

[14] <http://www.xerox.fr/ats/br/livead.html>

[15] http://www.xerox.fr/research/cms/eurocode_4.htm#HEADING3

As feedback information a video signal from an infra-red camera out of the Cyberstage room was sent to a monitor in the Blue Room. Thus, the person in the blue room was able to have a view of the virtual objects in Cyberstage and also of his own video image which was included as video texture within the virtual world.

The audio connections and equipment are shown in Fig. 6 with dotted lines. For the blue room one speaker and one wireless-microphone were used connected via an audio mixer. In the Cyberstage room AVOCADO's spatial audio system in conjunction with the 8 channel audio display and one wireless-microphone were used. Two receiver/transmitter pairs were used for the wireless microphones.

Fig. 7 shows a reconstruction of the Cyberstage view during the Schlosstag scenario demonstration. We integrated the remote person into an operating theater. The remote participant was giving some information and advice to the group of visitors currently present in the Cyberstage installation. To show the possibilities of our software framework, this operating theater includes a large set of interactive elements according to a design study of the German Max Dellbrueck Center in Berlin. The way to interact with these elements was explained to the visitors by the remote participant. A virtual projection screen showed a prerecorded laparoscopic treatment for additional information.

6. Conclusions and Future Work

We have presented a scientific approach for Immersive Telepresence in Projection-based Virtual Reality Systems. Our aim is to provide geographically dispersed groups of people the possibility to meet and work within Virtually Reality systems as if face-to-face. For this purpose we integrate stereo video images from remote participants into the same virtual space. Participants could be located at sites having different Projection-based VR systems. A first prototype of the approach was presented in GMD's open house event in October '97. The scientific approach requires camera calibration, in order to obtain the camera parameters, which are then used for integrating the stereo-video into the virtual space, while preserving the stereo-effect and the viewer's perspective. The results drawn by the use of this prototype have encouraged us to continue our research towards Immersive Telepresence in a number of different directions. In particular, we working on solutions on a number of open issues such as the use of more than one camera-pair, the real-time rectification algorithm, and different usage scenarios for different projective VR systems and different application areas.

REFERENCES

- [1] Breiteneder C., Gibbs S., Arapis C., TELEPORT- An Augmented Reality Teleconferencing Environment, Proc. 3rd Eurographics Workshop on Virtual Environments Coexistence & Collaboration, Monte Carlo, Monaco, February 1996
- [2] Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R., and Hart, J.C. The CAVE, Audio Visual Experience Automatic Virtual Environment. Communications of the ACM, June 1992.
- [3] Dai P., Eckel G., Goebel M., Hasenbrink F., Lalioti V., Lechner U., Strassner J., Tramberend H., Wesche G., Virtual Spaces - VR Projection System Technologies and Applications, Tutorial Notes of the 1997 Eurographics Conference, Budapest, 1997.
- [4] Dechelle, F., DeCecco, M., The IRCAM Real-Time Platform and Applications, Proceedings of the 1995 International Computer Music Conference, International Computer Music Association, San Francisco, 1995.
- [5] Haase H., Dai F., Strassner J., Goebel M., Immersive Investigation of Scientific Data, Scientific Visualization, IEEE Press, 1997
- [6] Hasenbrink F. and V. Lalioti, Towards Immersive Telepresence: Schlosstag'97, 2nd International Immersive Projection Technologies Workshop, May 10-11, Ames, Iowa, USA, 1998.
- [7] Ishii H. and Kobayashi M., ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact, CHI'92, May 3-7, 1992.
- [8] Johansen R., Groupware Computer Support for Business Teams, New York, London, The free press, 1988.
- [9] Krueger W. and Bernd Froehlich B., The Responsive Workbench, IEEE Computer Graphics and Applications, May 1994
- [10] Lalioti V., Garcia C., Hasenbrink F., Meet.Me@Cyberstage: towards Immersive Telepresence, Virtual Environments'98, SpringerVerlag, Vienna, 1998.
- [11] Lindemann E., Starkier F., Dechelle F., The IRCAM Musical Workstation: Hardware Overview and Signal Processing Features, In: S. Arnold and G. Hair, eds. Proceedings of the 1990 International Computer Music Conference. San Francisco: International Computer Music Association, 1990.
- [12] Wellner P., DigitalDesk, Communications of the ACM, Vol. 36, No. 7, July 1993.
- [13] Weatherall A., GroupSystems Electronic Meeting across the Enterprise and across the World, HICSS-30, January 97, Maui, Hawaii.

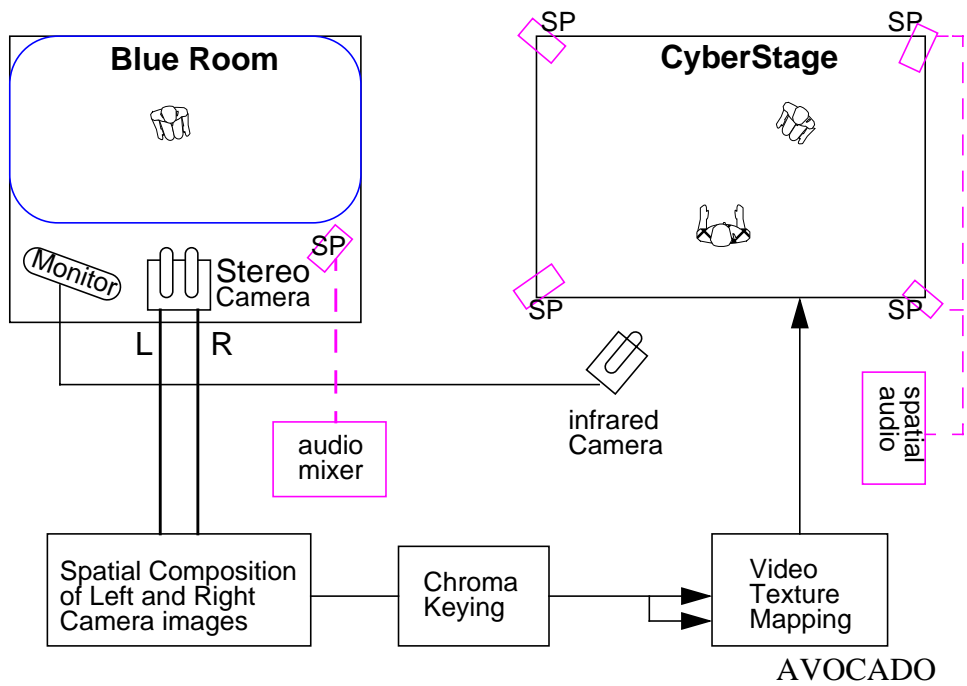


Fig. 6 Schlosstag'97 set-up for Immersive Telepresence

An overview of the set-up, is shown in Fig. 6. The stereo camera generated two video streams, one for the left and one for the right eye. These two streams were spatially combined. The resulting video stream, containing the even fields of each signal, was then send to an Ultimate keying system to be enhanced by an additional alpha channel, masking all background information. The resulting video stream was then send to an SGI Onyx with a 2 Pipe RE2 graphics subsystem, which is powering the Cyberstage installation. The Onyx has two Sirius boards which received the video stream as input and integrated it as a stereo video texture in the virtual scene.

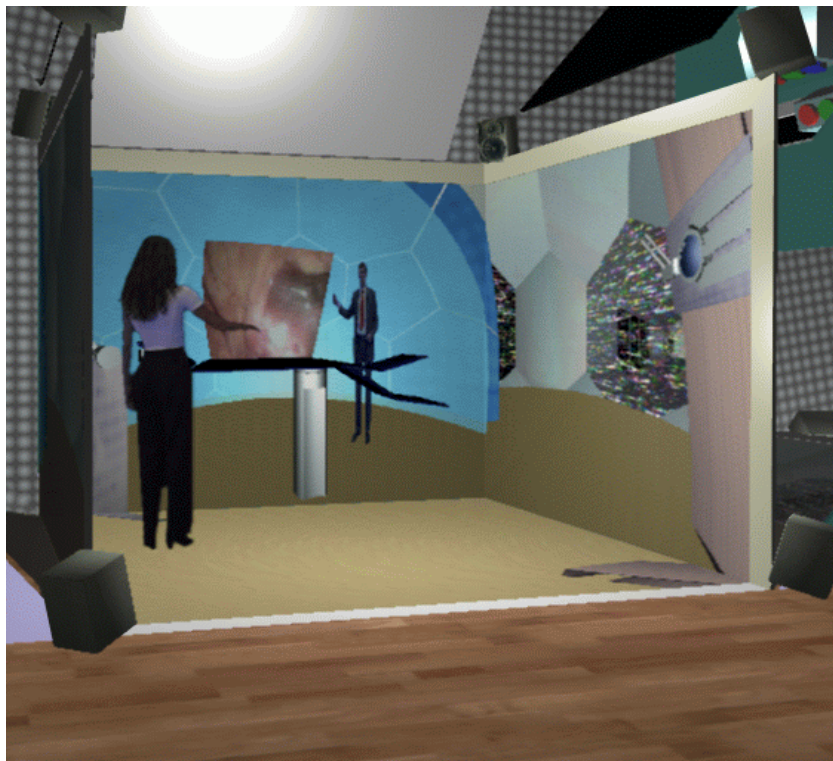


Fig. 7 Reconstruction of Schlosstag'97 Cyberstage View

remote person. In addition, the segmentation techniques of the TELEPORT system are also used in our approach. Therefore, we are able to determine the regions of the video signal that are of interest (i.e. the image of the remote participant) and combine this information into the original video signal, thus making the background transparent.

The Cyberstage viewer, must have the sensation of sharing the same virtual space with the remote participant. Therefore, when positioning the video planes we have to respect physical rules such as the size of the participants. In addition, when the image of a remote participant is positioned behind a virtual object the stereo impression for the Cyberstage viewer must confront to this (i.e. the stereo-effect should not result in the remote participant perceived as been within virtual objects). Thus, the stereo images have to be adjusted in terms of size, orientation and position, before being integrated into the virtual world. Automatic camera calibration and stereo image rectification is used to perform this task [10].

4.1 Stereo-video integration into Virtual World

In order to merge the stereo video of a remote participant into the Cyberstage, the rectified stereo images are mapped onto virtual planes. Fig. 5 shows the relation between Cyberstage, Camera and Virtual World coordinate systems. The world coordinate system (WC) is the coordinate system of the virtual world. We position the video planes, as well as any virtual objects, according to this coordinate system. The (0,0,0) of the Cyberstage coordinate system (CC) is at the center of the room 1.5m from the floor. The viewer's location is given according to the CC and we know the mapping from WC to CC at any given time (e.g. the Cyberstage is moving around the virtual world). On the other hand we have the stereo camera coordinate system (KC) which for matter of simplicity we assume that it is also the coordinate system of our real world.

When the mapped planes are projected onto the Cyberstage, the viewer must have the sensation of seeing the 3D image of the remote participant together with the 3D scene of the virtual world. For this purpose the video planes are centered around the corresponding eye of the viewer. There also have to be sized according to the distance of the viewer from the Cyberstage projection screens and to the distance of the video plane from the viewer. The camera parameters, obtained by automatic camera calibration, are then used for the positioning and scaling of the image planes.

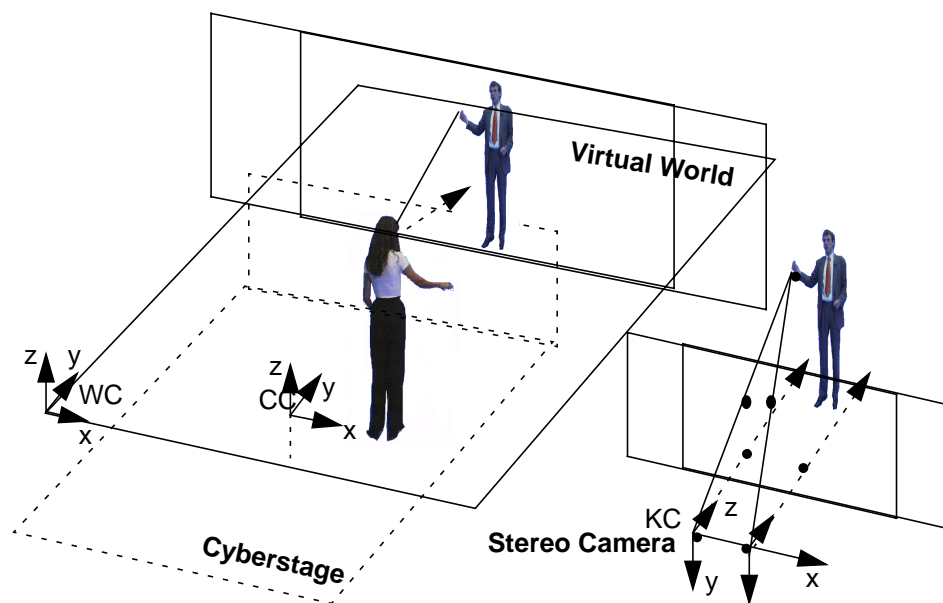


Fig. 5 Virtual, Camera and Real World Coordinate Systems

5. Schlosstag'97

For our open-house event in October 1997 (Schlosstag'97) we demonstrated a step towards Immersive Telepresence in Projective VR systems, by connecting our virtual studio facilities with Cyberstage. In particular, a person captured by a stereo-camera was keyed and integrated into a 3-dimensional virtual environment, allowing a fully immersed 3D virtual teleconference. The real-time stereo video image from GMD's Blue Room, was merged into the virtual world shown in Cyberstage, allowing immersive telepresence on the Cyberstage site.

Services

Services are neither field containers nor inherited from Performer classes, but they provide a functional API to unique system features. They are even more local than sensors and therefore must not be distributed either. A service would implement the access to an external device like 6 DOF-trackers for example, which exists only for a limited time and therefore should be used from only one location inside the application. A sensor may use this service to access a tracker and to maintain the data related to it to the modeling hierarchy via its fields.

Rendering

The different displays all have their own rendering mechanism applied to the modeling hierarchy. Only the visual rendering has a direct access through the Performer pipeline. The auditory and the tactile rendering are calculated on a second computer, connected to the "master" by a fast network. For the visual rendering the different tasks are divided into the already described processes of the pipeline. After the modeling hierarchy is updated to its actual state in the application process it is passed on to the culling process which strips all invisible objects. It is important to support this technique by dividing large geometry into smaller, cullable objects. The part of the scene left over after the culling is passed on to the drawing process where it is rendered to the screen with OpenGL. For configurations with more than one visual display system, the appropriate number of pipelines is used.

Auditory Rendering

Rendering the auditory scene has to take into account the position of the observer's head in the virtual world and in the auditory display as well as the characteristics of the auditory display (i.e. the loudspeaker configuration). The auditory rendering process is a two stage process. In the first stage a source signal is synthesized and in the second stage it is spatialized. In the first stage, only the sound model parameters are needed by the rendering process. In the second stage, the signals driving the auditory display are computed as a function of the distance between observer and sound source, the radiation characteristics of the source and the signature of the acoustic environment. With these signals the auditory display produces the illusion of a sound source emitted from a certain position in a certain acoustic environment shared by the observer and the source. The sound rendering is a dynamic process, that takes into account movements of the observer in the display, in the virtual world, and movements of the sound source. If these movements are faster than about 30 km/h, the pitch changes due to Doppler shift are simulated as well.

Tactile Rendering

The CyberStage display has a set of low-frequency emitters built into its floor. This allows vibrations to be generated, which can be felt through the feet and legs. There are two main areas of application of this display component. First, low frequency sound (which cannot be localized) can be emitted to complement the loudspeaker projection. Second, specially synthesized low frequency signals can be used to convey attributes such as roughness or surface texture. The vibration display is handled like sound in the rendering process. Sound models are used to generate the low-frequency signals. Sound synthesis techniques, generally referred to as granular synthesis, are very well suited to produce band-limited impulses that may represent surface features. Such features can be displayed through user interaction. For instance, a virtual pointing device can be used to slide or glide over a surface and produce vibrations. Additionally, higher-frequency sound can also be produced if necessary. Some of what can be felt usually through the skin of our fingers when sliding over an object is presented to our feet. This sensation can complement sound and vision dramatically.

The AVOCADO Framework has been operational since October 1996 and has been successfully used in a variety of projects.

4. Immersive Telepresence

Today's technology and advances in networking and multimedia systems stimulate a change in the way business is carried out, making it a globally distributed process, in which communication and collaboration of geographically dispersed groups is of vital importance. Teleconferencing and collaborative telepresence systems that provide high-degree of copresence give enough evidences that projection-based VR systems when combined with multimedia facilities, such as real-time video and audio, can greatly facilitate the communication and collaboration over distance in a variety of application areas. Immersive Telepresence creates an environment where remote participants not only meet as if face to face, but also share the same virtual space and perform common tasks. Multimedia datastreams, such as live stereo-video and audio, from a projection-based VR system are transmitted and integrated into the virtual space of another participant at a distant VR system, allowing geographically separated groups to be virtually present in the same virtual space, while maintaining eye-contact, gaze awareness and body language.

In the approach presented in this paper, we merge the stereo video of a remote participant into Cyberstage using the texture based effects of our AVOCADO Software Framework. In particular, each image of the stereo camera is mapped onto a simple geometry representing a plane. The AVOCADO Software Framework then displays the right camera image to the right eye of the viewer and the left to the left eye respectively. However, these two planes have to be fully defined in terms of size, position, aspect ratio and orientation in order to provide the Cyberstage viewer with the best quality stereo-video and to respect the appearance of the

Performer

Performer can be described from two points of view. First, there is the data processing organized in a pipeline and computed in parallel. This rendering pipeline consists of a set of optional units, for:

- a database connection,
- a user application,
- the visual culling of the scene,
- the intersection of objects,
- the drawing of the scene.

The second view focuses on the data structures used to describe the visual virtual world. There are different types of nodes available which can be connected by parent/child relationships to form a directed acyclic graph. Because Performer only supports visual displays, the nodes contain the information useful to describe the visual portions of the virtual world. AVOCADO extends Performer with a sound rendering and tactile feedback which can be configured through the scripting interface to meet the specific needs of different hardware installations.

The data structures in Performer have also been extended to meet the stated general concepts. These extensions are only available in the application process and have to be Performer-compatible, to be used in the processes involved in the pipeline afterwards. This compatibility is achieved by deriving all AVOCADO objects, which have to be rendered visually, from Performer node classes.

Fields

In AVOCADO every object encapsulates its attributes, which describe the state of the object, in fields. All objects are subclasses of a common base class. This base class provides a public interface which allows the retrieval of all field related information. Therefore all objects, even those which are added later as extensions, can be manipulated through a common API. This mechanism allows the effective implementation of the scripting interface, persistence, run-time loading of new object classes and distribution, as the work needs only be done for the common base class. Because there can be more than one field attached to a single object, every object is a field container. The set of fields available for a specific object is static, so it cannot be changed at runtime. Every field has a name and is typed so it can be easily accessed from the scripting interface. Performer 2.1 has a method based object API with getters and setters for all the different fragments of object state information. This is translated into our field API by subclassing all Performer object classes once using special Adaptor Fields to encapsulate method declarations. Further extensions are subclasses only from these adapted Performer classes. This approach ensures full Performer object functionality as a basis for extension development.

Fields come in four flavors. A Single Field holds a single, arbitrary typed value. A Multi Field holds any number of values of the same type. To adapt the Performer method based object API to our field based API a SingleAdaptorField and a MultiAdaptorField are used. All fields are derived from a single base class, and have methods to set and get field values. Fields can be connected to each other, i.e. field A connected to field B will receive B's value whenever field B is changed. This allows a data flow network to be constructed orthogonal to the object hierarchy. This data flow network is evaluated for each simulation frame. Loops are detected and handled properly.

Object representations

While fields are the basic model to represent data and the data flow in the AVOCADO system, at a higher level of abstraction (and inheritance) there are 3 types of object representations: nodes, sensors and services.

Nodes

Nodes are the already mentioned classes adapted to the Performer node classes for the description of the modeling hierarchy. They are field containers and their state is plainly described by their individual set of fields. In a distributed multi-user environment only nodes must be shared with other users. While Performer only knows about static visual objects, AVOCADO also allows audible and tactile properties to be defined.

Sensors

Sensors are field containers, but they are not inherited from Performer objects. They are used for data import and export from the AVOCADO system to the outer world. Sensors are not "visible", "audible" or "tangible" on the different displays and therefore don't have to be part of the modeling hierarchy. They implement the local features of an AVOCADO application and therefore must not be distributed. An example of a typical sensor object are the windows on a workstation screen used as visual display or a graphical user interface used to control some global parameters.

Ongoing projects, that take advantage of the new possibilities introduced by the CRWB, are in the area of geo science and in the medical field. In geoscience seismic data from oil fields and in the medical field volumetric medical data can be viewed as stereoscopic 3D virtual models and manipulated directly by the user. Additionally, two CRWB can be coupled and stereo-video of the remote participants could be exchanged. The collaboration metaphor used is that of a working situation where two specialists are virtually at two different sides of a desk and discuss a model present on the desk.

3. The AVOCADO Framework

AVOCADO is a software framework designed to allow the rapid development of virtual environment applications for immersive and non-immersive display setups like the CAVE, CyberStage, Responsive Workbench and Teleport. It supports the programmer in all tasks involved with this kind of applications. The development of AVOCADO was started in April '96 with the definition of the main goal of such a system:

- The system has to integrate a variety of different interface devices currently in use at GMD. (e.g. the Responsive Workbench, the CAVE, the Communication Wall and the Virtual Studio).
- It has to be sufficiently general purpose to support application development for all these devices.
- As new devices are likely to be invented, it has to be easily extensible and adaptable.
- The system has to be highly interactive and responsive.
- The system must support a rapid prototyping style of application development.
- It has to support the development of truly distributed applications.
- The System is targeted at high-end SGI workstations not less powerful than the Onyx Reality Engine and has to deliver every iota of performance these machines are capable of delivering.

To achieve the goals, AVOCADO supports the following concepts:

<i>Viewer</i>	All kinds of configurations of input and output devices can be assembled to so-called viewers. The viewer builds up the interface between the user and the virtual world. Typical elements of a viewer are the visual, auditory and tactile displays as output devices and spatial trackers, audio or video sources as input devices. In a multi-user environment every user configures his or her own viewer.
<i>Scripting</i>	All relevant parts of the system's API (Application Programm Interface) are mapped to an interpreted scripting language (Scheme). This enables us to specify and change scene content, viewer features and object behavior in a running system.
<i>Streaming</i>	All objects know how to read and write their state to and from a stream. This is the basic facility needed to implement object persistence and network distribution.
<i>Persistence</i>	Together with streaming support for objects this enables us to write the complete state of the system to a disk file at any time. An initial system state can be read from a disk file as well.
<i>Distribution</i>	All objects can be distributed. Their state is shared by any number of participating viewers. Object creation, deletion and all changes at one site are immediately and transparently distributed to every participating viewer.
<i>Extensions</i>	The System is extendible by subclassing existing C++ system classes. This concerns object classes as well as classes which encapsulate viewer features. Compiled extensions can be loaded into the system at runtime via DSOs (Dynamically Shared Objects)
<i>Interaction</i>	Viewers provide input/output services which can be mapped to objects in the scene. Objects can respond to events generated from input devices or other objects and can deliver events to output devices

3.1 Implementation

Because of the focus on the field of high performance rendering, the system is very hardware-dependent. Performer runs only on SGI machines and even there it is mainly designed to be used on the high end platforms. It uses multiprocessing to work on different tasks in parallel, and OpenGL for the visual rendering. OpenGL alone is available on different platforms but there are a lot of extensions only available on SGI machines. It is a very low level library, which does not implement any of the general concepts mentioned earlier and therefore is not used directly. User applications are mainly built using AVOCADO objects and Scheme. Sometimes it is useful to have access to Performer directly but most features are already wrapped by AVOCADO.

environment. A significant characteristic of the Cyberstage is the acoustic floor which allows to generate the sense of vibrations. The two existing CyberStage installations use a wooden skeleton to minimize noise for the electromagnetic tracking.



Fig. 3 Telepresence session in the TELEPORT room

TELEPORT is a synchronous collaboration system that provides high degree of co-presence [1]. The system is based around special rooms, called display rooms, where one wall is a “view port” into a virtual extension. The geometry, surface characteristics, and lighting match the real room to which it is attached. When a teleconferencing connection is established, video imagery of the remote participant (or participants) is composited with the rendered view of the virtual extension (see Fig. 3). The viewing position of the local participant is tracked, allowing imagery appearing on the wall display to be rendered from the participant’s perspective. The combination of viewer tracking, a wall-sized display, and real-time rendering and compositing, give the illusion of the virtual extension being attached to the real room. The result is a natural and immersive teleconferencing environment where real and virtual environments are merged without the need for head-mounted displays or other encumbering devices. The current system uses a 3m x 2.25m rear-projected *video wall* attached to a 3m square room. A *camera* is placed on a stand or a table and set at approximately eye height. The field of view is wide enough to take in a full upper body shot of the local participant. Two techniques are used for *segmentation* (for determining the regions of the video signal where a participant appears) chroma-keying and delta-keying. For audio, each participant wears a small microphone. The audio signals from remote participants are mixed together and sent to speakers mounted on either side of the video wall.

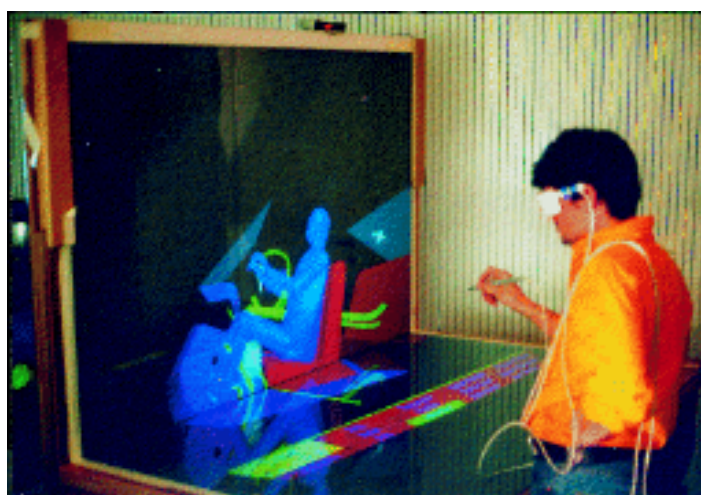


Fig. 4 Cooperative Responsive Workbench

The new Cooperative Responsive Workbench (CRWB) has a horizontal and a vertical display, which are smoothly adjacent. By extending the RWB with an additional horizontal screen, the viewing frustum is substantially increased and virtual objects can be observed at the user’s eye level, which was not possible before. Therefore, the CRWB can be used in more applications areas than the RWB. For example, applications that use high virtual objects, i.e. buildings, can take advantage of the larger viewing space of the CRWB. On the other hand, new functionality can be added to all applications, i.e. teleconferencing and Immersive Telepresence. This is achieved by introducing live stereo-video from a remote participant, which is projected on the vertical screen of the CRWB.

In the section that follows, we present some of our Projection-based VR systems. These systems and the VR applications of our group are based on AVOCADO, the Software Framework for virtual environments, which is presented in section 4. In section 5 the concept of immersive telepresence is introduced and our scientific approach for creating such an environment is described [10]. The demonstration in Cyberstage, of a prototype environment for Immersive telepresence, which took place at GMD's open-house event 1997, is summarized in section 6 [6]. Finally, section 7 concludes this paper with some of the open issues and future research directions.

2. Background

Projective Display Systems are the state of the art in high end Virtual Reality Environments [3]. Releasing the user from the heavy load and inconvenience related to head mounted displays, increasing resolution and rendering speed enables VR for serious applications [5]. Currently desk and room size installations are available, like the Responsive WorkbenchTM, the CyberStage, the Cooperative Responsive Workbench, or the TELEPORT[1]. All of them extend the real space by a virtual space providing a common world coordinate system, where the local and the remote participants are part of Fig. 1.

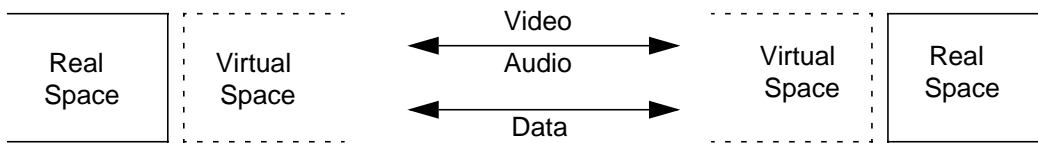


Fig. 1 Projective VR Systems

In the RWB^{TM1} concept [9] the user no longer experiences simulations of interesting procedures on the computer, but the computer is (invisibly) integrated into the user's world, Fig. 2. The virtual objects and control tools, displayed as computer-generated stereoscopic images, are projected onto the surface of a table. The user interacts with the virtual objects and manipulates them as if there were real. Only one viewer is tracked at the moment, while several observers can watch the operations simultaneously, by the use of shutter glasses.

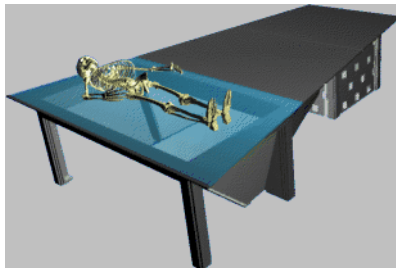


Fig. 2 Responsive Workbench

CyberStage is a CAVE^{TM2} like [2] four-side room-size stereo display system installed at GMD, which creates the illusion of immersion within a computer generated virtual environment. Users see large virtual spaces and hear spatially distributed sound. Projection systems like CyberStage allow a direct and body centered human interaction within virtual worlds as well as team work. Users immersed in a virtual world are physically standing within the display system. Three wall size rear projection systems are installed orthogonal to the floor projection, each with a size of 3x3 meters. An SGI 4 pipe Onyx 2 Infinite Reality generates eight user controlled images. Each pipe generates 10 million shaded triangles per second (peak rate) and is equipped with 64 MB of texture memory. 12 Mips R1000 CPUs are used in combination with 1.5 GB of RAM to compute VR applications. The user position is tracked with Polhemus Fastrak sensors. Crystal Eyes shutter glasses are used for stereo image perception. The display resolution is 1024 x 768 pixels at 120 Hz for each of the four displays. The eight channel-surround-sound system is fed by IRCAM's room acoustic software Spatilisateur [4][11] and provides support for localized sound sources within the virtual

1. CyberstageTM is a registered Trademark of the German National Research Center for Information Technology GMD
2. CAVETM is a registered Trademark of the University of Illinois

Immersive Telepresence in Responsive Virtual Environments

Vali Lalioti, Frank Hasenbrink, Henrik Tramberend and Martin Goebel

GMD - IMK.VMSD
53754 Sankt Augustin
Germany

{Vali.Lalioti, Frank.Hasenbrink, Henrik.Tramberend, Martin.Goebel}@gmd.de

Abstract

In the 10th year of research Virtual Environments now have come into age. Following the head-mounted technology of the first days, now industrial applications in the VR-Labs of the automotive industry are basically using stereo projection type systems like standard single screens, multiple screens, Responsive Workbenches and also cave-like-projection systems. As the globalization of markets requires also world-wide engineering, those industries are investigating into the next generation of an engineering workspace.

In recent years GMD has invented the Responsive Workbench (1993), set up the first European cave-like display (1996) and built a dual sided Responsive Workbench (1988) called Collaborative Responsive Workbench. All display systems have already been evaluated together with industrial partners for their use in engineering visualization. Current efforts at GMD concentrate on integrating video conferencing facilities into these Virtual Environment display systems to foster cooperative engineering visualization. The overall aim however is to develop immersive telepresence for VR displays to support collaborative applications in geographically distributed virtual teams. The paper will briefly review the Virtual Environment display systems and their acceptance at industry. The concept of immersive telepresence will be presented and the currently achieved results will be communicated.

1. Introduction

Virtual Reality is widely accepted as a promising approach to a better man-machine interface, overcoming the present limitations of desktop systems and adapting more closely to the user needs. Projection based VR systems are using metaphors, such as the blackboard or the desk for creating shared working environments that provide a more natural man-machine communication[9][12]. However, today's technology and advances in telecommunication rapidly change the way the business is carried out, making it a globally distributed process, in which communication and collaboration of geographically dispersed groups is of vital importance. VR systems are adapting accordingly, by providing not only a better man-machine interface, but also by facilitating human to human interaction and collaboration over distance. New challenges are introduced in terms of distribution and interaction. It is not only a question of solving the technical problems of gathering and transmitting multimedia datastreams with sufficient quality and speed, but also a question of addressing the specific needs of human communication. For example, facial expression, body language and eye contact are an integral part of this communication. In addition, different types of collaboration must be addressed:

- Synchronous or Asynchronous Collaboration (same-time, different-time)
- Symmetric or Asymmetric Collaboration (n to n, or 1 to n)

Synchronous and Asynchronous collaboration, as defined in CSCW, refers to collaboration taking place at the same time at the same or different place, or at different times respectively [8]. Symmetric or Asymmetric collaboration refers to the roles and degrees of communication between the participants during collaboration. This could reflect the participant's social or business role. For example, teaching or giving a speech could be defined as an asymmetric collaboration. In this paper, we are focusing on synchronous collaboration which could be both symmetric or asymmetric.

Teleconferencing systems that provide high-degree of Telepresence, such as [1], and collaborative co-presence systems such as [7][13][14][15], give enough evidences that projective VR systems when combined with Telepresence facilities can greatly facilitate the communication and collaboration over distance in a variety of application areas. The approach presented in this paper, creates an environment where remote participants not only meet as if face-to-face, but also share the same virtual space and perform common tasks, in order to reach a common goal. In particular, live stereo-video of remote participants is integrated into the virtual space of another participant, allowing a geographically separated group of people to collaborate while maintaining eye-contact, gaze awareness and body language. Participants could be using a wide range of Projective VR systems [1][9][3], resulting symmetric or asymmetric collaboration scenarios.