Nottingham University’s Jubilee Campus - projecting from the Computer Science building over the lake onto trees and on to the lake itself.
- In a quiet back street - projecting onto brick walls, the road, the van (inside and out), houses and people (ourselves).
- In a deserted retail park - projecting onto a ‘COMET’ (electrical appliance retail) shop from a distance of over 200 meters.
- In the Market Square - projecting onto various buildings.

Overall, all three projectors worked surprisingly well and we obtained some interesting effects. Black and white renderings seemed to work best. The long throw projector seemed able to cover a whole building from a good distance and the smaller projectors could also produce bright images projecting across a road. Masking out the edges of the images would probably help. The generator however was noisy and produced noxious fumes.

Figure 8.5. Examples of nighttime shadow projections. Left: Shadow of an avatar projected on to a brick wall from across the street. Middle: Shadows projected on to the side of a shop from a distance of over 200 meters. Right: The shadows on the shop as seen from the location of the projector

Wednesday

The main focus for Wednesday was to re-run Tuesday's failed test of using a laptop running the virtual world software with viewpoint driven by local GPS device in Nottingham's Market Square. This time we were more successful.

The test took the form of a simple game. A participant moved around the Market Square looking for clues. Each clue was a piece of text that had been ripped into two halves. One half was on a physical piece of paper hidden in an envelope in the square. The other was a texture map in the virtual world at roughly the same location. Reading a clue involved finding both halves, and each clue would guide the participant to the next clue. Technically, this worked quite well. The viewpoint in the virtual world seemed to follow the user's position within about four to about ten meters accuracy. This was probably more inaccurate nearer buildings, although this may just have been when the inaccuracies became more apparent to observers. It remains to be explored to what extent the inaccuracy was caused by the GPS or problems with the scale of the virtual model. Other issues were the weight of the laptop (too heavy for a long carry) and of course the weather (it rained again). In fact, our participant ended up being accompanied by quite an entourage, including an umbrella bearer. It was definitely more of a team-game.

We also did some early tests of WaveLAN signal strength at different locations in the Market Square.

Thursday

Thursday involved two key tests. In the afternoon we went back to the Market Square to continue our GPS/wireless device experiments. Our first experiment involved a participant moving round the square with a GPS compass attached to a laptop hidden in a rucksack. A second participant in a parked van monitored her progress on a second laptop that was displaying the 3D model of the market square. This offered a first person perspective of her avatar in the virtual Market Square. Whenever she approached close to a phone box the monitoring participant would ring its number and she would answer it. This worked fairly well providing that she was deliberately hunting for phoneboxes (as they produce relatively quiet rings).

Further experiments involved a more ambitious set-up for three participants:

- one with a GPS device attached to a laptop that in a rucksack. The laptop was equipped with a WaveLAN card that allowed it to transmit its GPS position back to …

- a second in a van parked in the Market Square who could see the progress of the first participant's avatar moving through the virtual Market Square on a laptop. In this case they saw both first person and bird's eye views. They were also using a mobile phone to talk to …
- a third participant who could receive instructions from them.

We attempted various tasks with this set-up. In one, the first participant was followed by the third who received guidance from the second. In another, the first was following the third who was receiving guidance from the second to help them avoid being seen. These tasks worked surprisingly well. The technology held up quite well and the experience was exciting for all involved.

Figure 8.6. Tracking games. Left: laptop for monitoring from the van shows first-person and bird's eye views of virtual Market Square. Middle: the bag with laptop and GPS placed next to the laptop used for monitoring. Right: deploying a 'LAN in a Van'

Perhaps the most interesting insight arose in the area of dealing with intermittent network connectivity. At one point we thought that the GPS was becoming increasingly spatially inaccurate – the physical user with the GPS seemed to be roughly 20 meters away from the position of their avatar. It turned out the problem was temporal not spatial synchronization. Following the loss of the network, the GPS transmitting laptop would buffer its updates and would then begin to send these out again once the network was available. This meant that the participant in the van would begin receiving positional updates that were a few minutes out of date (although still spatially accurate). Of course, this must be fixed by not buffering updates in this way. What was interesting however, was that we naturally mistook temporal inaccuracy for spatial inaccuracy. This suggests that our interfaces need to be more explicit about the temporal validity of the data being shown.

In the evening we went back to the center of Nottingham to try some more shadow projection tests, this time based around Nottingham's Broadway Media Centre in the Lace Market. We tried both long throw projection from a high vantage point onto buildings and onto the ground, as well as horizontal projections across the street into doorways. Again we tried both wireframe and black and white renderings, and the virtual world included a small crowd of moving avatars. The results seemed interesting, although it proved hard to align the projected viewpoint so as to create natural looking shadows and masking the projection to match the shape of a doorway would have helped.

Figure 8.7. Avatar shadow projections near the Broadway Media Centre (the foreground shadows are of actual observers who are standing in the beam of the projector). Left: shadows projected onto the floor from above using a long throw projector. Right: shadows of avatars projected into a doorway from across the road

Implications for future SHAPE work

Involvement in these Equator-led workshops has enabled SHAPE researchers to collaborate with others in order to brainstorm new kinds of mixed reality interfaces. From the perspective of SHAPE, the storytent seems to be a potentially interesting and useful prototype interface for a number of reasons:

- It embodies a novel twist on how to situate an immersive graphical display within a public space and could be extended to address several of SHAPE’s concerns. In particular, it provides a route to creating a boundary between public and more private or protected space while offering a view of interaction to both.
- Its size also makes it feasible to deploy a number of such interfaces throughout a larger space and to exploit relationships between them.
- There is also clear potential to use the tent as a way supporting object traversals, where participants bring artifacts that they have found into a virtual world, building on the approach of our Unearthing Virtual History (Chapter 2) and Invention Observatory (Chapter 6) demonstrators.
- Its physical scale suggests the possibility of integrating the kind of 3D sound work done with the Tonetable. Adding a form of 3D surround sound to the tent could be very powerful.
- The use of flashlights as an interaction device represents an interesting and potentially flexible form of multi-modal interaction.
- Finally, the tent metaphor may be highly appropriate for some museum contexts. One can imagine using it as part of a virtual archaeological dig, a virtual science field trip, a star gazing interface and so on.

The outdoors augmented reality workshop although perhaps less immediately relevant, also raised some interesting ideas for the future. Perhaps the most notable was the use of shadow projections. SHAPE had already experimented with projecting secondary public views from interactive displays (e.g., the projector attached to the Periscope in the Unearthing Virtual History demonstrator). It might be useful to experiment with more shadow-like representations of activity. Equally, with their associations with ghosts and spirits, shadow projections might be particularly suited to some kinds of public historical experiences (for example, city ghost-walks and exploring caves and dungeons).

SHAPE should throw these ideas and technologies into the melting pot as it moves towards planning its first ‘Living Exhibition’ in year two of the project.