GMD Report 75

GMD –
Forschungszentrum
Informationstechnik
GmbH

Wolfgang Strauss, Monika Fleischmann
Roland Ernst, Christopher Liesendahl
Rainer Liesendahl, Werner Magar
Jasminko Novak, Ute Suessbrich
Mette R. Thomsen, Steve Benford

Linking between Real and Virtual Spaces

Second Year Report

for the Esprit i3 project eRENA

Siena, Italy, October 20 - 22, 1999

December 1999

GMD –
Forschungszentrum Informationstechnik GmbH
Schloß Birlinghoven
D-53754 Sankt Augustin
Germany
Telefon +49 -2241 -14 -0
Telefax +49 -2241 -14 -2618
http://www.gmd.de

Anschrift der Verfasser/Address of the authors:
Wolfgang Strauss
Monika Fleischmann et al.
Institut für Medienkommunikation
GMD – Forschungszentrum Informationstechnik GmbH
D-53754 Sankt Augustin
E-mail: Wolfgang.Strauss@gmd.de
Monika.Fleischmann@gmd.de

ISSN 1435-2702

ABSTRACT

This paper is concerned with developing the concept of the Mixed Reality stage as a networked layering of physical space and virtual environments. These environment allow for interaction and communication of multiple users enabled through intuitive, free body interfaces. The deliverable discusses three main issues:

- A characterisation of Mixed Reality approaches and techniques uncovering the requirements for Mixed Reality stages,
- The technical development and implementation of a VRML 2.0 based multi-user system, which is enabled by an intuitive interface environment and supports Mixed Reality stages: the eMUSE - electronic Multi User Stage Environment,
- Trials and evaluation of the eMUSE prototype in a public setting with untrained users and in a theatrical setting with professional performers resulting in a networked interactive stage concept in year 3.

The innovative aspect of this approach is the combination of the VRML based multi-user system, the intuitive interface and the focus on the interactive experience in different settings. The goal is to produce an operating system for Mixed Reality stages through transdisciplinary exchange.

Keywords: Mixed Reality, multi - user environments, free body interfaces, extended VRML 2.0, eMUSE – electronic multi-user stage environment, interactive stage, Vision System, non-linear storytelling structures.

ZUSAMMENFASSUNG

Das Thema dieses Textes ist die Entwicklung des Konzeptes einer Mixed Reality Bühne als vernetzte Überlagerung von physikalischem Raum mit virtuellen Umgebungen. Diese Mixed Reality Umgebung erlaubt mehreren Benutzern Interaktion und Kommunikation ermöglicht durch intuitive „free body“ interfaces. Der Bericht erörtert drei hauptsächliche Fragestellungen:

- die Charakterisierung von Mixed Reality Ansätzen und die Definition von Anforderungen der Mixed Reality Bühne
- die technische Entwicklung und Implementierung des VRML 2.0 basierten Mehrbenutzersystem als electronic multi user stage environment (eMUSE) für die Mixed Reality Bühne ermöglicht durch eine intuitive Interface Umgebung
- öffentliche Proben und Evaluation des eMUSE Prototyps mit einem unspezifizierten Publikum und in einer Theater Anwendung mit professionellen Performance Künstlern mit dem Ergebnis einer vernetzen, interaktiven Bühne.

Der innovative Aspekt dieser Lösung ist die Kombination des VRML multi user systems, des intuitiven Interface und die fokussierung auf die interaktive Erfahrung in den unterschiedlichen Untersuchungs Anordnungen. Ziel ist durch interdisziplinären Austausch ein Betriebssystem für die Mixed Reality Bühne zu entwickeln.

Schlüsselwörter: Mixed Reality, multi - user environments, free body interfaces, extended VRML, eMUSE – electronic multi-user stage environment, Interaktive Bühne, Vision System, Non-lineare Erzählstrukturen.

This work demonstrates actual research undertaken within the the eRENA project (electronic Arenas for Culture, Performance, Art and Entertainment) which is one of thirteen European three years long-term research projects cooperating in the i3 net. The eRENA consortium brings together internationally known digital artists; experts in multi-user virtual reality, art & social scientists, and networking expertise from significant european institutions. eRENA focuses on developing information spaces in which all participants can be mobile and socially active. Members of the eRENA consortium are BT- British Telecom, EPFL - École Polytechnique Fédérale de Lausanne, GMD-IMK.MARS, KTH -Kungl Tekniska Högskolan Stockholm, Miralab - University of Geneva, ZKM – Center for Art and Media Karlsruhe.

Part 1: The concept of the Mixed Reality stage

1 Introduction

This report describes the work undertaken and results achieved in Task 6.2 „Linking between real and virtual spaces“. Workpackage 6, titled „Interaction: Displays, Navigation and Devices for Mixed Realities“, deals with new tools and interfaces for Mixed Reality shared environments that allow citizens to explore new forms of art, entertainment, performance and culture.

Task 6.2 explores the conceptual nature of Mixed Reality spaces. It seeks to build an awareness of the notion of connections and layers between real and virtual spaces in order to support interaction between people in shared spaces.

The focus of our work undertaken in the first year of eRENA was the examination of existing multi-user environments and development of new communication models for them, as well as the examination of fundamental issues of user representation in multi-user systems (Task 1.1). The research findings have been exemplified with a developed conceptual model for a VRML-based multi-user environment, 'Murmuring Fields'.

The results of that work served in the second year of eRENA as a basis for developing the notion of a Mixed Reality stage and for defining the requirements for its implementation. The eMUSE system (electronic multi-user stage environment) has been developed as a platform for realisation of Mixed Reality stage applications addressed to both performing arts and everyday life contexts. The basic concept of ‘Murmuring Fields’ has been significant reshaped and improved based on the results and evaluation of undertaken public trials.

Task 6.2 started in the second year of eRENA and is scheduled to run until the third year. In this period it involves three major public trials and demonstrations to explore different aspects of the development of Mixed Reality technologies from technological, social science and artistic perspectives. The eMUSE system and its two public trials and evaluations described in this deliverable investigate and demonstrate getting feedback from real world experiences and participation of a broad public in the research process.

1.1 Development goals and research method

This deliverable is concerned with methods and means of inscription and integration of virtual spaces into real, physical space. The goal is the creation of environments thatenable the participants in shared and remote physical spaces to interact and communicate through their natural senses: hearing, speaking, gesturing, touching, moving around. A particular challenge is creating a situation that suppresses the underlying computer system into participants‘ background awareness.

Addressing these concerns requires the investigation of the perception of individual senses and their concerted functioning as an integrated whole. A first concrete step in pursuing these goals is the development of a simple test platform as a laboratory for

exploring bodily awareness of space and communication in a combined real-virtual space.

Developing a virtual networked communication space related to real space means abandoning the idea of the mouse and keyboard based WIMP[1] interface paradigm. It leads to developing a design which allows us to see the computer not as a tool to be used, but as space that is to be entered: rooms, walls, corridors, furniture as an architecture that speaks, sings, and offers images as windows and doors to distant spaces. Imagine a physical space filled with data. Here the visitor can rely on the scale of the body and space while at the same time navigating through the "furniture" of data. Data is organised spatially and revealed as the visitor navigates the space. As a Mixed Reality environment the visitor's exploration of virtual space is connected to real space as well as to other participants' experiences.

This encompasses the development of natural and invisible interface environments for interaction and communication in Mixed Reality multi-user systems as well as the investigation of actually occurring communication behaviour. The interfaces to be developed are referred to as free body interfaces and the virtual environments serve the information and communication of participants in such environment. Development of appropriate forms of user representation is also addressed.

Such an approach challenges the researcher and developer with the following questions: What is info-communication space? What is a free body interface? How does one build systems where the mediating role of the computer is suppressed into background awareness? What new interface paradigms are required? How can the notion of „poetic displays" be understood and realised?

Addressing these questions in the context of everyday life and concrete, work-based applications is highly complex. For this reason our research focuses on the situation of a „stage", which is to be understood as an abstraction of the everyday life situation - therefore providing a more precise field of investigation. Relating the stage concept to the idea of digital information space comprises investigation of interactive storytelling methods and the design of non-linear narratives.

From a cultural and historical point of view, the abstracted situation of theatre and its narrative forms originates in everyday communication. This constellation motivates the decision to base the investigation of new digital forms of information and communication on the situation of the stage.

The concept of the Mixed Reality stage that we have developed reflects and exemplifies the embedding of these concerns in a concrete research process. We focus on two concrete test situations: developing the concept of a Mixed Reality stage in public setting with untrained users, and in performance space with professional performers.

Such a method of investigation follows both the overall goal of the development of new electronic arenas (stages) for culture and entertainment, and explores the relevance of its application in everyday life contexts.

[1] WIMP - Windows, Icon, Mouse, Pull Down Menues; refers to the current graphical interface paradigm.

1.2 Structure of the deliverable

This deliverable is divided into three parts:

Part 1 presents a characterisation of Mixed Reality approaches and techniques, seeking for requirements of a Mixed Reality stage and examples of stage presentation. This is followed by the discussion of the development and implementation of the multi user support system and the interface environment.

This is divided into following chapters:

- Chapter 2 introduces the fundamental concept of the Mixed Reality stage and distinguishes it from other approaches to the idea of Mixed Reality, such as augmented reality and Mixed Reality boundaries. This is complemented by a concrete example of a Mixed Reality stage presentation.
- Chapter 3 discusses main issues and requirements for the implementation of the Mixed Reality stage concept introduced in Chapter 2. This is accompanied by the description of how the developed eMUSE system supports the realisation of these requirements.
- Chapter 4 considers the development strategy of the Mixed Reality stage system eMUSE and introduces the rationale for public trials and evaluation described in the Part 2.

Part 2 describes trials and evaluation of the Mixed Reality stage prototype in a public setting with untrained users and in a theatrical setting related to the needs of a performance presentation:

- Chapters 5 to 10 describe the trial at the interactive media art festival Transmediale ’99 in Berlin, where the eMUSE system staging 'Murmuring Fields' was presented as a Mixed Reality stage in a public setting. Approximately 750 visitors tried out the system which was also presented and discussed in the Panel on Interactivity by the authors.
- Chapters 11 to 15 describe the trial at the Symposium “Perception of reality – reality of perception” within the scope of the German Puppet Theatre Festival FIDENA ’99 in Bochum. At this occasion, the eMUSE system staging ‘Murmuring Fields’ was tried out as a Mixed Reality stage in a theatrical setting for performance, and discussed by performers and theatre professionals, as well as by invited guest scientists from other disciplines (social studies, neuro-psychology, literature, art history, music). In particular the trial was observed and evaluated by the eRENA partners from the Work, Interaction and Technology Research Group at King’s College London.

Part 3 describes ethnographic research conducted at KTH into the production of ‘Murmuring Fields’ at the Fidena Festival in Bochum, April 1999. It complements this with further observation drawn from the ZKM hosted Blast Theory workshop and the Out Of This World experiment in inhabited TV, both of which are documented more extensively elsewhere (Deliverables D7a.1 and D7b.1). The emphasis of the chapter is on the practical work that artists and researchers have to engage in to produce interactive media art.

2 Developing the idea of the Mixed Reality stage

The term Mixed Reality is commonly used to refer to environments which combine real and virtual objects with visual representation of real and virtual space. The underlying paradigm is that of an information space merging components of physical and digital information, in different degrees of spatiality. The notion of simultaneous presence and communication between multiple participants in such environments is incorporated to a greater or lesser extent as an extension of this basic idea.

Different approaches demonstrating this understanding of Mixed Reality include augmented reality, augmented virtuality, tangible bits and Mixed Reality boundaries. In this chapter we introduce the concept of a *Mixed Reality stage.* This defines our approach to Mixed Reality as the interconnection of the real and the virtual that produces a new framework for communication and interaction possibilities. Real, physical space is filled with the virtual space and the user's exploration of virtual space is connected both to real space as well as to other users' experiences. This spatial arrangement of real and virtual is the means for creating a situation that connects the participants with each other.

2.1 Approaches to Mixed Reality

There has been a growing interest in techniques for combining real and virtual environments to create *mixed realities* – spatial environments where participants can interact with physical and digital information in an integrated way [Milgram94]. Mixed realities may be shared, enabling people who are distributed across multiple physical and virtual spaces to communicate with one another. A variety of approaches to creating shared mixed realities have been demonstrated, including augmented reality, augmented virtuality, tangible bits and Mixed Reality boundaries.

Augmented reality involves overlaying and registering digital information (e.g., text and graphics) onto a real world scene in such a way that the digital information appears to be attached to physical objects, even as they move about. The physical scene might be the local environment, with the digital information being introduced via a see–through head–mounted display (HMD). Alternatively, it might be remote, being viewed on a video display that is then enhanced with digital information. Early examples of collaborative augmented reality include the *Shared Space* system [BillKa99] in which users share virtual objects across a physical table top and *Studierstube* [Fuhr98], in which virtual objects are also displayed in a physical space between multiple users. Both of these systems utilise see–through head–mounted displays. Systems based on video views of remote scenes are inherently sharable as the video display is usually located in a shared physical space.

Another approach to a shared augmented environment using a physical table displaying virtual objects that can be manipulated by data glove or stylus, is the *Responsive Workbench* [Kru95]. Unlike Shared Space and *Studierstube*, the Responsive Workbench uses shutter glasses rather than HMDs, the table itself being a screen for stereoscopic back-projection.

In contrast, *augmented virtuality* [Milgram94] starts from a virtual world and then embeds representations of physical objects within it. These might take the form of textured video views, for example views of participants' faces on their avatars as in the *Freewalk* system [Naka96], or views of remote physical locations as in the 3–D media–

space interface of [Reynard, 1998]. The latter approach was explored in the first year of eRENA in Workpackage 4, where we introduced the mechanism of awareness driven video quality of service as a means of managing large numbers of embedded video views within a virtual world [Green98]. The projection of live video data of participants in a shared virtual environment into a virtual space was also used in the ACTS project DVP (AC089). Several CAVE™ and Responsive Workbench™ systems were linked via a transatlantic ATM connection in a shared virtual prototyping environment [Kindra98].

An alternative approach to embedding video views is to construct graphical, textual and aural representation of telemetry data that has been captured by remote physical sensors. The approach of *tangible bits* [Ishii97a] involves the use of graspable physical objects called phicons to interact with digital information, for example moving physical models across a table top in order to access a digital map that is projected onto it. This may be coupled with the use of ambient display media such as sound, light and airflow to provide more peripheral awareness of background information, for example, by showing the volume of network traffic as reflections of water ripples on the ceiling. A similar approach was presented earlier [Stra91] in the *Cybercity* system, where one could navigate through a wall projection of a 3D city model by moving a "virtual finger" through the streets of a map projected on a table.

The approach of *Mixed Reality boundaries* involves joining distinct virtual and physical spaces by creating a transparent boundary between them [Benf96]. With this approach, the spaces are not overlaid. Instead they are distinct but adjacent. The occupants of the shared physical space can see into the next–door virtual space and can communicate with its occupants (e.g. avatars within a collaborative virtual environment). In turn, the occupants of the virtual space can see back into the physical space. A distinguishing feature of this approach is that it places equal weight on physical and virtual environments, considering how each can be accessed from the other. It also offers the potential to use multiple Mixed Reality boundaries to bring together many physical and virtual spaces into a larger Mixed Reality environment in the same way that everyday boundaries such as doors, walls and windows are used to structure physical buildings.

A related, yet differently oriented system is the *Communication Wall* [Breit96] where two (spatially separated) halves of a room are joined by augmented reality and Virtual Studio techniques. Participants in a shared session can communicate like sitting face-to-face at different sides of a table, while the remote part is projected on a wall-size display, giving the illusion of a continuing room, respectively.

Mixed Reality may also be applied (or defined) as an extension to video conferencing systems through CSCW (Computer Supported Co-operative Work) and HCI (Human Computer Interaction) techniques [Pekko97]. According to the specific requirements and technical facilities of a particular teleconferencing scenario, a variety of communication channels may be integrated, ranging from real face-to-face communication over VR to Internet contacts and multimedia components. Similarly, one may approach Mixed Reality concepts with respect to application context, e.g. in collaborative interior design [Kli97], where the concept of augmented reality is understood primarily as a paradigm for user interaction and information visualisation.

Freeing the user from being tied to a stationary system is yet another way of understanding Mixed Reality With "Wearable Computing" Systems, one becomes mobile, remaining free to move and act in a real environment while staying informed via a wearable display system that stimulates one or both eyes [Mann98]. There is a

system developed by Sony [Reki97] that combines wearable as well as stationary computers to interactively create and store/retrieve virtual annotations to real objects.

Following Milgram's approach to define a taxonomy of Mixed Reality (MR) visual displays [MilKis94], there are six different classes of MR interfaces, ordered by increasing virtual component influence:

1. non-immersive, monitor-based video displays with overlaid virtual image components
2. immersive HMD-based video displays
3. see-through HMD video displays
4. virtual see-through HMDs via integrated video camera
5. primarily virtual display environments with overlaid video "reality"
6. completely virtual projection-based environments immersing user and surrounding reality as a whole.

Besides the ordering scheme used above, other means of distinction have been suggested, such as direct/indirect viewing of real/virtual objects (where "real" means "directly related to some physical existence" opposed to "virtual" as based on a computer-generated model), producing real/virtual images (i.e. images that do/do not occlude other images further down the viewing axis, respectively). Including the world of Virtual Studios (VS), which can also be regarded as a type of Mixed Reality, another, two-dimensional, classification could be made, based on the degree of immersion (or better: impression of reality) for (a) the person acting in the virtual environment and (b) an external spectator.

Using this classification scheme, digital video post-production and virtual TV-Studio production can easily be integrated with Milgram's Mixed Reality classes, placing post-production in one corner (least actor immersion, maximum spectator reality; changing but rather minimal virtual part), and completely immersive VR systems (using non-transparent HMDs) in the opposite one.

Common to all different approaches to the term "Mixed Reality" are two points:

1) The existence of a combined pair of a real and a virtual space (*Comris, Starlab, http://comris.starlab.org/comris-concepts.html*).
2) Employing the visual as the dominant mode of perception and integration of real and virtual space.

All the different approaches described in the section basically differ in the ratio between those two spaces, and the type of interface between them. While these investigations do research into complex problems such as 3D data retrieval, geometric data of layering provided by complicated tracking systems, special problems of video techniques etc. there is less work on networking issues. In the meantime much work had been done in this field, e.g. the development of VRML and interaction synchronisation and behaviour models in distributed virtual environments. At the technical level, contemporary research in Mixed Reality technologies for the broad public must focus on extending the open questions related to the VRML concept, especially in terms of multi-user communication and extension of computer graphic features provided by Java.

2.1.1 A review of shared space technologies

Current approaches to technologies which enable simultaneous presence of multiple geographically distant participants in a shared space can be classified into five categories:

- mediaspaces,

- spatial video conferencing,

- collaborative virtual environments,

- telepresence systems,

- collaborative augmented environments.

The notion of mediaspaces is used to refer to the "enhancement of existing workspaces with integrated audio and video communication". This differs from multimedia conferencing systems in supporting social browsing, peripheral awareness and the establishment and maintenance of long-term working relationships between physically separated people.

The term “spatial video conferencing” refers to video conferencing systems that attempt to introduce support for determining gaze direction. That means providing a way for participants to distinguish at whom one is gazing, which is normally indistinguishable when several people are presented with the image of someone looking at the camera.

The key concept of collaborative virtual environments (CVEs) is summarised as that of computer generated spaces in which each participant has his graphical representation and can control his own viewpoint and interact with other participants or various representation of data. Such spaces are usually referred to as shared virtual worlds. Typically nominated fields of applications are training, co-operative visualisation, simulation, design and entertainment.

The concept of telepresence is understood as "allowing remote users to experience a remote physical space through computer and communications technologies". Experiencing the space is understood as the ability to view the space, to navigate the space and to interact with objects in the space. A scenario where the remote participant controls a robot which in turn explores the physical space is nominated as a typical application.

The notion of augmented reality is understood as overlaying the views of a real world scene and the virtual scene with some level of dynamic linking between them. Besides using see-through head-mounted displays or overlaying graphics onto conventional video screens, some approaches explore the use of ambient display media such as sound, light and airflow for peripheral awareness [Ishii97b]. The latter is claimed to aim at providing "natural integration of digital and physical information and providing rich and multi-sensory experiences for users".

Relating the notion of interactive environments to the above classification places it across several categories: it involves and examines the concepts of mediaspace, telepresence and collaborative environments. In terms of the classification, interactive environments are mediaspaces which may provide elements of telepresence for multiple participants in a shared space. A major difference is that the notion of telepresence is concerned foremost with allowing remote participants to experience each other - not a remote physical space.

Regarding mediaspaces, the approach of interactive environments is not constrained to a particular scenario and is actually more interested in exploring public space than workspace contexts. It also departs with a much relaxed understanding of "communication", and doesn't necessarily assume geographically separate participants. The term mediaspace is understood as enhancement of physical space with different, most often computer-based, audio-visual media but also the “spaces” created through communication between participants using different computer-based media.

As to the idea of "shared virtual worlds", the notion of the interactive environment emphasises the idea of a shared world as a world of shared experiences through interaction of participants with each other, mediated by the situation that is created by the environment. Rather than interacting with objects in a computer-generated world, the focus is on different forms of interaction between participants or between a participant and his "inner self”. The latter refers to the recognition that the responses or actions that the situation provokes us to are motivated or determined by who we are as human beings and persons that cannot be parameterised and described as yet another "object" of the system.

2.2 The Mixed Reality stage

2.2.1 The basic concept: a room filled with data

The basic concept of the Mixed Reality stage is that of a room filled with data. The “room” stands for physical interaction space but the furniture of data is virtual and stands for an information space. It is a spatially organised information space in which data is revealed through users‘ movement in the combined real-virtual space, and through interaction with other users (Fig 1.). The physical space is filled with virtual space and extended with virtual space.

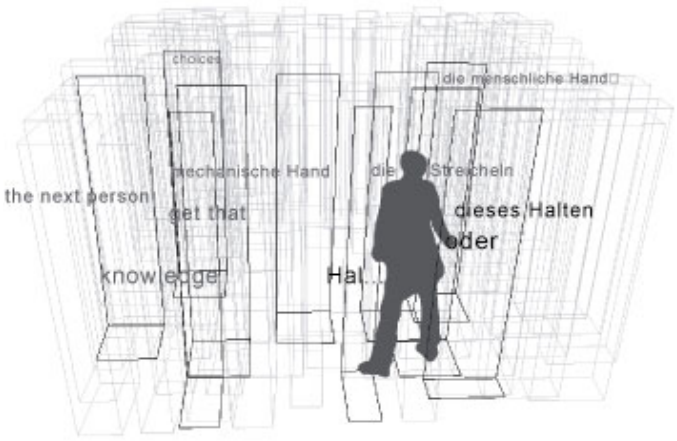


Fig.1 Movement reveals data from virtual space as if it were in the physical space

In this concept, the virtual space is realised as an interactive field of sounds which are triggered by users’ movement and emitted into the physical space. As it is the movement in physical space that causes the sounds, and as the sounds are heard in the physical space, the resulting impression for the users is that of an invisible field of

sounds existing in the physical space and revealed through their actions. If sounds are understood as data of an information space, this illustrates the realisation of the basic idea of using virtual data for filling real, physical space.

2.2.2 Levels of linking the real and the virtual

The concept of the Mixed Reality stage considers connecting the real and the virtual at three complementary levels:

- linking audio-visual manifestations of physical and virtual space,
- linking physical and virtual space through movement and bodily awareness of space,
- linking internal construction of individual experience with externalisation of experiences of others.

This is different from the approaches of augmented reality and augmented virtuality because they operate strictly in the image plane - overlaying images of real and virtual space on a video display. It is also different from the mixed-reality boundary approach where the real and the virtual space are not layered, but distinct and adjacent.

Linking real and virtual space through movement builds on the recognition that movement is an essential means of perception of physical space. Hence, in order to make the perception of virtual space function at the same level as that of the real space, we make movement the basic condition for perceivable manifestations of virtual space. This is why audio-visual elements of the virtual space are invoked only through user's movement in physical space.

The presence of users in the physical space is effected through their bodies, and the movement of bodies describes the spatiality of physical space. Each individual user perceives this in a twofold manner: 1) as a bodily awareness of one's own movement in space, 2) through visual and bodily awareness of the movement of other users' bodies. Connecting this level of real space perception with users' perception of virtual space is the reason why we annotate users' movement in real space with an audio-visual trace of movement in the virtual space.

The resulting manifestations of virtual space, the audio-visual traces of users' movement, and the movement of their physical bodies, provide an externalisation of the users' experience of the shared situation. The mental processes of individual users, such as interpretation, construction and imagination, are externalised and connected to their perception of the jointly constructed reality.

As a result, the nature of the created situation transcends that of a "space" understood in strictly physical terms of the word. What is produced is a new reality of perception as a framework for the communication of different, individual realities of the "same" situation – a situation of mixed realities.

3 Implementing the Mixed Reality stage : the eMUSE system

This chapter discusses main issues and requirements for the implementation of the Mixed Reality stage concept introduced in the previous chapter. This is accompanied by the description of how the developed eMUSE system supports these requirements. eMUSE (electronic Multi User Stage Environment) [Stra99] is the system developed to support the realisation and implementation of the Mixed Reality stage paradigm. It is a VRML-based platform for multi-user interaction and communication, interface, rendering and display organisation, in shared physical and virtual space.

The Mixed Reality concept is based on the following situation: one or several participants in the physical space are simultaneously present in an information space that is made visible only through participants' actions in the real space. The overlapping and integration of these two spatial situations – physical space filled with virtual information space - creates a new reality of perception.

The requirements that such a Mixed Reality model poses on the underlying technical system differ from the requirements of common approaches to multi user or collaborative virtual environments that focus foremost on representation and interaction of participants in a purely virtual shared space. They also differ from the requirements of the augmented reality approach which overlays images of real and virtual space [Kli97], as well as from requirements of the approach of Mixed Reality boundaries which treats real and virtual space as adjacent and distinct entities [Benf96].

The Mixed Reality stage paradigm requires support for different variations of four basic concepts:

- multi-user shared virtual environment, participants in a shared physical space,
- multi-user networked virtual environment, participants in geographically remote locations,
- multi-user mixed-reality environment, participants both in a shared physical space and in geographically remote locations,
- virtual environments supporting free body interaction.

This means that the system supporting the realisation of the described Mixed Reality model needs to combine elements of technologies for networked virtual environments and spatial immersive display technologies with support for free body interaction. Additionally, the modelling of the virtual space has to take into account that its primary purpose is that of materially filling the real space. The material is to be arranged and re-arranged through user's movement and interaction with other users, rather than serving as a predefined 3D scene to be observed and navigated. In particular, four main issues need to be addressed:

- Managing interaction of multiple participants,
- Transferring movement in physical space to navigation in virtual space,
- Connecting the participants with each other,

- Distributing virtual space in physical space – audio-visual display.

The next sections discuss these issues in more detail and describe how they have been resolved in the developed eMUSE system.

3.1 Basic structure of the eMUSE system

e-MUSE is built as a modular system providing independent levels of implementation for rendering and displaying the virtual space, for supporting multiple users in shared physical and virtual space, and for supporting attachment of non-standard, free-body interaction, interfaces.

It consists of three main parts: the external user interface driver, the multi-user driver and a VRML browser (Fig. 2). The external user interface driver allows attachment of arbitrary input devices for controlling movement and navigation in the VMRL scene.. It reads data from input device drivers, transforms it into appropriate VRML co-ordinates for controlling viewpoint and avatar movement, and passes it to the multi-user driver over a socket connection. The multi-user driver takes care of updating the user's local VRML scene accordingly. It runs as a JAVA applet inside the Web browser and uses the external authoring interface (EAI) to communicate with the VRML browser. Changes are propagated directly between the multi-user drivers supporting individual users, through TCP/IP peer-to-peer connections. The modular structure of e-MUSE enables the use of any VRML browser[2] as the rendering machine, provided it supports the external authoring interface.

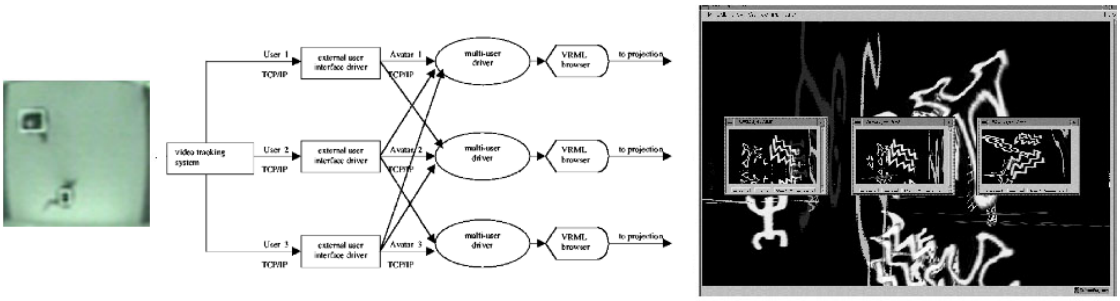


Fig. 2 Basic structure of e-MUSE

Such structuring of eMUSE is the result of focusing on the following functionality:

- support for multiple participants in a combination of shared physical and virtual space,
- support for attachment of arbitrary input devices,
- support for free body interaction,
- support for flexible and dynamic user representations connected to body movement,
- virtual space realised using VRML 2.0 as scene description language,
- system independent of browser choice,

[2] We have used the CosmoPlayer 2.0 as part of eMUSE.

- easy reconfiguration and extension of functionality.

Using VRML as the scene description format is important so that designing a dynamic virtual space requires little or no programming knowledge. Support for attaching different input devices is crucial for experimenting with different solutions to support free body interaction. Flexible and dynamic user representation is needed in order to explore different ideas and notions of (tele)presence.

As the VRML standard develops and new capabilities are implemented in different browsers, it is to be expected that at a certain point one may need to change the browser used. The module providing multi-user support also needs to be replaceable in order to be able to experiment with more sophisticated solutions as the need appears.

Because of the ease and flexibility of the underlying architecture, eMUSE can be seen as an experimental platform for prototyping and trying out new spatial configurations as well as interaction and communication concepts for shared mixed-reality environments. It is also suitable for experimenting with group interaction around common data or shared applications. The eMUSE system itself, along with the virtual environment, can be understood as an interface environment.

3.2 Managing interaction of multiple participants and consistency of the virtual space

Existing systems (DIVE, MASSIVE-2, SmallTool) for implementing shared and networked virtual environments provide different subsets of the requirements needed for the realisation of the Mixed Reality stage concept. The main problem targeted by these systems is the problem of the consistency of the virtual scene database and the related problem of synchronisation of remote participants‘ interactions. Their treatment of these issues provides much overhead for the basic Mixed Reality stage model focusing on multiple participants in a combination of shared physical and virtual space. This usually goes at the expense of the ease of reconfiguring and extending the system. Furthermore, these systems in general address the so-called „desktopVR“ or „window on the world“ approaches where the virtual space is displayed and perceived on a common computer screen.

In contrast to these systems, the most important requirements for managing the interaction of multiple participants on a Mixed Reality stage, with respect to their effects on the virtual space, are:

- dynamic behaviour and creation of objects in the virtual space invoked by user actions in the physical space,
- independent viewpoint control for multiple users connected to actions in physical space,
- independent organisation and display of an arbitrary number of viewpoints of the virtual space,
- support for free body interaction and attachment of different, non-standard interfaces.

No mechanisms for synchronisation of interactions are required for the basic model of several users in a combination of shared physical space and shared virtual space. However, in order to allow easy extension to support remote participants or connection

of several Mixed Reality stages, it is desirable that the system be organised in a distributed manner from the outset.

The simplest and most flexible solution in this case is to employ the homogenous replicated database model for the virtual space, with peer-to-peer TCP/IP for the propagation of changes caused by individual users, to all other participants. Such a distributed model also facilitates the assignment of individual viewpoints to individual participants and the distribution of resources for rendering and displaying individual viewpoints and for attachment of different interface devices. Restraining from sophisticated synchronisation mechanisms enables a simple and modular architecture that allows for easy reconfiguration and extension of capabilities for different scenarios.

This solution is implemented by the e-MUSE system. As can be deduced from Fig. 2, eMUSE implements the “replicated homogenous world” approach [Mac97] to maintaining the consistency of the virtual scene database. Rather than maintaining a single copy of the VRML scene on a dedicated server, every participant fetches a local copy of the VRML scene and of the multi-user driver applet by referencing a given HTML page.

The consistency of the users' copies of the scene is maintained by directly and immediately propagating information about actions of every participant to all users. This is achieved through TCP/IP connections of the eMUSE multi-user driver of each participant with the local external user interface, and with external user interfaces of all other users. This allows the immediate propagation of changes in movement and position of all users to the individual multi-user drivers in charge of updating the local VRML scene.

The multi-user driver updates the user's local VMRL scene by:

- updating position and orientation values of the objects representing the avatars,
- dynamically creating objects representing the traces of users' movement (trace-avatars),
- updating the position of all viewpoints grouped with each individual avatar.

This module is implemented as a set of Java classes invoked by a Java applet that is referenced as part of the HTML page in which the VRML file describing the virtual space is embedded. The access to the avatar nodes and other internal structures of the scene is realised through the external authoring interface. The communication between input device drivers and the external user interface is similar to that of the NPSNet system.

3.3 Mixed Reality situation as interface: connecting the participants with each other

“How can I produce a work that is to be grasped through the senses alone, without any need for spoken language or written instructions? How can I make the observer understand what he can do? What senses should be addressed? What cultural conditioning is to be incorporated or overcome?” [F&S97]

These questions reflect the concerns and requirements of an interface environment needed to support the realisation of the Mixed Reality stage paradigm. This chapter

describes how the Mixed Reality stage approach relates to realising systems that address the above questions.

In order to realise the Mixed Reality stage concept, the whole system - the medial staging of action – needs to be understood as the interface. Not only the connecting of man and machine, but the action-oriented situation in which the visitors become involved. This concept of "situation as interface" acknowledges and introduces the distinction between two different levels of an interface environment:

1) connecting the participants with each other through the situation created by the environment,
2) the technical devices needed for the realisation of an unobtrusive underlying computer system.

Participants should not be conscious about the connections between their actions and resulting effects on the environment. The role of the environment is to provide a situation which will make this happen. In order to achieve such a way of connecting the participants, the development of technical interface devices is tied to the development of "content" itself. Additionally, unobtrusive devices that allow free body interaction and engage different senses are required in order to keep invisible the participants' interaction with underlying technical devices.

In order to achieve this, different perception levels need to be layered and related in a coherent structure where individual layers are carefully constructed, scaled and suitably mapped to support and complement each other. Figure 3. illustrates how this concept is realised in the Mixed Reality stage model demonstrated by eMUSE with a given virtual environment. The interface environment functions through the interplay of following elements:

1) Vision system

A fundamental point of the Mixed Reality stage concept is connecting participant's bodily sense of being in physical space with that of being in virtual space at the same time. A straightforward approach to achieving this is using a video tracking system to connect participants' movement in real space with navigation in virtual space. The data provided by the tracking system is used to control the position and movement of the user in the virtual space, parts of which are "displayed" accordingly in appropriate locations of physical space. This supports the perception of virtual space reacting to users' actions as an integrated part of the physical space in which they are situated. As the capturing of participant's movement by the camera and the video tracking system is unobtrusive, the users need not be consciously aware of it.

2) Visual representation of content

The visual elements of the virtual environment serve as placeholders for sounds and an orientation aid for the participants. Without this visual reference, establishing the relationship between one's movement in physical space and the sounds triggered in the virtual space becomes much more difficult.

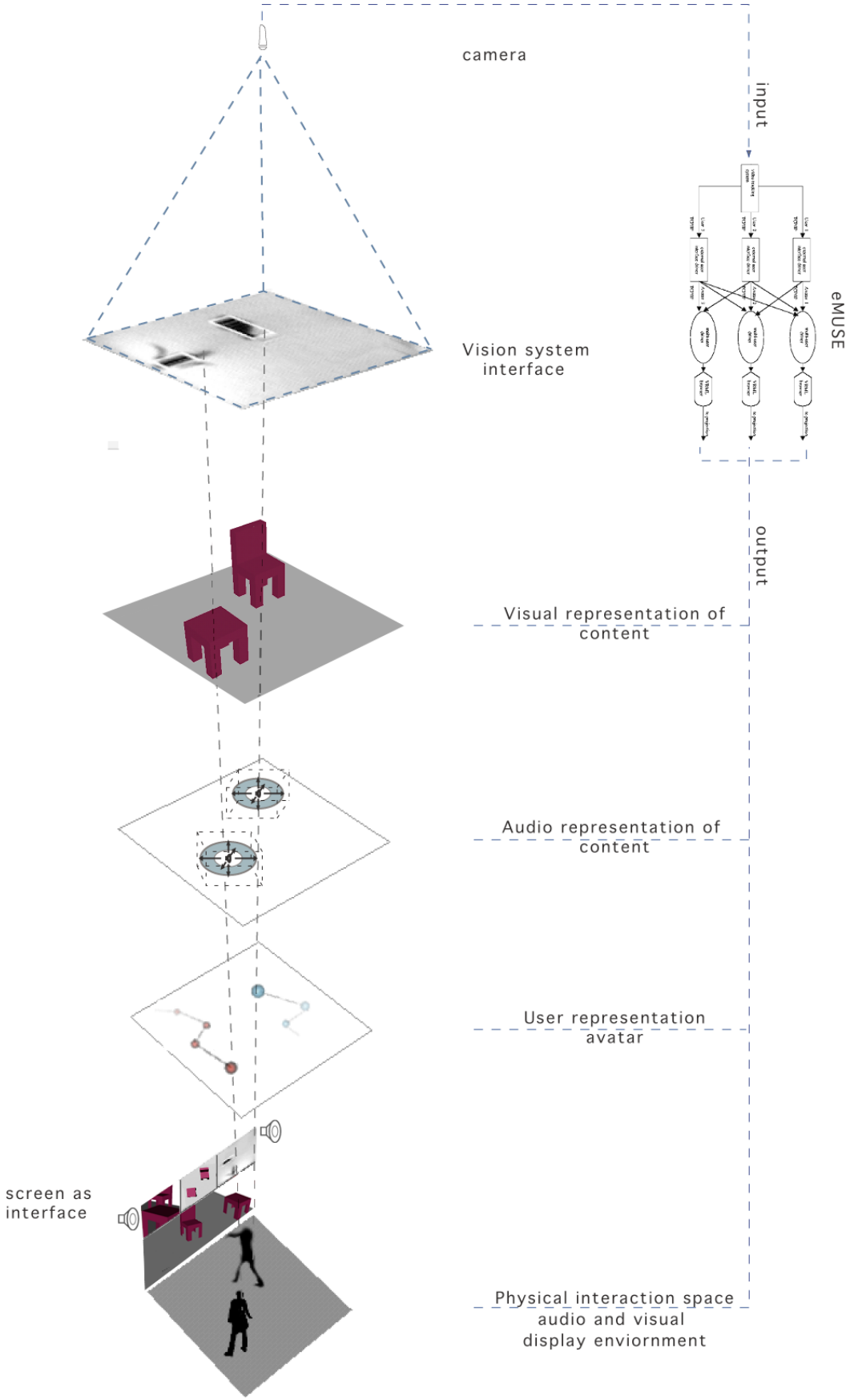


Fig. 3 Layers of the Mixed Reality stage

3) Audio representation of content

A participant's movement in physical space controls the creation of the corresponding trace-avatar in the virtual space. Trace-avatars of all participants trigger sounds in the virtual space that are emitted in the real space. These sounds inspire further movement and lead to a mutual play of the participants in producing sound patterns together – they connect the participants with each other.

4) User representation – trace-avatars

The trace of ones own movement – the trace-avatar – becomes the orientation field. Participants communicate through their virtual traces and determine thereby the action space as well. The trace-avatar becomes a graphically visible interface and a medium of communication. It shows the current position of the participants and triggers the sounds touching the sound objects. This "virtual touch" and the triggered sounds initiate a dialogue between participants that may develop into an audio-visual composition. In order to support arbitrary combinations of avatar-independent and avatar-bound views while at the same time exploiting VRML proximity sensors and collision detection scheme, in e-MUSE the avatar is separated from viewpoint control. This means that the current viewpoint is not necessarily tied to the user's avatar, as is the normal single-user limitation in VRML that imposes always viewing the world through the eyes of an imaginary avatar.

5) Physical interaction space

Unlike in systems where navigation in virtual space is achieved by directly manipulating some physical device, the virtual space of the Mixed Reality stage cannot be navigated without the existence of an appropriate physical space. Without moving in the physical space one is unable to move in the virtual space. In this way, the physical space becomes an essential part of the interface, connecting the participant to the virtual space. As a participant becomes immersed in the play of movement and sound, his awareness of being in space, and of the interaction with others, becomes increasingly a bodily one. As a result, the participants are connected through their bodily sense of each other and of their movement in space.

Neither of the above mentioned parts can be named as 'the interface'. Rather it is the whole situation of the environment and presence of the participants that connects the participants with each other as well as with the underlying computer system. The eMUSE system itself, along with the virtual environment, demonstrates the realisation of the situation as interface paradigm, as well as the deployment of physical space as interface to the virtual space. In doing so, the eMUSE system as an interface environment changes the usual computer environment of the WIMP[3] interface in two basic elements:

1) The interface becomes a spatial environment of attention.

   There is no separated action level. The development of action and communication happens in the space. One's own movement in space is made visible through an audio-visual display in space. The space between the camera on the ceiling and the

[3] WIMP - Windows, Icon, Mouse, Pull Down Menus.

projection surface on the floor creates an action space for the participants. This action space is a new inter-space, a new situation for spatial orientation and memory.

2) It is not us who hold the interface in the hand, but it is the interface (vision system) that senses us.

   We do not touch a mouse or a joystick to navigate. The camera eye “touches” us and relates our movement to the virtual soundscape. The direct interaction with the mouse is replaced through indirect touching of virtual objects. The distance to the interface – the video camera – is enlarged.

3.4 The Vision System as interface

In order to support the realisation of a Mixed Reality stage concept, the vision system used to detect users movement in physical space has to satisfy the following requirements:

- It must be capable of tracking the position, expansion, orientation, speed, and direction of movement of several people.
- It has to react in real time.

  This means that it has to react with no or at least little time delay to participants’ movement in the physical space.
- It has to be independent of the appearance of the users

  A system which has to be recalibrated for each user, is not acceptable for publicly accessible installations or theatre-like settings.
- It must be flexible in terms of illumination and spatial arrangement of the physical space.

  This is needed so it can be used in different, not predefined, environments. To achieve this, segmentation techniques that offer reliable results even in very low-light situations are required.
- It has to provide interactive runtime configuration and calibration.

  This is crucial as the system needs to be designed with no prior constraints concerning the environment and concrete scenario that it will be used in. Parameters to be adjustable include co-ordinate transformation values, number of tracked persons, usage of a Kalman filter, the size and total number of the users. Such flexibility most often results in reduced accuracy and reliability in comparison with highly specialised systems.

Such a video tracking system is suitable for gathering data describing users movement in the space at large. For finer sensing of users’ gesture or movement in local regions of space, additional devices are needed. Especially interesting for the Mixed Reality stage concept are devices based on electric field sensors as is investigated in Task 6.4. Using such devices would provide additional channels for connecting users movement and interaction in physical space with the virtual space.

eMUSE incorporates the Vision System that has been developed to enable connecting participants' movement in physical space with the virtual space in the introduced Mixed

Reality stage concept. The system is able to track the movement of several people simultaneously and currently provides the following functionality:

- It is able to monitor the shape, position, speed and direction of up to five people in real time.
- It offers different methods for image binarisation, which make it either usable in low light settings or in environments with few constraints according the appearance of the users, the room illumination and the spatial arrangement of the observed space. This makes it independent from the appearance of the participants and constraints on illumination conditions as well as from the spatial arrangement of the observed scene.
- It provides a simple GUI for adapting the system to a specific environment during run-time.
- Arbitrary applications can access the tracking data through a simple external user interface.

The tracking data provided by the system is used to control the position and the appearance of the avatars, and to control individual viewpoints in the virtual space (VRML scene) of the Mixed Reality stage. The co-ordinates of the user are mapped onto the co-ordinates of the virtual space allowing navigation in the virtual space. Position and movement of the users as picked up by the system can also be observed in the tracking window visualising the tracking process. The user appears as seen from above by the camera with the bounding box created by the system superimposed over his image (Fig. 4). The speed and size of the box is influenced by the movement and gesturing of the user which in turn determines the speed and size of the corresponding trace-avatar.

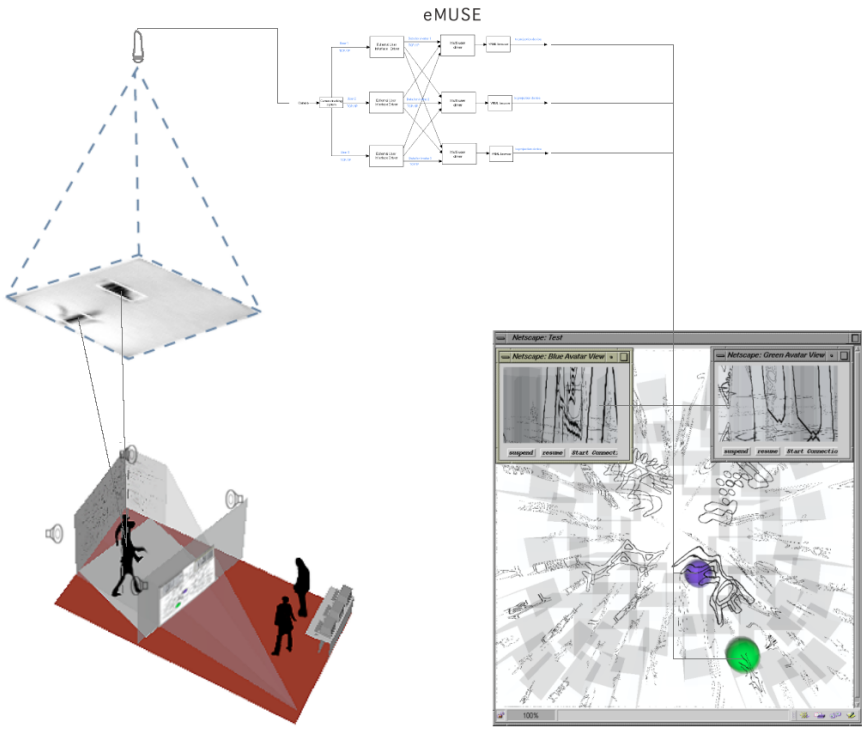


Fig. 4 Vision system controlling avatars and viewpoints

At present, the system uses one camera and hence measures the position, the size, the speed and the direction of movement in a 2D plane. This implementation enables the user to:

- create and position the trace-avatar in virtual space,
- manipulate the virtual space by creating patterns and compositions of sound,
- draw space as live input to change the existing virtual space,
- measure body movement in relation to virtual space (objective view)
- communicate through the trace avatar and movement in time,
- generate dynamic behaviour in the virtual space based on data of the tracking system.

Accommodating additional cameras and extending the system to extract full 3D information is possible but has not been of critical importance at this stage of development. The detailed process of development and the technical issues of the Vision system will be described in Task 6.4.

3.5 Navigation principles: mapping and scaling of data

The interpretation of the input from the interface determines how the visitor's body movement is transferred to the virtual environment. The navigational principle determines how data is taken from the interface and interpreted into movement in the virtual environment. This determines how the visitor perceives the virtual environment. Experimenting with different means of transferral is a way of exploring different 'ways of seeing'.

The 'Murmuring Fields' model is a generic environment that accommodates several navigational principles. Its construction does not prefer a singular navigational principle. This produces a basis from which we can compare and evaluate the different navigation principles.

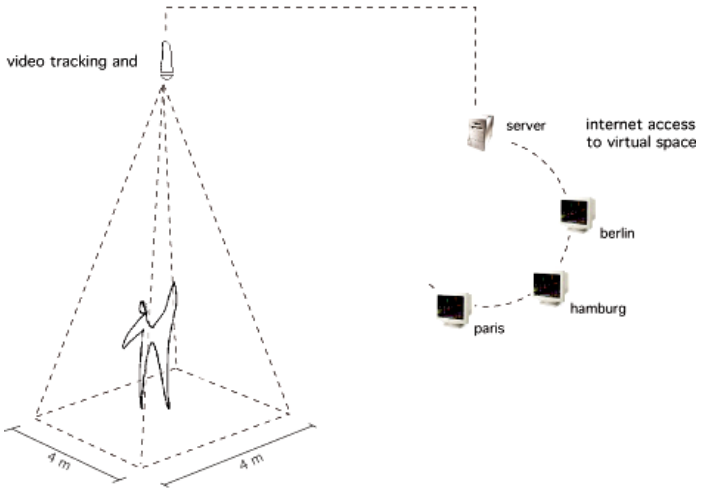


Fig 5 Boundaries of the tracking field

In contrast to navigating the 'Murmuring Fields' environment using the standard interface of the mouse, the camera as body tracking interface sets an invisible border within the virtual environment creating a restriction of interaction. The visitor is only tracked within the camera-tracking field. Movement can only be transferred from physical space to the virtual environments as long as it exists within the territory of the camera eye. The camera-tracking interface becomes the delineation of the virtual.

This invisible boundary changes the conceptual basis of the 'Murmuring Fields' environment. Having related to the context of scaleless infinity, the environment is now understood to have precise navigational boundaries. Virtual space is generally outside of a fixed notion of scale. Without the sense of a scaled body the flux of horizon lines and distance to objects establishes a fluid sense of scale.

Body-oriented interfaces depend on the transfer of body movement as navigational input into the virtual environment. This transferral infers a certain notion of body-related scale. The body-oriented interface of the camera-tracking relates the movement and size of the visitor's body to the finite size of the camera-tracking field. As the visitor's body inhabits the virtual environment it imposes a certain level of scaling.

The actual size of the eMUSE camera-tracking field varies with the different set-ups of the installation space. As a general size we have decided upon 4x4m.[4] By setting up a 1:1 scale relationship between the virtual environment and the camera-tracking field, movement in real space is directly transferred to movement in virtual space. Transferring interaction, movement and size in a 1:1 manor between virtual and real is a minimal solution that allows us to analyse the relationships between being in physical space and being in virtual space.

The re-sizing of the camera-tracking field profoundly changes the spatial relationships of 'Murmuring Fields'. Enlarging the camera-tracking field changes the scale relationships between virtual and physical. Virtual space is abstract and can be scaled to match the size of the camera-tracking field. But in doing so one looses precision in the scale relationships between the virtual and the physical space. The model must therefore optimally be built to match the precise size of the camera-tracking field. An environment modelled for a 10x10m tracking field would result in very different spatial relationships. Here it would be possible to sense a floor like condition and/or a space below. The space would change its intimacy requiring a rougher modelling for the re-collection of spatial situations, sounds and views.

In the following we describe the different navigational principles and their effect on the spatial setting of the Mixed Reality environment. Furthermore describe how these relationships determine a certain logic to the modelling of the virtual environment.

[4] At the Transmediale event of eMUSE in Public setting the camera-tracking field was significantly smaller (2.5x2.5m) while at the Bochum event of eMUSE in a performance setting the space was apporximately 4x5m.

3.5.1 Frontal navigational principle

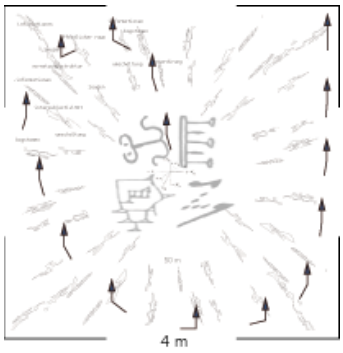


Fig. 6 Frontal navigational principle with fixed viewpoint direction

The frontal navigational principle maps movement in a 1:1 manner between the physical and the virtual space. The avatar view is directed in one direction making movement similar to that of moving forth and back within a theatrical space.

The frontal navigational principle allows for one directional vision within the virtual environment. Here, seeing means the freezing of the direction of the body towards one view. The space can be constructed in a manner equivalent to that of the traditional theatre stage. Screens can set-up spatial relationships that reveal and conceal elements as the virtual body of the visitor transgresses the screens.

3.5.2 Navigation using history

Using history allows navigation that directs the viewpoint of the virtual camera within the virtual environment. Here the visitor feels the sensation of moving his/her head.

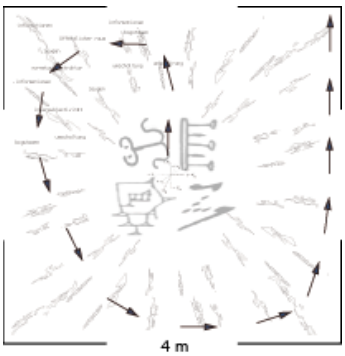


Fig. 7 Using history to control viewpoint direction

Steered by body movement the visitor can see multi directionally within the virtual space. Movement is interpreted in a 1:1 scaled relationship as the frontal navigational principle. Using history means that the user ‘sees with the direction’ of his/her body. This means that the display optimally should follow the body in its changes of direction either pivoting around an axis or be placed on the visitor as a portable display device.

Using history as the navigational principle brings out other notions of how to model the environment. The free movement of the visitors and the changing of the viewing direction mean that the visitors can look around objects determining whether objects are constructed as screens or whether they are volumetric.

3.5.3 Using scaled navigation

The scaled navigational principle uses the input of the visitors body movement to determine direction of movement within the virtual environment. The visitors body movement is interpreted as the 'setting in to motion' of movement. Taking a step to the left means the starting of leftwards movement only stopped when the visitor initiates a new direction.

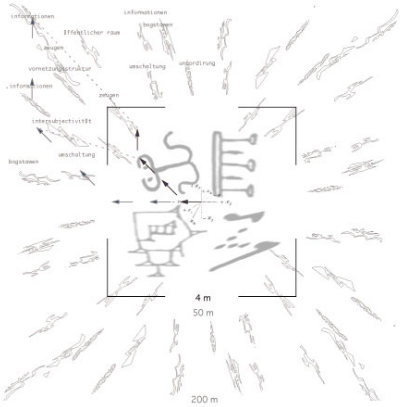


Fig. 8 Non-linear scaling between movement in real and movement in virtual space

Modelling for the scaled navigation principle means the negation of the invisible border of the camera-tracking field. Movement is freed from its 1:1 relationship between virtual and physical space and non-linear relationships between movement in physical space and navigation in virtual space can be accommodated.

3.6 The concept of avatar as a virtual body of representation

As the notion of avatar as a graphical representation of the participant in a virtual world became widely accepted, research [Benf95, And96, avpan97, Ant98, Thal97, Vil98, Olv98, Thal98] became concerned with how to make avatars that are:

- expressive – convey human emotions,

- individualised – reflect the preferences and/or personality of the participant,

- realistic – visual appearance resembles human appearance,

- or believable – provide a "believable" representation of a human user behind it.

What is often neglected is the search to understand what presence and representation of people in virtual space is really about. What is the purpose of the "representation"? In

task-based virtual environments the goal is usually to manipulate the objects in the environment and user representation is subdued to this purpose. In this case, rather than avatars, different symbolic representations are used, such as virtual hand, virtual pointer and the like [Poup98]. Avatars are the prevailing concept of representation in the multi-user virtual environments usually referred to as collaborative virtual environments. CVEs can be roughly divided into work-scenario environments and social environments. In the latter, the issue of user representation centres on avatars that provide means of verbal and non-verbal communication with other users. Communication is thereby understood as exchange of messages and the environment basically becomes a 3D chat space where communication occurs through exchange of textual or sometimes audio messages. This is possibly followed by rudimentary movement of avatar's extremities in order to convey gesture and rudimentary means for expressing emotions through facial changes. It is worth noting that both of these expression mechanisms are often "triggered' by clicking on a menu item or pressing a determined combination of keys. There are two basic instances of the approach to avatars as the virtual embodiment of the user:

- avatar as a virtual constructed self (1st person perspective),

- avatar as an externalised body of communication (3rd person perspective).

In the first approach the participant identifies himself with his virtual representation in the way in which we unconsciously "assume" that our body represents ourselves to others and reflects our actions. The avatar is a projection of the participant's mental space in the virtual world. The first person perspective of avatars makes them serve as masks. The participant speaks, hears, sees and acts through his mask. In a way he assumes a role, just like actors in theatre put their individuality into assuming a specific role requested by the play. This situation is mostly found in the so-called "social virtual environments" [Sch97]. Important to note is that this function of avatars in these environments is possible only because of a shared voluntary suspension of disbelief by the participants – the accepted awareness that "everybody is an attrappe" [eRENA98-1].

In the second approach the participant is mentally detached from his avatar. He is looking at the virtual world from outside manipulating his avatar as a puppet. The avatar is also detached from user's sense of his real body. Some researchers claim that linking the participant's real body to the avatar by, for example, having the avatar display movement of extremities when real body extremities are moved, reinforces the effect of immersion [Slater94]. Avatar as an externalised body of communication recognises the physical reality of the body as being outside of the virtual world. A participant's senses perceive a different reality from that of the virtual world and thus limit his immersion. The mental framework imparted by the third person perspective views the avatar as a separate body that replaces the participant's presence and serves him to communicate with others. From this approach develops the notion of an agent.

In both approaches, most of the existing models of avatars are prefabricated figures that adhere to the concept of the mask [Avpan97]. Usually, the user is offered a selection of choices from which to pick a figure to be his avatar. These figures are largely figurative and anthropomorphic visual representations often aimed at mimicking reality. Such approaches forget that the whole concept of avatars as masks - as another self - is possible only because of the silent agreement among the participants to accept the fact that avatars stand for themselves as their virtual personae. Consequently, there is no need for elaborate visuals aiming at providing this illusion. Research at the GMD's

MARS Lab experimented with the notion of the "extended avatar". This approach breaks with the notion of immersion implied by the concept of a virtual self but also refutes the concept of the avatar as a completely detached virtual body of communication. It introduces two points of departure [eRENA98-1]:

- the notion of user enactment,

- the notion of avatar as an extended body of communication.

The idea of user enactment replaces the notion of user representation while understanding the avatar as an extended body of communication resolves the conflict of immersion vs. detachment. In this approach the avatar is seen as a medium - an interface to other avatars and the virtual world. It is also a means of communicating with oneself as influenced by the experience of the virtual world. The notion of avatar as an extended body of communication is a first step in questioning the concept of avatar in itself. Through developing examples which implement this changed way of understanding avatars it became clear that even the concept of avatar in itself is not necessarily the best way of supporting user presence as interaction and communication. The following sections present the development of prototype avatar concepts following the ideas of user enactment and avatar as extended body of communication. Finally we consider an example of supporting participants' presence as interaction and communication without an underlying avatar-like representation.

3.6.1 The avatar as user enactment

The notion of user enactment reflects the recognition that user presence in an interactive environment is not the question of representation but the question of providing mechanisms of interaction and communication with other participants or with one's inner self. This approach develops the following guidelines for developing avatar concepts [eRENA98-1]:

- abstract the visual form,

- connect production of the avatar form to real body movement,

- provide existence in time, not only space.

These guidelines enforce the realisation that the visual representation in itself is not the critical way of achieving what is usually referred to as "individualised user embodiments". It also acknowledges the fact that perception of physical space works through a bundle of different senses [Leek98, Slater94], but is fragmented to individual ones in virtual environments. The major shortcoming of the majority of current approaches to virtual environments is the inadequate treatment of the problem of this sensory fragmentation. In most cases they still take the visual as dominant mode of perception which strongly limits the possibilities of the perception of the environment and interaction between the participants. These approaches also tend to focus on developing avatars that mimic real-world appearances and real-world physical interaction. The approach of avatars as user enactment abandons the visual as the dominant form of perception and representation. As a first attempt in developing new solutions it focuses on movement and gesture as two basic channels of presence and communication. Movement requests the provision of minimal visual information needed for localisation in space at time. Gesture requests that it be dynamically connected to participant's real body through senses other than the visual.

3.6.2 Avatar prototype concepts

Developing prototype concepts for the described approach went in a twofold manner:

- avatars as electronic abstract masks,

- avatars as an embodiment of bodily gestures.

The concept of avatars as electronic abstract masks uses abstraction of visual form as a way of abandoning the exclusiveness of first person perspective and of the illusion of purely visual immersion. The concept of avatars as gestural bodies heads away from the third person perspective by connecting the avatar to real body gesture and movement. The meeting point of these two concepts is the solution of trace-avatars as demonstrated in the interactive installation Murmuring Fields. The prototype example of this approach is the trace-avatar such as implemented in the interactive installation Murmuring Fields. Based on the real-time movement and gestures of the user in real space a trace is created in the virtual environment. The trace conveys real-time individual user presence in the virtual space – the trace is the avatar, the avatar is a trace.

In physical space, movement and gesture are perceived foremost through body senses such as balance, muscle tension, touch, force feedback etc. [Slater94]. The visual sense is only one among these many. The notion of trace and its connection to real body movement emphasise movement and gesture as basic channels of presence and communication. This concept satisfies all conditions for the avatar as extended body of communication and the paradigm of user enactment: its simple visual form provides the information needed for localisation in space and time, and it is dynamically connected to participants body movement through senses other than the visual. Trace-avatars also allow the same kind of presence for visitors in real space and visitors from Internet, although they are present through different interface channels. There are different possibilities for the technical realisation of trace-avatars. In Murmuring Fields a video tracking system that follows participant's movement in physical space is used to deliver information for the creation of traces in the virtual space

3.6.3 Relations of body, space and senses

“Our relationship to the world is a bodily one. Our body makes its experiences through movement in space. Through different senses we develop the perception of space. Sight senses the space through the movement of the eyes. Hearing provides spatial localisation and recognition of distance. Touching and hearing enables a threedimensional comprehension of space. Smelling makes the place recognisable. Through the variety and complexity of the senses we explore the space. This experience impresses the body as memory. It becomes visible as a body trace – unerasable.” [F&Sint] The major shortcoming of approaches following the idea of a “virtual body” as representation of the user is that they are overlooking two important factors:

- that human perception of space is foremost a bodily one [Slater94],

- the mechanisms of proprioception.

Proprioception is a set of mechanisms that induces an unconscious mental model of our body and its behaviour. It allows us to “know where our big left toe is, without looking” or to “touch our nose with our right forefinger, with closed eyes” [Slater94]. As a consequence, the ideal of immersion pursued by the “virtual body” approaches can never be fulfilled because while the sensory data from the virtual environment would

want us to identify with the virtual body, proprioception tells us that our real body is still there and different. Some approaches attempt to overcome this conflict by introducing "interactive techniques that maximise the match between proprioceptive and sensory data" as the one suggested by [Slater94]. The latter identifies immersion as a "kind of technology" and presence as "an associated state of consciousness" for which immersion is a necessary but not a sufficient condition. It is further observed that "in addition to the necessity of an immersive technology, the interaction techniques may also play a crucial role in the determination of presence". But the notion of this interaction remains tied to that of a virtual body as a graphical user representation following and displaying the actions of the real body as close as possible.

Our approach at MARS lab [F&S97] goes further realising that it is the very notion of a virtual body "outside" one's bodily sense of "being in" that poses the conflict. Instead, the focus is moved to the participants real body and metaphors of "representation" exploring ways of conveying a bodily sense of presence.

"Our interest is devoted to the virtual communication space in which the observer becomes an actor [active participant]. In the space of interactive systems the body is brought into movement through instinctive interfaces. The movement itself serves as a metaphor for representation of the actor in the virtual space." [F&Sint]

3.7 Distributing virtual space in physical space: audio-visual display

3.7.1 The display problem

The crucial component of the introduced Mixed Reality model are display devices which would enable the displaying of elements of virtual space in localised regions of physical space. Holographic display would be one possibility but it is not foreseeable that it could become feasible for such interactive real-time applications in near future. Another possibility is to deploy a number of small projection screens scattered around in physical space, but this poses problems with projector beam intersection by users as well as by individual screens. It also disturbs the overall perception of the combined space as the screens would represent tangible obstacles in the physical space.

There are two straightforward workarounds of this problem of localisation:

- using glasses with LCD displays that the user's would have to wear,
- using „chameleonVR" class of displays which are worn as hand watches and display a portion of the virtual space depending on their current position in real space.

The first solution is problematic since the goal is to have the users unencumbered and free from physical devices mounted on their bodies. It also poses the problem of wireless data transmission to the glasses and demands very high precision of the tracking system. The problem with the second approach is that it increases the users' attention to the technology employed for perceiving the virtual space, although it does represent an interesting possibility to be explored.

Using sound lends itself very well to the idea of real space filled with virtual space that is revealed through users actions and in localised pieces only. This is due to the following properties of sound:

- it is spatial in nature and there exist systems for spatialised reproduction of sound,
- human perception of sound works as succession of audible patterns in time,
- it conveys change in time i.e. dynamic action and not static representation.

In our work we have experimented with three solutions to the described problem:

1) Using sound as the founding element of the virtual environment,
2) Using a multiple screen projection of the virtual environment,
3) Using users' bodies themselves as projection surfaces for visual elements of the virtual environment.

These issues influenced the design of the virtual environment as a sound-space triggered by users movement in physical space, which is complemented by projection of visual elements for orientation purposes.

3.7.2 The virtual environment

The virtual space (VRML model) is primarily an acoustic architecture, a field of sounds. These sound objects are triggered by virtual proximity sensors activated by the movement of the user (Fig.9). As the user moves within the camera tracking field the data is transferred to the trace-avatar. By crossing the boundaries of the proximity sensors the trace-avatar triggers the sound files. The proximity sensors function as a touch sensor for virtual touches. The triggering of sound fields relates to the sensation of touch. As the user moves he creates a sound field by touching invisible borders of the proximity sensor with his body.

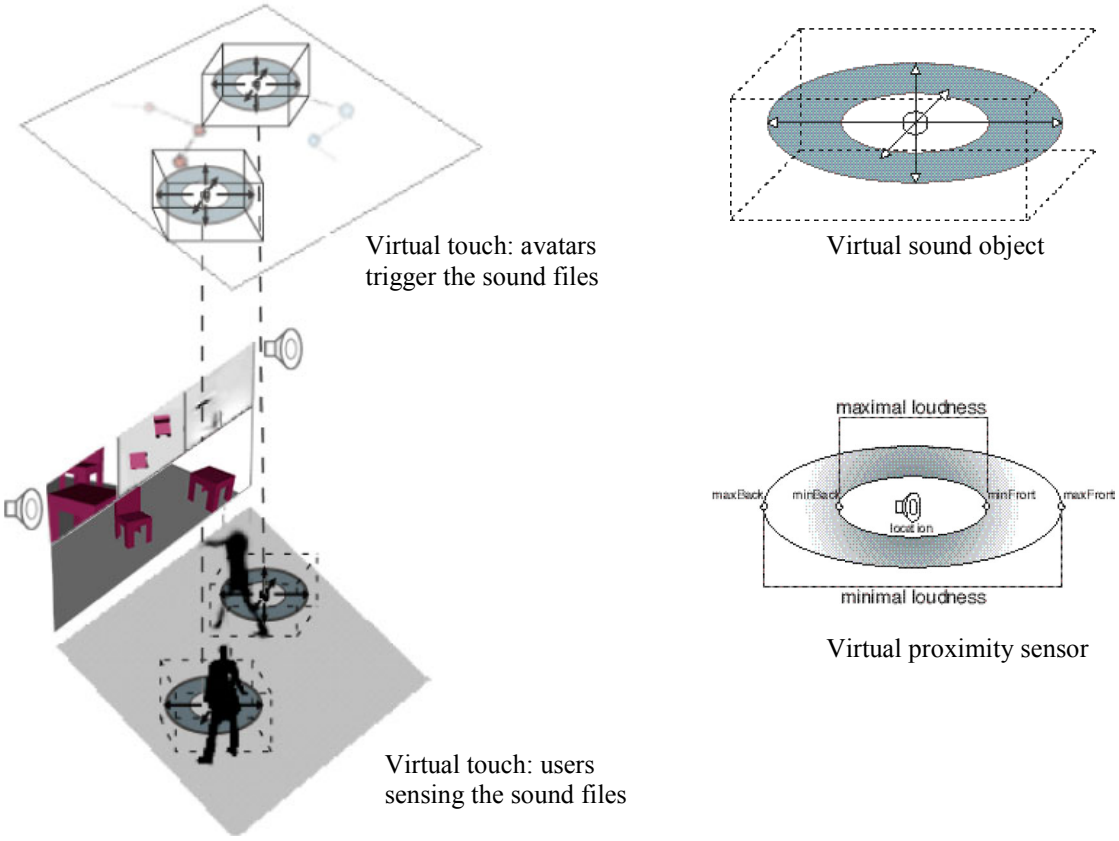


Fig. 9 Proximity sensors on sound objects in virtual space triggered through movement in physical space

This soundscape is structured in a manner that is parallel to the visual landscape. As sonification accompanies the navigation through space, the soundscape imposes a dynamic and temporal experience of space and data. The soundscape presents and reinforces spatial experience as the user enters the field of interaction and produces an individual sound collage. Users reactions are dynamic as they continue to move, recognising the data structure and its relationship to their movements. The structure and composition of the soundscape has the function of avoiding pre-fixed linear sound formations and to stimulate the visitors creation of an individual language of movement.

Such an architecture enables visitors movements to create a trace of dynamic sound events in real-time. The sound field translates a static, informational architecture (symbolised by the signs in the visual scape) into temporal configurations of individual movement. The soundscape transforms individual motion into a dynamic relationship between signs and sounds, visual and audible, into processes of narration and interaction.

Such modelling of the virtual space of the Mixed Reality stage results in the virtual space being perceived and experienced only through users‘ being in the physical space. It serves as content with which the real space is filled and is revealed as localised through users‘ movement in the real space.

This suggests that designing the virtual space of the Mixed Reality stage is not about designing a defined absolute form to be observed and navigated as in common approaches to 3D virtual space. Rather, the virtual space needs to be built as a density

filling the physical space. This density represents a framework of possibilities from which the participants themselves realise different spatial configurations through their movement and interactions. In the case of 'Murmuring Fields' as a model for the Mixed Reality stage, this density is primarily audible rather than visual content – the interactive soundscape.

4 eMUSE: strategy of development

In this work we have built the eMUSE system that extends VRML with multi-user support and attachment of arbitrary, non-standard interfaces. This enables the presence and communication of several participants in a shared virtual space. Their access to this shared space is supported through intuitive interfaces that mediate participants' body language. This includes tracking the presence and position of the participants in physical space and transferring it into a trace of motion in the virtual space (trace-avatars). These trace-avatars demonstrate the choice of an abstract, movement-oriented and action-oriented form of user representation that relates to users' spatial behaviour in time.

eMUSE is a theatre-machine for the creation of an interactive stage. An application for this stage is presented with 'Murmuring Fields'. This virtual audio-visual environment is overlaid with the real space in order to have the participants simultaneously present in real and virtual space.

Figure 10. depicts the initially planned presentation of 'Murmuring Fields'. The sound system emits the sounds triggered by users' movement. The image of the virtual stage (floor projection) overlays the real stage and facilitates orientation in the information space. The video tracking system surveys the entire (inter)action space while two projection surfaces display individual viewpoints of the two participants. Presence in the virtual space is mediated through the trace-avatars (coloured traces of movement) corresponding to the movement of individual participants. This provides a similar kind of presence for both participants in the shared physical space and the remote participants across Internet with different interface channels.

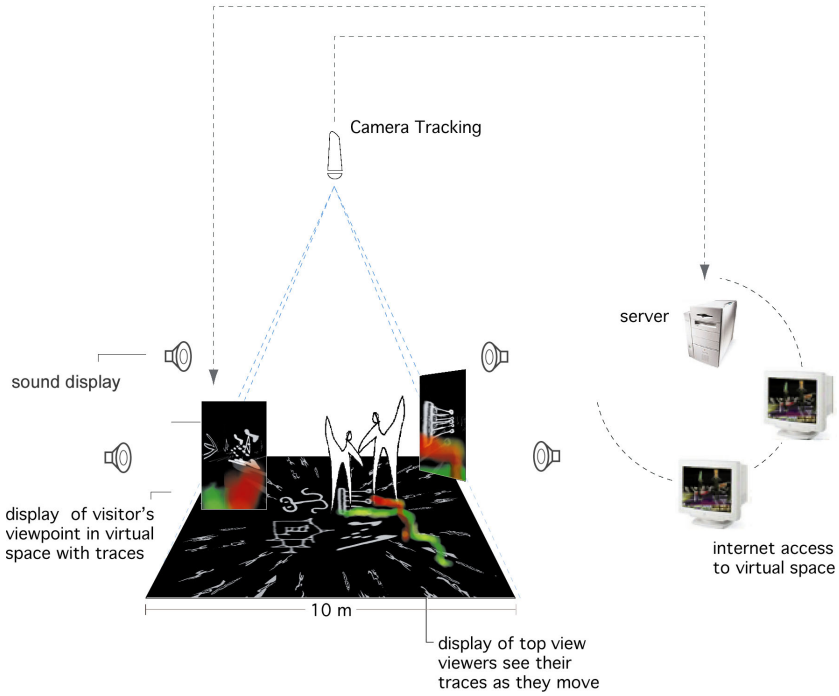


Fig. 10 Mixed reality stage system eMUSE with virtual environment 'Murmuring Fields': interactive, networked stage for real and virtual actors

With eMUSE we have built an instrument for the Mixed Reality stage, an instrument for theatrical play and learning in networked space. This instrument that consists of distributed components – vision system, spatial display and networked computers - enables the Mixed Reality stage. This is no stage to be observed, but a stage to be entered by oneself. Invitations to festivals, in theatre and into Internet, enable us to learn how to play this instrument together with accidental visitors as well as with invited guests, in a public rehearsal space.

We are developing this instrument in three stages. The first stage of development explores the conditions of public space, whereby public space can refer to a building with accidental visitors. We chose the media arts festival Transmediale '99 in Berlin for this occasion. The aim of the second stage is to explore the conditions of the Mixed Reality stage in the theatrical space with professional performers. The invitation to the symposium "Perception of reality – reality of perception" taking place within the framework of the German Puppet Theatre Festival FIDENA '99, seemed to us a suitable occasion. The third station is the Mixed Reality stage in networked space, which is planned for the third eRENA year, and we are considering different alternatives of a suitable occasion for its public trial and evaluation.

Part 2: Public Trials - Learning from the Real World

5 Mixed Reality in public space: Transmediale'99

'Murmuring Fields' – a multi-user Mixed Reality installation - was invited to be presented at the interactive media art festival „Transmediale'99" in Berlin.

With this installation the prototype setting of the newly developed VRML based electronic multi-user stage environment (eMUSE) system was demonstrated for the first time in public. In the Transmediale context several other interactive installations were presented, highlighted by a seven days' symposium with panel discussions and demonstrations. This installation was shown from February 12-15 in the exhibition space and presented in the Panel on Interactivity by the authors. The installation can be seen in the video material part1.

5.1 Context: eMUSE staging 'Murmuring Fields'

The strategic method of our approach is to develop interactive media concepts within a public realm that integrates the users' behaviour in the design process in order to enhance the users' media competence. The Transmediale revealed itself as an excellent place to test the eMUSE system with respect to its acceptance by the users, stability and functionality. It was the first step in exploring the system in a performance context in public space.

In interactive systems, the work and its content are created by public users. This often bears the problem, that the users find themselves acting in front of an audience. Such lack of intimacy inhibits unhindered interactivity. This becomes even more difficult in multi user environments where the users create both the content and their mutual communication at the same time.

6 Content and Structure of the Mixed Reality environment

The concept of the Mixed Reality environment is to identify and explore the notion of information space and its overlay with physical space. Different from games like the interaction scene in Desert Rain (see Blast Theorie in Task 7.b.1) the spatial concept supports the non linear method of interactivity in Murmuring Fields. The goal is to create a framework for individual narrations through an open structure of interactive material, instead of following a pre-programmed storyline. This method can be seen as a spatial storyboard principle.

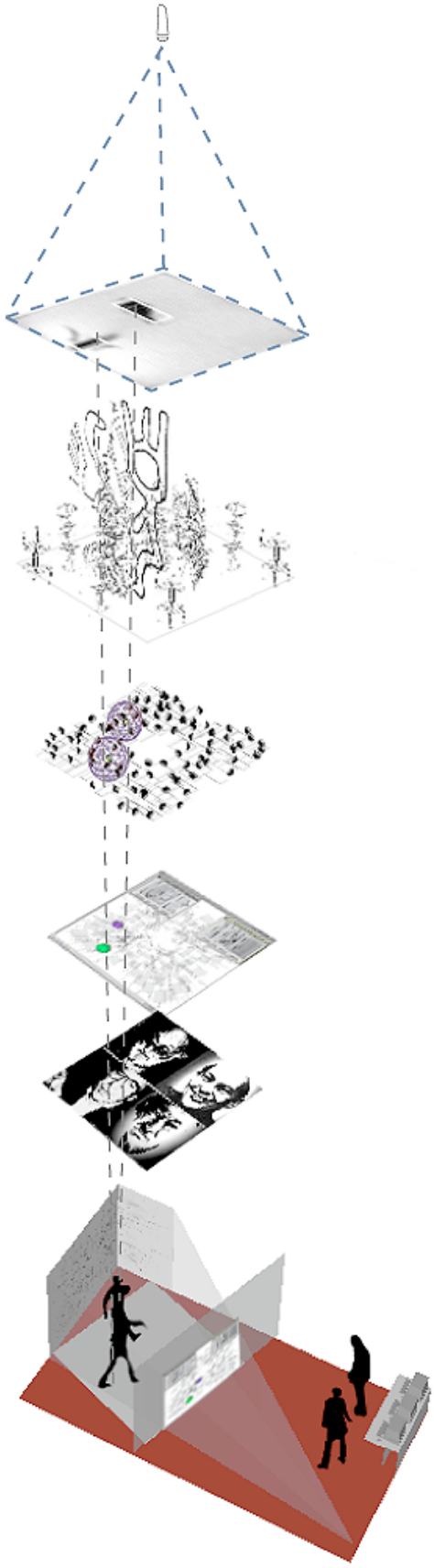


Fig. 11 The components of the Mixed Reality environment

6.1 Physical interaction space: environment set-up

Alongside the programming and the interface, the spatial set-up of an installation is an important factor in creating the framework for an interactive experience. At the Berlin Transmediale, the interactive installation 'Murmuring Fields' comprised three zones of different activities:

- the interactive zone for the active participants,
- the interpassive zone for the audience, and
- the corridor zone for discussions.

The audience arrives in the corridor street. Guided by a barrier leading to the entrance doors, people can choose to enter either one or the other black curtain doors: the *interactive zone* or the *interpassive zone*. The visitors are free to decide how they start their exploration of the installation, but the handout with a short description is laying on a pedestal at the *interactive zone*.

The size of the interactive zone and the number of participants are dependent on the size of the space. The active space of 'Murmuring Fields' at Transmediale was 3 x 3 metres. The number of active participants was therefore limited to two, which is also advantageous from the observation and evaluation point of view. The interpassive zone is the same size, and can accommodate up to twelve observers, plus two or three operators.

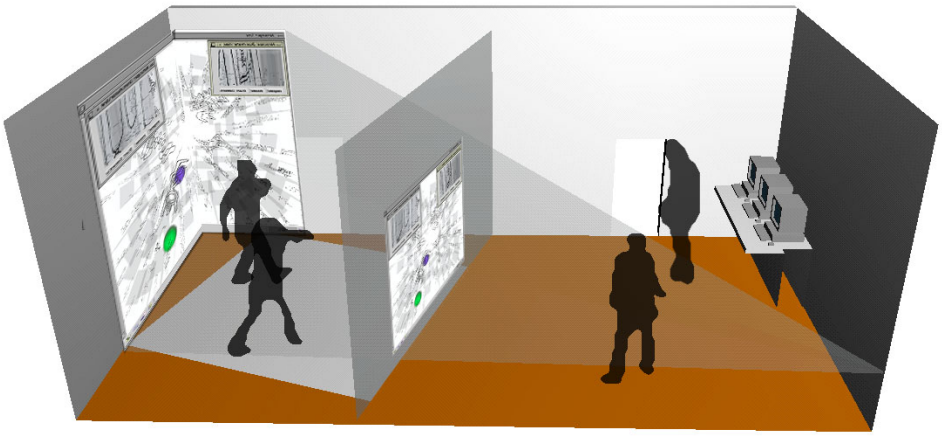


Fig. 12 Spatial structure of the exhibition space with double projection screens

The exhibition space is set up in a dual mode. The main element of the installation is the gauze screen as a transparent boundary between the interactive and the interpassive zones. The visual appearance of the interactive zone is determined by the white light of the projection. The actors are embedded in a double projection field between the gauze screen and the background wall. The projected images appear on the screen but also on the surrounding walls. The transparent boundary screen serves as a skin to be perceived from both sides.

The interactive zone gives the impression of intimacy, but in fact, it allows inspection from the interpassive zone. The audience in the interpassive space perceives the virtual

environment inhabited by the interactors. They become viewing participants or voyeurs. Looking from the backside of the canvas integrates the viewers as part of the installation since the actors see the audience as background shadow figures within the installation. The transparent screen absorbs the presence of the audience to the actors. People in both spaces don't disturb each other. Viewers and interactors share one space divided by the transparent screen displaying the central image visible from both sides.

Another important feature of the installation is to allow the understanding of the technical setting. The equipment is a visible part of the installation and enables the visitors to understand its functioning. The team is at disposal to the public to answer questions.

6.2 Content of the Mixed Reality environment

Virtual living stage: 'Murmuring Fields'
Protagonists: Four voices
Sound picture: Statements of the four thinkers Vilèm Flusser, Marvin Minsky, Paul Virilio, Josef Weizenbaum and background sounds samples

The virtual environment of 'Murmuring Fields' is a space of language, sound and image addressing the senses of hearing and seeing. Language is present on several layers:

- through the soundscape and the dissolution of text into fragments,
- through the visual environment made of archaic signs that refer to meanings now unknown to us,
- through the creation of a borderline made of hand-written words.

The mapping of the language consists of four zones - the houses of the four media thinkers which address the audience in different languages. The zones are subdivided in different sectors: with full sentences, with words, with phonemes. The underlying sounds emphasise the particularity of each field. The layers of language and sound intersect and weave around the interaction of the visitor, generating an environment of communication. Language fluctuates from the readable to the indecipherable.

6.3 Visualisation of presence : the avatar

In interactive computer games, the player is represented by a game figure, the avatar. In 'Murmuring Fields' the avatars take the minimalist form of small spheres. As the visitor enters the interactive zone, he is tracked by the camera and represented by an avatar: a virtual object which reacts directly to his movement and position in real space. Individual participants are represented by different colour avatars which act as their visual interface for the sound navigation: when a participant's avatar approaches a sound symbol in the virtual space it triggers the sound related to it. The avatar can also change in shape and size according to participant's gestures. If the visitor stretches out his arm, for example, the avatar changes in shape from a sphere to an ellipsoid. This elongation means that the visitor stretches out further in the virtual space and activates more sounds. This direct relationship between the movement of the avatar and the participant gives the visitor a feeling of presence.

Each avatar produces its own sound space and its own view space in the 3D environment, it serves as a participant's virtual camera for image production. Accordingly, the image of the virtual environment is composed of numerous views seen by the participants' avatars. These views are integrated as windows into the virtual environment. During the public test setting we limited the participation to two players so that the communication can be analysed more precisely. The players can better understand who controls or touches what, who triggers the sounds, and who takes on which perspective.

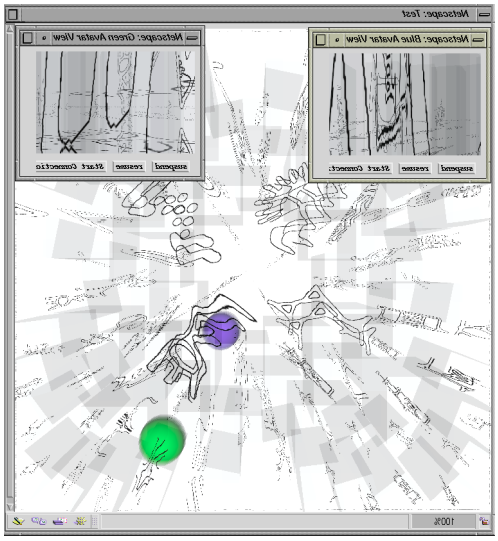


Fig. 13 Top view with avatars and individual viewpoint windows

6.4 Visualisation of content : space of image

'Murmuring Fields' is constructed as a spatial drawing. The signs are drawn on virtual planes in the space. These structures are not conceived as three-dimensional objects but as abstract signs floating somewhere in between the two-dimensional and the three-dimensional.

The visual environment of 'Murmuring Fields' is composed of four sets of hieroglyphics from ancient cultures, assigned to four areas of speech. The symbols relate to ancient forms of communication that cannot be easily decoded today.

The structure of the virtual environment uses lines as a stylistic tool. The lines stand as flat figures in the virtual space. The basic colours are white for the background, black for the symbols and grey for shading.

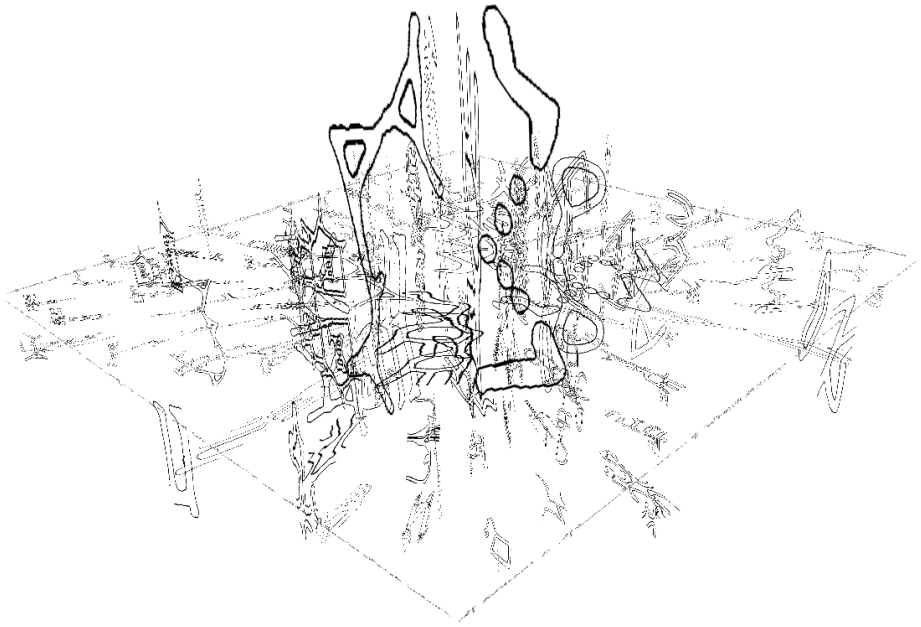


Fig. 14 Isometric view of the visual environment

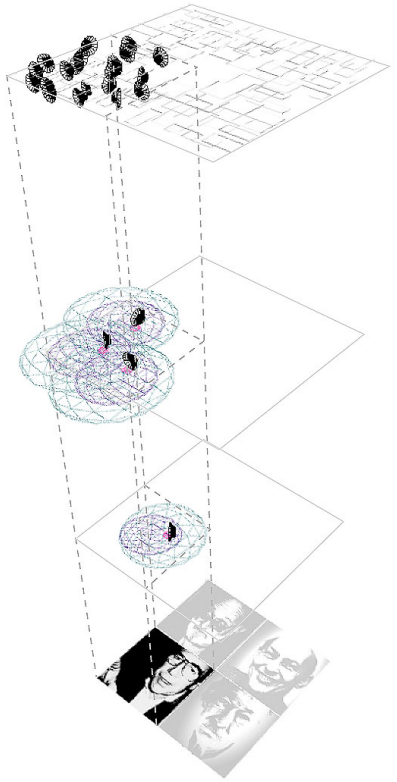
The lines build up the system of notation for the sound space. Dominant lines mark the different areas of speech, and lighter lines are arranged to form structures, interweaving to create a network of background sounds. The resulting topographical network establishes the connection to the sound environment.

The shapes and patterns of movement create an interactive environment that resembles a data-filled, virtual sculpture that can be entered. As the participants move around, the spatial references, the shape and the course of the lines change. This means that the virtual sculpture is charged with kinetic energy.

6.5 Audio representation of content: space of sound

The structure of the 'Murmuring Fields' environment is built up around the concept of a dense area of overlapping statements. Moving outwards the statements disentangle, and are then succeeded by the field of words; single words cut up from the statements and finally the field of phonemes; dissected from the individual words.

The basis of the sound material is recordings of statements of the media thinkers: Marvin Minsky, Joseph Weizenbaum, Vilèm Flusser and Paul Virilio. The largest part is spoken text from which single word clips and phonemes are deduced, as well as synthesised text fragments and musical sound bits. These clips are supported by a series of background sounds such as rhythmic noises (breathing) and isolated instrumental sounds. The background sounds create a weave of sound that gives individual character to the four different quarters.



Foreground sounds: Field of words and broken words. The files are not looped

Background sounds: A background weave that gives individual character to the four different quarters. The files are looped and heard when the users is within the sound field.

Centre: Statements of the media thinkers. The files are looped and heard as the user enters the sound field.

Content: example of Paul Virilio‘s quarter

Fig. 15 Centre field of sounds

The statements of the media thinkers are placed as the centre core of the environment. Through the superimposition of the statements different conversations are generated. The looped speeches are heard by the user as he enters the sound fields (the field in which the volume of the sound can be heard). When the user leaves the sound field the sound continues although not heard by the user. As the user re-enters the sound field he enters at a random point within the text so that the text always appears in a new context.

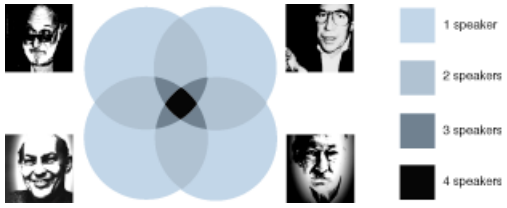


Fig. 16 Centre field of the soundscape - overlay of speeches

Foreground sounds

Moving outwards the statements in every quarter disentangle. They are succeeded by a field of words, short phrases or fragments from the central statements.

Here the user can play with single words and make new combinations of words and meaning. The word clips are not looped and play when triggered. When the proximity sensors are touched tentatively the words are broken up as fragments. When the user re-triggers the sound file the sound is played over again which creates a sensation of

breaking the words. A rhythmical repetition of movement can produce effects that are similar to scratching in rap music. The user can dance with the words. Moving from the centre to the border, the user finally encounters a sphere of phonemes, dissected from the citations and fragments of the text passages.

Background sounds

The foreground structure relies on a background weave of sounds giving an individual character to the quarters. Every quarter has a background comprised of three sound clips. The sound clips are generated from the character and content of the individual statements. They build a metaphorical and emotional world of illustrations and comments to the content of the speeches. The words in the foreground give rhythmical structure connected to the users movement while the background sounds give a harmonic backdrop.

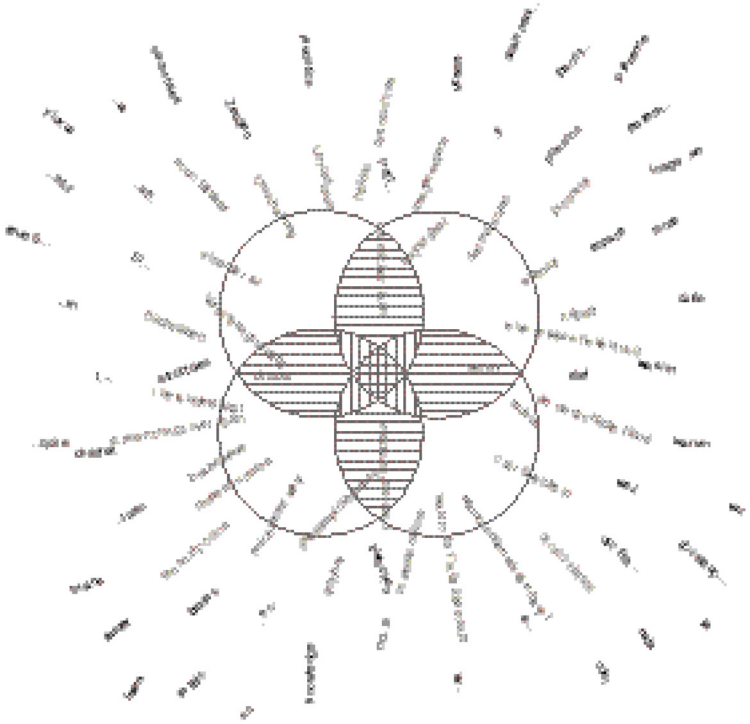


Fig. 17 field of words: background and foreground

The Statements

«L'espace temps ancien était un espace extensif. C'était un espace où la longue durée prenait le pas sur la courte durée. La courte durée était considerée comme un mal, comme quelque chose de péjoratif. Peu durée c'était le ne-pas-être-présent, c'était négatif. Aujourd'hui, nous entrons dans un temps intensif. C'est-á-dire, les technologies nouvelles nous font découvrir l'équivalent de l'infiniment petit dans le temps.»

Paul Virilio

«Was ist denn der Unterschied zwischen dieser Hand und der menschlichen Hand? Warum sollen wir überhaupt einen Unterschied suchen? Warum die Frage überhaupt? Naja, weil die menschliche Hand eben mit dem Menschen verbunden ist; weil die menschliche Hand gehalten werden kann, sagen wir, von einer schönen Frau. Und dieses Halten – nich' - oder Streicheln hat für den Menschen eine gewisse Bedeutung - nich' - aber für diese mechanische Hand überhaupt keine.»

Joseph

Weizenbaum

« Wir sind Zeugen des Untergehens des politischen Bewußtseins, und an Stelle dieses Bewußtseins tritt etwas anderes auf: Ich will es - mangels eines besseren Namens - 'Intersubjektivität' nennen. Die Buben und Mädel, die da vor den Terminalen sitzen und miteinander dank reversibler Kabel in Verbindung sind, die wenden der Politik den Rücken und einander, sie wenden sich einander zu. Und das ist eine neue Struktur, die nicht mehr Politik ist, sondern eine Vernetzungsstruktur.»

Vilèm Flusser

«Its a waste of time, to learn things over again, when other peoples brains have those structures, - so the children will acquire knowledge in this much faster way and then they will make small improvements to themselves and then the next person can get the choices.»

Marvin Minsky

Structural principles of the sound space

The central position sounds and the background sounds are looped. The single fragments and phonemes are not looped. The looped sounds are repeated as long as the user is within the sound field. As the sound fields superimpose, a complex structure of sound is produced which is dependent on the users‘ interaction. Their individual movements as well as their focus on certain layers of information allows a differentiation between the rhythmical and the narrative structures. This is a new way of conceiving sound fundamentally different to linear composition.

6.6 The Vision System as a connective structure

The Computer Vision System consists of the Vision System software and the camera suspended from the ceiling above the stage over viewing the whole area of action. The camera picks up the contours of the participants and passes them to the computer. The Vision System traces the path of the visitor's movements as time-based gestures, extracts positional data from them, and uses this data to control the virtual image (avatar), thereby placing the movement in relation to the virtual environment.

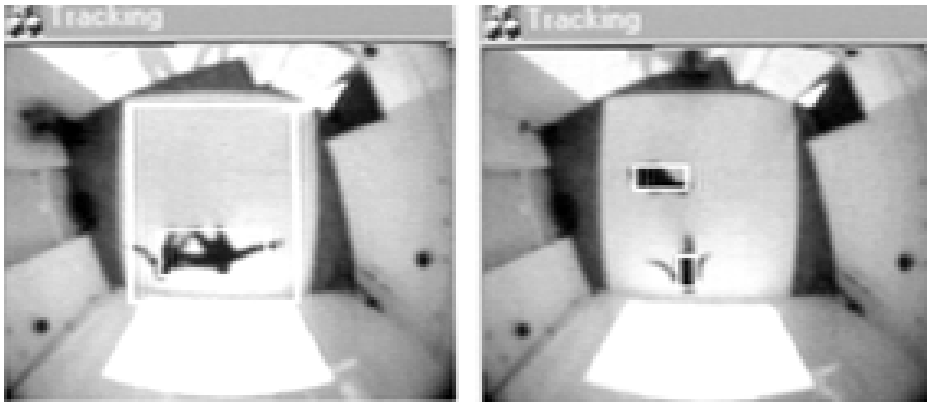


Fig. 18 Vision System tracking window

At the Transmediale in Berlin the system worked with the default image resolution of 384 x 288 pixels which proved too little for an accurate calculation of position and movement direction needed for the soundscape environment employed. The lighting conditions were determined by the white floor and white projection walls. This was acceptable since most people wore dark clothes, although the few visitors wearing light clothes were difficult to track under these conditions. These cases made it necessary to have a human operator manually adjust the threshold of the tracking system. Compared to experience with other camera based tracking systems, e.g. the Very Nervous System, changing light conditions and colours of users cloth are one the major problems. [Wink97]

As we expected a field for interaction of about 4 x 4 meter the virtual environment was modelled in the size of 4 x 4 meter for 1:1 scaling of real and virtual space. In fact the actual size of the interactive field was only 3,33 x 2,5 meter. We had to scale down the virtual environment accordingly. Unfortunately we could not scale down the users. The soundscape appeared to dense in relation to the body size of the users resulting in a difficult mode of navigation. One had to move very careful and slow. For example moving one meter in real space means moving through a lot of sounds in the virtual environment which would need double as much space. The result of this malfunction is a jittering of images from the virtual environment as can be seen in the video for the deliverable.

The interplay of different scale factors in the system needs to be co-ordinated as precise as possible. A problem causing inaccuracy is the scaling of the tracking image and the fact that not all of the tracking data is used for rendering the viewpoints. Every third point is thrown away by the tracking software because of rendering time needed by the VRML client. Also the scale factor of the avatar movement in relation to the virtual environment has to be revised; we need to experiment with different scale factors. The precision of the bounding box is also an important factor for the tracking. The movement of the avatar corresponds exactly to the bounding box movement. The chain of events – which has to be as precise as possible - is the following: the person is tracked by the bounding box, the position data is used for avatar movement control (speed and direction) and the rendering of the viewpoint in the 3D environment. We observe the accurate mapping of bounding box and avatar, but we have to consider the unstable condition of the bounding box tracking system relating to the users' position. Even if a person stands without any movement the bounding box moves slightly resulting in a "breathing" avatar, which looks quite nice and lively.

We assume the jittering images are caused by the movement of a walking person which is not smooth and regular movement like mechanical movement. We tested this situation. Replacing the person with a LEGO Mindstorm moving robot has resulted in smooth viewpoint rendering. We are therefore seeking for appropriate algorithms able to smooth the users path of navigation for better viewpoint control.

Different to the approach of "interfacing real people" (Task 6.3.3) where specific events are triggered due to the movement of the tracked person and reduces to position and the change of position, in our approach the Vision System is not restricted to specific events and the volume of the body and some body gestures are used to define the appearance of the avatar. Comparing the Vision System in 6.3.3 which uses the input from body movement for control of predefined events, our system supports the underlying structure of the interactive material and the interactivity in the Mixed Reality environment.

7 Participation and interaction in the Mixed Reality environment

7.1 The transition from observer to active participant

By participating in an interactive installation, the visitor places himself in an exposed position. He is no longer an observer, he becomes a participant. This presupposes a certain courage for self-portrayal. The usual arrangement for an interactive environment is a closed room – the black box. This is required to be able to assemble and use the necessary display technology. At the same time the black box confronts the acting viewer with an extract of reality or shows virtual environments quite plainly: It isolates the viewer from the real environment and expatiates a clinical clean-room, a sort of isolating block.

Admittedly it is only a technical necessity but the black box provides a certain intimacy which makes it easier for the participant to act. Nevertheless the interactor is acting forefront, in direct contact with the installation, while other spectators, the waiting or the passive audience, is looking over his shoulder.

Offering just a picture frame contradicts the basic thinking behind virtual reality, which propounds the frameless image and the principle of immersion as an experience of reality. If the frame imposed by the black box is not included in the thematic context as a part of the installation, but is instead incorporated only for technical purposes of projection, then the display effect is in the centre, but the interaction thinking is not. In the black box the interactor is at the very front, and the audience is at the back. Depending on the mentality of the participant, this situation affects the quality of the interactive experience. The other extreme of interactive installations provides a kind of waiting room situation with a waiting list to ensure that one person at a time has the experience space to himself in an exposed laboratory situation.

Our Mixed Reality concept, on the other hand, aims to blend or overlay virtual experience situations into the every day space of perception. The question of how the transition between the real and the virtual space is handled is of central importance. The contradiction lies in the fact that a threshold situation must be created, but must not be noticeable as such.

The difficulty lies in creating the transition from the real situation, to the other situation, the virtual. In Luc Courchesnes' work, “Portrait no.1” (1990), the manner in which the display appears in the room serves as an unnoticed transition between real space and virtual discussion partner. The display consisting of a semi-transparent mirror, whose magic effect is reversed by operation of the trackball interface required for the interaction. The interface highlights the threshold again all too clearly. The question is how this threshold is defined in poetic terms and whether the threshold situation should not already be perceived as an important part of the interaction. In search of an adequate design we see the design principles “Form follows Function” confronted with a design of the framework of rules, the discovery of thresholds and rules – “Form follows Structure”.

Above all, if the interface is no longer visible, but only felt through its effect, the problem of transition arises. Once again, an invisible interface creates a sudden surprise,

which raises the participant's awareness, as can be observed with the Computer Vision System in ‘Murmuring Fields’.

7.2 The screen as an interface: types of display

The projection screen acts as a transparent boundary between the interpassive and the interactive zone, and can be observed from both sides. For the observer it is part of the visual offer of perception, and for the active participant it is the visual interface. The screen shows a room-high Netscape window with two integrated smaller Netscape windows showing three individual points of view. The following views are presented in standard setting:

- top view : Plan of space (2 D), overview of virtual environment and avatar movement
- 3D insight view 1: participant 1 subjective view as seen through the eyes of his avatar,
- 3D insight view 2: participant 2 subjective view as seen through the eyes of his avatar, perceived by participant 1 as another persons view

Fig. 19 Mixed reality screen with Netscape windows

The top view gives an overview of the visual environment, while the insight view provides a three-dimensional spatial view – the view seen through the eyes of the avatar. There is no way of choosing between the two types of display. During the presentation the operators switch between views manually. The aim is to examine the different degrees of legibility with reference to 1) the spatial understanding of the three-dimensional view, 2) the understanding of the navigation, and 3) the ability to monitor the presence of the participant in the virtual environment.

Top view (objective camera)

The projection of the large Netscape window shows the overview plan of the virtual scene for the movements of two participants and their avatars. The perspective view of the two participants is provided by two small windows integrated into the large window.

Insight view (subjective camera)

The projection of the large Netscape window shows the view of a participant in the three-dimensional virtual environment. The image of the second participant is shown in a small window, and the overview plan with the avatar movements is displayed in another small window.

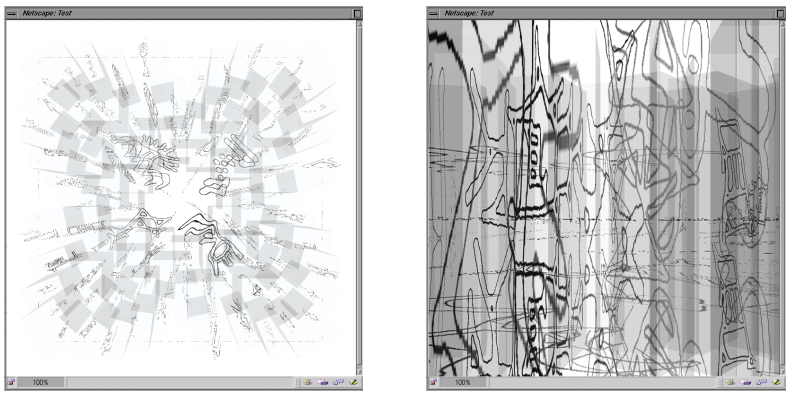


Fig. 20 Separated top view (objective camera) and insight view (subjective camera)

7.3 In-between the screens

The interactive zone lies between the transparent screen and the opposite white projection wall, where the shadow figures from the virtual environment overlap with those of the participants. The objective of the spatial projection is to prevent one central view from being dominant. The fact that the projection appears twice makes the participants feel as though they were part of an image.

The dual projection is a response to the CAVE by creating a display-like environment without stereo projection and intrusive cabling. In the dual projection the synaesthesia of hearing, sight, feeling and movement creates a temporally and spatially overlapping flow of activity between the displays in the virtual environment. Instead of a stereoscopic display intended to trigger spatial perception, here the action is visualised in space, and perception of behaviour in space is therefore supported.

The projection space shows two image areas: the display of the virtual environment and the representation of the participants by their avatars. The background projection area opposite the main screen shows the superposition of the shadows of the participants with the avatar objects and the virtual scene.

The interaction is supported by the special spatial conditions described above. On the first screen the participant controls the avatar, which triggers sounds. In the interactive

zone in the centre the participants are occupied with the avatar game and with themselves. On the second image area the participants see themselves communicating with each other as shadows. The avatars provide a commentary.

Fig. 21 Shadow figures and avatar

8 Evaluation and user feedback

8.1 The audience and the location

During the three-day presentation in the Podewil in Berlin, around 750 people visited the 'Murmuring Fields' installation. This figure is based on the fact that we have received over 1,000 information sheets and on our own estimates. 10 members of the MARS Team were on-site in an open workshop situation. They were in charge of setting up, giving presentations and lectures and conducting evaluation. The visitors were an interested general public as well as open-minded professional audience.

The evaluation of the Mixed Reality installation in a public setting aims to investigate the process whereby an exhibition audience uses the installation over a period of a few days. Observations took the form of interviews, video and tape recordings, and free discussions.

250 visitors were questioned about their experiences and impressions of the installation for around 10 minutes each. They wanted to talk about their experiences and were keen to be interviewed.

We identified two groups of visitors:

- interested lay persons with some prior knowledge of interactive installations.
- media professionals who have some interest both in the digital production of sound and image and in the interactive movement and communication space.

Statistically speaking:

- 20 % were interested in the Mixed Reality concept of the sound space
- 30 % were interested in the technology
- 50 % were interested in the spatial/visual environment.

8.2 Evaluation criteria

The aim of the presentation was to create a public laboratory situation in which to observe people dealing with a Mixed Reality installation. The objective of the evaluation was to assess the interaction process in the Mixed Reality environment for:

- Usability,
- Audio-visual orientation in image space and in sound space,
- Play behaviour as interactive strategies,
- Identification of aesthetic structures.

Interactive works give the viewer the task of making the whole system function. In return the viewer is given power over the work, its form and structure depending on his

decisions. In this process the viewer becomes the user. As a participant he shapes the semantic communication options of the work. Without him, nothing happens. With him a range of perceptive and experimental activities take place. Every time he moves, something reacts as though it were alive itself. The fact that his actions produce different manifestations of space leads him to sound out the visual and acoustic possibilities. There is therefore the opportunity of developing relevant aesthetic criteria for the new medium. Setting up the Mixed Reality stage requires the following questions to be considered:

- What attracts people to enter the Mixed Reality space?
- What is the initiating moment to go into the installation?
- How do the people explore the modes of interactivity?
- Do they play without further explanation?
- How do they play together?
- What is their engagement?
- Is the soundscape of 'Murmuring Fields' engaging or confusing?
- Is the visual environment irritating or helping for orientation?
- Does navigation have to be instrumental and what is the artistic value of experiential browsing?
- Is the minimalist user representation accepted as a figure to play with?

9 Evaluation results

9.1 Comparing the two different modes of presentation

Two different presentations of the visual setting, their associated forms of access and their different forms of interaction were compared.

In the *top view,* a two-dimensional overview map of the virtual environment is displayed on the projection wall. The *top view* is like a bird's eye view. The participant therefore sees his avatar as a flat, wandering trace of colour. Through the reciprocal relationship between the action of the participant and the reaction of the avatar, the participant recognises his simultaneous presence in both spaces – the physical and the virtual.

The *insight view* places the participant in a virtually rendered, three-dimensional environment which changes according to his movements. The avatar is not visible, since it occupies the viewpoint of the participant. Therefore the section of the virtual environment which the avatar is showing to the participant is covered by the view of the participant. A reciprocal relationship develops between the action of the visitor and the reaction of the image space. The virtual scene appears to the participant as an imaginary image and sound sphere that he enters through his movements. This gives the participant the impression of being in the centre of the action. One observer describes his impression as "walking through a forest". The participant's own string of movements

brings in the term "process", and hence the notion of time. Seen in these terms, his interaction creates a period of time which only takes place in the „here and now“ for the duration of his visit.

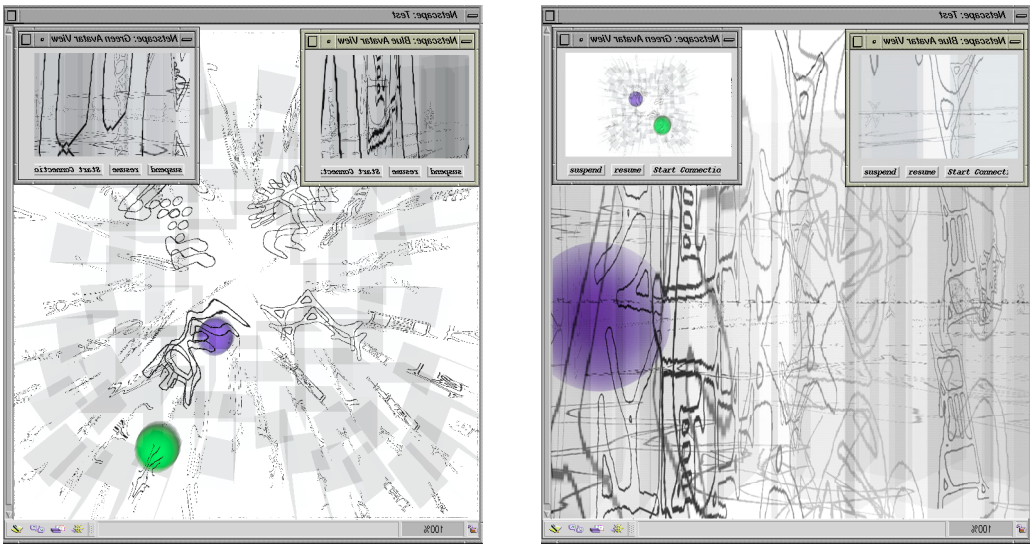


Fig. 22 Top view and insight view

In film terms, the *top view* is an overview of the long-shot, while the *insight view* is that of the subjective camera. It has been observed that a 2D scene (top view) is enough for the viewer to work out the rules of the game. This version can be compared to locating a country on a map. It equates to an intellectual examination of the Mixed Reality environment.

The 3D scene (insight view), on the other hand, places the participant in the centre of a virtual space that he wanders around and discovers like a traveller. This brings in a game-like component that is clearly more attractive to most of the visitors and offers them new experiences.

The play behaviour in communication with the other participants develops from the game material on offer. As in the game "tag", for example, one avatar can capture the other and merge with it. As soon as the two participants move apart again, the avatar figures separate. This may cause the participants to exchange the large avatar field for the smaller one, and vice versa. This happens, for example, when the participants change direction in order to adopt the position of each other. This game revolves around identifying the characteristics of the avatars with the objective of gaining control over image and sound.

9.2 Interaction: space of movement

Interactive environments, such as 'Murmuring Fields', are based on the principle of free play. The important factor is the mechanism that incites participants to act but remains invisible in itself, making it seem as if it happened spontaneously.

The performance situation is determined by the action of the participant, who actively helps create the work through touching, manipulation, shouting, jumping, etc. In 'Murmuring Fields' the participants trigger sounds and voices through their movements.

As they try to find out who is triggering which voices and which sounds, their perception changes from strictly visual to a multi-sensorial and synaesthetic one. The participants are released from a position of an observer that merely stands in front of the work and observes it; in contrast, they find themselves in a space which invites them to enter into communication.

As the participant in an interactive installation develops a dynamic relationship with the work of art, events are torn apart over and over again and then reassembled. In this way the participant identifies the structure of the system, which is built of overlapping components. The non-material is eventually brought into the physical space of action via the circuitous route of the way it is organised. This spatial system of notation makes it possible for all users to structure and play with the sound material, and gives them the opportunity to create new sound images.

We can identify four different phases in the way a visitor approaches the interactive environment: access, selection criteria, examination and experience. According to our observations, the approach to the interactive environment for the access phase and the selection criteria phase is the same in both depictions – the top view and the insight view. When it comes to the examination and experience phases, the tendencies are different.

At the entrance to ‘Murmuring Fields’, the visitor is invited to enter the interactive play space by an acoustic signal – "drrrrh". At this stage the visitor is unable to determine whether the system is functioning at random or whether it is reacting to his presence. At this initial stage most visitors try to discover how the interaction mechanism works. They look for sound sensors by clapping their hands or speaking loudly, and check for pressure sensors by stamping their feet. Neither type of sensor is present.

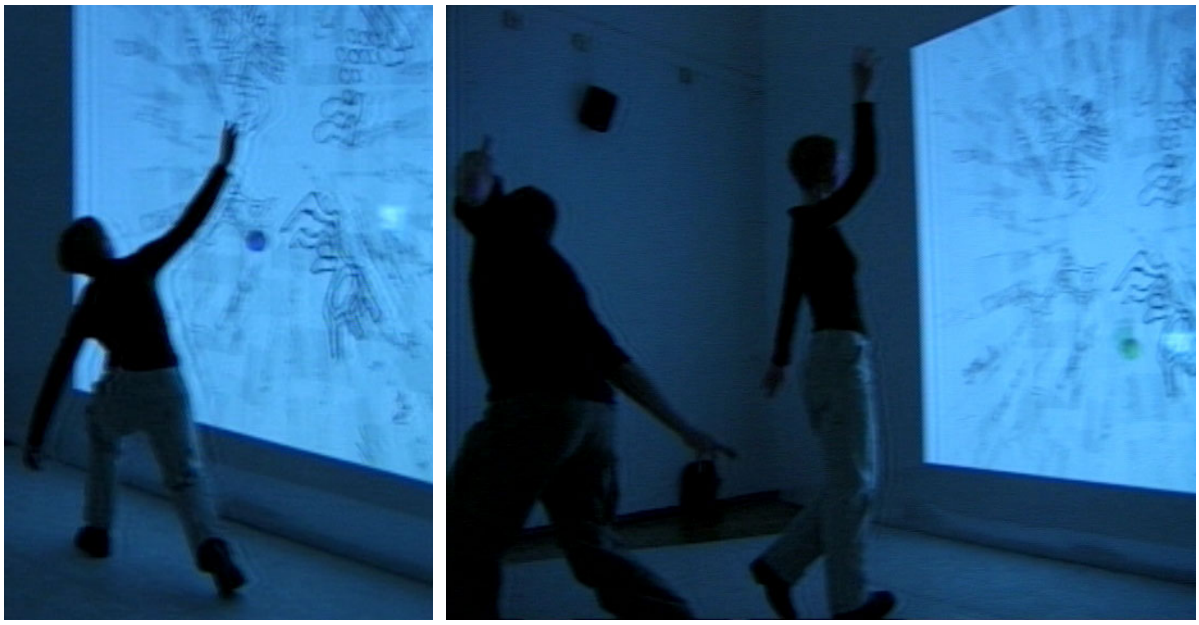


Fig. 23 Visitors play of movement

Once these methods have proved ineffective, most visitors start to observe their own movement, and thus understand how the interaction works – their movement generates a moving image. The influence exerted by participants in top view is limited to correlating their own movement with that of the avatars. The self is no longer presented

as a reproduction, but by a representative; a representative not of the visitor but of his movement. It is a live symbol of movement in an otherwise static system. This symbol is therefore taken to be a game figure.

When trying to identify the characteristics of the avatar – what can it do? – its parameters are discovered: stretching out the body generates density of sound. The sound is no longer a mere multi-media addition, and the visitor understands its function as a sound navigator which he can control by movement. He now turns his attention to playing with movement.

We have observed that the top view incites a different process of movement from the insight view. The two-dimensional display (top view) tends to encourage quiet and concentrated movement processes. Actions are objective-based – intended to catch the other avatar or merge with it. This means that an intensive form of communication develops between the participants. The top view opens up a movement and communication space between the participants.

In contrast, the three-dimensional display in the insight view produces dance-like, autarkic movement processes which are centred on the participant's own body and show a desire to take control of space. It opens up a haptic space of emotion, which leads the participants to interact with the image space rather than to communicate with each other. Here the interaction is between the virtual events in the space and the participant. The participants often develop imitative behaviour in which certain sequences are repeated. Moreover, the observers in the interpassive zone of the environment are encouraged to move into the interactive zone.

Fig. 24 The shadow projection diminishes the dominance of the main screen

The inclusion of shadow projection on the rear wall shifts the action – away from the dominance of the visual projection screen – and into the sound space. The insight view facilitates an experience of the sound and image landscape. Here improvisational gestures and attempts to dynamically change sound and image are predominant. Each participant concentrates fully on himself and, for this reason, his game with the narrative structures can be observed. The participants follow the sound changes in a meditative listening mode over long periods. They navigate through the sound space

with their eyes closed. They attempt to maintain a balance between a passive reaction to the sound image and its active composition through changes in position.

9.3 Interaction: image space

The image space serves primarily as an orientation help for the participants. It is the visualisation of the otherwise invisible interface of the sound space. Because of the abstract symbols that are projected onto a flat screen, the two-dimensional image is only perceived as a three-dimensional image space at second glance.

The graphic symbols in the image space may sometimes disorient the visitor, but most visitors find that the mysterious shapes are visually stimulating and full of excitement. They regret the low frame rate of the rendered images, which results in a hesitant image sequence.

Visitors find the map-like top view clearer. They grasp the navigational situation more quickly thanks to the orientation aid provided by the avatar. The insight view appears less clear to the visitors at first, but they perceive their room to manoeuvre to be larger because movement takes them into the space.

9.4 Interaction: sound space

In his "Philosophie der neuen Musik" ("Philosophy of new music"), Theodor W. Adorno stresses the immersion effect created by the concentrated listening process that takes place in sound installations as a central paradigm of listening psychology. Adorno distinguishes between two types of listening behaviour. In one case the immersive effect of rhythmic-temporal processes is preferred, while in the other case preference goes to the figure-space perception (sound colour – concurrent sound – melody).

These two types of listening behaviour can also be identified in this trial. Some visitors are more interested in the rhythmic structures, while others are more attracted by the harmonious sounds. While the rhythmic structures of interaction and the visitor's own desire to create are an encouragement, the harmonious sound images trigger a sensory attitude which circumvents the active relationship to the sound image initiated.

One third of those questioned were technical, image or sound experts, and were particularly interested in the concept of the sound space and its spatial notational principle. Composers, music experts and sound engineers saw the advantage of the Vision Interface (camera) and the unhindered navigation that it permits. The fact that, with this interface strategy used in a stage situation, it is not the sound that defines the dance movement, but instead the movement which triggers the sound, is considered an innovative production technique.

Music theorist Barthelmis sees the sound space in the installation as "innovative, both in its artistic and its technical meaning". It is very densely structured and offers numerous possible combinations, which the visitors assume to be poetic structure. This causes most visitors to spend a fair amount of time (up to 30 minutes) in the installation.

Conclusions:

The room made available at Transmediale was not suitable for sound installations. Because of the short set-up time, it was not possible to make any acoustic improvements. The room tends to carry sound away, which is disadvantageous to the

visitors and to their interaction and listening experience, and interferes with the perception of the foreground and background of the sound image.

Suggestions for further development of the sound space:

- Algorithmic correspondence between sound and image
- Sound-based presentation instead of the dominance of the visual
- Rhythmic structures instead of harmonisation
- Sound structures instead of language
- Generative interactivity instead of prescribed sound elements

10 Lessons learned so far

The design and evaluation of the eMUSE system in a public setting, is related to basic research in the field of mediated perception. In the light of the necessity to find new ways of man-machine-interaction which contributes to the development of interaction as cultural practice, media arts comes close to the design of advanced telecommunication practices. It shows the importance of aesthetic consideration understood in its initial sense, i.e. as examination of the basic principles of perception. The aesthetics of communication becomes a highly relevant research goal that resolves traditional categories like art and design and proves the social impact of aesthetic experiments with new communication systems.

10.1 Questions of spectatorship

At the Transmediale, the Mixed Reality environment could be experienced from two different zones: the *interactive zone*, the action field to perform in, and the *interpassive zone*, the place for visitors to watch both the virtual world and the active participants through a semi-transparent projection screen. The monumental projection plane addresses two needs: on one hand, it supports the feeling of immersion through the size of the plane. On the other, it also functions as a visual display for the audience. In this way the question of how to intensify the aesthetic experience of the audience is also taken into account. The setting in Berlin has shown that the design of the *interpassive zone* strongly determines the overall acceptance of the system.

Presenting different viewpoints, especially the simultaneous transmission of the two avatars' insight views, underlines the dynamics of perception, as the image of the virtual object becomes a dense field of different viewpoints. For the audience, the virtual environment appears as a dynamic sculpture that is not defined by its static appearance but by its ever-changing viewpoints. The mental image of the environment becomes an individual combination of actually experienced and memorised impressions. With this kind of display arrangement we prevented the public from falling back into traditional modes of spectatorship which are defined by passive observing as for example in cinema. Besides the question of how to immerse the participants to ensure a convincing interactive experience, strategies to involve the audience are of major importance of the successful implementation of Mixed Reality environments as electronic arenas.

10.2 The invisible interface and the support of intuitive gesture

The question of accessibility to the interactive system becomes highly relevant when the relationship of its location is examined. The interface of eMUSE does not force the user to learn a new technique but refers to their normal physical activities, which become the central part of the interactive structure: the dynamics of intuitive movement are symmetrically translated into the virtual environment. Gestures and movements that are, either consciously or not, elemental to everyday spatial experience are the fundament of using the system.

The application of the eMUSE system at Transmediale shows that eMUSE lowers the threshold of user access. Its invisible interface and the capture of intuitive movement and gestures avoids direct manipulation of physical devices and refers to unconditioned gestures as they are common in everyday life. The users show a higher degree of attraction to the system, as they are not forced to learn techniques. As a result, their focus is on spontaneous movement in the action space and the corporeal actions show a high degree of freedom.

The soundscape forms the main part of the set-up and has a special function in the Mixed Reality environment: it guarantees and reinforces physical spatial experience. When the user enters the action field, he immediately produces an individual sound collage. Navigation through space is accompanied by the perceptual process of listening. The users move dynamically through space and tend to continue moving. This means that the soundscape not only parallels the visual landscape but also forces dynamic experience of the spatiality of the physical surrounding. The spatialisation of sound appears to be highly effective in transforming neutral space into a space of dynamic experience.

10.3 The informational landscape of 'Murmuring Fields': immersion and physical interaction

In most virtual reality systems, the spatialisation of data follows common concepts of space as a container where – for the sake of comprehensibility - single, easily identifiable objects are situated at fixed locations. In contrast to that, one of the main characteristics of the virtual environment in 'Murmuring Fields' is that data is not presented as clearly defined single units but as a dynamic visual and acoustic structure:

- neither the virtual nor the acoustic landscape shows close frontiers or easily identified landmarks of orientation,
- the soundscape offers few landmarks of orientation - the philosopher's voices - and the users movement in space results in a continuous collage of sounds.
- the visual landscape follows this artistic principle which explores an aesthetic of fusion and dissolution: abstract shades diffuse and dissolve while the user virtually moves along and through them.

The renouncing of easy navigational structures forces the users' attention to the relationship between his movements and the environment: space is experienced as dynamic through the involvement of the users' actions. Through the definition of the perceptual process as continuous intersection between data streams and users' unhindered movements, the environment provokes a high level of activity. The

behaviour of users at the Transmediale shows that in an environment which does not offer pre-fixed structures, the users' part in experiencing the virtual scape is extended. As the evaluation shows, the insight view leads to a higher level of physical activity than the top view.

We conclude that the insight view leads to a mental immersion into the visual landscape. One may compare the two different viewpoints, insight view and top view, with the everyday experience of urban space. While the top view corresponds with the reading of the city map, the insight view can be compared with the actual experience of urban space where the body is immersed in a situation of various, constantly changing sensual stimuli.

The insight view gives up the perceived distance between the viewer and the virtual environment. As the top view establishes a virtual distance, the insight view forces a feeling of immersion which itself leads to a higher degree of interaction with the virtual environment. Visitors of the exhibition claimed to feel "stroked" or "washed around" by the dynamic components of the environment as experienced from inside.

In Berlin, ‘Murmuring Fields’ was presented with a set-up that combines top view and two insight views, the one user‘s view and view of his interaction partner. These views were simultaneously presented on the projections surface. The perceptual process is characterised by a continuous shift between the user's different viewpoints (insight view or top view) or from the own insight view to the insight view of another participants. As one can experience different viewpoints of the same environment simultaneously, the passage through the virtual environment becomes accompanied by elements of distraction. With these slight moments of irritation the users become aware of the presence of other participants and their viewpoints.

From the experience at the Transmediale we draw the conclusion that the insight view forces a higher level of experiential interaction. Especially dancers preferred the playful investigation of the environment through the insight view. This is in opposition to many virtual spaces that prefer a rigid distance between viewer and environment through the construction of perspective (e.g. as in conventional digital reproduction of architecture). In contrast to this instrumental approach to navigation, we defined virtual space as dynamic and mutable, and hence inseparable from user's presence. Although the distance between real and virtual space remains established (as opposed to full body immersion), users project themselves into the virtual environment. The high level of physical activity demonstrates this mental immersion, i.e. the user accepts the virtual surrounding as action field. The demonstrated design of the virtual environment proves to be a powerful tool to provoke feelings of immersion even if the distance between physical and virtual space remains.[5]

We regard the high level of physical activity as an indicator of the acceptance of the virtual environment by the users: as they interact they define virtual space as action space and thus accept digital space as part of the experienced reality. Movement can be defined as continuous and dynamic physical interaction with the world. In ‘Murmuring Fields’ movement is directed towards immaterial phenomena.

[5] As the experience in Berlin has shown, a high level of physical interaction does not only lead to satisfaction of the users, but also for the audience who watched the spontaneous dance performances. This point is of course of importance for the acceptance of the system's entertaining qualities for passive perceptants.

We conclude that the acceptance of a virtual environment is not exclusively dependent on its design of an illusion. Obviously artificial environments, such as the one of 'Murmuring Fields', become familiar if they are able to involve the user in a circle of intuitive interaction. The body, its gestures and movements, form the link between real and virtual space. The acceptance of both as an action space defines these two worlds as equivalent and furthers the successful implementation of the Mixed Reality installation in a public setting.

10.4 Mixed Reality as contemporary public space

The major goal in developing eMUSE is the design of a publicly accessible system that offers new kinds of spatial experience to the user. The question of bodily awareness is a major topic in contemporary media theory. Many scholars speculate that face-to-face interaction will be substituted by connected communication in cyberspace. The dematerialization scenario is based on the idea that "electronic forms of communication and a range of tele-based services simply displace the need for physical movement between home and work, while urban functions will no longer have a physical presence, as services are delivered in electronic form."[Mar97].

With its exploitation of Mixed Reality conditions, the work described contributes to the development of a new social space by exploiting basic questions as navigation and corporeal experience in physical and virtual space. It forms a fusion where the experience of the physical space is as important as the one of virtual space. Virtual environments do not replace real space. Rather, eMUSE interlocks them and examines the basic condition of their mutual dependency. With the installation of eMUSE as 'Murmuring Fields' at Transmediale, the focus was on the setting of a Mixed Reality System in public space, combining passive spectatorship and interactive experience to a new form of an electronic arena.

For this reason we have to investigate the nature of interactive experience of the participants as well as the experience of an audience which, in contrast to most interactive installations, is regarded as an important element for the question of user acceptance. Spectators form an essential part of an installation which is meant to supply new environments for public events, and hence heavily depends on the acceptance of visitors which are not involved interactively.

11 Mixed Reality in performance space: FIDENA '99

The second public trial of the eMUSE system took place in the framework of the symposia „Wahrnehmung der Wirklichkeit, Wirklichkeit der Wahrnehmung" funded by the German Ministry[6] of NRW for culture. The symposium is part of Fidena'99[7] – (Figurentheater der Nationen) an annual international theatre, puppetry & new media festival in Bochum, exhibiting and discussing actual productions in the field of theatre, performance and installations. 'Murmuring Fields' was invited to be presented and discussed as a work in progress with the possibility of a five days' rehearsal. These five days were planned as 3 days for experimenting with the setting of the stage and 2 days for working with two performers and a stage director. The presentation took place in the theatre of the Museum Bochum on April, 25. Amongst others, the Australian artist Stelarc held a performance lecture about the relations of the body in his implications with the interactive media. The basic statement of Stelarc concerns the changing condition of the body shaped by digital culture. The performance can be seen in the video material part2.

11.1 Context : eMUSE in theatrical setting

The productivity of the theatre rests on the reflexive connection of inner human perceptions and the materials in the outside world. The Mixed Reality stage reverses this situation – the inner perceptions of the performers become visible through their dealings with the virtual traces on the Mixed Reality stage which embody an invisible outside world. On the Mixed Reality stage, the stark difference between perception and action normally found in the theatre is dissolved. Several performers interact simultaneously in the virtual space and can thus generate a new situation at any time. The question therefore arises as to the effects of this fact on the theatrical interaction. In the setting of the interactive Mixed Reality stage in a traditional theatre context the following main issues are explored:

- the usability of technical conditions for the performer in relation to the Mixed Reality system
- the possibilities of development of interactive storytelling methods.

The Mixed Reality stage presentation focuses on the influence of an interactive stage to the individual actor and the interplay amongst several actors. The question for the performing artists is to find interactive strategies within the environment. We also question the reaction of the audience when confronted to an interactive stage and the influence of interactivity on the performance:

- What do we learn from performers?

---

[6] FIDENA is supported by: Ministerium für Arbeit, Soziales und Stadtentwicklung, Kultur und Sport des Landes NRW, den Beauftragten der Bundesregierung für Angelegenheiten der Kultur und der Medien, der Stiftung Kunst und Kultur NRW.

[7] Biennale des deutschen Forum für Figurentheater

- How do professionals perceive Mixed Reality stages?
- How do they work with body centred intuitive interfaces?
- How does the interactive element become apparent?
- How does the audience influence the performance?

To answer these questions, a team of specialists from different disciplines followed the event from the first set up and the rehearsals to the final presentation.

- Hinderk Emrich: Neurologist
- Ulrike Hass – Theatre studies expert
- Dirk vom Lehm: Cultural scientist
- Danièle Perrier: Art historian
- Walter Siegfried: Sound artist

The invited performers were: Maya Brosch and Martina Leeker, directed by the stage director Lambert Blum.[Blum95] They worked before several times with David Rockeby and his „Very Nervous System“. [Rock93] They therefore have experience with interactive systems in performance situations.

12 Content and structure of the Mixed Reality environment

For the Bochum event the virtual environment was remodelled to accommodate the different situation of Mixed Reality in a performance setting (Fig. 25). At Bochum the experience of the environment was differentiated between the experience of the performers and the experience of the audience.

12.1 Physical space: environment set-up

An open stage setting on a raised platform is planned in conjunction with the organisers of the symposium. The audience should not have a fixed place, but should be encouraged to walk around the stage. Therefore seating is not to be provided. The Mixed Reality stage is to be set up between two transparent projection screens, and is to make the action on the stage visible on three levels:

- The avatars are displayed on a gauze screen in the foreground of the stage.
- Behind the screen, the performers, dressed in white, are illuminated by the projection beam and act as shadow figures in the background. Their shadows appear larger than life on the stage backdrop.
- The small avatars overlap with the larger-than-life shadows and the life-size silhouettes of the performers. A game therefore develops between the different levels.

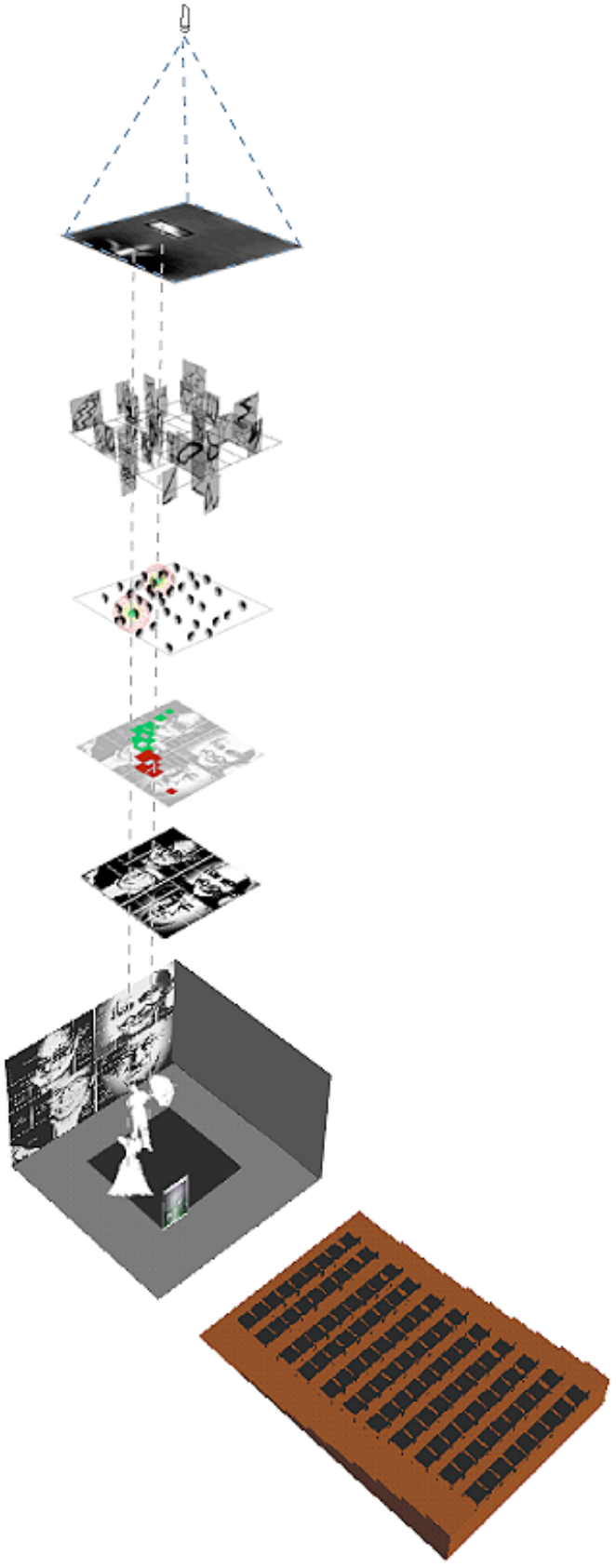


Fig.25 The components of the Mixed Reality stage

Because of the requirements of other performances we were faced with a situation far different from the one we were expecting. The compromises that had to be made because of the tight timetable completely changes our planning. During the short three day set-up period, we therefore not only have to try out different situations on an ad hoc basis, but also completely change the complicated technical setting. In order to give the audience an insight into the ongoing operating process, we originally planned to place the computers and their operators at the side of the stage, so as to allow a question and answer game to develop between the team and the audience. Instead of this the technicians now have to communicate with each other, the performers, the producer and the writers from two separate locations in the room.

Given the altered arrangements, we now have the following situation. The audience has to sit on rows of seats as in a normal theatre. They therefore only experience the environment from a frontal position and find themselves looking at something resembling a television. Since the projection screen provided is not sufficiently see-through, the front screen is removed for the presentation. The virtual environment is projected in black/white on the rear wall of the stage area. The avatars of the two performers appear as red and green traces of colour. In the foreground of the stage is a small projection wall showing the live video image from the Vision System, which allows the audience and the performers to follow the movements of the figures on the stage in real time from the perspective of the tracking camera – i.e. a bird's eye view.

The performers are dressed in white and look like mobile projection surfaces against the black stage background. The performers pick out individual image components – sentences or words – which are heard as a sound at the same moment. The activated virtual sound object becomes audible and at the same time visible on the body or in the space.

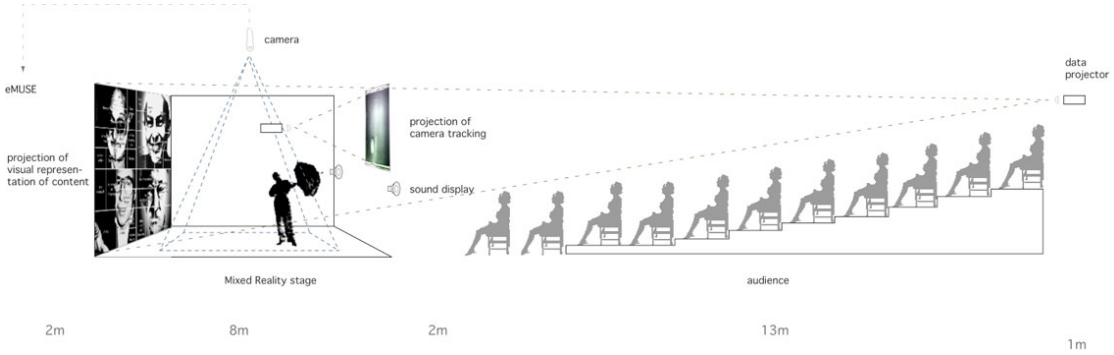


Fig. 26 Spatial set-up

12.2 Content of the Mixed Reality environment

Given its comparability with the situation in Berlin, the concept of the speech spaces and their structure is used for the performance situation in Bochum too. It is an attempt to create a speech-based navigational space in which text montages are created through interaction with the virtual environment. Texts are arranged spatially and then continually presented in various contexts by the different positions of the performers in the space. Speech positions are interwoven and appear as a network of dialogue. Four opposing speech spaces are arranged in the corners of the virtual model, and thus on the stage as well. The spatial arrangement is intended to support the spatial memory of the performers.

The underlying conceptual question is that of structures for information space. Rethinking the role of antique mnemotechniques, the arrangement of the interactive stage is an experiment of spatial order of information space. The concept of mnemotechnique attaches memory to physical space, as for building a mental map. A historic source can be found in the theory of Ramus (Pierre de la Ramée 1515-1572) [Ong97]. Ramus wanted to "transplant the logic of Aristotle from the sphere of contemplation into the sphere of action" and make "an instrument for gaining, processing and transmitting useful knowledge" out of it. The main characteristic of Ramus' logic is the reworking of the teaching of the *Örter* ("locations") of thinking (Greek *tópoi,* Latin *loci*).

This teaching, the Topic (from the Greek *Topika*), deals with points of view from which it is possible to *erörtern* ("discuss") various theories. The Topic became very important in the 16th century, since the amount of available knowledge had snowballed in the age of discovery and letterpress. The Topic served as a method of classification, since with its help it was possible to place new knowledge into a systematic relationship with existing knowledge, in other words to intellectually *verorten* ("displace") it. Ramus saw the concept of the Topic as an "instrument of unity of thought and speech". He divided this instrument into two parts, to which he gave the rhetorical names *inventio* and *indicium*. *Inventio* means systematically walking around the *Örter* ("locations") in an intellectual space in which proof or arguments for the thesis under investigation are to be "found" (Lat. *invenire* = to locate, to find). By addressing standard questions to the object of the investigation, these "locations" make it possible to "find" relevant statements or sentences which join together to produce a "discourse" about that particular object.

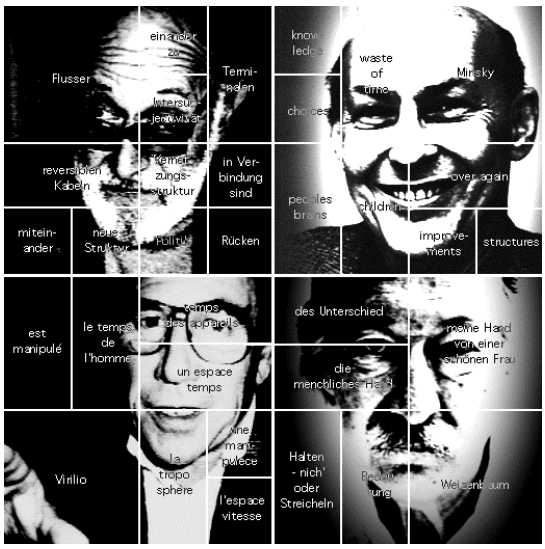


Fig. 27 Fields of language map

The second part of the Ramus instrument is the *indicium*, the evaluation of the sentences and structure of the discourse. This process involves "methods", which are thought operations dealing with the linking of individual sentences to form larger units. Each term is defined and then broken down into sub-terms. These are defined and broken down again, and the process continues until nothing more can be defined and broken down. These terms can be visualised using linking lines to overview tables or branching tree diagrams. A further aspect of the *indicium* is that of linking all the discourses,

sentences and terms together – or networking them (*coniunctio artium omnium et ad Deum relatio*).

'Murmuring Fields' only offers a simple model of the speech spaces, since it operates with statistical, unchanging material. The provision of the speech spaces is an attempt to analyse their utility for the performance. The current question about the creation, visualisation and structure of information space concerns the extent to which spatial memory techniques and classification systems can deliver a model for the creation of a dialog-based communication space.

12.3 The avatar as part of the stage

The original plan for the Mixed Reality stage in Bochum was to use additional floor projection with the map of the sound locations which would have served as a map of orientation for the performers. The avatar was to appear on the floor of the stage, the floor therefore becoming a kind of mirror giving a view into the virtual. However, the floor projection caused an imprecision in the tracking system, and it was therefore decided to project onto the back wall of the stage. To the audience, the avatar display on the transparent screens appears in the way described in the chapter entitled "The spatial set-up". In the stage performance actually presented, the avatar display plays only a minor role. The avatars navigate around the projected stage background in accordance with the position of the performers. The projection is a two-dimensional map of the speech landscape which shows the avatars as traces of the movements of the performers.

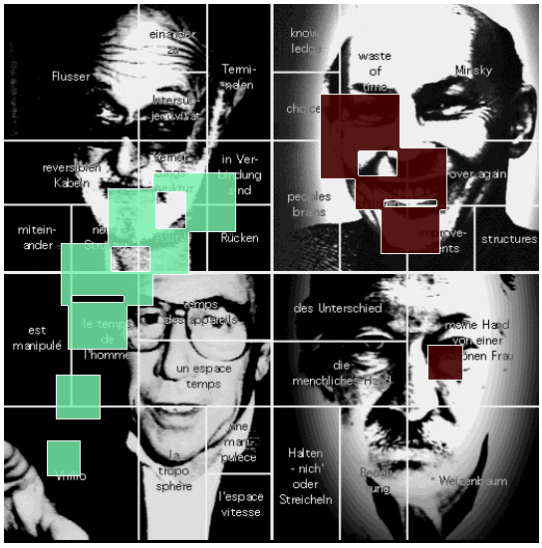


Fig 28 Map of word fields and avatar traces

The avatar is like a shadow, a symbol of presence and movement in time. The camera positioned above the stage records movement from a bird's eye perspective. The space that the body occupies when seen from above determines the size of the avatar. The avatar therefore changes direction, speed and size in relation to the movements of the performer and the expansion of his body. If the performer makes a stretching movement, he triggers a number of sounds. On the other hand, small movements near the body produce a small avatar.

12.4 Visual representation of content: space of image

When the Mixed reality space is used as a theatre space, the relationships between experience and action change. The audience experiences the matter of the Mixed Reality second hand. Removed from the experience of action and re-action they are presented the Mixed reality space through the actions of the performers. It becomes the task of the performers to "show" the spatial construct – to express the spatial structure.

For the Bochum event a series of visual as well as sound environments were created. The environments follow the same basic structure of the orthogonal structure.

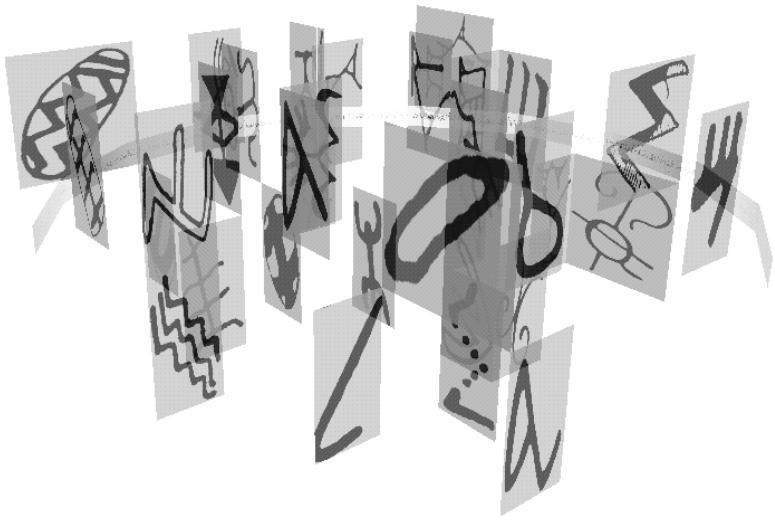


Fig. 29 Isometric view of the visual environment

The visual environments inherited the language of the Transmediale model. The signs taken from extinct languages were placed as textured polygons extruded from the ground plane of the orthogonal structure. The result was a similar weave of language and lines as the Transmediale model although radically simplified. The differentiation between the models was the inversion of black signs on white background and white signs on a black background that allowed flexibility during the set-up towards lighting of the stage and the sensitivity of the camera tracking system.

The virtual map gives an overview of the sound structure. It is a map of content - a top-view of the four quarters identified by the portraits of the media thinkers superimposed by the structure of subdivision and the naming of the words that the sound clips contain.

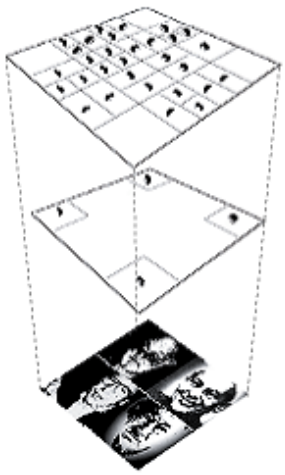
During the set-up of the Bochum event it became clear that the intended set-up was too complicated. As a minimal solution we decided to use the virtual map as a reduced visual representation of the environment. Confronted with the virtual map as the only visual representation of the virtual environment meant that the focus on the soundscape and the camera tracking window was accentuated.

As a testing environment and to allow our selves to compare and analyse the setting we also set up the Transmediale model during the rehearsal period and after the Performance event when the audience were invited to enter. Using the Transmediale model in the Bochum setting allowed us to see the environment under far better conditions. The action field was enlarged to 5,20 x 5,20 meters, scaling up the

relationship between the users' body size and the size of the proximity sensors, thereby lessening the complexity of the soundscape and slowingdown the movement in the visual scape.

12.5 Audio representation of content: space of sound

Two independent soundscapes were presented at the Bochum event, both following the basic orthogonal structure. The soundscapes are variations of the original sound material derived from the four statements of the media thinkers.



"Positions": sound clips (not looped) with words and short phrases.

"Talk": Looped full statements. The individual speeches are positioned in the corners of the sound field.

"Content": The four quarters of the media thinkers with speeches, phrases and words as elements of a discourse.

Fig. 30 Layers of the space of sound

Scene 1: Speech-Play

The main goal of the soundscapes is to accentuate the interaction and to make the situation of a discourse. Therefore the environment has the character of a "speech-play", developed as a simplified version of the structure in the Berlin Transmediale environment.

Fragments of the statements build the configuration in the centre. As the performers move through the environment they make new configurations of phrases and single words. By changing the succession of the sound clips they create poetical speeches with concise rhythmical movements.

To keep the situation as clear as possible, no loops in the background structure are used. Furthermore, there is no superimposing of sound clips. The characters of the quarters are defined by the structure of narration and the media thinkers' voices. They are differentiated by the number of fragments and the length of the cuttings. The scope of the sensors is defined by the length of the speech fragments and the scope of the body. The smallest sensor is 50 by 50 centimetres. A longer fragment is 50 by 100 centimetres. A longer fragment gets a bigger sensor, a small word gets a smaller one.

Scene 2: Rhythm play

The second soundscape offers a more dance-like situation as well as a more fragmented narrative. Following the same orthogonal structure as scene 1, the soundscape differs through its intense use of sonorous and rhythmical forms demonstrating different ways of transforming spoken language. This environment allows new modes of playing, ways to interact with speech as with a musical instrument.

In the first environment the user deconstructs single words by fast, rhythmical repetitions of small movements. The sound-clip always starts at the beginning. As a result, the user can produce patterns such as: "Politik … popoppolipolitik.“, or "Politiktiktik-k-k-k…". The repetition of syllables invites the actor to change the whole word. By scratching (repetition of movement) the whole word can be deconstructed: "P-p-p-po-po-poli-poli-tiktik" etc.

The sonorous changes are done through the transformation of syllables: setting them in a polyphonic chorus, changing the internal rhythm, inverting the syllables etc. Other groups of transformations change pitch, colour and speed of language. Words are transformed into little melodies, like in childrens' songs.

The characters of sound and rhythm in Flusser‘s and Weizenbaum's quarters remind one of different strategies of rapping or "slam poetry", although Weizenbaum's quarter contains more melodic elements. Minsky‘s quarter demonstrates transitions to the methods of minimalism. Virilio's emphatic language is transformed into technoid or trippy sounds, similar to the vocals of rave-music.

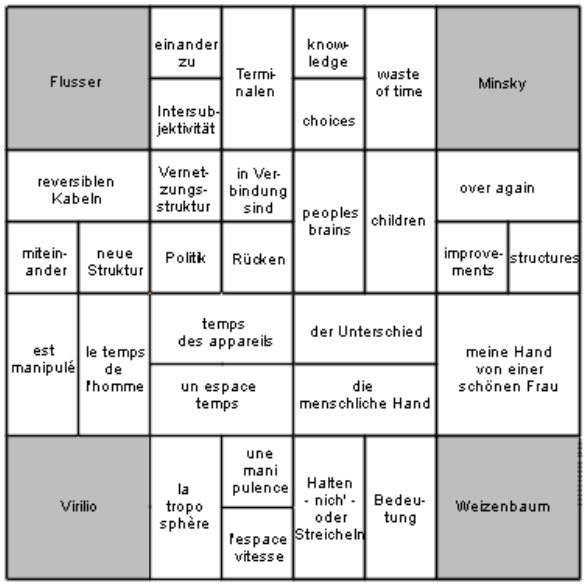


Fig. 31 Structure of the soundscape for scene 1 and scene 2

12.6 The Vision System as a connective structure

During the FIDENA presentation we used an extension of the original tracking system. The greater height of the room in Bochum allowed the use of a video camera with a lower visual angle, with the effect of reduced optical distortion and better image quality. The problem with camera lenses with very high visual angle is the increasing spatial distortion when moving form the centre of the image to a border. A straight movement

in the room in the border area results in a curved movement in the camera image. Accordingly, the size of the user is decreasing from the centre to the border of the image. We also observed the effect of decreasing lighting from image middle to the corners. The corner pixels were nearly black, which made them indistinguishable from a human user for the computer. Besides taking a lens with lower visual angle, in Bochum the floor was black and the actors had to wear white to overcome the problem of decreased lighting.

The enhanced image resolution (768x576 pixels) was used in order to increase the number of pixels representing the users in the image. However, due to falsely classified pixels the calculation of position and movement direction could not be done more accurately.

At the Bochum event, the environment was tried out by selected, trained people. This permitted us to dress them in specially chosen costumes with high contrast to the background. The performers had to wear white costumes and umbrellas, which made each of them be bigger in relation to the overall image. This simplified the tracking of a human body in the camera image.

The biggest difference between the set-up in Berlin and in Bochum was that the actors in Bochum had a direct view off the tracking results produced by the vision system. The audience and the performers were able to watch the tracking camera image showing the person encircled by a bounding box. This enabled the users to control their action and the response of the tracking system visually. Therefore it was easier to bridge the gap of being present in both spaces the real and the virtual. Using the control window of the interface system as part of the stage design makes the technology more transparent to the performers and the audience. Conceptually, this raises the issues of presence.

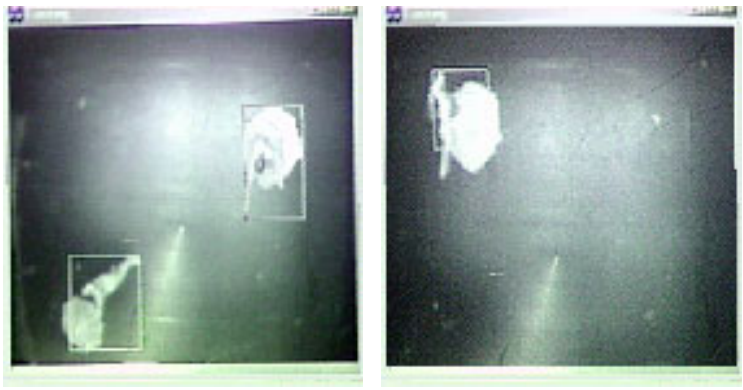


Fig. 32 Vision System tracking window

The use of the video camera as a tracking system can be understood as a new variation of Albrecht Dürer's perspective window ("Handbuch des Malers", 1525), which taught us how to convert three-dimensional nature into a two-dimensional image. Anne Friedberg describes this technique of perspective as follows: "The world is perceived and understood as an image. The subject stands outside and imagines reality." [Fried95] Participants see themselves acting on the screen window of the observation camera.

'Murmuring Fields' offers a paradoxical situation to the users. In the tracking window of the surveying camera, one can see ones' self acting. The user sees himself from an external point of view in the context of his environment. In other words he takes the position of an objective point of view. In the Mixed Reality environment the subject becomes the object assuming its own shape of appearance. The subject watches itself

from a distance as a digital ego. The distance between the digital ego and the subject dissolves. This, along with the subjects concentration on the tracking window, leads to a different perception of the self. The experience of this artificial framework determines the participants‘ behaviour. It is exactly this paradox of „being in“ and „being out“ at the same time that characterises Mixed Reality perception.

Looking at the camera tracking image transforms the action in real time. This media transformation opens the way to a general virtuality, which sets an end to reality by nearly anticipating that end at every moment. Heideggers conclusion stated in “Zeit des Weltbildes” (1938) becomes true: “The basic procedure of modern times is the conquest of time as image.”

13 Participation and interaction

13.1 Interactivity as a tool in stage production

In comparison with the Mixed Reality environment in Berlin, the structure of the Mixed Reality stage serves to formulate interaction in the context of the performance. The two performers Maya Brosch and Martina Leeker, and the literary and artistic director Lambert Blum, who are familiar with the installation from Berlin, are to examine the interactive stage and its possibilities in detail with the aim of developing game ideas. There is limited time in which to rehearse together and to sound out the system, but it is the only opportunity to work with performers – something that we consider important to the continuing development of the work.

Having set up the computer and the stage, installed the system and tuned the real and the virtual environments to each other, there were only two days left to deal with the theme, the content, the dramaturgy and the choreography of a short workshop presentation. Normally it is the theme and choreography that are developed before the stage itself. Here, the procedure is reversed. The framework of the Mixed Reality stage is prescribed and serves as a basis for the performers, in a similar manner to the text in a traditional play or the choreography in a ballet. In this test phase, the tracking system is tuned more precisely to the movement of the performers and the light conditions on the stage. For the performer it is essential to learn the possibilities of the audio-visual interactive game environment and the interface to its mechanisms in order to develop a choreography from the opportunities which exist and the imposed limitations.

For the performers, the art of dealing with the given framework lies in the need to adjust their technical repertoire to suit the requirements of the system. Traditional forms of movement, such as sign language or gestures, have no meaning at this stage, since the Vision System is not based on gesture recognition. Instead the Vision System registers the simple stretching movements of the body and its movement through the space. The rehearsals therefore serve to thematize the limitations that arise from the view of the performance, and to contribute to the development of specific methods of training physical expressions for the media performance. The notion of interactivity is applied and examined in an experimental manner as a tool in stage production.

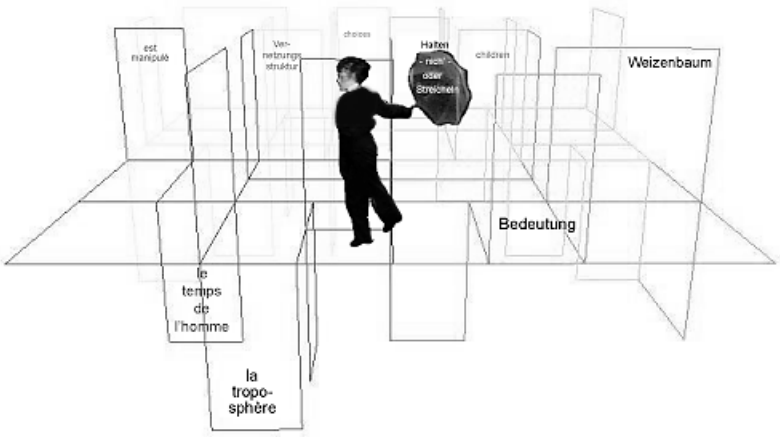
In the rehearsals the interactivity is directed primarily at the performers and at developing theatrical structures. In this case, the professional audience are simply observers for the subsequent discussion. In the stage situation we are dealing with various forms of perception: the perception of the performers, the perception of the

audience, the perception of members of the audience who become performers. The interactivity of the audience is not of prime importance, but one of the professional observers is asked to take part in the action on the stage as a non-professional performer.

In an initial three-hour rehearsal, we established that the performers find it difficult to detach themselves from their professional habits, and tend to work against the system rather than with it. The interactive stage suddenly becomes a threatening machine against which they try to defend themselves. Despite prior discussion and their common interest in the development of the Mixed Reality stage, their professional instinct proves stronger, causing them to apply their bodies, rather than to use them as an instrument. In contrast to the traditional theatrical concept of presenting oneself on the stage, the concept of Mixed Reality requires them to see themselves in relation to the environment. Therefore, on the second day, the literary and artistic director of the group takes the possibilities he can see, and develops the concept of dividing the performance into three parts, which relate to the three virtual 'Murmuring Fields' scenarios which have already been prepared. The rehearsal also gives the technicians a chance to determine and co-ordinate the scene sequence. 13.2 Interaction as a method of story boarding

The first part provides for the investigation of the purely linguistic, virtual environment, which is demonstrated in concentrated, slow movements. The virtual speech network is arranged at the four corners of the stage and is introduced to the audience as game components. Investigating the location of the speech fields leads to a contrast between silence and sound that corresponds to the process of movement. The scenario is characterised by slow-motion-like movements aimed at a slow, sequential discovery of the system. The performers therefore cross zones of silence and islands of words. If a performer touches the words on a virtual level, she automatically pauses, which breaks the momentum of her exploratory movements for a moment. The first scenario is stopped by the literary and artistic director after about 5 minutes' improvisation. In the background the second scene is loaded onto the computer. It uses the same visual backdrop, but this time with a more complex audio network.

Fig. 33 Snippets of the virtual become visible on the costumes and props of the



performers.

The second scene differs from the first in the number of performers and in the sound. Two performers try to construct a rhythmic dialogue with the denser sound network.

The white costumes of the performers serve as a projection surface for the virtual environment. Snippets of the projection, which is otherwise only a background element, now appear in the centre of the stage. The second sequence is determined by the rhythm of the synthetically processed language material and by the dynamic strings of movement which extend the sounds until they form a murmur. The different body language of the two dancers creates a sound collage. The exciting situation is stopped suddenly by a command from the choreographer.

The third audio-visual model differs from the others in the complexity of its sound space and in the visual display. It shows an abstract landscape of symbols that look like trees of knowledge. The trees start moving and form a moving network as soon as somebody enters the stage and activates the word and sound network. A professional observer is asked to investigate the scene using a process of improvisation. He commences a detailed investigation of the Mixed Reality stage by observing the effect of his movements in the tracking window on the front edge of the stage. He becomes aware of the subtle sounds being triggered by the dancer. During his patrol of the sound landscape he approaches the dancer in a certain rhythm. After a stage-managed dialog between the speech spaces, the two performers move into the same location. The image freezes and the sound ends in the prolonged sound of wind.

These three scenic displays are particularly interesting from the point of view of developing new, networked communication routes. The dramaturgy of a Mixed Reality environment primarily focuses on defining an abstract form of display containing elements suitable for interactive narration.

14 Evaluation : user feedback and audience feedback

The context of the performance situation in Bochum lent itself to thematic observation of the learning process of perception, and to better identification of the dynamic action required to organise the Mixed Reality stage. The relatively small group of performers and professional observers allowed for targeted evaluation. During the rehearsal, after the performance and in the subsequent processing phase, two members of the MARS Team conduct a series of interviews.

On the hunt for the structure, form and content of the Mixed Reality stage, the first questions are as follows:

- What experiences does the acoustic and visual environment have in store?
- What goes on in an atmosphere of uninterrupted stories and changing positions?
- What factors trigger which events?

During a live performance the questions asked mainly relate to the persons producing the atmosphere – the performers. Instead of acting, the performers produce the atmosphere that in turn has an effect on them. What happens to the performers during the process of production?

In contrast to the form of the installation in Berlin, which we were able to plan, there were numerous unforeseeable difficulties with the rehearsal stage in Bochum. Since the projection and light conditions necessary for the camera tracking unfavourable, the stage décor was greatly reduced. It was therefore not possible to effectively achieve the desired effect of bright projection in a dark space. While in Berlin the light conditions were better and the acoustics of the room were much worse because of the longer echo time, the situation in Bochum was the opposite. From this we conclude that an optimal presentation of the Mixed Reality stage requires professional light conditions and sound equipment. It is also necessary to have a selection of different projection screens in order to be able to respond to different room and light conditions.

14.1 Evaluation criteria

Given the altered context of the performance, the criteria chosen for the Berlin demonstration had to be changed for Bochum. The criteria relate to the visual appearance of the installation, the stage set-up and the dramaturgy developed in the rehearsal.

Here, as before, our general questions to participants concerned their impression of:

- usability,
- audio-visual orientation in the image and sound space,
- game opportunities as strategies for interaction,
- identification of aesthetic structures.

The user group was composed of performers, professional observers and the members of the audience, and their comments were based on the areas we are investigating – the space of movement, space of image and space of sound. The performers' evaluation of the interactive stage concept took place during the rehearsal. The rehearsal setting answered the question of the quality of use that the system offers the performers. During the discussion after the performance, the audience got the chance to voice their opinions and to go on the stage themselves. The professional observers made comments directly after the performance and structured the discussion with the audience.

Different participants had different views of the exercise. In the opinion of the authors, the presentation was a success from the point of view of the overlapping of the virtual and the real space and the adjustment of the tracking to the size of the stage. Relationships between the scaling factors of a virtual environment and the real space could be made more precise. With respect to the camera tracking system, the influence of lighting factors in the environment was recognised and allowances made for this in subsequent development.

From the point of view of the performers, questions arose which are to be dealt with in a subsequent symposium. Their main criticism was that rehearsal time was insufficient. In the opinion of literary and artistic director Lambert Blum, a minimum of a week was necessary rather than the six hours made available. So far the performers had been unable to gain more than a superficial understanding of the systems and its potential.

The audience, made up of theatrical professionals and communications scientists, had different critical opinions. They found it difficult to follow the interactivity, and certainly did not experience it in any sensorial way because it was seldom identifiable from the presentation made by the performers. One member of the audience commented, "The presentation would have worked better if it had taken place in two different locations.“

The professional observers raised valuable questions and made some important criticisms and suggestions that are expressed in the interviews published afterwards. If one considers how seldom the opportunity arises to test the system in a demonstration, observation and discussion forum, the presentation in Bochum was very positive overall. It became apparent that, given the complexity of its technology, the interactive eMUSE system can support new forms of theatre for professional performers and dancers, but that these require a radical change to the discipline of the professionals. This is the most important conclusion of the presentation.

The authors, performers and professional observers were able to compare the aim they envisaged for the performance during the rehearsals with the actual performance on the following day. The short rehearsal time mentioned above and the occasional failure of the technical system made it difficult to stage the performance as planned. In the overall concept insufficient attention had so far been paid to the audience, since the focus had been on the performers' handling of the Mixed Reality stage. The most important suggestion made by the audience was that care should be taken to ensure that the interactivity creates a narrative momentum.

14.2 Comparing the performance with the rehearsals

1st scenario: The performer is in a sound space with only speech elements. First, all the external sensors are activated in turn, and then the inner sensors one by one. The first

performer, Martina Leeker, makes concentrated movements and navigates using the sounds. She uses the flowing movement of her white dress to repeat the sound elements in a dynamic process of development. The bending and stretching of her body create clear differences in volume that places movement, posture and sound into context. The game with the word groups and the marking of the different speech locations works very well in this phase. At the suggestion of the professional observers during the rehearsal, she moves more slowly so as not to disturb the image too much. The result is less satisfactory than in the rehearsals when she spontaneously investigated the system, since the relationship between sound and silence appears unbalanced.

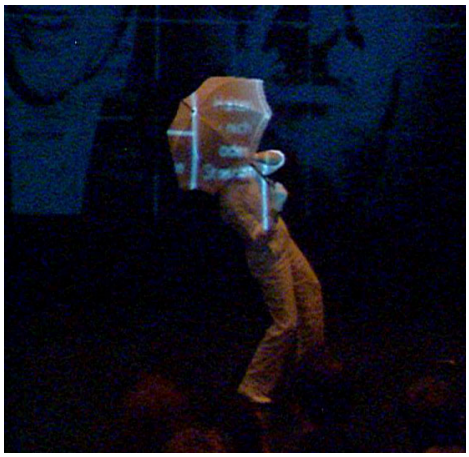


Fig 34 Performer with word space

2nd scenario: The rhythmic structure of the speech space is presented by two performers in turn. They use dance-like movements to create the sound landscape, which alternates between clearly comprehensible words and their transformation. A game-like situation is created by the contrast of sound and text and movement. Unfortunately the second avatar fails at the beginning of the scene. Consequently only one performer can be tracked and the other remains "silent". The result is that the dance-like improvisation of the first performer becomes the focus. She uses the sound production as a crude system of co-ordination. Intuitively she develops the skill of improvising with the system, and presents a successful combination of interactive creation and professional body language.

3rd scenario: The performer and a male observer develop a joint narrative sound scenario. The dancer enters the stage from the left and triggers rhythmic sounds. Shortly afterwards the male observer enters from the right and moves through a structure of words. The dancer looks for rhythms and uses her hooped skirt for the purpose. The man sounds out his surroundings for words using a white umbrella, and tries to get closer to his partner, which he manages to do.

For the audience, the presentation serves as an introduction for their own exploration of the scene. After the discussion with the audience and the professional observers, the Mixed Reality stage is open for all to try out.

Fig 35 Performing in word space

14.3 Feedback and Interviews about 'Murmuring Fields' as presented at Fidena'99

This chapter presents the most interesting excerpts from six interviews with the professional observers and performers from 'Murmuring Fields' who have been invited to Fidena'99.

Taking part in the interview:

Martina Leeker – Performer

Lambert Blum – Literary and artistic director

Danièle Perrier – Art historian

Hinderk Emrich – Neurologist

Ulrike Haß – Theatre studies expert

Walter Siegfried – Sound artist

14.3.1 Comments from the performers

Martina Leeker – Performer

Tangible sounds …

In Martina Leekers' opinion, there are various levels that overlap during the performative interaction of participants with the installation. One level is the desire to deliver a performance, and the other is to have a synaesthetic experience. The Mixed Reality concept presents difficulties when it comes to implementing choreography. "Because it's not about me presenting something great, it's about me perceiving and experiencing something." The behaviour observed on the Mixed Reality stage is a different ball game to conventional stage work. There, action models can run like an "internal film in your head" which shows co-ordination of movement."

Professional knowledge of performative forms of expression has shown itself to be a hindrance. The Mixed Reality system requires only a few, slow movements with clear gestures. "It requires people to listen to the sounds triggered by their movement. If this listening exercise is successful, it creates a very fragile type of perception which is

independent of the outside world. It's as if I'm moving through sounds and these sounds are tangible to my body. "

For Martina Leeker, a key experience in 'Murmuring Fields' is the influence of tactile elements on orientation in the sound space. She uses the hooped skirt as a "tool that makes possible certain movements which belong to the sounds." She sees the lace hem of the skirt "as something very tactile, like lots of small fingers or hairs … like an animal that makes very small movements and touches sounds as it does so ".

Lambert Blum – Literary and artistic director

Habits of expression …

Lambert Blum is interested in new forms of interaction between the performers. "Forms which do not (yet) exist in everyday life." In his work with 'Murmuring Fields' he stresses the difference between it and the conventional understanding of movement in the theatre, which is heavily influenced by symbols. "The difficulty for actors who are used to expressing themselves lies in the necessity to reduce themselves on the Mixed Reality stage. They have to hold back." The sound space prescribes the movements. This means that, from a stage direction point of view, it is less the scope of movement of the dancers and more the Mixed Reality installation itself which becomes the object of observation and demonstration. "It was our task to explain this installation to the audience through practical work." Blum says that the problems experienced with presentation were mainly due to a lack of time.

Tracking down forms of interaction that can lead to mental changes in imaginary images requires a long testing period and a "new key" for dealing with interactive time. If such conditions exist and given sufficient time for testing and reflection, Lambert Blum expects to see new communicative and meditative game and display forms emerging from Mixed Reality concepts. If media productions for the Internet and everyday life are to be investigated, Mixed Reality stages will need to be available on a long-term basis, since they require great willingness to co-operate from producers, literary and artistic directors, actors and audiences. Lambert Blum therefore calls for electronic arenas to be made publicly available for purposes of theatrical experimentation.

14.3.2 Comments from the professional observers

Danièle Perrier – Art historian

Chance as an artist …

With his sound installation "33 1/3", John Cage encouraged visitors to play an active part in the creation of his sound collage. The installation consists of thirteen record players and one hundred unlabelled records. The music is only generated through interaction. This is a sound collage which obeys the laws of chance – it all depends on the number of participants, the records they blindly select, the number of records played, the time at which each record starts playing. Added to this is the factor of "whether", "how" and "when" visitors participate or don't participate. This implies that periods of silence are also to be considered as an important element of Cage's composition.

There are clear overlaps between Cage's sound installation "33 1/3" and the Mixed Reality set 'Murmuring Fields'. The environment consists of a stage equipped with an

optical sensor system. The members of the audience become performers who by their movements initiate a virtual discussion between the four philosophers Vilèm Flusser, Paul Virilio, Joseph Weizenbaum and Marvin Minsky. Here, as with Cage's creation, the viewer must become a performer in order to activate the mechanisms of the installation, with chance playing its role as the third factor in the equation. The performer brings in his own personal notes through the manner in which he moves and interacts with other people on the stage.

Nevertheless, he discovers the verbal collage which he has initiated and tries to influence it, albeit blindly. In contrast to Cage's installation, the interaction here is also between the real and the virtual space. The real actions of the performers cause an avatar to appear on the screen, i.e. a visual trace of the active viewer. The performer interacts with his own virtual trace as though it were an imaginary partner. The image of his body in the virtual space is determined by an aesthetic of behaviour. The result is an overlapping of the real and the virtual, which not only opens up a new level of communication, but also redefines the notion of a stage and a performance in terms of the theatre. The spatial stage on which the performers move is not identical to the virtual stage used for the performance. The participants, on the other hand, are linked together in real time on the Internet and on the stage.

Hinderk Emrich – Neurologist

A sensitive system …

Talking about the Mixed Reality stage, Hinderk Emrich stresses the importance of an "anthropology of disillusionment" which depends on the fact that people constantly live in an "inner mental space" and an "outer physical space". These two levels of reality are normally in harmony. According to Emrich, it is necessary to disillusion this harmony when a person ventures out of his small space and seeks communication. Emrich identifies one way of facilitating this experience in interactive zones such as those which ‘Murmuring Fields’ offers its users.

He derives this theory from the following thinking: ‘Murmuring Fields’ is a signal for the displacement of things, since the perceived impressions are no longer in harmony with the order of everyday life. This causes disillusionment with the accepted notion of harmony between the inner and outer worlds, and causes the individual to consider this displaced state of affairs. The electronic data space stands for the inner reality, while the physical space can be equated with the outside world. What we are looking for are the "inner bonds" in the Mixed Reality interaction.

Hinderk Emrich describes the experience in ‘Murmuring Fields’ as an interactive process in several stages. Firstly he mentions the "enormous sensitivity of the system", and then refers to the shift of the "me perspective", which results in a re-evaluation of the way in which the user sees himself. "I don't see myself in my usual perspective, I see myself from a second perspective. This is a phenomenon that can also be used for therapeutic purposes. We never live only in the outside world, we also constantly live in self-generated imaginary worlds. And perhaps the installation has the important task of making clear to people who don't realise this that we are actually constantly living in mixed realities." Investigating the rules of the system is a precondition for orientation in ‘Murmuring Fields’: how do you make contact with other participants?

Hinderk Emrich advises inviting people who enjoy games into the installation rather than professional performers, whose training will have shaped them in particular ways.

"In order to experience Mixed Reality spaces, we need the mental ability to find out new things. That means becoming "Homo Ludens", the playing man."

Ulrike Hass – Theatre studies expert

Removing the limits of narration …

The Mixed Reality stage makes everyone a bit more of an actor, and at the same time reduces the difference between theatre and real everyday life. Ulrike Hass sees a similar striking change for theatrical narration. The boundaries between the symbol and narration are removed. An extension of the theatrical space between physical action and perception takes place. The special thing about this Mixed Reality stage is that it remains relatively free of images. In 'Murmuring Fields', pre-pictorial, audio experiences become the object of interaction. But this also presents a difficulty for those watching. There are few places where images are generated. Hass describes three key points about 'Murmuring Fields' and the Mixed Reality stage:

- The incongruent relation between the two-dimensional effect of the virtual space and the three-dimensional nature of the real stage space,
- The enhanced status of the space compared to the image,
- The Mixed Reality space as an extension between physical action and perception.

The real processes between perception and action are deconstructed and inner orientation patterns are expressed in narrative elements. "That automatically has something to do with the fact that (during a sensory experience) we dive into earlier worlds. And not into fantasy worlds, but rather – and this has become quite clear to me here – acoustic space, pre-optical space, in other words the mother-child dyad – touching one hurts the other."

Walter Siegfried – Sound artist

Desire for variety …

The fact that the lens on the tracking camera reduces the movement parameter transforms the "spatial shape of the moving person" into simple image elements, and then translates the body itself into a three-dimensional acoustic image. It is the reconstruction of the movement process that is of particular importance here.

In his symposium lecture "Irritations to perception as attention stimuli", Walter Siegfried puts forward a game theory supported by ethnological principles. Against this background he sketches the following typical process of experience: the central element in 'Murmuring Fields' is initially movement. With this in mind, he stresses the two-dimensional structure of the sound landscape. 'Murmuring Fields' works like a surface on which movement triggers sound. The performer is surprised by the noise that he triggers. By repeating the movement, he tries to work out the system and use it to make particular sound shapes. This can become the key situation in the playing behaviour. If the performer is successful, he immediately develops the desire for variation.

A different process may lead to interactive games with other performers. Rhythmic repetitive movement attracts attention. Other performers are encouraged to copy it, and an understanding of their shared movement process develops. The dialogs that arise in this way are like improvised rhythms. The dance-like momentum that leads to the release of inner images and suspends the everyday world raises interest further.

The movement in ‘Murmuring Fields’ opens the way to complex communication and interaction processes. Walter Siegfried credits the new technologies with the potential to change the senso-motoric handling of reality, which allows a particular form of sensitisation of the body, and of the present being. "If technology can lead us to value this momentary experience once again, then why not use this method?"

14.4 Comments from the audience point of view

Dirk vom Lehn, Christian Heath
Work, Interaction & Technology Research Group
King's College London

During the two hour event visitors of Fidena'99, were first part of a 'traditional' audience sitting in an auditorium facing a stage on which actors would perform. After the performance and a brief discussion of the installation members of the audience were allowed to enter the stage and explore the (virtual) space For the understanding of 'what has been going on' during the exhibition, in what follows we will take into consideration two perspectives through which visitors of the event could experience the exhibit. First, we will investigate how visitors of event experienced the installation from the audience perspective. And second, we will describe how visitors themselves explored the installation. With respect to both sections we will rely upon observations and video recordings made during the event and informal interviews with members of MARS. The discussion presented here does not attempt to capture visitors' (subjective) experiences of the exhibit but rather the organisation of the event and of the installation from the visitors' viewpoint. Thus, we hope to understand the installation as it has been explored by the visitors, either as 'audience' or as 'participants'. Before we begin to describe these two perspectives we will give a brief account of how the performance was introduced to the audience in order to illuminate how members of the audience have been informed about the installation and its features.

Introduction to the performance.

After the members of the audience have entered the theatre and taken their places in the auditorium, a presenter gives an introduction to the performance principally comprised of two sections: a brief presentation of the artists and the research group MARS followed by the mentioning of one of a few so called 'professional observers' who have been invited to give a 'professional comment' to the installation. He then hands over to one of the artists who introduces the installation and the idea that has led to its development.

The artist explains that 'Murmuring Fields' has been produced by MARS an interdisciplinary research group. Deviating from other projects, 'Murmuring Fields' is a co-production of technicians and artists which relies on the co-operation between specialists from different disciplines and is not based upon the ingenious invention of an individual artist. Then, he changes topic and begins to introduce the installation itself. 'Murmuring Fields' consists of a two distinct spaces, two realities as the artist says. One is the empty stage: a black coloured floor is extended to the back by a huge picture showing photographs of four philosophers. On the stage to the left of the audience scaffolding is placed on which a projector is located, broadcasting an image to a small screen on the left top corner of the stage. This screen gives the audience (in the

auditorium) and participants (on stage) an image of the movements on the stage as tracked by the computer-system. The other space is "furnished with data", which is symbolised by the image in the back of the stage. The images of the four philosophers onto which a pattern of squares is laid represent the map of this virtual space. Small squares that move across the four images are 'avatars' representing the users of the installation as they move across the physical space. The artist explains that what we will see is an example of 'Mixed Reality', two different types of reality that overlap: the physical space which is 'here' - the artist points to the area behind him - and the virtual space which is also 'here' "but you can't see it". In order to give further explanations he compares the experience one can make in the 'virtual space' with the way a blind person experiences real spaces. "You feel like a blind person. The space has to be felt, listened to and sensed."

After this introduction to the installation the artists hands over to the performers. The performance will consist of three scenes. At the moment the system has been prepared to make it possible for real individuals to develop forms of interaction in real space mediated by the technology. It can however be imagined (and the system is prepared for that next step) to connect the system with the Internet and thus to facilitate forms of interaction between remote participants. Then, the spotlight that illuminated presenters and artists is switched off, the performance begins.

'Murmuring Fields': The audience perspective

The audience sits in an auditorium facing the stage. During the performance, the audience only watches the occurrences on the stage without actively participating or intervening in the accomplishments by the performers. The first performer, a woman in a white costume, enters the stage and begins slowly to move along the left hand side of the black floor. After about 20 seconds when she reaches the top left corner of the stage, sound is produced by the loudspeakers. An utterance can be heard: "reversiblen Kabeln" whose meaning remains unknown for the audience. A couple of seconds later the performer has moved slightly forward when a longer utterance comes from the loudspeakers including the two words heard earlier. In the further course of the event the performer continues to move slowly across the stage, sometimes stepping backwards or stopping for a moment and swinging her body forward and backwards. For the observer it remains unclear what guides her movement. While moving her gaze is mainly directed forward. However, it remains unclear whether she actually 'sees' something and co-ordinates her movement with her gaze. Her body quasi floats across the floor, sometimes stopped by invisible boundaries, turns around and continues its way through the dark space. Her movement thereby is accompanied by occasional sounds, syllables or utterances which sometimes seem to reoccur and possibly are related to the performer's activities.

After the first performer has left the stage, the scene is reset and the second one enters the stage. Also dressed in a white costume and using a white umbrella, the second performer also moves across the stage. In comparison to the first performer this one makes use of an artefact, an umbrella, while navigating the stage. Sometimes the umbrella is held high over the head of the performer; sometimes she stretches out her arm to her side, thus keeping a distance between the umbrella and her body. The meaning of this movement in relation to the installation or her movement does not become clear. She also appears to play with her body when stopping on occasions swinging it back and forth. Slowly the members of the audience become aware of the relation between body movement and sounds. It seems as if the sounds and utterances

are triggered by certain movements of the body or when the umbrella is used. This relation becomes obvious when the performer stops at the right side of the stage and moves her lower limbs forwards and backwards thus (seemingly) repeatedly triggering a certain sound.

After a couple of minutes the second performer is joined on stage by the first one. The two performers then first independently move across the stage. Apparently they are not interested in the other's activities and only produce movements and thus trigger sounds. For the observers (audience) it remains unclear whether their movements rely upon some organisation or whether the two aim for some common objective when moving across the stage. Only towards the end of their performance do the actors turn their upper bodies towards each other and begin to co-ordinate their movements with each other. While up to this moment their gazes either were 'blind' and did not orient to anything on stage, now they appear to orient to each other and make use of each other's position in the 'virtual space'. Nevertheless, it remains difficult for the observer to recognise the organisation of the 'virtual space'.

The impression of the relation between body movement and sounds is enhanced when in the third phase of the event one of the 'professional observers' is asked to enter the stage and to explore the space. While standing on a spot from where he can see himself on the tracking-screen he begins to move his arm up and down, then opens the umbrella and moves it into various directions. In the continuance of his performance he keeps his gaze directed towards the tracking-screen and apparently co-ordinates his accomplishments with the movements he can see on the screen. A few minutes later the 'professional observer' is joined on stage by one of the actors. While at first the two move independently across the stage, the performer with a 'blind gaze' and the 'professional observer' monitoring the tracking-screen', after a few minutes the man turns his gaze towards the performer and when his 'look' is answered moves close by her side holding the umbrella over her head. In the continuance of the scene the two begin a kind of dance on stage whereby the man keeps his gaze directed to the tracking screen. The interaction between the two on is produced with regard to the images shown on the tracking screen. For the audience, it remains difficult however to understand the relation between tracking-screen, body movements and sounds triggered by both of their movements. Sometimes, when both actors simultaneously move on stage it is hard to attribute sounds to movements of the individuals.

'Murmuring Fields': The participant‘s perspective

For the participants, in the beginning 'Murmuring Fields' presents itself as a space which one can navigate and walk through. Within the space there is nothing of interest to see that the user could possibly turn his interest towards. A few locations on the floor of the space are marked by white stripes, thus producing some kind of spatial boundaries whereby it remains unclear what is separated through the stripes. The scaffolding to the right of the space has no meaning for the user as the artist has explained in his introduction to the performance, and the image on the wall in the background of the space is quartered and shows four philosophers, each in one corner: Vilèm Flusser, Marvin Minsky, Paul Virilio and Joseph Weizenbaum. All four images are divided into regions by various straight lines. Also embedded within the image is a small red square which moves across the image according to the changes in position the user undertake in the space. Obviously, the user is tracked by a device which is located at the ceiling of the space.

Apart from making available the position of the user in space with regard to the image of the four philosophers, the tracking-process is made visible on a screen which is located on the top left corner of the space. This screen makes it possible for both the users of the installation and the audience to monitor body movements on the black floor. In order to facilitate the tracker to follow users' movement participants either wear white clothes or use a white artefact. The performers in the three scenes described above wore specially selected white costumes and used a white umbrella. The umbrella also allowed the 'professional observer' later on in the performance to use the system effectively and to explore the 'virtual space'.

Relying on those devices the relationship between user and system can be summarised as follows: when moving across the black floor, the user is tracked by a device located on the ceiling. By its connection to a computer system, the tracking data are made publicly available via a screen, translated into a red square moving across the image on the back wall, as well as into sounds audible for users as well as audience. Through their movements, users are supposed to explore the 'virtual space', a space that is "furnished by data", as one of the artists stated in his introduction. By taking regard either of the image on the back wall or the tracking-screen users can locate themselves in relation to the system. By relating their position to the sound they have just heard before taking that position they can begin to explore the order of the 'virtual space'. For example, when users observe their movement on the back wall they will see the red square crossing one of the line and just afterwards hear a sound, an utterance or the like.

Body movements or movements of the umbrella back and forth across the line results in the repetition of the same piece of sound as often the line is crossed in the same direction. At the same time, the red square on the back wall will repeatedly cross a line in the image. Gradually users are supposed to understand the order of the 'virtual space' by virtue of their own movements, and when being in the space with someone else, to explore the organisation of the 'virtual space' in collaboration with each other. It can be argued that the order of the 'virtual space' is produced through body movements in 'real space', which by repetition in the same area result in repetitions of sounds. The spatial organisation is created out of experienced redundancies of sound produced by the users.

Users as well as the audience can recognise the relations between the elements of which the virtual space is comprised only by repeating movements across the black floor. By moving, for example, up and down or by repeatedly lifting up the umbrella at the same position; they find out about the two dimensionality of the 'virtual space'. The 'virtual space' indeed is like structured like a flat land. It can be imagined as the image of the back wall with the vertical and horizontal lines mapped onto the black floor. Movement then within the various regions of the image cause sounds that can be attributed to one of the four philosophers in which quarter one moves in.

When another person joins the first on the floor, sounds are triggered by two persons at the same time. The attribution of sounds to movements becomes problematic. Thus, it becomes possible to understand the order of the 'virtual space' by monitoring at the same time one's own and the other's movements in relation to the produced sounds. The tracking screen thereby becomes a means to support this reflexive engagement. The 'professional observer' who was part of the performance in the third part of the event supports this argument when in the discussion he states that "On the screen I see myself in a situation from a different perspective than my own". This device can be considered as a monitoring system that by a further development can become useful to make possible social interaction in the virtual space. The subjective experiences of the (up-to-

now) isolated actors on stage can make use of such a monitoring device and thus socially co-ordinate their individual activities with each other.

Conclusion

In concluding, we want to summarise our observations of the 'Murmuring Fields' exhibit and thus produce a starting-point for a possible comparison to other 'electronic exhibits' and 'eRENAs'. In the discussion about the performance after the event, a number of audience members appeared to be disappointed about their experience. One compared 'Murmuring Fields' with a clumsy radio-like device where you have to search for senders (sounds) by moving around your entire body. Others noticed that indeed the 'virtual space' is not a three-dimensional space but rather a two dimensional flat land.

It has, however, to be acknowledged that, firstly, the given presentation only was considered by MARS as an experiment in which the public stage served as 'natural laboratory' to find out how the installation works. And secondly, in order to understand the 'nature' of the installation one has to go into the space and experience it oneself. The audience perspective rather undermines the exhibit's potential and conceals its effects.

In concluding, we want to compare very briefly the observation we made in Bochum with those observation we have made with regard to other, more 'traditional' exhibits. Thereby, we will also draw on current literature in Visitor Studies concerned with visitor behaviour in museums and with what often is called the 'museum experience'.

Recent publications on visitor behaviour principally concentrate on two questions: how can the visitor's 'museum experience' be enhanced, and how can the visitor's 'learning' from exhibits be enhanced. With respect to both issues, researchers have come to the conclusion that both experience and learning in museums are influenced by multiple variables. However, it is more and more acknowledged that the social interaction between visitors has a major impact on whether or not an exhibition is considered to be 'effective' and 'successful'.

With regard to more 'traditional' exhibits we have made the following very general observations:

- Visitors explore exhibits together with others (e.g. group and family members);
- Exhibits and their features are rendered noticeable for each other in and through social interaction. Thus, visitors begin to make sense of their environment in and through social interaction;
- Visitors are sensitive to and peripherally aware of activities accomplished in the local environment;
- Visitors produce the context in which they experience exhibits, in and through interaction with others, both those they are with and those who just happen to be in the same local environment;
- The structure of the museum visit is not pre-produced by the exhibition design but the physical environment is embedded within the social interaction and thus the 'museum experience' and 'learning in museums' is produced in interaction with co-present others.

Based on these observations the following preliminary conclusions can be drawn from the descriptions of the event in Bochum (see previous sections):

- Visitors[8] got their first understanding of the exhibit from their observation of performers. However, the exhibit's features were not entirely intelligible which made it difficult for audience members to understand 'what actually was going on' and for participants to 'know what one can do with and in the installation'. The lack of intelligibility makes explanations about the features of the exhibit and possible experiences unavoidable. Thus, the experience is possibly 'blown' before one can make it;
- Only those features of 'Murmuring Fields' were rendered noticeable for the visitors which either were mentioned in the verbal descriptions and explanations by MARS or were made 'perceptible' through the 'live' performances;
- Social interaction with other visitors, firstly, was restricted to verbal discussions about the exhibit between members of the audience. Secondly, as pointed out above social interaction between participants within 'Murmuring Fields' only is possible to some extent. The experience of the exhibit is widely dependent on subjective experiences and 'interaction with the exhibit';
- The exhibit was mainly experienced from an audience perspective. The context of the experience therefore was mainly pre-produced by the performers on the stage. From the participants' perspective the exhibit could individually be explored. The interactive features of the exhibit largely remained concealed and invisible;
- Social interaction between the participants on the stage of 'Murmuring Fields' was hardly possible. The structure of the museum visit is not pre-produced by the exhibition design but the physical environment is embedded within the social interaction and thus the 'museum experience' and 'learning in museums' is produced in interaction with co-present others. The tracking-screen only is an initial device that allows visitors to reflect on theirs and other's activities in relation to the 'virtual space'. This device would need to be improved and its features to be extended.

15 Lessons learned so far

- **Development of interactive environments is an inherently iterative process**.

Availability of tools to rapidly implement ideas and developed concepts is of crucial importance. The implementation of the system is only the beginning of more work. Observation and analysis of the behaviour of participants in the environment is critical to understanding how successfully the envisaged ideas and their implementation relate to "real world practice". Observed results often lead to drastic changes of initial concepts.

- **The psychological and cultural conditioning of the users**

[8] Please note that here 'visitors' refers to both members of the audience and participants who actually used the installation.

The whole situation created by the environment, its connection to the psychological and cultural conditioning of the users, is as much the "user interface" as are the underlying technical interface devices. It is important that the developers can work with a mental model of "interface" as that which connects the participants with each other, or with one's inner process of imagination, not with the underlying computer system. It is not the threedimensionality or the existence of a predefined virtual space that matters in itself. Rather, it is how the virtual space is connected to the actions of the participants and the interaction between them.

- **Simple modes of interaction as a necessary condition for a "successful" system.**

Focusing on the interactive soundspace as the defining element of Murmuring Fields made the interaction concept much more natural and intelligible for the participants. To this end the visual design of the virtual space in Murmuring Fields was simplified and the number of viewpoints displayed in real space reduced. But these issues were discovered only after the system had been set-up and tried out by the public. Focusing on mechanisms for interaction and communication connected to participants' proprioceptive awareness of body and space is crucial for supporting a bodily sense of (tele)presence in a combination of shared physical and virtual space. Video tracking can be used to get information on participants' movement in space at large but additional devices (such as electric field sensors) are needed for finer sensing of gesture in local regions of space.

- **A replicated scene-database with UDP or TCP/IP peer-to-peer communication is suitable for a small number of users.**

As the number of users increases multicast is the communication method of choice. For interaction concepts involving object manipulation and simulated autonomous behaviour, synchronisation and causal ordering mechanisms need to be incorporated. For large-scale virtual environments, spatial partitioning techniques need to be implemented to reduce the communication load of the system.

- **The possibility of reconfiguring the functionality of the technical platform, and of incorporating own custom-made modules, is more important than having a closed platform offering "everything".**

This consideration was one of the main factors determining the design of the e-MUSE platform as a bridge between existing toolkits and general-purpose systems for realising virtual environments . It is reflected in the fact that eMUSE doesn't offer "all" common features of similar systems but is structured as a set of independent modules which can be configured, extended or exchanged with custom-made or other available modules at will. Murmuring Fields can be understood as a conceptual basis for researching and developing new models of communication. The main interest has been the development of a communication space in which participants experience a mixed-reality condition. A related goal has been to facilitate this experience through the development of unobtrusive physical interface devices. The communication model of Murmuring Fields can be placed in the context of different applications dealing with information spaces as well as architecture, theatre and performance. Understanding these fields as meta-communication models for everyday life can help in gaining new knowledge on human performance in media spaces. The biggest problem with developing Murmuring Fields was the lack of an existing platform to try out prototype concepts and ideas. Existing system didn't offer the flexibility needed, the support for non-standard input devices, and the support for user representations that could change dynamically based on users'

interaction and movement in space. As a result, the system to enable the realisation of envisaged concepts had first to be designed and developed. This slowed down the development of the prototype and many ideas turned out to be unfeasible once tried out.

The work on Murmuring Fields also demonstrates how figurative user representations are neither the only nor the best approach for providing the sense of presence for participants in shared physical and virtual space. The simple visual form of trace-avatars proved sufficient for supporting participants' relationship to the visual aspects of the virtual space. Its scarceness didn't distract the users' "immersion" into the interactive soundscape. This turned out crucial for supporting the development of a bodily sense of one's movement, the space and the presence of other participants. A final realisation is that development of interactive environments requires a team of people with a wide range of backgrounds and cross-profession capabilities, led by an experienced and very open-minded "director". The similarity with the idea of a crew producing a film is not a coincidence. Developing spatial concepts, interaction models, forms of user presence as well as concepts of interface and underlying physical devices asks for computer scientists and hardware specialists as well as for architects, anthropologists, artists and theatre professionals. Ideally, they would all share some insight into each other's fields of profession.

In reality, team members must be ready to learn that everyone is equally an "expert" and a layman in this kind of joint undertaking, as well as to appreciate the differences of views and approaches of different professions. The leader of the team faces the task of ensuring that all the individual members can communicate their work to each other and understand the approach to the development of the system as a whole. The essential capability is to communicate how the very different ideas from very different disciplines and approaches fit into the picture of the system as a whole. The development of Murmuring Fields and the eMUSE system involved a team of roughly ten people, varying in size, demonstrating this point.

15.1 Installation versus stage

In interactive installations there is a harmony between production and communication, to the extent that sender and recipient are identical. This contrasts to conventional stage productions where production, and communication come into conflict.

In a system in which the interactive observer participates "directly" in the action by activating sounds and thereby influencing the break-down of the images and the progress of the narration, the role of the audience has become superfluous, since the tangibility of the interactivity cannot be conveyed to those outside. The excitement is therefore lost on the audience. At this moment the system is only of meaning for the performers. The question therefore concerns how the interactive process can be conveyed to the audience. In the third scenario the perspective of a performer was integrated in the stage by projecting the three-dimensional insight view into the space. This display was visible simultaneously on the wings of the stage and on the white "surfaces" of the performers. It created a fabric that reacted to the movements of the performers. The synchronised movement created the impression that the performers and the image space were merging together. From the point of view of the appearance of the stage, the three-dimensional moving image was more exciting than the two-dimensional static display of the navigational map of the speech spaces.

The interactive installation in Berlin created the impression that ‘Murmuring Fields’ was a cinematic interactive space whose audio-visual subject was interactively controlled by the visitors. In contrast, the interactive stage for only two actors set up in the museum in Bochum proved to be a test of patience for the 50 members of the audience. They were only able to follow the interactive experience once they had the chance to go onto the stage themselves.

15.2 Five steps of perception towards an interactive experience

In interactive systems the body reforms itself. Interactive media are time-based. This means that the action occurs in real time and the performers bring their own rhythm and their own subjective concept of time into the action. They can stop or alter the course of events at any time, and therefore structure perception themselves. This is highlighted particularly clearly by Hinderk Emrich in his observations on the use of the Mixed Reality stage. He describes the various steps in the perceptive process of interaction as follows:

1. The participant tries to identify the structures and rules
2. The participant plays with the structures and rules
3. The participant considers what his action looks like / how it sounds to the audience
4. The participant becomes aware of the other player(s)
5. The participant tries out communication with the other(s)

The overlapping of the real and the virtual space creates a new legibility, a different usability and experience of space. The possibility of the simultaneous presence of several participants in the real and in the virtually networked space opens up new ways of making contact and interacting. The eMUSE system we have developed creates a framework for action for the interactive Mixed Reality stage. ‘Murmuring Fields’ is a play developed to test out this stage. The structure of the Mixed Reality stage and its temporal processes must be investigated by the performers in a process of improvisation. There are no game rules and learned patterns of action do not apply. In ‘Murmuring Fields’, the extended performance plays with the presence and the representation of the user. The Mixed Reality space is a condition of networking that is understood as a communication environment.

The discoveries made through the Mixed Reality stage in terms of patterns of behaviour go far beyond the bounds of the theatre, and could perhaps be used in a more highly developed form in the field of psychiatry. What type of contact can be made there? How is the virtual space created? What makes it a communication space? The moment is the unit of time that must be perceived within its situation. The moment – the present – is an opportunity that lasts for 70 – 600 milliseconds, depending on how one defines it. Networking means communicating in the present. The real time in the virtual reality of the Mixed Reality environment represents simultaneity as an appropriate form of perception – thinking in doing. It is not about giving oneself over to acceleration, since interactive time is reversible. It is about using experimentation in an attempt to escape the situation "think first, then act".

References

[And96] Anderson D.B. et al .; A Social Virtual Reality System with 3D Animation, Spoken Interaction, and Runtime Modifiability, Technical Report at Mitsubishi Electronic Research Laboratories, Cambridge, 1996

[Ant98] Antonio, A.; Carrazoni, G.; Aguado, G.; Segovia, J.; Nonverbal communication and believable interaction in virtual environments, in Abstracts of the i3 Annual Conference, Nyborg, Denmark, July 1997

[Avpan97] SIGGRAPH 97 Panel , Putting a Human Face on Cyberspace: Designing Avatars and the Virtual Worlds They Live In, in Computer Graphics Proceedings of SIGGRAPH 97, Los Angeles, California, August, 1997

[Benf95] Benford S. et al.; User Embodiment in Collaborative Virtual Environments, in Proceedings of CHI'95, p. 242-249

[Benf96] Benford, S. et al; A Review of Distributed Architectures for Networked Virtual Reality, Virtual Reality: Research, Development and Application, pp 155-175, 1996, Virtual Press.

[BillKa99] Billinghurst, M.; Kato, H.; Mixed Reality – Merging Real and Virtual Worlds, in Proceedings of the First International Symposium on Mixed Reality (ISMR '99), Springer Verlag, Berlin, 1999, p. 261-284.

[Blum95] Blum, L.; Theater als literarisiertes Medium, in Flamboyant 2, Schriften zum Theatered studio 7, Köln 1995, pp 23 – 34

[Breit96] Breiteneder, C. et al; ATM Virtual Studio Services", Proc. 6th Intl. Workshop on Network and Operating System Support for Digital Audio and Video NOSSDAV'96, pp 63-68.

[DIVE] Carlsson, C. et al; DIVE – A Multi-User Virtual Reality System, in Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'93), Seattle, September 1993

[eRENA98-1] Strauss, W.; Fleischmann, M.; Novak, J.; Thomsen, M.; VRML based extended galleries – reconsidering the concept of a multi user environment, i3-eRENA 1st year deliverable, Bruxelles, May, 1998, p. 1-30

[F&S97] Fleischmann, M.; Strauss, W.; Images of the Body in the House of Illusion, in Art&Science, Sommerer C. (ed), Springer, Wien/New York, 1997

[F&Sint] Fleischmann, M.; Strauss, W.; Virtualität und Interaktivität als Medium oder vom Zuschauer zum Akteur, GMD.IMK-MARS, internal paper, June, 1998

[Frie95] Anne Friedberg; Gerahmte Visualität: Das virtuelle Fenster", in: Der Sinn der Sinne, Schriftenreihe Forum/Band 8, Kunst-und Ausstellungshalle der Bundesrepublik Deutschland, Bonn and Steidl Verlag, Göttingen

[Fuhr98] Fuhrmann, A.; Helwig, L.; Schmalstieg, D.; Gervautz, M.;Collaborative visualization in Augmented Reality, IEEE Computer Graphics and Applications, July/August, 1998, p. 54-59

[Green98] Greenhalgh, C. M. et al;., Supporting Rich And Dynamic Communication In Large Scale Collaborative Virtual Environments, Presence: Teleoperators and Virtual Environments, MIT Press 1998.

[Ishii97a] Ishii, H. et al; Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms CHI 97, http://tangible.www.media.mit.edu/groups/tangible/projects.html

[Ishii97b] Wisneski, C.; Ishii, H.; Dahley, A.; Gorbet, M.; Brave, S.; Ullmer, B.; Yarin, P.; Ambient Displays: Turning Architectual Space into an Interface between People and Digital Information, in Proceedings of International Workshop on Cooperative Buildings (CoBuild '98), Darmstadt, Germany, February 1998, Springer Press, pp. 22-32.

[Kindra98] Kindratenko, V et al; Distributed Virtual Reality (DVR) – Evaluation Report, http://viswiz.gmd.de/DVP/Public/deliv/deliv.445/d445.html, March , 1998.

[Kli97], Klinker, G. J. et al; Confluence of Computer Vision and Interactive Graphics for Augmented Reality, Massachusetts Institut of Technology, August, 1997

[Kru95] Krüger, W. et al.; The Responsive Workbench: A Virtual Work Environment, Virtual Environments, Computer, IEEE, July 1995.

[Leek98] Leeker, M.; From literate to postliterate to bio-electronic mime: Forms and functions of mime from antiquity to Decroux, lecture at the Center for Performance&Research, Wales, UK,1998

[Mac97] Macedonia, M.R.; Zyda, M.J .; A Taxonomy for Networked Virtual Environments, IEEE Multimedia, January-March 1997, p. 48-56

[Mann98] Mann, S. Wearable Computing as means for Personal Empowerment, ICWC-98. http://wearcam.org/icwckeynote.html.

[Mar97] Marvin, S. ; Telecommunications and the Urban Environment. Electronic and Physical Links, in: Peter Droege(ed.), Inteligent Environments. Spatial Aspects of the Information Revolution, Amsterdam 1997, p.179-197, p. 183.

[MASSIVE] Communications Research Group, Department of Computer Science, University of Nottingham, http://www.crg.cs.nott.ac.uk/research/systems/MASSIVE-2/, 1999

[Milgram94] Milgram, P; Takemura, H.; Utsumi, A.; Kishino, F.; Augmented Reality: A class of displays on the reality-virtuality continuum, SPIE Vol. 2351, Telemanipulator and Telepresence Technologies, 1994

[MilKis94] Milgram, P; Kishino, F.; A taxonomy of mixed-reality visual displays, IEICE Transactions on Information Systems, Vol E77-D, No.12 December 1994.

[Naka96] Nakanishi, H. et al; FreeWalk: Supporting Casual Meetings in a Network, International Conference on Computer Supported Cooperative Work (CSCW-96), pp. 308-314, 1996.

[Olv98] Olveres, J.; Billinghurst M.; Savage, J.; Holden, A.; Intelligent, Expressive Avatars**,** in Proceedings of the First Workshop on Embodied Conversational Characters (WECC '98), October 12-15, 1998, Lake Tahoe, California

[Ong97] Ong, Walter J., Ramus, Method, and the Decay of Dialogue. From the Art of Discourse to the Art of Reason, Cambridge (MA) 1958, quoted in: Justin Stagl, Die Entdeckung des Fliegens und Schwebens im Rahmen der ramistischen Wissenschaftskonzeption des 16. Jahrhunderts, in Fliegen und Schweben; Hg. Bauer Dieter R., Behringer Wolfgang, dtv München 1997

[Poup98] Poupyrev, I.; Weghorst, S.; Billinghurst, M.; Ichikawa, T; Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques, Computer Graphics Forum [Eurographics Conference Issue]

[Pekko97] Pekkola, S, et al; Collaborative Virtual Environments, Real-Time Video and Networking, International Conference on Virtual Systems and MultiMedia 1997 (VSMM'97) September 10-12, 1997, Geneva, Switzerland.

[Reki97] Rekimoto, J. et al; Perceptual Surfaces: Towards a Human and Object - Sensitive Interactive Display, Workshop on Perceptural User Interfaces (PUI'97), 1997.

[Rock] The Very Nervous System" is developed since 1986 by David Rockeby. It uses video cameras, image processors, computers, synthesizers and a sound system to create a space in which the movements of one's body create sound and/or music: http://www.interlog.com/~drokeby/vns.html

[Sch97] Schroeder, R.; Networked Worlds: Social Aspects of Multi-User Virtual Reality Technology, Sociological Research Online, Vol. 2, No. 4, 1997
http://www.socresonline.org.uk/socresonline/2/4/5.html

[Slater94] Slater, M.; Usoh, M.; Body Centered Interaction in Immersive Virtual Environments, in N. Magnenat Thalmann and D. Thalmann (eds.), Artificial Life and Virtual Reality, John Wiley and Sons, 1994, 125-148

[SmallTool] http://orgwis.gmd.de/projects/VR/#Pub

[Stra91] Strauss, W. et al; Telecomputing im Bereich Stadtplanung, Telemedien für Morgen, Vistas Verlag, Berlin 1991.

[Stra99] Strauss, W. et al; Staging the space of mixed reality - reconsidering the concept of a multi user environment, Proceedings VRML99 – Fourth Symposium on the Virtual Reality Modeling Language, February 23-26, Paderborn, Germany, pp. 93-98

[Thal97] Capin T.K.; Pandzic I.S.; Noser H.; Magnenat Thalmann N.; Thalmann D.; Virtual Human Representation and Communication in VLNET Networked Virtual Environments, IEEE Computer Graphics and Applications, Special Issue on Multimedia Highways , March-April 1997

[Thal98] Capin, T.K.; Pandzic, I.S.; Magnenat-Thalmann, N.; Thalmann, D.; Integration of Avatars and Autonomous Virtual Humans in Networked Virtual Environments, in Proceedings of ISCIS 98, IOS Press, 1998

[Vil98] Vilhjalmsson, H.; Cassell, J.; BodyChat: Autonomous Communicative Behaviors in Avatars, In Proceedings of the 2nd annual ACM International Conference on Autonomous Agents, Minneapolis, 1998

[Wink97] Winkler, T.; Creating Interactive Dance with the Very Nervous System, in proceedings of the Connecticut College Symposium on Art and Technology, 1997