

Networked Virtual Environments and Rehabilitation

Robert V. Kenyon and Jason Leigh

Electronic Visualization Lab, Department of Computer Science, University of Illinois at Chicago, Chicago, IL USA.

Abstract—Virtual environments (VE) have been shown to be a powerful tool for various forms of rehabilitation. Currently, VE has been used in psychological training, motor recovery, visual relearning, and pain reduction during physical therapy. VE's use is also being explored in stroke rehabilitation coupled to robots, posture control in labyrinthine deficit patients, tele-rehabilitation and in other areas. Expanding VE's influence is its coupling with high-speed networking [Tele-Immersion]. Current experimental networks have speeds approaching 100 Gb/sec. At these speeds the network bandwidth rivals that of the internal connections of a PC (i.e., bus speed). Accordingly, these new networks will permit various peripherals attached to computers on this network to be connected together and act as fast as if connected to a local PC. This innovation may soon allow the development of previously unheard of networked rehabilitation systems. This presentation will discuss various forms of VE that are currently available for rehabilitation. Furthermore, we will explain the characteristic of these new networks. Finally, we will examine how such networks might be used for extending the rehabilitation clinic to remote areas.

Keywords—Networking, Rehabilitation, Virtual Reality

I. INTRODUCTION

Virtual environments (VE) or virtual reality has taken a foot hold in rehabilitation with dramatic results in some cases. Some applications have the patient wearing VE systems to improve their ability to get around [1]. Others bring the VE technology to the patient to improve much needed rehabilitation [2]. The ability for us to use this technology outside the area of research labs and bring these systems to clinics is just starting. However, the cost is high and the applications that can best be applied to rehabilitation are few and far between. The cost of such systems might be mitigated if this technology allowed therapists and patients to interact more frequently and/or resulted in better patient outcomes. Such issues are under study now at several institutions. This brings us to the idea of tele-rehabilitation, which would allow therapy to transcend the physical boundaries of the clinic and go wherever the communication system and the technology would allow [3]. For example, at some location remote from the clinic a patient enters a VE suitable for rehabilitation protocols connected to the clinic and a therapist. While this idea is not new, the kind of therapies that could be applied under such a condition is limited by the communication connection and facilities at both ends of the communication cable. In this paper, we explore what we consider should be the future of networks and rehabilitation but beyond the use of telephone lines or even T-1 connections. Here, we are considering the use of

high speed, low latency connections that are currently experimental networks. Given the characteristics of VE, physical therapy protocols, the communications system and rehabilitation devices, what kind of signals must be transmitted between sites and what is the needed response time?

II. BACKGROUND

What is known as VE has also been redefined in some cases when it is applied to rehabilitation. Therefore, let us first define what we consider VE and consider the signals that need to be transmitted for such a system to operate remotely (TeleImmersion). VE is immersion of a person in a computer generated environment so that the person experiences stereovision, correct perspective for all objects regardless of their motion, objects move in a natural fashion with subject motion. To achieve these characteristics, certain technology must be brought into play. To provide stereovision, slightly different images must be presented to the right and left eyes with little if any cross talk between the two images. In some systems this is provided by using field sequential stereo in combination with liquid crystal shutter glasses (StereoGraphics, Inc). In this system the right LC lens is clear while the left is opaque and the scene generated on the screen is that of the right eye. Then the left eye lens is clear and the right is opaque and the left eye view is displayed. This method of producing stereo has found its way into projection based systems [3] and desktop system also known as "fish tank VR" [5]. In other systems the person wears a head mounted display (HMD) where the right and left eye each see a dedicated display so that the computer generates a left and right eye image and each image is connected to the corresponding monitor. Such systems have used miniature CRTs, Liquid Crystal Displays, and Laser directed light to create the image on the retina [6]. In contrast to the above mentioned systems, an auto-stereographic system displays stereo images to the person without the aid of visual apparatus worn by the person [7]. The person merely looks at the screen(s) and sees stereo images as one might in the natural world. Because of their ease of use by the subject and their versatility these new and experimental systems have the potential of becoming the ultimate VE display when large motions of the subject are not needed.

Regardless of the system used, to keep all the stereo objects in the correct perspective and to keep them from being distorted when the person moves in the environment, it is necessary to track the movements of the person so that

the computer can calculate a new perspective image given the reported location of the person's head/eyes. The tracking systems that are used to do this are varied. The most prominent of these are the magnetic tracking systems (Ascension, Inc and Polhemus, Inc.). With these systems a small sensor cube is placed on the subject and the location of the sensor within the magnetic field is detected. When the sensor is placed on the head or glasses of the person the orientation of the head and therefore the location of the eyes can be presumed. Other systems use a combination of acoustic location to delineate position and acceleration detection to get orientation. The combination results in 6 degrees of freedom for the location information (InterSense, Inc). Tracking is necessary if the subject is to physically move in the environment when using a VE system.

So far we have confined our discussion to visual objects and have not considered the use of haptic or other forms of information to be integrated into the VE system [8]. To provide a realistic haptic experience to the subject objects must be rendered at 1000 times per second. While a local haptic system such as that produced by Sensable Inc. and others can provide such high speed communication with local systems. When such information is floated over the network the issues of bandwidth and latency of the network are paramount to consider. While experimental networks have significantly increased the bandwidth of the network, our ability to move information over these networks is currently fixed by the speed of light. While prediction and other methods can be employed to help reduce the effective latency (Handshake Technologies, Inc), this characteristic will continue to pose a problem for many conditions that we would like to use in tele-rehabilitation.

III. NETWORKED VIRTUAL ENVIRONMENTS

In networked VEs several types of data need to be transmitted between collaborating sites: 1. the main data-set itself (this often consists of 3D geometry); 2. the changes to the data-set (these occur when collaborating users modify the geometry in some way- perhaps by moving the object or deforming it); 3. the virtual representation of the remote collaborator (this often referred to as an avatar); 4. the video and/or audio channel (that facilitates face-to-face conversation.)¹

The common model for data sharing in networked VEs is to have most of the main data-set replicated across all the sites and transmit only incremental changes. Furthermore the main data-set is often cached locally at each of the collaborating sites to reduce the need to have to retransmit

¹ Video has limited use in stereoscopic projection-based VEs because the large shutter glasses that the viewer uses to resolve the stereo tends to hide the viewer's face from the camera. Furthermore most stereoscopic projection systems operate in dimly lit rooms which are usually too dark for effective use of video.

the entire data-set each time the application is started. Classically TCP (Transmission Control Protocol- the protocol that is widely used on the Internet for reliable data delivery) has been the default protocol used to distribute the data-sets. TCP works well in low-bandwidth (below 10Mb/s) or short distance (local area) networks. However for high-bandwidth long-distance networks, TCP's conservative transmission policy thwarts an application's attempt to move data expediently, regardless of the amount of bandwidth available on the network. This problem is known as the Long Fat Network (LFN) problem [9]. There are a wide variety of solutions to this [10], however none of them have been universally adopted.

Changes made to the 3D environment need to be propagated with absolute reliability and with minimal latency and jitter. Latency is the time it takes for a transmitted message to reach its destination. Jitter is the variation in the latency. Fully reliable protocols like TCP have too much latency and jitter because the protocol requires an acknowledgment to verify delivery. Park and Kenyon [11] have shown that jitter is far more offensive than latency. One can trade off some latency for jitter by creating a receiving buffer to smooth out the incoming data stream. UDP (User Datagram Protocol) on the other hand transmits data with low latency and jitter, but is unreliable. Forward Error Correct (FEC) is a protocol that uses UDP to attempt to correct for transmission errors without requiring the receiver to acknowledge the sender. FEC works by transmitting a number of redundant data packets so that if one is lost at the receiving end, the missing data can be reconstructed from the redundant packets [11]. FEC however is not completely reliable. Hence to achieve complete reliability (at the expense of an infrequent increase in jitter) FEC is often augmented with an acknowledgment mechanism that is only used when it is unable to reconstruct a missing packet.

The virtual representation of a remote collaborator (avatar) is often captured as the position and orientation of the 3D tracking devices that are attached to the stereoscopic glasses and/or 3D input device (e.g. a wand). With simple inverse kinematics one is able to map this position and orientation information onto a 3D geometric puppet, creating lifelike movements [13]. The 3D tracking information is often transmitted using UDP to minimize latency and jitter- however since the data is mainly used to convey a user's gesture, absolute delivery of the data is not necessary. Furthermore since tracking data is transmitted as an un-ending stream, a lost packet is often followed soon after (usually within 1/30th of a second) by a more recent update.

Audio and video data are similar in property to the avatar data in that they usually comprise an un-ending stream that is best transmitted via UDP to minimize latency and jitter. Often video and audio packets are time stamped so that they can be synchronized on the receiving end. When more than two sites are involved in collaboration it is more

economical to send audio/video via multicast. In multicast the sender sends the data to a multicast address and the routers that receive the data send copies of the data to remote sites that are subscribed to the multicast address. One drawback of multicast is that it is often disabled on routers on the Internet as one can potentially flood the entire Internet. An alternative approach is to use dedicated computers as “repeaters” that intercept packets and transmit copies only to receivers that are specifically registered with the repeater. This broadcast method tends to increase the latency and jitter of packets, especially as the number of collaborators increases.

IV. QUALITY OF SERVICE (QoS)

QoS refers to a network’s ability to provide bandwidth and/or latency guarantees. QoS is crucial for applications such as networked VE, especially those involving haptics or tele-surgery, which are highly intolerant of latency and jitter. Early attempts to provide QoS (such as Integrated Services and Differentiated Services) have been good research prototypes but have completely failed to deploy across the wider Internet because telecommunications companies are not motivated to abide by each others QoS policies. Some researchers have argued that QoS is unnecessary because in the future all the networks will be over-provisioned so that congestion or data loss that result in latency and jitter, will never occur. This has been found to be untrue in practice. Even with the enormous increase in bandwidth accrued during the dot-com explosion, the networks are still as unpredictable as they were a decade ago. Ample evidence is available from the online gaming community which often remarks about problems with bandwidth, latency and jitter during game sessions [14]. These games are based on the same principles that govern the design of networked VEs and therefore serve as a good metric for the current Internet’s ability to support tightly coupled collaborative work.

V. CUSTOMER OWNED NETWORKS

Frustrated by the lack of QoS on the Internet, there is growing interest in bypassing the traditional routed Internet by using the available dark fiber in the ground. Dark fiber is optical fiber that has not yet been lit. Currently it is estimated that only about 5-10% of the available fiber has been lit, and each fiber has several terabits/s of capacity. The dot-com implosion has made this dark fiber and wavelengths of light in the fiber, very affordable. The newly emerging model is to construct a separate customer-owned network by purchasing or leasing the fiber from a telecommunications company, and installing one’s own networking equipment at the endpoints. A number of federally supported national and international initiatives have been underway for the last few years to create customer-controlled networks explicitly for the scientific

community. These include the National Lambda rail [15], StarLight [16], and the Global Lambda Integrated Facility [17]. By creating dedicated fiber networks, applications will be able to schedule dedicated and secure light paths with tens of gigabits/s of unshared, uncongested bandwidth between collaborating sites. This is the best operating environment for tightly coupled networked, haptic VEs.

VI. CONNECTION CHARACTERISTICS FOR REHABILITATION

The ability to use this technology for rehabilitation is a function of cost, availability, and the kind of applications that can best utilize the network and provide rehabilitation services. While current research has been exploring use of low speed and inexpensive communication networks this eliminates the possible implementations that new technology may have albeit in a few locations. Let us consider the case where a high speed network connects a rehabilitation center and a remote clinic. The question is what kind of services can be provided remotely.

The scenario that we envision is one where patients are required to appear at the rehabilitation clinic to get therapy. This condition we have a therapist with VE, a dedicated high speed network, haptic device, video connection, and software to help analyze the incoming data (i.e., data mining). As is displayed in Figure 1, the therapist station has several areas of information that connects him/her to the remote patient. The VE (in this case Varrier) provides them with a representation of the patient and the kind of trajectory that will be needed for this training session. Notice that the use of Varrier removes the need for HMD or shutter glasses being worn by the patient or therapist. While this may seem small but now the patient and the therapist can see each other eye to eye. The video connection allows more communication (non verbal or bed side manner) to take place between the two linked users of this system. The haptic device is used to feedback the forces from the patient’s limb to the therapist and the feed to the patient the forces that the therapist wishes to act on the patient. Here instead of giving the desired motion of the affected limb we provide a task that uses the affected limb so that learning and coordination is encouraged. Other possibilities include having the robot apply forces to the patient appendage so the actions of the patient cause adaptation and aid in the healing of the nervous system [8]. This scenario could also allow the patient to see both the virtual limb and their own limb so that objects are not obscured by their own motions. As can be seen from Figure 1, the bandwidth and latency requirements change as a function of the kind of information that is being transmitted. A system as described above is possible today although expensive. The network characteristics that would be needed for each information channel would be as follows. A high bandwidth connection would be needed for video and audio streamed to the plasma displays at each location, in addition to the high bandwidth a low latency and jitter connection would be needed for the

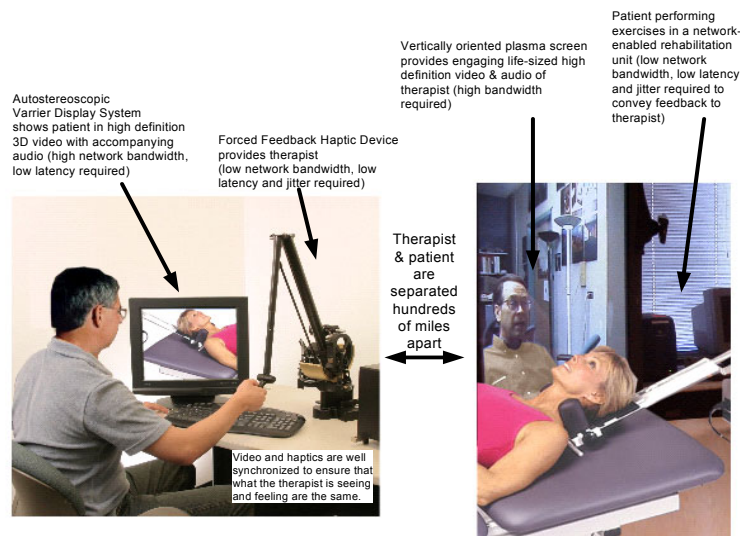


Figure 1: Possible Tele-Rehabilitation Scenario Facilitated by High Bandwidth Networking

Varrier Display system (VE). For a force feedback haptic device communicating between the patient and the therapist, a low network bandwidth could be used but the latency and jitter need to be low.

VII. CONCLUSION

The ability to provide rehabilitation services to locations outside the clinic is emerging as an important option for clinicians and patients. Effective therapy may best be supplied by the use of high technology systems such as VE and video, coupled to robots, and linked between locations by high-speed, low-latency, high-bandwidth networks. The use of data mining software would help analyze the incoming data to provide both the patient and the therapist with evaluation of the current treatment and modifications needed for future therapies. Further research and funding in this area will be needed to answer these question and more questions that arise from the use of these technology.

ACKNOWLEDGMENT

This work is supported by grants DC05235 from NIH-NIDCD and AG16359 from NIH-NIA, H133E020724 from NIDRR and NSF grant # ANI-0225642.

REFERENCES

- [1] T. Riess, S. Weghorst, Augmented reality in the treatment of Parkinson's disease. In *Proceedings of Medicine Meets Virtual Reality III* (San Diego, CA), pp: 298-302, January 1995.
- [2] H.G. Hoffman, D.R. Patterson, G.J. Carrougner, S.R. Sharar, The effectiveness of virtual reality-based pain control with multiple treatments. *The Clinical Journal of Pain*. 2001 17:229-235.
- [3] M.J. Rosen, Telerehabilitation, *NeuroRehabilitation*, special topic issue on Technology in Neurorehabilitation, ed. by M.J. Rosen, 12(1), April, 1999.
- [4] C. Cruz-Neira, D.J. Sandin, T.A. DeFanti, R.V. Kenyon, and J.C. Hart, The CAVE: Audio Visual Experience Automatic Virtual Environment, *Communications of the ACM*, Vol. 35, No. 6, June 1992, pp. 65-72.
- [5] C. Ware, K. Arthur, K.S. Booth, Fisk tank virtual reality. In Ashlund S, Mullet K, Henderson A, Hollnagel E, White T (eds) *INTERCHI '93 Conf Proc*, pp. 37-42., 1993.
- [6] M. Tidwell, R.S. Johnston, D. Melville, and T.A. Furness The Virtual Retinal Display - A Retinal Scanning Imaging System. In *Proceedings of Virtual Reality World '95*, pp. 325-333, 1995
- [7] D.J. Sandin, T. Margolis, G. Dawe, J. Leigh, T.A. DeFanti, The Varrier™ Auto-Stereographic Display, *SPIE*, vol. 4297, no. 25, San Jose, California, 2001.
- [8] J. Patton, G. Dawe, C. Scharver, F. Mussa-Ivaldi, R.V. Kenyon, Robotics and Virtual Reality: A Perfect Marriage for Motor Control Research and Rehabilitation, *J. Assistive Technology*, (in press).
- [9] W. R. Stevens, "TCP/IP Illustrated," vol. 1: Addison Wesley, 1994, pp. 344-350.
- [10] E. He, J. Leigh, O. Yu, T.A. DeFanti, Reliable Blast UDP: Predictable High Performance Bulk Data Transfer, *Proc. IEEE Cluster Computing*, Sept, Chicago, Illinois, 2002.
- [11] K. Park, R.V. Kenyon, Effects of Network Characteristics on Human Performance in the Collaborative Virtual Environment. *IEEE Virtual Reality '99 Conference*, March 14-17, 1999, Houston Tx.
- [12] J. Leigh, O. Yu, D. Schonfeld, R. Ansari, et al., Adaptive Networking for Tele-Immersion, *Proceedings Immersive Projection Technology: Eurographics Virtual Environments Workshop (IPT/EGVE)*, May 16-18, Stuttgart, Germany, 2001.
- [13] K. Park, Y. Cho, N. Krishnaprasad, C. Scharver, M. Lewis, J. Leigh, A. Johnson, CAVERNsoft G2: A Toolkit for High Performance Tele-Immersive Collaboration, to appear in the *Proceedings of the ACM Symposium on Virtual Reality Software and Technology 2000*, Oct 22-25, 2000, Seoul, Korea, pp. 8-15
- [14] Available: <http://xbox.ign.com/articles/377/377303p1.html?fromint=1>
- [15] Available: [http://www.nationallambdaail.org](http://www.nationallambdarail.org)
- [16] Available: <http://www.startup.net/starlight>
- [17] Available at: <http://www.glif.is>