Tele-Immersive Collaboration in the CAVE Research Network

Andrew Johnson & Jason Leigh

Electronic Visualization Laboratory
University of Illinois at Chicago

aej | spiff @ evl.uic.edu
www.evl.uic.edu/cavern/

In this chapter we will discuss our work to enable effective collaboration between remote participants within immersive virtual reality environments. We will discuss several of the environments we have created, the lessons we have learned, and the issues that are driving our current research. Our focus is on high-quality collaboration among a handful of participants connected by high-speed, high-bandwidth national and international networks in areas such as scientific visualization, education, and design. Sharing these virtual environments with your remote collaborators as well as the topic of your collaboration may be better than collaborating with them in person, as you can now collaborate while standing inside your data, sharing the context of your discussions.

X.1 Introduction

In 1992 the Electronic Visualization Laboratory (EVL) developed the CAVE. Now in the year 2000, with more than 200 CAVE and related projection-based VR environments around the world, there is a community that is eager to collaborate. The CAVE™ Research Network (CAVERN) is an international alliance of research and industrial institutions equipped with CAVEs, ImmersaDesks™, and high performance computing resources, interconnected by high-speed networks. This high-end visualization hardware combined with high bandwidth networks allows us to explore new research problems and applications of this collaborative technology without being hindered by the limits of the existing Internet.

The CAVE [Cruz-Neira et al, 1993] is a virtual reality (VR) system where the display is a 10 foot-cubed room that is rear-projected with stereoscopic images, creating the illusion that 3D objects appear to co-exist with the user in the room. A user dons a pair of lightweight liquid crystal shutter glasses to resolve the stereoscopic imagery, and holds a three-button 'wand' for three-dimensional interaction with the virtual environment. An electromagnetic tracking system attached to the shutter glasses and the wand allows the CAVE to determine the location and orientation of the user's head and hand at any given moment in time. This information is used to instruct the Silicon Graphics Onyx that drives the CAVE to render the imagery from the point of view of the viewer. This way, the user can physically walk around an object that appears to exist in 3D in the middle of the CAVE. If the viewer wants to look behind a virtual object, he walks around to the back. If the viewer wants to look under an object in the CAVE, she crouches down and physically looks under the virtual object. The wand contains three buttons and a joystick that can be programmed for different features depending on the application. Typically the joystick is used to navigate through environments that are larger than the CAVE itself, such as architectural walk-throughs. The buttons can be used to change modes, or bring up menus in the CAVE, or to 'grab' a virtual object. Speakers are mounted to the top corners of the CAVE structure to provide environmental sounds from the virtual environment and audio from the remote participants. CAVEs and CAVE-like devices are currently in use in universities, research labs, design centers such as General Motors and Caterpillar, and museums such as Ars Electronica in Austria, and the Foundation for the Hellenic World in Greece.

In 1994, EVL developed the ImmersaDesk [Czernuszenko et al, 1997] – a smaller, drafting-table-like system with a 6' by 4' angled screen that is also capable of displaying rear-projected stereoscopic images. Originally designed as development and testing stations for CAVE applications, the ImmersaDesk, and its successor the ImmersaDesk2 quickly found a place as immersive display devices in research laboratories and as a convenient way to 'take VR on

the road' to conventions for presentations. Whereas the CAVE is very well suited for providing panoramic views of a scene (particularly useful for walking through architectural spaces), the ImmersaDesk better suited for displaying images that fit on a desktop (for example, CAD models.)

The CAVE and the ImmersaDesk can run the same VR applications, providing their users with different views of the same virtual worlds. These applications are typically written in C or C++ using OpenGL or SGI Performer to render the computer graphics and the CAVE library to deal with interfacing with the VR hardware. The CAVE library has traditionally been run on SGI computers, but a Linux version is now available and a Windows NT version is under development that will take advantage of the new generation of fast personal computers, and broaden the accessibility to these virtual worlds. The CAVE and ImmersaDesk are shown in Figure 1.

It is common to have several people standing in the CAVE or in front of the ImmersaDesk at the same time. While only one person has the correct stereo viewpoint and the ability to interact with the environment, the other viewers can still see the virtual world in 3D. We have found that the ability to talk with co-workers who are standing next to you is very important, and since the CAVE and the ImmersaDesk do not isolate the user from the real world, it is convenient to have these interactions. Our current work focuses on tele-immersion - making remote collaboration just as easy, and in fact <u>better</u> than standing next to your collaborator because now each user can stand within the shared virtual environment seeing a view of that environment that is customized to their interests and experience.
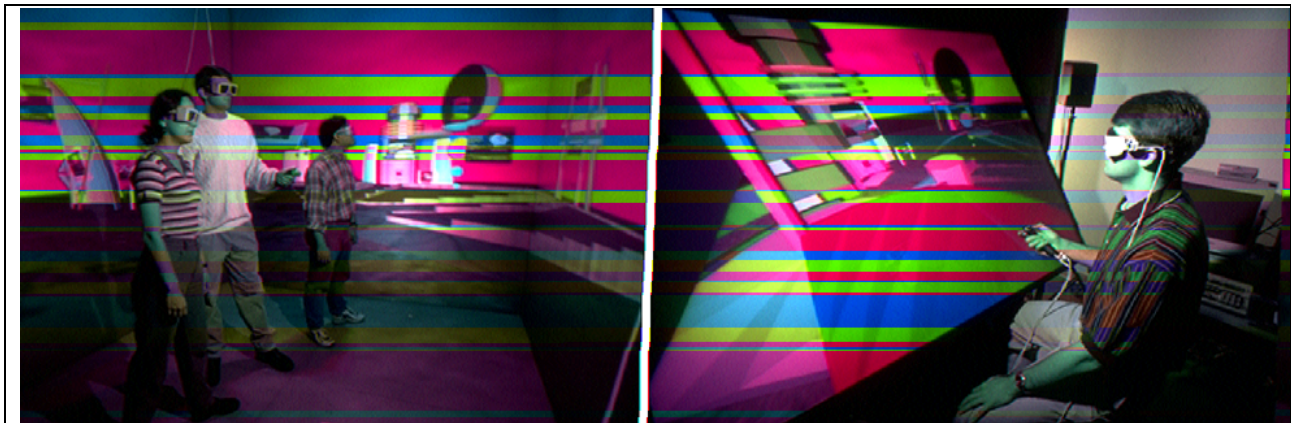
Figure 1: The CAVE and ImmersaDesk. The left image shows three people in the CAVE. All three wear shutter glasses to see the virtual world in stereo, but only one has a tracked pair of glasses and carries the wand. The right image shows a single user sitting in front of the ImmersaDesk wearing the same tracked glasses and carrying the same wand as in the CAVE. The CAVE and ImmersaDesk users can interact with the same virtual worlds from different perspectives.
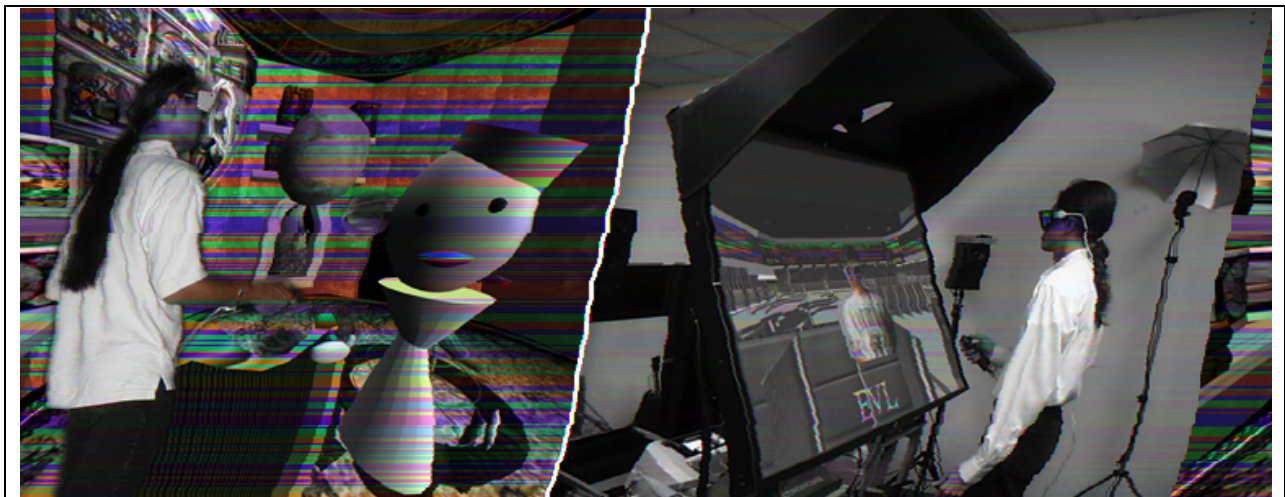
X.2 Tele-Immersion

Tele-immersion connects users of high-end VR equipment, such as CAVEs and ImmersaDesks, together over high-speed high-bandwidth networks. The focus of tele-immersion is supporting high-quality interaction between small groups of participants involved in design, training, education, scientific visualization, or computational steering. The ultimate goal of tele-immersion is not to reproduce a face-to-face meeting in every detail, but to provide the 'next generation' interface for collaborators, world-wide, to work together in a virtual environment that is seamlessly enhanced by computation and access to large databases. While the goal of audio and video teleconferencing is to allow distributed participants to interact as though they are in the same physical location, tele-immersion allows them to interact as though they are the same immersive virtual environment. This way they can interact with each other and the objects in their shared environment.

This shared environment may be the design of a new car, a visualization of climatological data, or other three-dimensional environments that do not physically exist, or cannot be physically visited. The participants are not

talking about a thunderstorm, they are standing inside it; they are not looking at a scale model of a new car design, they are standing inside the full size engine block. By transmitting gestures as well as audio and video between the collaborators, we believe these shared virtual environments give their users a greater sense of presence in the shared space than other collaborative mediums. By encouraging collaboration and conversation within the data, these environments may become the preferred place to work and interact even if more traditional face-to-face meetings are possible. However, tele-immersion is not going to replace email, phone calls, or existing tele-conferencing systems. They each have their strengths and uses. Just as word processing documents, and spreadsheets, and white boards are shared across the Internet to put discussions into their appropriate context, sharing a virtual space with your collaborators as well as the 3D design being considered or the simulation being visualized, puts these discussions into their appropriate context.

A typical tele-immersive space will be a persistent virtual environment maintained by a computer simulation that is constantly left running. The space exists and evolves over time. Users enter the space to check on the state of the simulated world, discuss the current situation with other collaborators in the space, make adjustments to the simulation, or leave messages for collaborators who are currently asleep on the far side of the planet. For example, in a computational steering application, a supercomputer may be running a very large simulation that takes several days to complete. At regular intervals the supercomputer produces a 3D snapshot of the current data, perhaps a visualization of cosmic strings. A scientist can then step into the CAVE and look at the 3D data that has been produced to see if the simulation is progressing correctly, or whether the simulation needs be tuned, to focus on particular details, rather than waiting for the simulation to complete.

Presence in the virtual world is typically maintained using an avatar, or a computer generated representation of a person. These avatars may be as simple as a pointer that depicts the position and orientation of the wand in the virtual world. However, having representations of the physical bodies of the collaborators can be very helpful in aiding conversation and understanding in the virtual space, as you can see where your collaborators are and what they are looking at or pointing at. Tracking the user's head and hand position and orientation allows the computer to draw computer-generated characters representing each of the remote collaborators. These articulated characters move along with the remote user, and are able to transmit a decent amount of body language such as pointing at objects in the scene, and the nodding or tilting of the head. They are very useful in task oriented situations, but do not work as well for negotiation. Seeing high quality live video of a person's face can improve negotiation. Video avatars, full-motion full-body video of each user, allow very realistic looking collaborators in the space which improves recognition but require much higher network bandwidth and careful setup to achieve high quality. Figure 2 shows a user interacting with an articulated avatar in the CAVE and a user interacting with a video avatar at the ImmersaDesk2.

Figure 2: Remote participants in tele-immersive sessions can be seen in the shared virtual space in several ways. The user on the left is interacting with the articulated computer generated avatar of a remote participant in the CAVE. The user on the right is interacting with the live video avatar of a remote participant on the ImmersaDesk2

In light of the complex interaction of computer graphics, networking, databases, and human factors involved in tele-immersion, developing these applications can be a daunting task. To address this issue, EVL has been developing CAVERNsoft - a software infrastructure that supports the rapid creation of new tele-immersive applications and eases the retrofitting of previously non-collaborative VR applications with tele-immersive capabilities. Software frameworks for collaborative virtual environments have been under development for some time by several research groups [Kessler et al, 1998]. Their focus however, has been largely based on supporting existing low-bandwidth Internet infrastructures or massive connectivity involving thousands of participants at the same time (as in military simulations or Internet based computer games.) Our focus on the use of VR for manufacturing, scientific and information visualization has had a different set of requirements. We are building systems for small working groups (typically no more than 7 collaborators at a time) but with large data distribution requirements. This is motivated by the need for high fidelity audio and video communications, and the sharing of large engineering and scientific data-stores that are connected over high speed nation-wide (vBNS - very high speed Backbone Network Service, MREN - Metropolitan Research and Education Network) and world-wide (STAR TAP, TransPAC, EUROLINK) networks.

The CAVERNsoft framework was developed based primarily on our experiences creating two tele-immersive applications, CALVIN and NICE, in 1995 and 1996 respectively, which are shown in figure 3 and described next.

X.2.1 CALVIN

CALVIN (Collaborative Architectural Layout Via Immersive Navigation) [Leigh et al, 1996] was a collaborative virtual environment that allowed multiple users to synchronously and asynchronously experiment with architectural room layout designs in the CAVE. Participants were able to move, rotate, and scale architectural design pieces such as walls and furniture. Participants could work as either `mortals' who see the world life-sized (classically known as an "inside-out" view), or as `deities' ("outside-in" view) who see the world as if it were a miniature model. Deities, by virtue of their enlarged size relative to the environment, towered above the scene and were better at performing gross manipulations on objects. Mortals on the other hand were in the same scale as the environment, and were hence better able to perform fine manipulations and see how the layout felt from the perspective of someone walking through it. The roles of 'mortal' and 'deity' we not fixed. A user could decide to change their role, and their size, at any time.

The avatars for the mortals and deities in CALVIN were human-shaped, but with minimal detail and exaggerated features. Our main inspiration for them was the Muppets. Kermit the Frog is not very detailed or realistic looking, but he seems quite alive through his motions and his actions and reactions. We wanted to try emphasizing realistic motion over realistic appearance and use the exaggerated features of these articulated avatars to make those motions easier to see. For verbal communication, ambient microphones at each participating site were connected through a telephone conference call. At the time, late 1995, the networks were not fast enough to stream multi-participant high-quality audio reliably, and the conference calls gave us high quality audio with multiple participants at very high reliability. If only two sites were collaborating, the participants could also connect over an ATM network, enabling them to bring up a window in the virtual space to see live video of the other participant. This allowed the collaborators to see each other face-to-face and aided in negotiation.

Asynchronous access to the current design allowed users to enter the space whenever inspiration struck them, rather than requiring them to wait to schedule formal meetings with their collaborators. Participants were able to work on their own, and save different versions of the design as the collaboration progressed. They could then hold a collaborative session and discuss either the current design or any of the previous designs in a collaborative session.
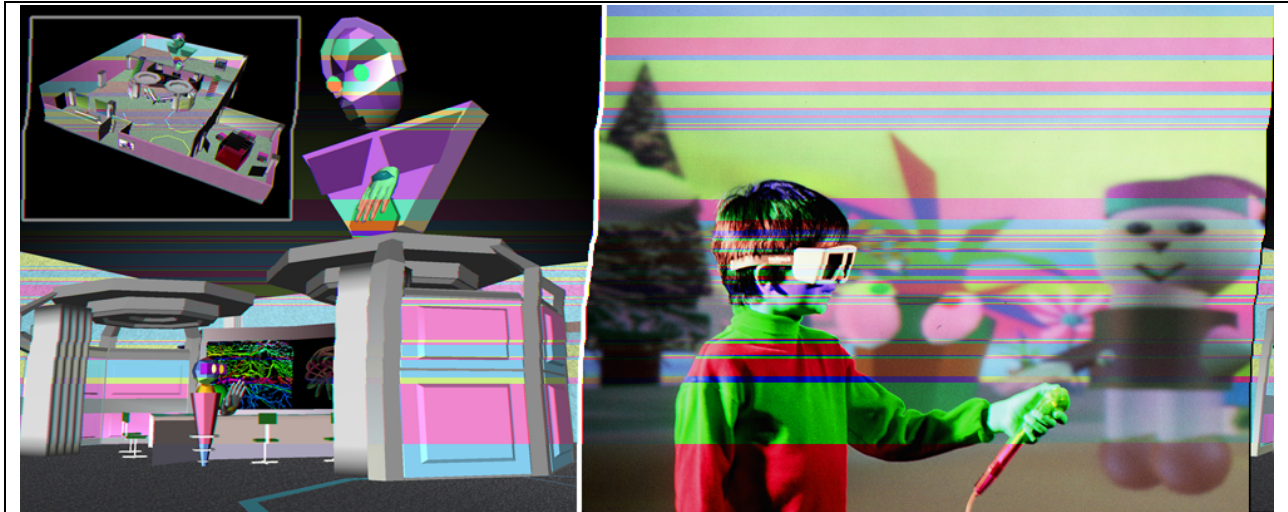
Figure 3: CALVIN and NICE. The left image shows the CALVIN design space with the mortal working within the scene and the diety towering overhead. The right image shows Eddie interacting with Jim, depicted as an avatar, in the NICE garden.

X.2.2 NICE

When the CALVIN project ended, we were still very intrigued by this idea of asynchronous collaboration and heterogeneous views, and wanted to investigate them further in other application areas. The NICE (Narrative Immersive Constructionist / Collaborative Environments) project was a collaborative environment in the form of a virtual island for young children [Roussos et al, 1999]. In the center of this island the children could tend a virtual garden. The children, represented by articulated avatars, could collaboratively plant, grow, and pick vegetables and flowers. They needed to ensure that the plants had sufficient water, sunlight, and space to grow, and kept a look out for hungry animals which might sneak in and eat the plants. The idea was to help teach children simple concepts about plant growth, though the application turned out to be much more successful as a social space and a testbed for tele-immersion than as an educational space.

With a new generation of graphics computers available, the avatars in NICE became more sophisticated than those in CALVIN, though our emphasis was still on realistic motion more than realistic features. In NICE we also included avatars that were clearly non-human. Participants could be birds or bunny-rabbits, or bees. Since the world is virtual, it can be rather limiting to always be represented as a humanoid avatar, and there may be advantages to looking like something else. In NICE we allowed adults to share the virtual space with the children without appearing as obvious adult authority figures. Instead of having the avatar of a man or a woman, the adult could appear as a friendly owl. Continuing with our interest in heterogeneous perspectives, the children in NICE could shrink down to the size of a mouse and walk under the garden to see the root system. This also allowed a full-size child to pick up one of the shrunken kids and carry them around. Using the same audio connection as CALVIN, the children could talk to each other using ambient microphones connected through a conference call.

In CALVIN the virtual world only changed when one of the users made a change. In NICE we wanted a more sophisticated world that could evolve on its own, even when all the participants had left the environment and the virtual display devices had been switched off. Unlike CALVIN where the server that maintained the state of the virtual world was started before each session, NICE's virtual environment was, and still is, persistent; the server remains running. The plants in the garden on the island keep growing, and the autonomous creatures that inhabit the island remain active. The current garden has been growing since spring of 1996. Taking the VR equipment and these worlds to various conferences like SIGGRAPH and Supercomputing made us acutely aware of the issues involved in connecting to these shared worlds conveniently. Leaving a server running at a constant location made it easy to go on the road and connect back to the server at home. This also made it easier to add new collaborators since we could

send the NICE client software to a remote site and they could just run it, knowing that the server would always be available.

Because of this flexibility, we could take NICE with us to more places and watch more people experience it. At SIGGRAPH '96, the visitors found they wanted a friendly way to greet new visitors to the space, and began using the flowers in the garden. When a new person appeared on the island, one of the older visitors would offer the new visitor a flower. At Supercomputing '97 we had 17 visitors from three continents on the island and we used the avatars to look at the speed of the network connections. We had everyone stand in a circle and we did the hokey-poky. Watching all of the users 'put their right arm in, take their right arm out' etc gave us a simple but very effective method of seeing which users had fast connections, and we still use this method today as a quick initial check of the networks.

It was not uncommon for members of the NICE team to meet each other in the virtual world, before meeting in the real world on a given day, since the Electronic Visualization Laboratory is spread out over several floors. Sometimes it was easier to find members of the team in the virtual world than in the physical laboratory. Allowing easy access to the island from several remote sites meant that we would occasionally encounter other people on the island and not know who they were or where they were from. We quickly implemented avatar 'nametags' to make identification easier.

Beyond, this, we also needed a convenient way to see who was currently in the virtual world without having to start up a CAVE or an ImmersaDesk. Since the WWW is pretty ubiquitous, it was natural to monitor the NICE garden with a web page. So just as people would set up a 'web cam' in their office, we set up a virtual web cam to monitor our virtual world. Now that we could monitor the NICE garden from the web, we also wanted to look at interacting with it from the web as another form of heterogeneous perspectives. We created a 2D Java applet where the desktop user's mouse position was used to position an avatar in the 3D virtual world, and the locations of the bodies of the VR users were used to position 2D icons on the desktop. At the time Java was still too slow to make this work effectively, but we continue to work on CAVE to desktop collaboration.

Whereas in CALVIN we wanted to see what was possible, in NICE we wanted to see how we could improve the quality of the collaboration and expand access to the shared world.


X.3 Current Applications

Scientists, engineers, and educators are interested in using tools that help them do their work more conveniently and efficiently. When we first introduce scientists to the CAVE, their typical reaction is low-key. Its often hard for them to make the intellectual jump from whatever virtual world they are experiencing to seeing how they could apply this technology in their own field. But when these same people see their own data visualized in the CAVE, then the ideas for how this can be applied come pouring out.

Current tele-immersive work is priming the pump with collaborative projects with interested domain scientists to create these tools, deploy them, evaluate them, and then generalize their effectiveness. Candidate problems for tele-immersion are those that, firstly, can benefit from visualization in an immersive environment; and secondly can benefit from a collaborative solution. Several of these projects are described below, and shown in figure 4. These are compelling applications because, not only do they enhance problem solving in their respective domains, but they provide real world test cases for observing how collaborators work in these high-end visualization environments..

X.3.1 General Motors Collaborative VisualEyes

General Motors has developed VisualEyes- an application that allows designers to import 3D CAD models into the CAVE for quick visual inspection and design reviews at 1:1 scale. This initial use of CAVE-based technology has generated considerable interest in other General Motors sites around the world, some of which have their own CAVE installations. This prompted General Motors to work with EVL to further extend VisualEyes to allow General Motors' trans-globally situated design and manufacturing teams to collaborate in remote design reviews. The goal is to allow designers to both synchronously and asynchronously access a design that persists and evolves over time.
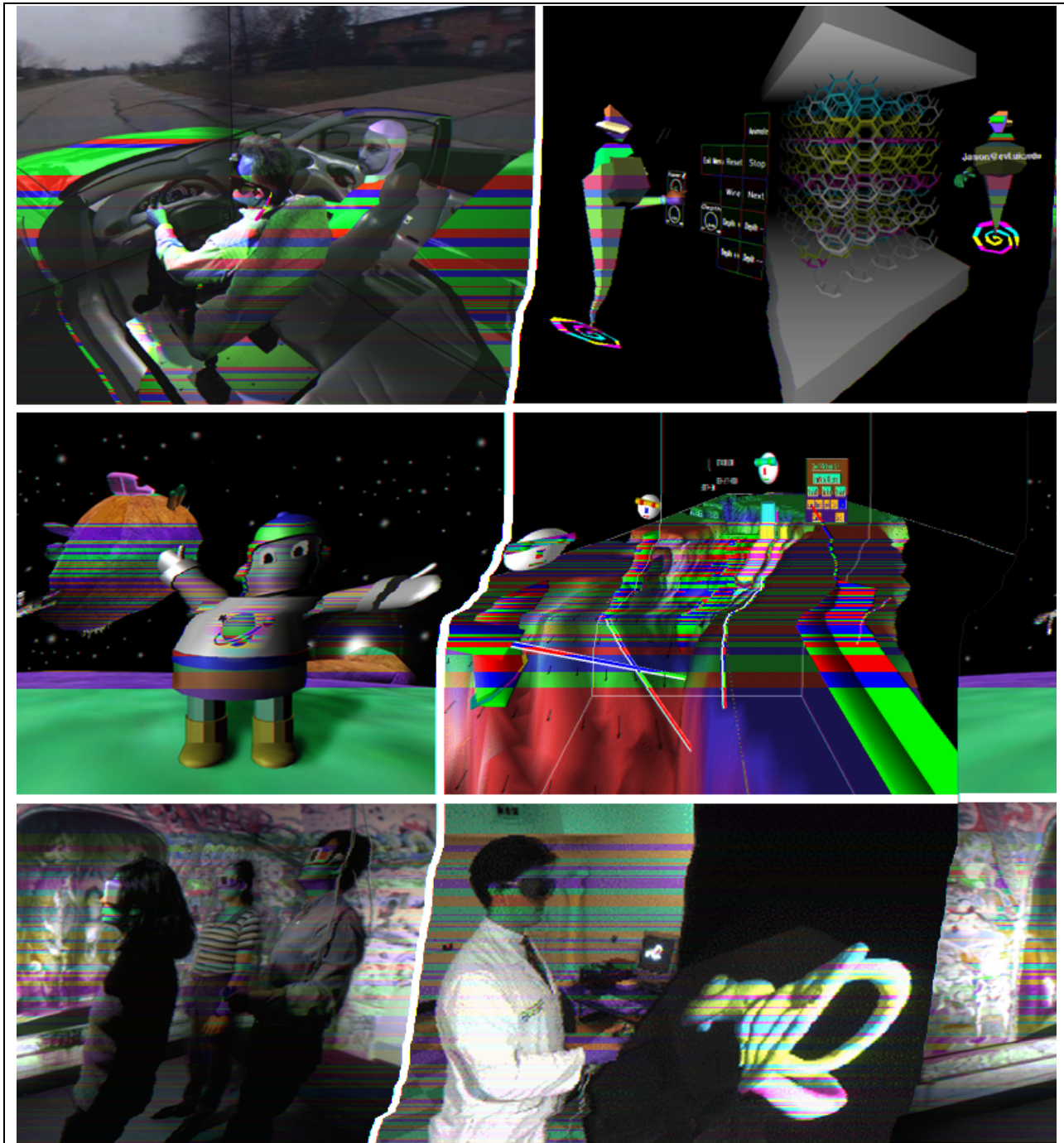
Figure 4: Several current tele-immersive worlds. Clockwise from top left, GM's Randy Smith sits in the driver's seat and a remote participant sits in the passenger seat evaluating a new vehicle interior using GM's VisualEyes. Two users investigate the compression of a sponge lattice in TIDE – the Tele-Immersive Data Explorer. Three remote users, represented as avatars, join the local user in exploring the mouth of the Chesapeake Bay using CAVE6D. A doctor conducts a tele-immersive training session on the structure of the inner ear using the Virtual Temporal Bone. Three users explore one of the cave shrines in the Mogoa Grottos in China within the CAVE. A pair of children collaborate in the Round Earth Project - one child explores the surface of a small spherical asteroid while a second child guides the first from an orbital view - the children must integrate their two views to complete their mission.

A typical working scenario involves a team evaluating the current design in the CAVE and then asking the designer to make modifications on a workstation using a three dimensional modeler such as Alias. Changes in lighting and materials are then propagated automatically to the networked virtual environment allowing all collaborating participants to see the changes simultaneously. They are then able to critique the design and suggest changes to the designer who can do so immediately at the CAD workstation.

X.3.2 CAVE6D- a Tool for Tele-immersive Visualization of Environmental Data

CAVE5D [Wheless et al, 1996] from Old Dominion University, is a configurable VR application framework supported by Vis5D, a very powerful graphics library that provides visualization techniques to display multi-dimensional numerical data from atmospheric, oceanographic, and other similar models. In collaboration with EVL, CAVE6D emerges as an integration of CAVE5D with CAVERNsoft to produce a tele-immersive environment that allows multiple users of CAVE5D to jointly visualize, discuss and interact with data sets in the environment. Avatars possess long pointing-rays that can be used to point at features of interest in the data set while they converse. Visualization parameters such as salinity, circulation vectors, temperature and wind velocity slices which can be visualized by CAVE5D, have been extended in CAVE6D to allow participants to collectively manipulate them. As collaborators are able to operate these visualization tools independently they are able to effectively reduce the individual load of interpreting the data by distributing the multiple observable dimensions between them. We are currently in the process of studying the conditions under which collaborative independent or coordinated view customizations can accelerate or hinder the data interpretation process.

X.3.3 The Virtual Temporal Bone

The Virtual Temporal Bone [Mason et al, 1998] is a tele-immersive educational application from the VR in Medicine Lab and EVL allowing a remotely located physician to teach medical students about the three-dimensional structure and function of the inner ear. In this environment the students and instructor may point at and rotate the ear to view it from various perspectives. They may also strip away the surrounding outer ear and the temporal bone to view the inner anatomy more clearly. Audio from their voice conference is used to modify the flapping of the eardrum to illustrate its function. This application is effective because it leverages the stereoscopic capabilities of the CAVE and ImmersaDesk system to disambiguate the spatial layout of the various structures in the inner ear- something ordinarily difficult to do on standard flat images in medical textbooks.

X.3.4 Silk Road Cave Shrines

The cave shrine project is an ongoing collaboration between historians, artists and computer scientists at Northwestern University and EVL to create a virtual cultural and artistic exhibit of the Mogoa Grottoes of Dunhuang, China. The grottoes, one of western China's ancient cultural sites, are considered the gateway to the well-known Silk Road - the East-West trade route between Asia and Europe. They consist of 492 caves with murals covering 25,000 square meters, wall fresco paintings, and more than 3000 painted sculptures. These caves were built over a period of one thousand years, from the forth century to fourteen century. VR is important here because physical access to these caves is limited, and photographs do not accurately convey the relationships between the murals that surround the visitor. Tele-immersion will allow historians to take remote visitors on tours of the caves.

X.3.5 Tele-Immersive Learning Environments

The successor to the NICE project was the Round Earth project. While NICE was successful as a tele-immersion testbed, it was not as successful as an educational space because it did not have a focused educational goal. The Round Earth project has a very focused educational goal - using tele-immersion to help teach young children the concept of a spherical Earth [Johnson et al, 1999]. Two children collaborate in exploring a small spherical asteroid. One child acting as an astronaut explores the surface of the asteroid. The other child, acting as mission control, guides the astronaut from an orbital view that shows the spherical asteroid and the astronaut's position on its surface. VR helps situate the astronaut on the surface of the asteroid where he or she can experience circling the globe and coming back to the same place, not falling off of the 'bottom' of the asteroid, and seeing objects appear over the horizon top first. VR gives mission control an obviously spherical world to monitor. The two children share the same

virtual environment, but see it in different ways. The children experience both roles, and must integrate these two different views to complete their mission – searching for enough fuel cells on the surface to allow their stranded spaceship to leave. Through integrating these two views they learn to map from the 'flat' surface of the sphere to its true spherical shape.

We typically run this application using heterogeneous VR devices, emphasizing the heterogeneous roles. In the lab, the astronaut explores the asteroid in the CAVE while mission control monitors from an ImmersaDesk. In a local elementary school where we have set up an ImmersaDesk, the astronaut uses the ImmersaDesk, while mission control uses a 19" stereo monitor.

X.4 Lessons Learned

We have learned many lessons from building, using, and watching others use these virtual worlds. Several of the more interesting lessons are described below.

In all these applications a single hand (wand) and head were tracked hence each avatar possessed a movable head, one movable hand, and a body. For the most part this was sufficient, as it allows a user to see where another user is standing, looking or pointing with their tracked hand. The tracked head also conveys a certain amount of body language, from nodding yes or no to a confused tilt. These motions are also recognizable. We frequently record our tele-immersive sessions on video for later review. One day while reviewing one of these tapes without the audio, one of the students in the lab walked by the monitor, looked at the avatar gesturing on the screen, and said "Oh, that's Andy." While the body was different, the body language was the same.

However having only one hand made it difficult to express concepts of size such as "it is this big" which usually requires the use of two hands. The correspondence between the wand and the hand was not always clear to new users. Hence when users waved at each other or pointed, they often used the non-tracked hand, forgetting that the position and orientation of their other hand weren't been seen by the remote users. We are currently moving towards a standard set of four trackers to allow both hands to be tracked. Ideally each of the hand trackers should be worn on the back of the palm as an obvious affordance to their function.

The avatars are also useful in alerting other users about a person's next actions. For example, the declaration: "I'm going to move this chair" combined with the visual cue of your avatar standing next to a chair and pointing at it, alerts other users that you about to grab that chair. As in real life we know that if we also try to grab that chair that it will be awkward. We have tended to rely on these social cues rather than a computer-based locking scheme to discourage users from trying to jointly modify a shared object.

Different avatar forms are useful in different virtual worlds. In certain situations, such as pointing at an object on the ImmersaDesk, pointers can be better than full avatar bodies, because a remote user's avatar body may block your view, if you are standing close together – like in real life. For small objects, short pointers work well, but in larger spaces it helps to have a long beam allowing the users to point accurately at objects a hundred feet away. When articulated avatar bodies represent the users, each user can see what the other participants look like and often want to know what they themselves look like. In NICE we solved this problem by allowing the children to look into the sea to see their reflection.

Avatars with highly stylized bodies are much easier to differentiate within the environment but may not be appropriate for all types of users. First time users tend to laugh the first time they meet of the articulated avatars, as they are greeted by what appears to be a living cartoon character, but they quickly adapt once this character begins interacting with them, and have no trouble treating that character as a living person. Other users desire a more 'serious' representation of themselves in the virtual world. Using photographs to generate avatar heads that look like the actual user is a way to help bridge between recognition in the virtual world and the real world. However as in educational environments such as NICE, having avatars that do not look like the actual user may be beneficial. The avatars can help to 'equalize' people in the virtual world. In one session with NICE, Eddie who was 6 years old was collaborating with Jim, a 6' tall graduate student. Within the virtual world they were roughly the same size and had identical powers. When the virtual session was over, Jim came over to Eddie and introduced himself, again, this time in person. Eddie craned his neck to look up at Jim, and it was clear that their relationship changed at that point.

For international collaborations, when English has been used as the default language, foreign speakers whose first language is not be English, may find it difficult to converse naturally, and hence these participants tend to be less vocal. While the tracker is able to transmit gross gestures, it is harder to spot more subtle gestures. What is normally considered a clear nod in the real world usually amounts to a suggestion of a nod in the virtual world. Cultural differences also impact the degree with which a participant gestures. For example Americans tend to gesture considerably while speaking, whereas the Japanese tend to gesture very little. In these situations it may be useful to include video to help mediate discussions so that the faces of the participants can be clearly seen.

Audio is the most important communication channel to maintain the shared collaborative space. If the audio deteriorates then the collaboration quickly falls apart. When multiple people are in the space it can often be difficult to tell who is speaking if the audio does not come from a particular point in space. Adding mouth movements, even crude lip synching to the avatars is useful to give cues about who is speaking. We previously used telephone

schedule synchronous meetings, so asynchronous work may be the most appropriate mode for trans-oceanic tele-immersion. Asynchronous work also has the advantage that geographically distributed teams can work on the same problem around the clock, by passing the work off at the end of the day to another team who is just arriving at work in their morning. Lessons can be learned from the sharing of more traditional text and image information on remote computers through such tools as Lotus Notes, Microsoft Net-meeting, and MUDs / MOOs [Schiano and White, 1998; Bruckman and Resnick, 1995], and from the sharing graphical, audio and video data through media spaces [Bly et al, 1993].

Email is a very successful tool supporting asynchronous work. However, in international collaborations there is typically a 1-day turnaround time to get responses, so collaborators can easily waste days clarifying the work to be done and making instructions clear. When working in a virtual environment this is even more difficult as it is hard to use text, speech, or even 2D images alone to describe work to be done, or discoveries that have been made in a dynamic 3D environment. It is important that the messages between the distributed team members be clear, to reduce misunderstandings. In a virtual environment it is important to be able to put these messages into their appropriate context – the context of the virtual world itself.

One of the advantages of doing design or scientific visualization in an immersive environment is the ability to have geographically distributed participants sharing space with each other and the objects under discussion. This allows the participants to point at specific objects in the scene or set the parameters of the simulation to specific values to clarify what they are saying. It gives the users a common context for their discussions. Especially in international collaborations where the language barrier can be a large hurdle, being able to gesture relative to the environment (pointing at the red box, turning you head to look at the green sphere) helps to clarify the discussion. The ability to hand off work quickly and accurately is also of great importance. A user stepping into the virtual environment needs to know what work has been done since he / she was last there, and what new work may need to be done. We are currently investigating several approaches: v-mail, virtual post-it notes, and the VR-VCR.

When asynchronously working in a virtual space, email and telephone conversations are not enough to describe changes that need to be made or observations that need to be verified because the email and the telephone conversation are not in the context of the virtual space. V-Mail allows a user to record themselves as an avatar talking and gesturing within the virtual space. Thus the user in the recorded V-Mail appears just as they would if they were in a synchronous collaborative session. In fact, during our initial studies of the system we found that the user retrieving the V-Mail would often talk back to the pre-recorded avatar [Imai et al, 1999] as though they were conversing. These recording preserve the important head and hand gestures, should help clarify the process that went into creating the artifacts in the world, and help support orientation in the space.

Like V-Mail, Virtual Post-it Notes will help explain changes that were made or need to be made in a collaborative design environment, specify where interesting phenomena were discovered in scientific visualization environments, and give researchers an easy way of creating meta-data recordings from within the virtual space. Most computational scientists agree that a crucial part of the knowledge crystallization process [Card et al, 1999] includes the creation of snapshots and annotations to track the progress of the exploration and to record discoveries [Springmeyer et al, 1996]. In desktop environments these annotations (meta-data – data about the data) are typically entered in text windows, however this common mode of data entry is problematic as well as limiting in an immersive environment. Current VR displays lack the resolution to display text clearly in a virtual window. Recording and replaying audio messages has been used to try and circumvent these problems. Adding in the avatar of the person recording the message complete with their gestures allows those audio messages to be put in the proper context. We have also used the Post-it Notes to create pre-recorded tour guide avatars for visitors to our virtual worlds such as the Silk Road Shrines. If there is no live expert tour guide available, the world can still introduce itself and take a user on a tour, giving information at various points of interest.

The VR-VCR allows users to record the entire VR experience for playback and analysis. Since all of the state changes to the collaborative virtual world come through the application's networking layer, we can store the time-stamped sequence of changes and play back the virtual experience as an immersive movie and experience it from with the virtual environment, watching the action unfold around us. This allows a person see everything that they would have seen if they were present in the immersive space at the time the recording was made. With a video recording of the collaboration, the user is limited to seeing the action from the position of the camera, here the user can walk (fly) through the space while the collaboration is underway, watching the action from whatever location

seems most interesting. These recordings could also be shared tele-immersively during playback, allowing geographically distributed participants to collaboratively watch the recording of a previous collaborative session. We first recorded these VR movies while several groups of children were interacting with the NICE environment. During their sessions, we could watch their interactions with the VR hardware, later we could replay their interactions within the virtual world. Currently we have been using simple VCR-like controls for playing back this data, but more sophisticated querying of this data is also possible dealing with spatial relations between the users and the world, temporal relations between actions, and phonetically searching the collaborator's dialogue.

X.5.2 Heterogeneous Representations

Environments such as CALVIN, CAVE6D, and the Round Earth allow and in fact <u>emphasize</u> giving each user a different view of the shared space. These heterogeneous views allow us to leverage the capabilities of a shared virtual space to allow each user to customize their view to their needs.

Collaboration can improve the overall efficiency of the data analysis process. One aspect that we are particularly interested in developing is the concept of Multiple Collaborative Representations (MCR.) Individuals who are trying

TIDE's current focus is on visualizing massive data-sets from the Department of Energy and weather data from the National Oceanic and Atmospheric Administration. TIDE's main research foci are: to develop techniques for managing and mining massive data-sets for visualization and navigation; to develop collaboration and network-aware visualization tools which can adapt to changing network conditions; to develop tools for creating three-dimensional annotations and recordings of discoveries during collaborative sessions; and to perform user studies of collaborative data exploration to evaluate the effectiveness of these tools

We are currently in the process of performing user-studies to observe how users take advantage of multiple collaborative representations, and in what situations they are effective or a hindrance. These studies should also allow us to form ideas on the kinds of tools that will be needed to help participants coordinate the representations. What we have seen so far re-affirms that previous CSCW findings are applicable to tele-immersion:

- There is a need for individual pointers to allow collaborators to point at shared data items. However these pointers can become a source of distraction and users should have the ability to toggle them on/off [Lauwers and Lantz, 1990; Tang, 1991; Stefik et al, 1987].

- It is useful to have some cue of which region of space a user is manipulating [Patterson et al, 1990; Roseman and Greenberg, 1996].

- Even in a fully shared environment, participants found the need to work with localized views [Roseman and Greenberg, 1996].

- There is a frequent transition between parallel / independent and coordinated activities [Stefik et al, 1987].

- In a fully shared WYSIWIS system, frequent usage collisions will occur [Stefik et al, 1987].

- The user-interface should be considered part of the visualization so that collaborators can gain greater awareness of their collective actions as they manipulate the visualization [Gutwin and Greenberg, 1998].

In fact our work most mirrors the work of Gutwin and Greenberg [Gutwin and Greenberg, 1998] except that we are exploring these issues in the context of a fully immersive environment in which co-presence as well as computer-mediated and augmented interaction is simultaneously possible.

X.5.3 New Devices

Even with mirrors, a CAVE takes up a 30' by 20' by 15' space and requires a couple days to setup and align. The ImmersaDesks were designed for laboratories, and take up a large amount of space in a typical office. The ImmersaDesk screen is large to present a wide field of view, but the desk structure itself is large because the rear-projection distances are significant and the desk uses a large and heavy projector. EVL's current work is focusing on desktop systems that are more suited to an office environment. Fish-tank VR using CRT monitor technology has been used for several years, but these systems typically have very small fields of view. Increasing the field of view without filling the office with equipment means an increased reliance on flat-panel display technology. Instead of leaving your office to walk over to a CAVE or an ImmersaDesk to enter a tele-immersive space, the goal is that the surface of your desk becomes a display device, the walls become display devices, perhaps even the floor and the ceiling. In effect, your office becomes a CAVE on demand.

The CAVE, ImmersaDesk, and ImmersaDesk2 are commercial products. Several companies offer well-designed, non-tracked displays for the office and showroom. Others have products similar to the CAVE and ImmersaDesk. The goal of EVL's research is not to compete with the commercial sector, but to investigate and inspire new display and tracker technologies for tele-immersion. Given that affordable bright wall-sized high-resolution border-less displays with high refresh rates and fast decay rates do not yet exist, we are prototyping these systems using available components: existing flat panel technologies, or simulating flat panel systems with rear-projection hardware. These more office-friendly systems may change the way that people interact with virtual environments, and perhaps more importantly may change the way people think about interacting with virtual environments.

X.6 Summary

Our goal isn't simply to make tele-immersive work possible, but to make it convenient: convenient access to the VR hardware, convenient access to the virtual world, convenient synchronous or asynchronous access to collaborators, and tools which make it convenient to share your data tele-immersively. We believe the best way to reach this goal is to build tele-immersive environments for a variety of domains and evaluate their effectiveness, thinking not in terms of how to recreate reality, but how to leverage virtual reality to make the collaboration better than being there in person.

Acknowledgements

References

Ainsworth, S., Wood, D., Bibby, P. Evaluating principles for multi-representational learning environments. 7th EARLI conference, Athens, 1997.

Bibby, P., Payne, S. Internalization and the use specificity of device knowledge. Human-Computer Interaction, 1993; 8(1): 25-56.

Bly, S., Harrison, S., Irwin, S. Media spaces: Bringing people together in a video, audio, and computing environment. Communications of the ACM, Jan. 1993, pp. 28-47.

Bruckman, A., Resnick, M. "The MediaMOO project: constructionism and professional community," Convergence, 1995; 1(1): 94-109.

Card, S., Mackinlay, J., Schneiderman, B. Readings in information visualization, using vision to think, Morgan Kaufmann, San Francisco, California, 1999, pp. 10-12.

Cruz-Neira C., Sandin D., DeFanti T. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In Proceedings of ACM SIGGRAPH '93, New York: ACM Press, 1993, pp. 135-142.

Czernuszenko, M., Pape, D., Sandin, D., DeFanti, T., Dawe, G., Brown, M. The ImmersaDesk and Infinity Wall projection-based virtual reality displays. Computer Graphics, 1997; 31(2): 46-49.

Gutwin, C., Greenberg, S., Design for Individuals, Design for Groups: Tradeoffs Between Power and Workspace Awareness, Proceedings of the ACM CSCW'98 Conference on Computer Supported Cooperative Work, 1998, pp. 207-216.

Imai, T., Johnson, A., Leigh, J., Pape, D., DeFanti, T. Supporting transoceanic collaborations in virtual environment, In the proceedings of the 5th Asia-Pacific Conference on Communications / 4th Optoelectronics and Communications Conference - APCC/OECC '99, Beijing, China, October 19-21, 1999, pp. 1059-1062.

Johnson, A., Moher, T., Ohlsson, S., Gillingham, M. The round Earth project: collaborative VR for conceptual learning. In IEEE Computer Graphics and Applications, 1999; 19(6): 60-69.

Kessler, G., Hodges, L., Ahamad, M. Ravel, a support system for the development of distributed multi-user VE applications, in Proceedings of the Virtual Reality Annual International Symposium (VRAIS) 98, 1988, pp. 260-267.

Larkin, J., Simon, H. Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 1987; 11: 65-99.

Lauwers, J., Lantz, K. Collaboration awareness in support of collaboration transparency: requirements for the next generation of shared window systems. In Proceedings of the ACM CHI'90 Conference on Human Factors in Computing Systems. ACM, 1990.

Leigh, J., Johnson, A., Vasilakis, C., DeFanti, T. Multi-perspective collaborative design in persistent networked virtual environments. Proceedings of IEEE Virtual Reality Annual International Symposium '96. Santa Clara, California, Mar. 20 - Apr. 3, 1996, pp. 253- 260, 271-272.

Mason, T., Applebaum, E., Rasmussen, M., Millman, A., Evenhouse, R., Panko, W. The virtual temporal bone, in Proceedings of the Medicine Meets Virtual Reality Conference, 1998.

Park, K., Kenyon, R. Effects of network characteristics on human performance in a collaborative virtual environment, in the Proceedings of IEEE VR `99, Houston TX, March 13-17, 1999.

Patterson, J., Hill, R., Rohall, S., Meeks, W. Rendezvous: an architecture for synchronous multi-user applications. In A. Press, editor, Proceedings of the ACM CSCW'90 Conference on Computer Supported Cooperative Work, 1990, pp. 317–328.

Roseman, M., Greenberg, S. Building real-time group-ware with groupkit, a groupware toolkit. ACM Transactions on Computer-Human Interaction, 1996; 3(1): 66–106.

Roussos, M., Johnson, A., Moher, T., Leigh, J., Vasilakis, C., Barnes, C. Learning and building together in an Immersive virtual world. In Presence 1999; 8(3): 247-263.

Salzman, M., Dede, C. Loftin, B., Ash, K. VR's frames of reference: A visualization technique for mastering abstract information spaces. In Proceedings of the Third International Conference on Learning Sciences, Charlottesville, VA: Association for the Advancement of Computers in Education., 1998, pp. 249-255.

Schiano, D., White, S. The first noble truth of CyberSpace: people are people (even when they MOO) in the proceedings of CHI '98. April 18-23, 1998, pp. 352-359.

Smith, G., Mariani, J., Using subjective views to enhance 3D applications. In ACM Symposium on Virtual Reality Software and Technology, Lausanne, Switzerland, 1997. ACM, ACM Press. Pp. 139-146.

Springmeyer, R., Werner, N., Long, J. Mining scientific data archives through metadata generation, in the proceedings of the First IEEE Metadata Conference, April 16-18, 1996, NOAA Auditorium, Silver Spring, Maryland.

Stefik, M., Foster, G., Bobrow, D., Kahn, K., Lanning, S., Suchman, L. Beyond the chalkboard: computer support for collaboration and problem solving in meetings. Communications of the ACM, 1987; 30(1): 32–47.

Tang, J. Findings from observational studies of collaborative work. International Journal of Man-Machine Studies, 1991; 34(2): 143–160.

Wheless, G., Lascara, C., Valle-Levinson, A., et al. Virtual chesapeake bay: Interacting with a coupled physical / biological model. IEEE Computer Graphics and Applications, 1996; 16(4): 52–57.