

Towards Immersive Telepresence

SCHLOSSTAG'97

Frank Hasenbrink, Vali Lalioti
GMD German National Research Center
VMSD
53754 Sankt Augustin
Germany
{Frank.Hasenbrink, Vali.Lalioti}@gmd.de

Abstract

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 group is of vital importance. Teleconferencing and collaborative telepresence systems that provide high-degree of copresence give enough evidences that projective 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. 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. Multimedia datastreams, such as live stereo-video and audio, from a projective VR system are transmitted and integrated into the virtual space of another participant at a distant VR system, allowing geographically separated groups to meet in a common virtual space, while maintaining eye-contact, gaze awareness and body language.

1. INTRODUCTION

Today's technology and advances in telecommunications lead to sophisticated multimedia systems which combined with virtual reality can provide a high degree of copresence and co-working for geographically dispersed groups [2] [12]. Projective VR systems are adapting accordingly, by providing not only a better man-machine interface [10][14], but also by facilitating human to human interaction and collaboration over distance. New challenges are introduced in terms of multimedia integration in distributed virtual reality environments 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, refers to collaboration taking place at the same time at the same or different place, or at different times respectively [9]. Symmetric or Asymmetric collaboration refers to the roles and degrees of communication between the participants during collaboration. These different types of collaboration require different uses of multimedia information, thus demanding different solutions in terms of storing, transmitting and manipulating multimedia datastreams. In this paper, we are focusing on synchronous collaboration in projective VR environments, which can be both symmetric or asymmetric.

Teleconferencing systems that provide high-degree of copresence, such as [2], and collaborative telepresence systems such as [8][14][15][16], give enough evidences that projective 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. 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 and audio 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 [2][10][5], resulting symmetric or asymmetric collaboration scenarios.

In the section that follows, we present some of our projective VR systems and their capability of incorporating multimedia streams. In section 3 the scientific approach and the demonstration in CyberstageTM¹, of a prototype environment for Immersive telepresence are described. Finally, section 5 concludes this paper with some of the open issues and future research directions.

2. BACKGROUND WORK

Projective Display Systems are the state of the art in high end Virtual Reality Environments [5]. Releasing the user from the heavy load and inconvenience related to head mounted displays, increasing resolution and rendering speed enables VR for serious applications [7]. A common characteristic of the projective VR systems, is that they all extend the real space by a virtual space providing a common world coordinate system, where the local and the remote participants are part of, as shown in Figure 1.

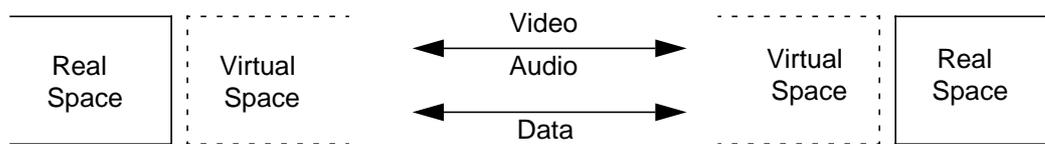


Figure 1 Projective VR Systems

The mapping of virtual space to real space allows us to characterize them as desk or room size installations. Currently desk and room size installations are available, like the Responsive WorkbenchTM, the CyberStageTM or the Teleport [2]. In addition, the specific installation and the application, determine the type of multimedia datastreams that can be used during a VR session. For example, spatially distributed sound is used in Cyberstage, live audio and video is used in Teleport.

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

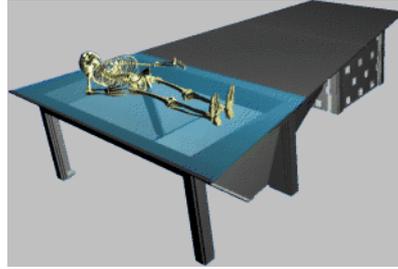


Figure 2 Responsive Workbench

In the RWB concept [10] the user no longer experiences simulations of interesting procedures on the computer, but the computer is (invisibly) integrated into the user's world, Figure 2. The virtual objects and control tools, displayed as computer-generated stereoscopic images are projected onto the surface of a table. The computer screen is changed to a horizontal, enlarged desk top version and replaces the two-dimensional flat screen. 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.

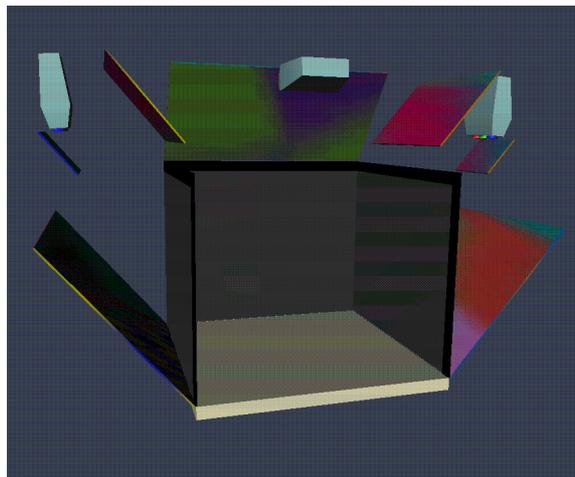


Figure 3 Cyberstage installation

CyberStageTM is a CAVETM¹ like [4] four-side room-size stereo display system installed at GMD, which creates the illusion of immersion within a computer generated virtual environment, Figure 3. 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 Spatialisateur [6][11] and provides support for localized sound sources within the virtual

1. CAVETM is a registered Trademark of the University of Illinois

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



Figure 4 Telepresence session in the TELEPORT room

TELEPORT is a synchronous collaboration system that provides high degree of co-presence [2]. 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 Figure 4). 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.

3. TOWARDS IMMERSIVE TELEPRESENCE

Our software framework for distributed virtual environments, AVOCADO, was developed in parallel to the Cyberstage installation and became the main platform for research and development concerning virtual environments at VMSD, GMD’s research department for Visualization and Media Systems Design. One of the main features of AVOCADO is its easy extensibility which facilitates the expansion of the system. Therefore, numerous effects have been added to AVOCADO including a set of texture based effects. These effects mainly deal with the dynamic exchange of the texture image and are capable of showing live video input, play movie files from disk or slice through a 3d image volume.

In the approach presented in this paper, we integrate live stereo video and audio of a remote participant into Cyberstage using the texture based effects and the spatially distributed sounds effects of our AVOCADO Software Framework. For live stereo video, 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. 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.

3.1 SCHLOSSTAG'97

For our open-door event in October 1997 (Schlosstag '97) we demonstrated for the first time our approach 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.

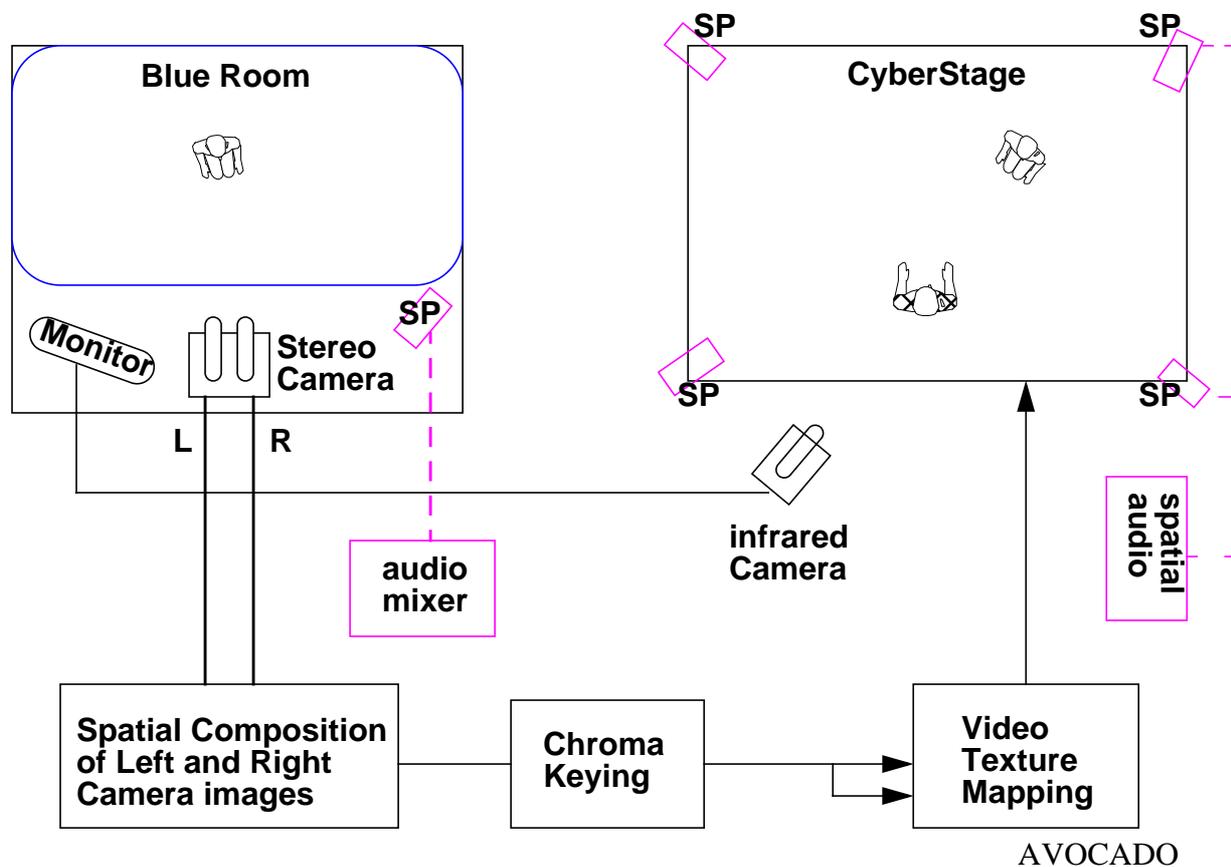


Figure 5 Schlosstag 97 set-up for Immersive Telepresence

An overview of the set-up, used during Schlosstag '97 for the demonstration of the virtual meeting scenario, is shown in Figure 5. 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.

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 Figure 5 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.



Figure 6 Reconstruction of Schlosstag'97 Cyberstage View

Figure 6 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. In addition, a virtual projection screen showed a prerecorded laparoscopic treatment for additional information.

REFERENCES

- [1] Bly S.A., Harison S.R., Irwin S., MediaSpaces: Bringing People Together in Video, Audio, and Computing Environment, CACM, Vol 36, No. 1, pp. 29-47, January 1993.
- [2] 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
- [3] Buxton W., Telepresence: integrating shared task and person spaces. Proc. Graphics Interface '92, 123-129.
- [4] 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.
- [5] Dai P., Eckel G., Goebel M., Hasenbrink F., Laloti 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.
- [6] 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.
- [7] Haase H., Dai F., Strassner J., Goebel M., Immersive Investigation of Scientific Data, Scientific Visualization, IEEE Press, 1997
- [8] Ishii H. and Kobayashi M., ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact, CHI'92, May 3-7, 1992.
- [9] Johansen R., Groupware Computer Support for Business Teams, New York, London, The free press, 1988.
- [10] Krueger W. and Froehlich B., The Responsive Workbench, IEEE Computer Graphics and Applications, May 1994
- [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] Magnenat Thalmann N. and Thalmann D., Virtual Worlds and Multimedia, Wiley Professional Computing, John Wiley and sons, Chichester, England, 1993.
- [13] Okada K., Maeda F., Ichikawa Y., Matsushita Y., Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design, Proc. CSCW'94, pp. 385-393
- [14] Wellner P., DigitalDesk, Communications of the ACM, Vol. 36, No. 7, July 1993.
- [15] Weatherall A., GroupSystems Electronic Meeting across the Enterprise and across the World, HICSS-30, January 97, Maui, Hawaii.
- [16] <http://www.xerox.fr/ats/br/livead.html>

