A Review of Tele-Immersive Applications in the CAVE Research Network

Jason Leigh[1,2] (spiff@evl.uic.edu), Andrew E. Johnson[2], Thomas A. DeFanti[2], Maxine Brown[2]
Mohammed Dastagir Ali[2], Stuart Bailey[8], Andy Banerjee[9], Pat Banerjee[9], Jim Chen[5], Kevin Curry[11],
Jim Curtis[12],Fred Dech[9], Brian Dodds[12], Ian Foster[3], Sarah Fraser[5], Kartik Ganeshan[8], Dennis Glen[5], Robert Grossman[8],
Randy Heiland[1], John Hicks[6], Alan D. Hudson[12], Tomoko Imai[2], Mohammed Ali Khan[2], Abhinav Kapoor[2], Robert V. Kenyon[2],
John Kelso[11], Ron Kriz[11], Cathy Lascara[4], Xiaoyan Liu[7], Yalu Lin[2], Theodore Mason[10], Alan Millman[10], Kukimoto Nobuyuki[13],
Kyoung Park[2], Bill Parod[5], Paul J. Rajlich[1], Mary Rasmussen[10], Maggie Rawlings[2], Daniel H. Robertson[6],
Samroeng Thongrong[2], Robert J. Stein[1], Kent Swartz[11], Steve Tuecke[3], Harlan Wallach[5],
Hong Yee Wong[7], Glen H. Wheless[4]

1. National Center for Supercomputing Applications, Univ. Illinois at Urbana-Champagne (NCSA);
2. Electronic Visualization Laboratory, Univ. Illinois at Chicago (EVL);
3. Argonne National Laboratory (ANL); 4. Center for Coastal Physical Oceanography, Old Dominion University (CCPO);
5. Northwestern University (NWU); 6. Indiana University Purdue University Indianapolis (IUPUI);
7. Institute of High Performance Computing, National University of Singapore (IHPC);
8. National Center for Data-Mining, Univ. Illinois at Chicago (NCDM);
9. Industrial Virtual Reality Laboratory, Univ. Illinois at Chicago (IVRL);
10. Virtual Reality in Medicine Laboratory, Univ. Illinois at Chicago (VRML);
11. Virginia Tech (VT); 12. Searle; 13. Tohwa University

Abstract

This paper presents an overview of the Tele-Immersion applications that have been built by collaborators around the world using the CAVERNsoft toolkit, and the lessons learned from building these applications. In particular the lessons learned are presented as a set of rules-of-thumb for developing tele-immersive applications in general.

1. Introduction

Tele-Immersion (TI) is defined as the integration of audio and video conferencing, via image-based modeling, with collaborative virtual reality (CVR) in the context of data-mining and significant computation. The ultimate goal of TI is not merely to reproduce a real 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 large databases.

When participants are tele-immersed, they are able to see and interact with each other and objects in a shared virtual environment. Their presence will be depicted by life-like representations of themselves (avatars) that are generated by real-time, image capture, and modeling techniques. The environment will persist even when all the participants have left it. The environment may autonomously control supercomputing computations, query databases and gather the results for visualization when the participants return. Participants may even leave messages for their colleagues who can then replay them as a full audio, video and gestural stream.

In 1992 there was a single CAVE in Chicago, in 1998 there are over 80 CAVE and ImmersaDesk installations around the world. CAVERN (the CAVE Research Network) is a community of participating industrial and research institutions equipped with CAVEs, ImmersaDesks, and high-performance computing resources all interconnected by high-speed networks for the purpose of supporting tele-immersive- engineering and design; education and training; and scientific visualization and computational steering.

This paper will begin by providing an overview of CAVERNsoft, our architecture for achieving the technical goals of TI. This will be followed by a brief description of some of the most recently developed applications of CAVERNsoft. The purpose is to illustrate the breadth of tele-immersion activity that is currently being engaged by domain scientists, engineers and artists around the world. From this experience are distilled rules-of-thumb for developing tele-immersive applications.

2. CAVERNsoft: An Approach to Tele-Immersion

The CAVERN research agenda consists of three foci: 1. to conduct research and development of algorithms and tools for supporting Tele-Immersion; 2. to conduct evaluative studies of human-factors issues and network performance issues in TI; and 3. to deploy and assist domain specialists in the development and use of tele-immersive applications. Each focus will naturally contribute results that will feedback into the other foci.

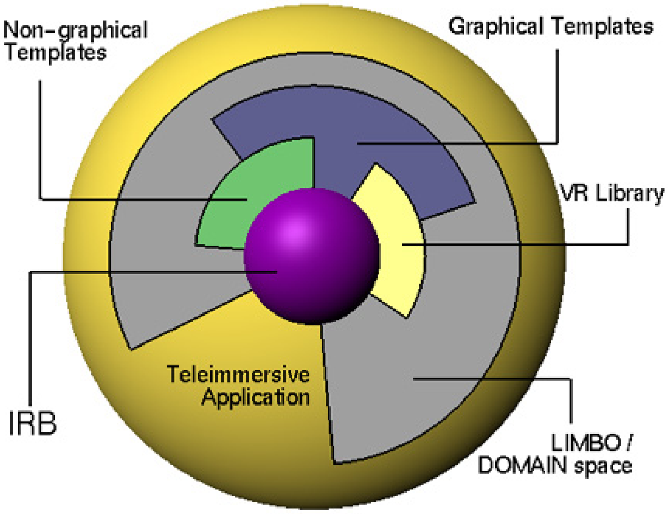

Figure 1. CAVERNsoft

Within the limitations of this paper we will provide an overview of two aspects of the CAVERNsoft effort; the tools, and the applications being developed.

CAVERNsoft[4, 5] is an ongoing effort to develop a TI infrastructure that supports the rapid creation of new tele-immersive applications, and the retrofitting of previously non-collaborative VR applications with tele-immersive capabilities. The primary difference between CAVERNsoft and previous approaches is its focus on integrating collaborative VR with supercomputers and terrabyte 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- Science, Technology And Research Transit Access Point) networks[1], TransPAC[9]. This affords us the ability to explore new research problems and applications of this collaborative technology without being hindered by the limits of the existing Internet.

The CAVERNsoft architecture diagrammed in Figure 1, consists of a central structure called the Information Resource Broker (IRB) surrounded by layers of supporting software. The inner layers facilitate the construction of new components, as well as the retrofitting of existing applications. The outer layers facilitate the rapid development of new tele-immersive applications.

The IRB (inspired by CORBA ORBs but with lower overhead) is a relatively low-level merging of networking and database capabilities that is completely separated from graphics- so that it can be placed into any software application regardless of its ability to render graphics. This permits graphical applications to communicate with non-graphical applications and it also allows existing non-collaborative applications to achieve networking capabilities with minimal disturbance to their existing graphics sub-system. The IRB is essentially a merging of a distributed shared memory (DSM) with a message passing system (MPS) and a persistent data store. However unlike previous architectures that combine DSMs with MPS'[11], CAVERNsoft allows the selection of custom networking protocols (TCP, UDP, Unreliable Multicast,) and soon quality of service, over which data is shared. These networking capabilities are supported by Globus[2]- a software suite that simplifies and optimizes security, scheduling and data distribution amongst heterogeneous distributed computing resources.

The IRB is client/server symmetric- this means that all client programs are automatically server programs without the need for any additional programming effort. This symmetry allows a collection of clients to easily form arbitrary collaborative topologies.

At a layer above the IRB reside additional layers of software that are still much under development. The development is being steered primarily by the needs of the CAVERNsoft users. The non-graphical template libraries support the coordination of avatars, as well as audio and video data compression algorithms. On top of this is a higher-level layer that consists of graphical versions of the previous layer, like OpenGL, Performer, and Video avatar templates. Work is actively under way now to develop image capture and modeling techniques to produce more life-like avatars.

From almost a decade of experience with technology transfer we have learned that it is not enough to simply provide a suite of tools to our domain users. We have found that a higher level organization of these tools, in the form of an application, must first be provided to allow them to conceptualize how these tools might be used in a domain-specific scenario. LIMBO was developed for that purpose.

A LIMBO space (Figure 2) provides a canonical set of capabilities needed for Tele-Immersion- it supports audio conferencing, avatar rendering, model importing, distribution, and manipulation. By using a basic LIMBO space, collaborators can work in a virtual space without programming. The client/server symmetry of CAVERNsoft's IRB allows individual LIMBO clients to act as servers making it easy to set up multi-way connections. And, if desired a dedicated server may be launched to act as a permanent entry point into a persistent environment hence supporting asynchronous collaboration.

LIMBO users are able to quickly import 3D models (e.g., car designs, scientific data-sets) into the space and LIMBO will ensure the proper distribution of the model to all remote participants. Once objects are distributed, participants may collectively modify them. Finally, application developers may use the available source code of LIMBO to jump-start the development of their own domain-specific tele-immersive applications. As more domain-specific applications are developed, a growing library of CAVERNsoft-based reusable components will emerge. These can be added to the library of existing templates and collected to build DOMAIN spaces that are specializations of LIMBO spaces. For example, this will enable a designer to build a Tele-Immersive design application starting with an existing DOMAIN space equipped with collaborative tools specifically for collaborative design, rather than starting from the basic LIMBO space.

3. CAVERNsoft Applications

In a year since CAVERNsoft was first introduced, approximately a dozen CAVERNsoft applications have been built. These applications span the full spectrum of domains: scientific visualization, data-mining, industrial design and manufacturing, grade school education, training, history and art. The page limits of this paper will not permit detailed explanations of each of all these applications. Instead the following is a brief summary of each of the projects. Longer descriptions may be found at the CAVERN Web page[6].

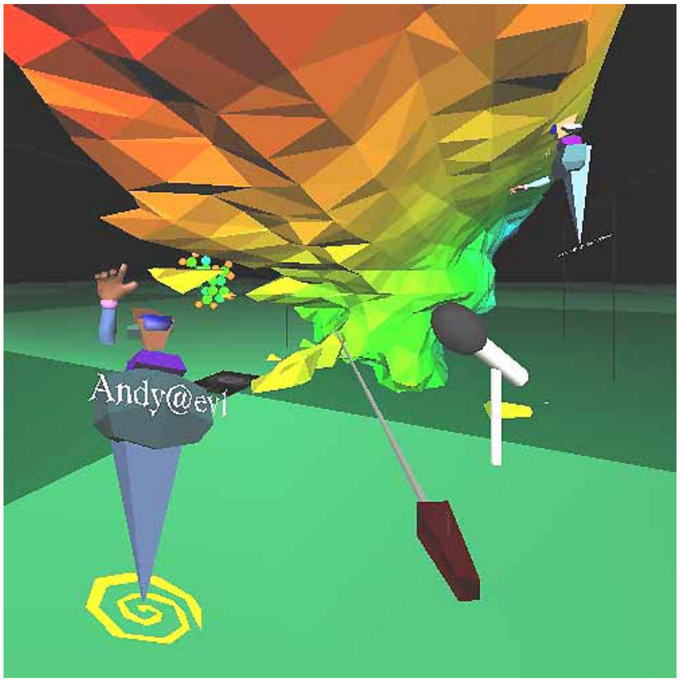

Figure 2. Participants engaged in a LIMBO session

Figure 3. General Motors- VisualEyes. Shown is the inside of a virtual CAVE. By selecting this CAVE the participant will be teleported into the environment where the interior of a GM car is being designed.

3.1. Retro-fitted Applications

These consist of applications that have already been developed but that have been augmented with collaborative capabilities using CAVERNsoft.

Tele-Immersive VisualEyes
Randall C. Smith, David A. Brown- General Motors Research and Development Center; Jason Leigh- NCSA/EVL

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. This initial use of CAVE-based technology has generated considerable interest in other GM sites around the world, some of whom are planning their own CAVE installations. This has prompted GM to further extend VisualEyes to allow GM's 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.

VE was retro-fitted with tele-immersive capabilities by incorporating one of CAVERNsoft's avatar templates. Then the scene graphs of each participating VE client were linked using CAVERNsoft's shared memory services allowing the change of one part of a design to be automatically shared amongst all the other participants. To facilitate both synchronous and asynchronous work, CAVERNsoft's database services were used to save the state of the environment before all the participants left. As CAVERNsoft is client/server symmetric, VE clients were able to connect to each other or to a central broadcasting server (which itself was a CAVERNsoft program.)

The Virtual Temporal Bone
Alan Millman, Mary Rasmussen, Theodore Mason- VRML; Mohammed Dastagir Ali - EVL

The Virtual Temporal Bone is a tele-immersive education program to allow 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 temporal bone to more clearly view the inner anatomy. Audio from the voice conference is used to modify the flapping of the ear drum to illustrate its function.

This application originally used the CAVE library's limited set of networking capabilities (either point-to-point UDP or multicast only) to coordinate the collaboration. To replace this networking with CAVERNsoft, the CAVE2CAVERNsoft library was developed that allowed a one-to-one replacement of the CAVE's networking API with those that use CAVERNsoft. The CAVE2CAVERNsoft library was developed so that existing CAVE applications that use the CAVE's networking API can easily transition to CAVERNsoft without significant re-implementation. This transition will allow an application to further extend its capabilities by leveraging the additional networking and database facilities provided by CAVERNsoft. Furthermore CAVE2CAVERNSoft can itself be used independently as a library to develop tele-immersive applications.

CAVE6D- a Tool for Tele-Immersive Visualization of Environmental Data
Glen H. Wheless, Cathy Lascara- CCPO; Abhinav Kapoor, Jason Leigh, NCSA/EVL

CAVE5D[18], co-developed by Wheless and Lascara from

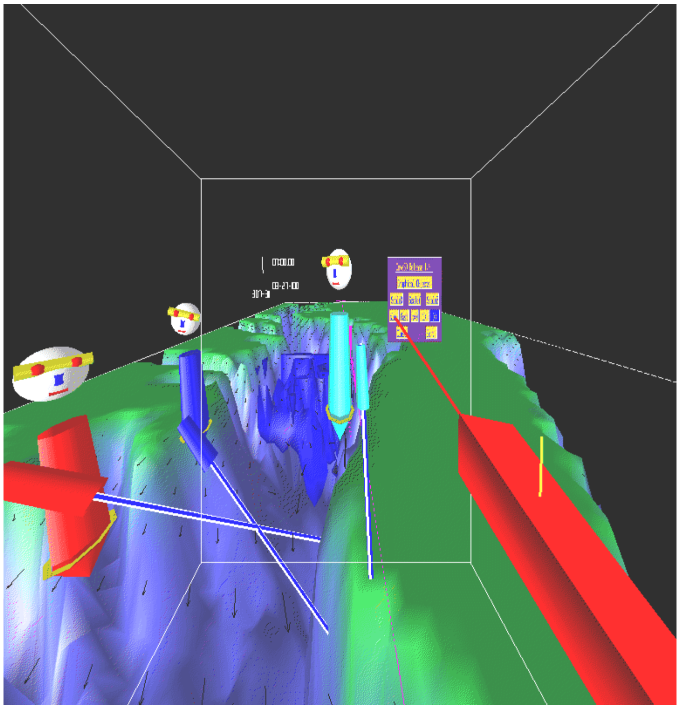

Figure 4. A number of participants in CAVE6D.

ODU and Dr. Bill Hibbard from the University of Wisconsin-Madison, is a configurable virtual reality (VR) application framework. CAVE5D is supported by Vis5D[3], a very powerful graphics library that provides visualization techniques to display multi-dimensional numerical data from atmospheric, oceanographic, and other similar models, including isosurfaces, contour slices, volume visualization, wind/trajectory vectors, and various image projection formats.

CAVE6D, emerges as an integration of CAVE5D with CAVERNsoft to produce a TI 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 over a telephonic conference call. Visualization parameters such as, Salinity, Circulation Vectors, Temperature and Wind Velocity Slices, larval fish distributions etc. that can be visualized by CAVE5D, have been extended in CAVE6D to allow participants to collectively operate them (Figure 4.) Participants may turn any of these parameters on or off, globally (affecting everyone's view) or locally (affecting only one's own view.) This affords each participant the ability to individually customize and/or reduce the "cluttered-ness" of the visualization; one participant may be viewing the relationship between one subset of the total dimensions of the data while another participant may be simultaneously correlating a different subset. Time on the other-hand is globally shared and hence participants view all time-varying data synchronously. This constraint was enforced so that even though the participants may be viewing decidedly disparate dimensions of the data they are all bound by at least one common dimension.

3.2. New CAVERNsoft Applications

V-Mail

Tomoko Imai- EVL

Traditionally E-mail has been the most effective asynchronous collaboration tool available to us. The goal of V-mail is to provide an equally, if not more, effective asynchronous collaboration tool for TI. In particular V-mail was designed to facilitate collaborations between participants that may be geographically located at distant parts of the world where timezone differences prevent routine synchronous collaboration.

When recording a V-mail message both audio and gesture from the collaborator are captured in addition to the surrounding scenery. When the message is played back, the avatar of the original sender of the message materializes to re-enact the message. After viewing the message the viewer may reply as in a standard e-mail system.

By recording speech, gesture and the surrounding scenery, collaborators are able to create richer messages than those traditionally possible via standard e-mail or voice mail. Participants are able to see the sender's avatar naturally gesturing while describing an observation in the virtual environment. This is particularly effective when the concept being described has an intrinsicly spatial quality.

Early implementation of a prototype of this system has led us to expand the application of this technology to also include the use of this recording system for guided tours in TI. In this situation pre-recorded sequences may be scripted and played at appropriate moments that are triggered by user responses.

Motorola Drop Test Visualization

Hong Yee Wong, Xiaoyan Liu- IHPC; SiewEng Quah- Motorola, Singapore; Jason Leigh- NCSA/EVL

One of the areas of research that Motorola actively engages in is the drop testing of their products- in this case one of their pagers. Physical drop tests typically produce few useful results as they only identify whether the product did or did not survive a fall. There is no easy way to observe the fall at the moment of impact. In particular there is no easy way to observe the effects of such a crash from within the device to identify key weaknesses in the product design.

However, by simulating the data on supercomputers, Motorola is then able to visualize the effects of the impact from within the CAVE. The immersive capabilities of the CAVE affords a more natural interface to view the data. In this particular application a cutting plane can be pushed through the pager revealing the inner circuitry and printed-circuit board (PCB) housing. As the pager is rendered larger than life in the CAVE, the observer can physically walk inside the pager to observe how the PCB dislodges from its housing at the moment of impact. From this the designer can devise supports to further strengthen the housing. Furthermore by adding tele-immersive capabilities to this visualization tool, Motorola engineers can discuss their findings with their collaborators around the world.

3.3. LIMBO-based Applications

LIMBO/VTK

Jason Leigh- NCSA/EVL; Paul J. Rajlich, Robert J. Stein- NCSA

LIMBO/VTK combines LIMBO with the Visualization Toolkit[14] to allow application developers to use the rich set of visualization tools built into VTK to generate sharable three-dimensional objects in LIMBO. VTK was integrated into the CAVE environment and consequently into LIMBO using the **vtkActorToPF** library[12].

The provided LIMBO/VTK distribution consists of a tele-immersive environment to load and generate isosurface visualizations of volume data. VTK is used to load the volume data and generate the polygons of the isosurface. A user may instantiate a persistent copy of the isosurface at any time by pressing one of the CAVE-wand's buttons. This persistent copy is distributed amongst all the remotely connected participants who are then able to pick and move the objects for inspection. Other participants may also join in changing the isosurface threshold to produce more persistent instantiations (Figure 2).

The Virtual Whiteboard

Kukimoto Nobuyuki- Tohwa University

The Virtual Whiteboard extends LIMBO's capabilities by providing a shared writing slate within the collaborative environment. Rather than having the slate permanently fixed at one position, as in most implementations of virtual whiteboards, this slate is carried along with the participant as he/she traverses the virtual space. When the slate is activated by a participant the slate is drawn as part of that participant's avatar to indicate to others that the participant has his/her slate in sight. This slate can then be used to sketch diagrams or write kanji using the CAVE's wand.

Data-Mining Visualization

Stuart Bailey, Kartik Ganeshan, Robert Grossman- NCDM; Jason Leigh- NCSA/EVL,UIC

The Health Care field is data rich, but information poor. Data exists from admissions, laboratory, radiology, and pharmacy systems, as well as from third-party systems, which consolidate data from a variety of sources. LIMBO is used as a visualization tool for viewing the outcome of a class of data mining classifiers called decision trees. The visualization is similar to the tree-viewer in Mineset (an SGI product) except that multiple participants can load and view together. Future improvements would include the ability to activate the classifiers from within the virtual environment.

Although this was originally intended as a tool for understanding the decisions made by data-mining algorithms, it has currently found a greater use in debugging the algorithms. The added dimension of stereoscopic 3D graphics allows one to visualize more data than is normally possible in a tree drawn on a flat display. The collaborative capabilities allows one to discuss these trees with other researchers.

Virtual Reality Tele-collaborative Integrated Manufacturing Environment (VRTIME)

Alan D. Hudson, Brian Dodds, Jim Curtis- Searle; Fred Dech, Pat Banerjee, Andy Banerjee- IVRI

Searle, a Monsanto Company, is a research-based company headquartered in Skokie, Illinois (USA) that develops, manufactures and markets prescription pharmaceuticals and other healthcare solutions worldwide. VRTIME is a tele-immersive system targeted at industrial process design, operation, and control. Searle's operations in Britain, United States and Russia will be used for long distance TI case studies. In the industrial process planning phase, VRTIME will be used for facilities design, regulatory review, operational review and training. This is a particularly good application for TI because these reviews often involve geographically separated participants: architect, engineers, consultants, external and internal reviewers, federal, state and local planning officials, and environmental interest groups.

The tele-immersive nature of the training sessions means that they can be guided by remote teachers and undertaken before the new plant is operational, enabling a rapid start-up phase when the plant becomes operational. We are currently applying some of these techniques to a new plant being built in Moscow, Russia. The use of LIMBO in this training system has allowed us to jump start development by allowing us to quickly build a tele-immersive environment by importing existing three-dimensional models of the plant's production units.

Future Camp 1998

Daniel H. Robertson, John Hicks, Gordan Fricke- IUPUI

The IUPUI Future Camp[13] is a one-week multidisciplinary, virtual reality, day camp in its second year. The 18 campers, grades 9-11, attended camp from June 22 - 27 between the hours of 8:00AM - 3:00PM. The three projects for this years camp were: Virtual Art Gallery, Virtual Ocean Colonization, and Virtual Indianapolis Zoo. The Virtual Art Gallery and Virtual Ocean Colonization were accomplished on an ImmersaDesk using the CAVE libraries and LIMBO at the IUPUI Virtual Reality/Virtual Environments (VRVE) Lab. In both projects the 3D objects were created in Alias Wavefront and loaded into the LIMBO worlds. In LIMBO the participants collaboratively designed the layout of the artifacts. Once a satisfactory layout had been created, the coordinates of the artifacts were extracted to hard-code their positions permanently in LIMBO. LIMBO was then used as a browser for participants to collaboratively tour the space.

The CAVE Collaborative Console

Kevin Curry, Kent Swartz, John Kelso, Ronald Kriz- Virginia Tech

The CAVE Collaborative Console extends LIMBO by providing users with a graphical heads-up display of icons that enable the use of key tools needed for an efficient and effective collaborative session. This includes awareness tools such as a radar that shows where the participants are in a space; navigational tools such as a teleporter that allows one to transport to different points in a space; and annotative drawing tools for drawing primitive 2D as well as 3D objects.

The Silk Road Cave Shrines

Sarah Fraser, Bill Parod, Jim Chen, Dennis Glenn, Harlan Wallach- NWU; Samroeng Thongrong, Jason Leigh, Maggie Rawlings, Mohammed Ali Khan, Yalu Lin - EVL

This project involves a collaboration between historians, artists and computer scientists to create a virtual cultural and artistic exhibit of the Mogoa Grottoes of Dunhuang. Dunhuang, one of western China's ancient cultural sites, is considered the gate way to the well-known Silk Road- the East-West trade route between Asia and Europe. The Mogoa Grottoes are located in the Gobi desert. It consists of 492 caves with murals covering 25,000 square meters (approx. 269,097 square feet), 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.

Each cave and each mural has its own story.

The virtual exhibit of the Grottoes is now under its earliest phases of development. Three-dimensional models of the Grottoes are accurately being created using data from researchers at NWU and the Silk Shrine Foundation in China. These models are then imported into LIMBO as a rapid collaborative evaluation and brainstorming tool. Once completed, this piece will be the first collaborative, virtual exhibit of a historic site in the world.

4. Lessons Learned

Many of these applications have had a significant impact on their respective fields and others have been developed as feasibility studies. The reader is encouraged to refer to the CAVERN Web page[6] for detailed information about these applications. Below, we will present a compilation of the lessons learned from developing these applications that we believe can be generalized as rules of thumb for building tele-immersive environments. Many of these rules are based on our observations- and hence in some cases will require further proof by rigorous experiments. However it is presented here nevertheless as a way to stimulate ideas for future research.

4.1. Representation of Self: Avatars

1. In all our applications a single hand (wand) and head were tracked hence each avatar possessed a head, one hand and a body. The bodies orientation was locked to the front of the CAVE. For the most part this was sufficient however having only one hand made it difficult to express concepts of size as in "it is *this* big," which usually requires the use of two hands. Also the correspondence between the wand and the hand was not always clear to the users. Hence when users waved at each other they often used the non-tracked hand. Ideally each of the hand trackers should be worn on the back of the palm as an obvious affordance to their function. Finally, since no trackers are attached to the body the avatar's body was always locked to the front orientation of the CAVE. As the CAVE is navigated through a larger space by using a joystick, the avatar's body would also move through the space. This poses a problem in six wall CAVEs where the most natural means of navigation was to simply turn ones body in the direction one wishes to walk and then push a button to command the CAVE to move in that direction. However when a user faces the back wall of the CAVE the avatar's head will be oriented in the opposite direction of its body.

2. For international collaborations, when English is used as the default language, foreign speakers, whose first language may not be English, may find it difficult to converse naturally. Hence these participants tend to be less vocal. In these situations it would be useful to include video to help mediate discussions so that the faces of the participants can be clearly seen. Subtle facial cues can be used to determine if a listener understands what a speaker is talking about. Furthermore head tracking does not completely capture gestures such as the nodding or shaking of the head. What is normally considered a clear nod in the real world usually amounts to a very small suggestion of a nod in the virtual world. This is usually because the tracking systems are not always sampled frequently enough to reflect subtle gestures. Cultural differences may 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.

3. Use of video is cumbersome in TI as, in the case of the CAVE, stereoscopic shutter glasses are worn by all the participants hence hiding their eyes. For collaborators using HMDs, video may not be useful at all. Also the near-black conditions in the CAVE require extremely-low-light (or infra-red) cameras to be used.

4. Attaching a pointing ray to the end of each avatar is useful to accentuate what each avatar is pointing at. Whereas in the real world it is often difficult to refer a distant object to an observer this limitation can easily be overcome in the virtual world by extending a long pointing ray from each avatar's hand. This was found to be useful in the CAVE6D application, however it was noted that having the long pointing ray visible at all times was distracting and hence the rays should only be activated when it is being used specifically for pointing.

5. Adding mouth movements (even the crudest lip synching based on audio amplitude) to avatars is useful to give cues about who is speaking without the need to use spatialized sound.

6. Collision detection is useful to maintain social comfort, to prevent users from running through each other. However in tight regions where precise navigation is necessary, collision detection can be a hinderance[17].

4.2. Supporting the Collaborative Process

1. In most collaborative VR applications synchronous collaboration is given precedence over asynchronous collaboration. In some cases, as in international collaboration, the most convenient mode of long term collaboration is asynchronous.

Moreover, building an application that allows asynchronous collaboration can increase the opportunities for collaboration[7]. To support asynchronous collaboration one typically keeps a server constantly running as a persistent, known entry point into the environment. In addition annotation and recording tools are needed to allow participants to leave messages for one another and to track the changes that have occurred in the environment in their absence. Perhaps an agent acting as your surrogate can be left in the environment to track the things most important to you. The surrogate can send you a message over non-3D channels (email for example) when those changes have occurred, or when a participant has entered the environment (as in ICQ- a small tool for PCs that allows one to track the presence of their colleagues online.)

2. When multiple participants are working together in the same space, it may be more useful for each participant to have independently customizable views of the space. For example in CAVE6D one participant may view one set of visualization parameters while another participant watches over a different set. The implication here is that perhaps the viewing of multi-variate data can be partitioned across participants to augment the overall interpretation of what is visualized.

However, this capability should not necessarily be extended to the graphical interface. In our experience we have found that when expert users are working with novice users who may be struggling with the use of a menu, it is helpful for the expert user to be able to see the novice's interface to allow the expert user to offer assistance.

4.3. Human Communication

1. If audio communication fails then the collaboration degrades dramatically unless something else can mediate dialogue transfer (such as a chat session.)

2. Ambient microphones are useful to capture the remarks of an audience that may be standing in the CAVE at the same time. However, if only single users are collaborating it is better to use personal head-set microphones with earphones to eliminate audio feedback.

3. Telephone conferencing works well however it is not possible for participants to conduct private conversations. Sending audio digitally allows the tele-immersive environment to filter audio based on proximity.

4. When giving a tele-immersive demonstration, the demonstration is likely to be more successful if the *remote* participant explains the project rather than a local aide. When an avatar explains the project to a visitor, the visitor is more naturally drawn into the experience. It also demonstrates the effectiveness with which the medium can convey instruction over long distances.

4.4. Fundamental Networking Issues

1. It is well known that unreliable UDP is frequently used to deliver avatar tracker data as the loss of a single packet will quickly be followed by another[8]. However broadcast UDP is in practice preferable over multicast (even though most of the previous collaborative VR literature recommend multicasting.) For tele-immersive worlds involving hundreds of participants, multicasting can indeed reduce the number of messages that need to be sent by each client. However for smaller work groups, especially for highly mobile groups, broadcast UDP is more convenient to deploy. Multicasting usually requires the prior configuration of multicast tunnels to collaborating sites or to the M-Bone (requiring system administrator intervention.)

2. Reliable TCP is commonly used to guarantee that the state of the environment is consistent amongst all the collaborators. When sharing data use a database-centric approach rather than a command-centric approach. That is, share the state of the environment (e.g. an object is at x,y,z) rather than send commands to cause actions remotely (e.g. wand button 2 was pressed.) By distributing state information, new collaborators just joining in can naturally reconstruct their local environment from the states reported by other collaborators or servers. Employing a command-centric approach will require the new clients to independently ask for the current state information as well as interpret new incoming commands. Furthermore, by maintaining a server that holds state information one encourages the building of persistent applications.

3. High speed networks between continents are commonly characterized by high bandwidths and high latency (also referred to as Long Fat Networks (LFNs- pronounced elephants).) The theoretical capacity of a network is given by the bandwidth-delay product:

$$capacity = bandwidth \times roundTripTime$$

As an example, in the case of our US to Singapore (STAR-TAP/SingAREN) connection the network has a capacity of approximately: 12Mbps x 0.25s = 3Mbits.

When delivering data over reliable TCP connections the maximum speed per flow is limited by:

$$speed = TCPwindowSize/roundTripTime$$

The window size is the amount of data TCP will send before waiting for the first acknowledgement. For a typical TCP window size of 16Kbytes (varies depending on operating system implementation,) the speed achievable by a flow to Singapore is approximately: 131072 bits / 0.25s = 524288 bps which significantly under-utilizes the 12Mbps bandwidth.

One solution to this is to simply *not* use TCP, but instead to deliver data over UDP (if possible) since UDP requires no acknowledgements. Another solution would be to increase the send and receive window sizes to match the capacity[16] (if supported.) Yet another solution would be to deliver data over multiple parallel flows.

4. Jitter in the network will more greatly impact collaborative coordination than latency. Higher latencies with low jitter will still allow collaborators to make reasonable predictions of how an environment will behave (albeit overall task performance will decline.) However high jitter reduces predictability and hence collaborators are forced to employ a purely sequential interaction strategy (such as turn-taking)[10].

4.5. Application Development

1. When building a new application, it is easier to architect the system with collaborative capabilities early rather than later. That is, it is easier to build a new collaborative application than to retrofit a legacy application for collaboration. Part of the problem facing retrofitting is that it often involves the integration of software components that may be highly incompatible as they may have been designed with entirely different programming philosophies or are inherently incompatible at the operating system level.

2. In cases where objects may be co-manipulated at the same time, take advantage of social protocols to avoid having to build highly restrictive distributed mutual exclusion capabilities. Often when participants see that another participant is about to manipulate an object they will naturally refrain from manipulating it also. However, in cases where some form of consistency is required, determine if delayed consistency- with the promise of eventual consistency is acceptable.

3. When building collaborative systems do not necessarily assume that all clients, on separate computers, will connect to a separate dedicated server. In some situations resource limitations may prevent such a luxury in which case one of the clients may have to assume the role of a server. Hence it is of some value for applications to be able to act as a client as well as a server. This has been one of the motivations for making CAVERNsoft's IRB client/server symmetric.

4. Always assume that your system will crash. Design the application to allow collaborators to continue working even when other collaborators have left or crashed.

5. Take advantage of persistence (store more data locally) to reduce the need to distribute a lot of data every time a collaborative session is begun. In CAVERNsoft the IRB couples networking with a persistent data store allowing all collaborative applications to possess persistent capabilities.

5. Ongoing Work

As we mentioned earlier, the CAVERN research agenda consists of three foci: development of tools for supporting TI; evaluation of human-factors and network performance issues; and deployment of the results and products of this research to domain specialists for the development of applications.

The development of the many tele-immersive applications presented here has helped direct the research and development of CAVERNsoft. Many of these applications were recently demonstrated at the Supercomputing'98 conference in Florida, where collaborators from Singapore, Japan, Australia and the United States were able to teleport into the various persistent collaboratories to demonstrate their applications to each other using all the newly erected high speed international networks (STAR TAP and TransPAC).

In the future we intend to focus deeply on three sub-areas of research:

1. the development of techniques to allow multiple collaborating participants to tailor their views based on their expertises and to determine how these heterogeneous perspectives[7] may be used to augment the process of interpreting multi-variate data and collaboration as a whole.

2. the development of techniques to support long-term, asynchronous work amongst collaborators at the end points of our international connections. These will include tools for creating annotations in TI and especially tools for capturing, replaying and querying TI sessions. This is important because we believe most of our collaboration with overseas participants will be via this mode of operation.

3. experimenting with the network Quality of Service (QoS) tools that have been developed by our collaborators at the Nippon Telephone and Telegraph Optical Network Systems Laboratory, Sony Computer Science Laboratory and Keio University[15]. The goal is to make CAVERNsoft the first TI library to routinely use QoS to allow us to study the effectiveness of QoS under varying tele-immersive conditions. In preparation for this we have instrumented CAVERNsoft to allow us to gather networking bandwidth, latency and jitter information. This allows us to automatically instrument any application built with CAVERNsoft.

Additional details about the CAVERNsoft project, its capabilities, and implementation; as well as a copy of CAVERNsoft, LIMBO, CAVE2CAVERNsoft can be downloaded from the Tele-Immersion WEB site at www.evl.uic.edu/cavern.

Acknowledgements

Major funding is provided by the National Science Foundation (CDA-9303433.) The virtual reality research, collaborations, and outreach programs at EVL are made possible through major funding from the National Science Foundation, the Defense Advanced Research Projects Agency, and the US Department of Energy; specifically NSF awards CDA-9303433, CDA-9512272, NCR-9712283, CDA-9720351, and the NSF ASC Partnerships for Advanced Computational Infrastructure program. The CAVE and ImmersaDesk are trademarks of the Board of Trustees of the University of Illinois.

References

[1] T. A. DeFanti and S. Goldstein. The STAR TAP web site, http://www.startap.net, 1998.

[2] I. Foster and C. Kesselman. The GLOBUS web site, http://www.globus.org, 1998.

[3] B. Hibbard and B. Paul. Vis5D http://www.ssec.wisc.edu/~billh/vis5d.html, 1998.

[4] J. Leigh. *CAVERN and a Unified Approach to Support Realtime Networking and Persistence in TeleImmersion*. PhD thesis, University of Illinois at Chicago, 1997.

[5] J. Leigh, A. E. Johnson, and T. A. DeFanti. CAVERN: a distributed architecture for supporting scalable persistence and interoperability in collaborative virtual environments. *Journal of Virtual Reality Research, Development and Applications*, 2(2):217–237, 1997.

[6] J. Leigh, A. E. Johnson, and T. A. DeFanti. The CAVERN web page: http://www.evl.uic.edu/cavern, 1998.

[7] J. Leigh, A. E. Johnson, C. A. Vasilakis, and T. A. DeFanti. Multi-perspective collaborative design in persistent networked virtual environments. In *Proceedings of IEEE Virtual Reality Annual International Symposium '96*, pages 253–260, Apr. 1996.

[8] M. R. Macedonia, M. J. Zyda, D. R. Pratt, P. T. Barham, and S. Zeswitz. NPSNET: A network software architecture for large-scale virtual environments. *Presence*, 3(4):265–287, 1994.

[9] M. McRobbie and S. Goto. The TransPAC web site, http://www.transpac.org, 1998.

[10] K. S. Park and R. V. Kenyon. Effects of network characteristics on human performance in a collaborative virtual environment. In *Proc. IEEE Virtual Reality Conference*. IEEE, IEEE, Mar 1999.

[11] J. Protic, M. Tomasevic, and V. Milutinovic. *Distributed Shared Memory: concepts and systems*. IEEE Computer Society, 1997.

[12] P. Rajlich, R. Stein, and R. Heiland. The vtkactortopf library. http://hoback.ncsa.uiuc.edu/~prajlich/vtkActorToPF, 1998.

[13] D. Robertson and J. Hicks. Future Camp '98. http://www.science.iupui.edu/future_camp98, June 1998.

[14] W. J. Schroeder, K. M. Martin, and W. E. Lorenson. The design and implementation of an object-oriented toolkit for 3D graphics and visualization. In *IEEE Visualization*, pages 93–100, 1996.

[15] A. Shionozaki, K. Yamashita, S. Utsumi, and K. Cho. Integrating resource reservation with rate-based transport protocols in AMInet. In *Proc. of Worldwide Computing and its Applications'98 (WWCA'98)*, Mar 1998.

[16] W. R. Stevens. *TCP/IP Illustrated Vol. 1*, chapter 24, pages 344–350. Addison Wesley, 2 edition, 1994.

[17] J. Tromp, A. Bullock, A. Steed, A. Sadagic, M. Slater, and E. Frecon. Small group behavior experiments in the coven project. *IEEE Computer Graphics & Applications*, 18(6), Nov/Dec 1998.

[18] G. H. Wheless, C. M. Lascara, A. Valle-Levinson, D. P. Brutzman, W. Sherman, W. L. Hibbard, and B. Paul. Virtual chesapeake bay: Interacting with a coupled physical/biological model. *IEEE Computer Graphics and Applications*, 16(4):52–57, July 1996.