Immersive High Quality Communication

Andreas Gräff and Norbert Schiffner

Fraunhofer Institute for Computer Graphics, Rundeturmstrasse 6, D-64283 Darmstadt, Germany

Abstract

This shows our vision of advanced high quality tele-communication which can be achieved by combining large-scale displays, real-time audio/video transmission, virtual reality and global, high bandwidth computer networks.

1. Introduction

In natural communication, people have visual contact while talking to one another. In this way, they receive more information than just what they hear: they can see and interpret movements, gestures and mimic. This supplementary information influences their reactions and their behaviour as well as their further decisions.

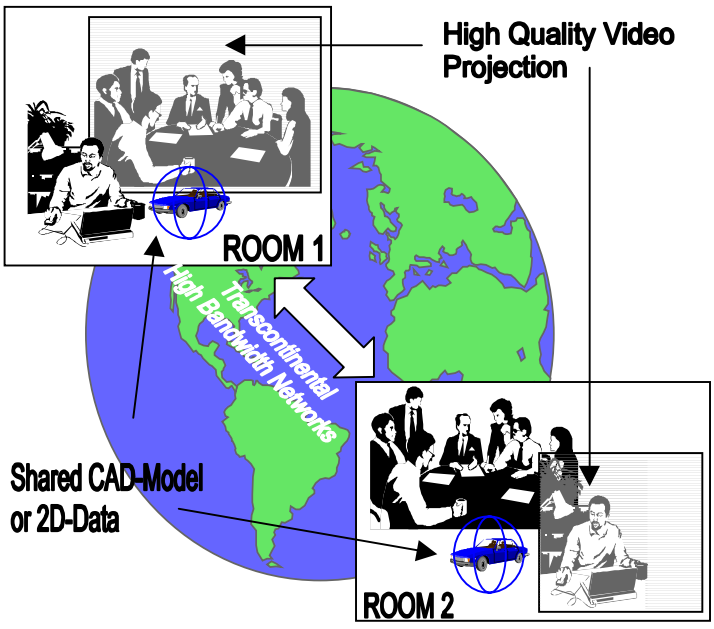
In order to integrate this type of information in tele-communication over large distances it is necessary to adapt already existing technologies (e.g. large-scale displays, spatial audio, high bandwidth networks) to realistic scenarios as well as develop advanced forms of audio/video encoding.

2. Visions of Communication

The installation described in this section (Picture 1) allows natural communication – as described in the Introduction - between people situated at distant locations.

High resolution, large-scale back projection screens can be used in an office or living room: they can either be put up on walls or integrated into horizontal surfaces (e.g. tables, sideboards). Real-time video images from other rooms located around the world can be projected into these screens. In this way, the real office or room can be extended by one or more virtual rooms. The smooth transition from the real into the remote virtual rooms is achieved through the correct positioning and perspective of the remote video captures [1].

Additionally, it is possible to overlay video objects with computer-generated graphical objects (e.g. the two-dimensional output of standard word processors or three-dimensional models of CAD and medical visualisation programs) [2]. This enables the users to communicate and to share data simultaneously in a very natural way. In a conference, a virtual object can be projected on top of the table between the conference partners. All participants can see the object from their perspective and



Picture 1: Immersive High Quality Communication

each participant can make annotations or manipulate the object by way of an input devices.

For an optimal immersive impression of a video extended room all participants must see in stereo. In order to achieve this effect, video capturing has to be made by stereo cameras and all participants have to wear active shutter glasses or polarised goggles. Active stereoscopic displays which allow the users full three-dimensional perception without any kind of on-body equipment is under development and will be available in the near future. This new technology is based on standard display techniques merged with an active image-splitting system and controlled by an optical head-tracker.

For high quality immersive audio connections between all sites a spatial audio system for recording and playback has to be installed. Special attention has to be paid to the echo cancellation for keeping the users free of disturbing devices as head-speakers and microphones.

One input device - e.g. a keyboard - should suffice for the whole system. All remaining input (which cannot be handled by a keyboard) should be controlled automatically through motion capturing, speech and gesture recognition and autonomous agents.

3. Real-time Audio/Video Data

A stable audio and video toolkit with excellent quality and low transmission delay is an essential part of any telecommunication solution.

The human audio sense has a sensitivity extending from ca. 10Hz to 20kHz which implies an audio sampling rate of 44kHz (CD quality) or more. The human visual sense has a resolution of ca. 8400x5400 pixel at a repetition frequency of ca. 50Hz which leads to a 54Gbit/s data stream (24bit true-colour) for every single video source. But data streams of this size cannot be handled by any existing real-time system. Therefore, new mechanisms for real-time data reduction and compression have to be developed. The discrete wavelet transformation (DWT) is one possible solution for the encoding of audio and video data [3] Together with optimised quantisation and encoding technologies this method has already produced excellent results for audio encoding and also good results for video encoding [4].

4. High-Bandwidth Networks

Even if the real-time encoding mechanisms are very effective, a high-quality communication system will produce network traffic of up to 10Mbit/s between all participants. To guarantee the reliable transmission of this data over the Internet, new advanced network protocols and services have to be developed and implemented into the transport layer of the communication applications.

Multicast protocols will help to reduce bandwidth in communication scenarios with more than two participants.

QoS filters will support heterogeneous groups. Generally, the perception of video and audio quality is user-dependent and for many multimedia applications degradation of play-out quality can be tolerated to some extent [5]. QoS filters can be used to adapt continuos media streams to specific QoS requirements. A QoS filter removes a certain amount of information from the stream while preserving the most important parts.

The resources reservation protocol RSVP [6] has been designed to request a specific quality of service from the network. It is proposed to be part of the Integrated Service Architecture [7] of the future Internet. In this context, RSVP is intended to request resources like bandwidth or processing time on behalf of the receiver.

References

[1] H. Fuchs, Building Telepresence Systems: Translating Science Fiction into Reality, Eurographics 1997
[2] Rob Koenen, MPEG-4 Overview, ISO/IEC JTC1/SC29/WG11, 1998
[3] Amara Graps, An Introduction into Wavelets, IEEE, 1995
[4] M. Isenburg, Transmission of multimedia data over lossy networks, ICSI, Berkeley, 1996
[5]J.Pasquale, G.Polyzos, and V.Kompella. The multimedia multicast Problem. Technical Report TR CS93-313, UCSD, 1993
[6] l.Zhang, S.Deering, and, D.Estrin. RSVP: A new resource Reservation Protocol. IEEE network, 7(5) Sept. 1993
[7] R.Braden, D.Clark, and, S,Shenker. Integrated services in the Internet architecture: an overview RFC 16633, ISI, June 1994