Collaborative Work with Volumetric Data Using Augmented Reality Videoconferencing

István Barakonyi Tamer Fahmy Dieter Schmalstieg
Vienna University of Technology
{bara|tamer|schmalstieg}@ims.tuwien.ac.at

Karin Kosina
Systems in Motion
kyrah@sim.no

Abstract

The Augmented Reality Videoconferencing System is a novel remote collaboration tool combining a desktop-based AR system and a videoconferencing module. The novelty of our system is the combination of these tools i.e. superimposing AR applications on live video background displaying the conference parties' real environment, thus merging the advantages of videoconferencing (natural face-to-face communication) and AR (interaction with distributed virtual objects using tangible physical artifacts). We demonstrate the system's collaborative features with a volume rendering application that allows users to display and examine volumetric data simultaneously and to highlight or explore slices of the volume by manipulating an optical marker as a cutting plane interaction device.

Keywords: Augmented Reality, Computer Supported Collaborative Work, Videoconferencing, Volume Rendering.

1 Introduction

We developed an AR videoconferencing tool that runs AR applications superimposed on live video background. The system merges the advantages of videoconferencing and AR, combining natural face-to-face communication with the capability to interact with distributed virtual objects using tangible physical artifacts. The conference parties' real environment is recorded with a video camera, allowing them to see each other and enhance their communication with non-verbal cues. Virtual objects, i.e. the "conference material", are bound to physical optical markers that can be freely positioned in the user's environment.

Our application is built on the Studierstube collaborative AR platform [5]. Studierstube supports multi-user interaction on various configurations including our desktop-based setup. The interaction and object manipulation is done with the Personal Interaction Panel (PIP), a two-handed AR interface provided by the Studierstube platform. The PIP consists of two tracked, handheld physical props for each user: a panel augmented with virtual interaction widgets (buttons, sliders, etc.) and a pen for user input (i.e. clicking, dragging etc.) on the widgets. Tracking data for the interaction props can come from either optical markers or keyboard commands.

Each conference party sees two application windows on the screen: one representing the local user and the other representing the remote party. The application is currently limited to two collaborators. The local user's application window displays the virtual objects, which are interaction props and the "conference material" objects, overlaid over the local video. The position and orientation are calculated from optical markers that are extracted from the local video sequence. The application window representing the remote party displays the remote video sequence, which is encoded and compressed on the remote computer and sent over the network via the videoconferencing module. At the same time, the remote user's tracking data is extracted and streamed to the local machine using multicast or TCP/IP connections. Position and orientation information of the optical markers have already been processed on the remote client's machine while the actual rendering and overlay over the video data is done on the local machine.

Our system is superior to a video-only solution since the image quality is significantly better. Tracking data calculations are not duplicated. The tracking information is more precise, since it is extracted from the higher-quality local video, and it is possible to interact with the virtual objects.

Advantages of our system over an interactive application sharing approach are the much higher speed and lower bandwidth, and the fact that no extension module for proper handling of real-time video is needed.

2 Volume Rendering Application

We chose a volumetric rendering application for demonstrating the capabilities of our tool since the users can jointly and interactively examine a shared set of volumetric data, which nicely illustrates the collaborative features of our system and serves as an appealing visualization software, particularly in the medical and geological domain. We are using Systems in Motion's SimVoleon [2] library, which visualizes volumes by using 2D texture slices. SimVoleon loads VOL-format files and can handle arbitrary volume dimensions, i.e. non-power-of-two dimensions and different dimension along the axes. It supports mixing of volume data and polygonal geometry, picking operations on voxels, and changing of the transfer functions in real time.

In our application the collaborators independently choose a volume that they can examine as a whole or in slices. Different settings of the volume like e.g. axis, interpolation, color maps etc. can be adjusted. Optical markers are used to either move the volume as a whole or to allow the user to view arbitrary slices of the volume.

The system makes it possible to remotely and interactively explore and discuss data sets. Potential users of such an application are e.g. surgeons examining parts of the human skeleton acquired by Computer Tomography or geologists studying processed seismic data sets acquired by onshore / offshore surveys using a series of geophones.

3 Implementation

One of our major goals was to keep a simple, low-cost setup that is affordable for everyday use, avoiding solutions that use expensive tracking systems, high-bandwidth networks or costly proprietary videoconferencing software. Our setup for each client consists of a 1.5GHz PC with 512MB RAM and an NVIDIA Quadro4 graphics card, a flat-panel LCD monitor, and a lightweight FireWire camera. We use the ARToolKit [4] software for getting tracking information from the optical markers, which can be easily created.

The videoconferencing module is based on the OpenH323 [1] software, an open source protocol stack incorporating a set of communication protocols developed by the International Telecommunications Union (ITU). We chose a subset of communication protocols that relies on the call control and media control protocols and the H.261 video codec for low-bandwidth video transmission. The required bandwidth does not exceed 150 kbps.

4 Demonstration

We use two PCs (as specified in section 3) for the demo setup. Each of them has a camera flexibly mounted on the top of the monitor, pointing at the respective user in front of the screen. Interaction with the system is done with optical markers, the keyboard and the mouse. Users wear a headset for the voice communication.

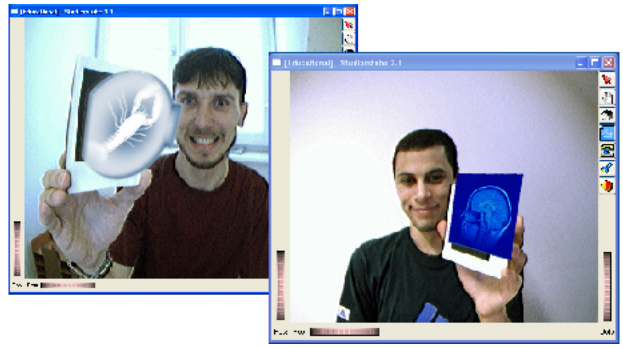

Figure 1. Screenshot from our demo

5 Future work

Our future plans include the following:

- Handling more than two conference parties. This might cause some communication conflicts (e.g. who speaks and when in the audio channel) and implicit object locking, therefore a collaboration protocol needs to be designed and utilized.
- Experimenting with other codecs for higher image quality and better compression.
- Using other features available in the SimVoleon software for other ways of volume exploration.

Acknowledgments

This work was sponsored by the Austrian Science Fund *FWF* (contract no. Y193) and the European Union (contract no. IST- 2001-34204). Systems in Motion [3] kindly provided us with their Volume Rendering library SimVoleon[2].

References

[1] Openh323 project website. *http://www.openh323.org/*.
[2] Simvoleon. *http://www.sim.no/products/SimVoleon/*.
[3] Systems in motion. *http://www.sim.no/*.
[4] H. Billinghurst, M. Kato. Collaborative mixed reality, in proc. ismr 99. Technical report, 1999.
[5] D. Schmalstieg, A. Fuhrmann, G. Hesina, Z. Szalavri, M. Encarnao, M. Gervautz, and W. Purgathofer. The studierstube augmented reality project. In *PRESENCE -Teleoperators and Virtual Environments)*. MIT Press, 2002.