Integrated Voice and Haptic Support for Tele-Rehabilitation*

Margaret McLaughlin, Roger Zimmermann, Leslie S. Liu, Younbo Jung, Wei Peng,
SeungA Jin, Jill Stewart, Shih-Ching Yeh, Weirong Zhu and Beomjoo Seo
Annenberg School for Communication, Computer Science Department
and the USC-UT Consortium for Interdisciplinary Research
University of Southern California, Los Angeles, CA 90089
[mmclaugh, rzimmerm]@usc.edu

Abstract

In recent years, rapid technological advances have changed how many existing tasks are performed. For example, the combination of virtual environments and the Internet is presenting a powerful opportunity in the area of medical training and rehabilitation. Here we present our design and preliminary experience with a tele-haptic environment augmented with a voice conferencing system. Initial tests are encouraging and reveal the promise of our system.

1   Introduction

Haptics involves the modality of touch and the sensation of shape and texture an observer feels when virtually 'touching' an object with a force-feedback device. Our work and that of colleagues in this rapidly developing field is introduced in our book, *Touch in Virtual Environments: Haptics and the Design of Interactive Systems* [11]. In recent years we have been studying how to integrate this sensory modality into aware collaborative environments in such application areas as medical training and simulation, in collaboration with researchers at USC's Keck School of Medicine [14]. We have recently partnered with the USC-UT Consortium on Interdisciplinary Research to address the problem of functional recovery from stroke and our group has been awarded funding from the National Institutes of Health. Our role within the larger project is to develop collaborative virtual environments (VEs) that include different levels of haptic sensory feedback and to evaluate the effectiveness of these applications for neurorehabilitation training. The tasks to be performed within these VEs span a range of activities from everyday functional tasks to game-like activities designed to motivate specific motor action. The virtual rehabilitation environments which we have designed will serve stroke patients in the subacute phase. Our particular objective is to develop task-specific virtual exercise environments which will reinforce the compensatory

brain mechanisms that facilitate recovery from stroke.

Within this virtual exercise environment a number of modalities must be conveyed for an effective therapy session to take place. In addition to the training task to be performed with the haptic device, a communication channel must exist between the therapist and the patients. With the expanding capabilities of the Internet to handle digital media streams it is now possible to transmit all the relevant information. Recently, broadband Internet connections present a viable and cost-effective way to send and receive audio, video and haptic streams. Our AudioPeer project aims to provide a foundation for collaboration tools that will enable novel interactive applications. AudioPeer, which is a multi-party voice conferencing system, is designed to allow individuals and groups to communicate. Our goal is to provide users with an interactive audio experience based on a platform that is scalable, practical, integratable and extensible with new features.

The integration of a tele-haptic system with remote voice conferencing presents a powerful opportunity to provide high quality and effective healthcare in a cost-effective manner. Here we introduce our technology and present preliminary results from our first experiences with this integrated technology platform. We survey the related work in Section 2. Next we detail our approach in Section 3. Finally, Section 4 presents some concluding thoughts.

2   Related Work

Tele-Haptic Environments   Broeren et al. [3] used a 3D computer game to promote motor relearning in a patient suffering from a left arm paresis. Connor et al. [6] tested the efficacy of haptically-guided errorless learning (EL) with a group of patients with post-stroke visuoperceptual deficits. EL training with haptic guidance benefited some patients, but not all. Boian et al. [2] developed a rehabilitation environment using a haptic device called the 'Rutgers Ankle'. The patient uses the Ankle as a foot joystick. Variations in the exercises can be achieved through manipulating device-level impedance and stiffness. Boian et al. reported some gains in gait speed and muscle strength for study participants.

Although the use of VR environments and haptics for stroke rehabilitation has been studied in several disciplines [9], we have produced a number of technical innovations which will prove extremely valuable [12, 10]: (1) we have created robust techniques for recording and playing back the activities of the user during a haptics-enhanced VR session regardless of device type, useful when a trainee is attempting to carry out a series of skilled movements in the manner of an expert trainer. (2) we have implemented Kalman filtering algorithms which allow us to anticipate incorrect motion trajectories by the user and head off errors before they occur, a technique which may be useful when implicit memory processes are invoked in the training protocol. (3) we have developed a spoken dialog system fully integrated with our haptic systems which permits voice control of the force feedback devices. (4) we have made it easier for the user to locate objects in the virtual environment and to remain "in contact" with them, and made "snap-to" features which move the haptic cursor to digital objects with voice commands. (5) we have created new techniques for collaboration between users of both like and heterogeneous haptic devices; this "hand-over-hand" guidance capability creates a sense of immediacy and presence in virtual rehabilitation environments and facilitates rapid demonstration of correct movement and correction of error.

Multi-Party Voice Conferencing Among the existing audio mixing based systems, hierarchical architectures [13] place all participants at the leaf nodes and locate the mixers at non-leaf nodes. A mixer relays the audio coming from its children to the parent after combining the streams. A root mixer broadcasts the merged audio to all participants through the IP multicast network. Several commercial products such as the ClickToMeet conferencing product, formerly CUseeMe, construct an efficient delivery path (full-mesh or multicast) among the mixers. But because most of these products are integrated with video conferencing, they have a high network overhead and cannot fully take advantage of audio mixing. In an end-system mixing architecture, one of the participants is chosen as a mixer. However, all the above systems do not describe how to build efficient delivery paths among mixers and participants.

The unsatisfactory deployment of IP Multicast has motivated researchers to propose a number of *application-level* multicast protocols. Such protocols only make use information provided at the application level to transmit data. Many-to-many protocols can be classified into mesh based (e.g. Narada [5]), tree based (e.g., Yoid [7] and our proposed ACTIVE), cluster based (e.g., NICE [1]) and Distributed Hash Table (DHT) based architectures (e.g., Pastry/Scribe [4]). All the application-level multicast protocols fail to recognize the feasibility of constructing a more efficient topology if the aggregation of audio packets is allowed along the delivery path.

Figure 1. The tele-rehabilitation therapist is communicating with the patient via the AudioPeer system while monitoring his actions and progress on the screen.

3 Approach and System Architecture

The goal of our tele-rehabilitation platform is to provide effective task exercises to remote patients. A tele-haptic collaborative environment provides a shared exercise space. In addition, voice conferencing allows for effective therapist-patient communication. We will describe both components.

3.1 Tele-Haptic Environment

We are currently developing several task environments, using different types of haptic feedback devices, which can be host to a progressive set of training tasks from precise fine motor movements to reaching movements that involve full arm, shoulder and torso activity. Here we describe exercises that make use of the PHANToM haptic device.

The PHANToM is a small, desk-grounded robot that permits simulation of single or two-fingertip contact with virtual objects through a thimble or stylus; its actuators communicate forces back to the user's fingertips as it detects collisions with virtual objects, simulating the sense of touch. The PHANToM device is being used for a series of tasks to move a ball through a maze-like tube, select a small object from an array of other objects and barriers, and pick up and squeeze a deformable cube to put it through a small aperture (using two PHANToM thimbles worn on thumb and forefinger). These applications can display the objects and motor action on a standard PC monitor that can be set for both mono and stereo viewing.

Figures 1 and 2 show an example environment for both the therapist and her patient. The patient (Figure 2) runs the exercise application on his local computer and performs the exercise task with the attached PHANToM. Concurrently the AudioPeer conferencing system allows the patient to communicate with the therapist via either a headset or a combination of a microphone and desktop speakers. The therapist (Figure 1) can monitor both the actions and

2

Figure 2. The remote patient is using the PHANToM to grasp and maneuver a small cube through a three dimensional maze of spheres ('Ball Array' exercise).
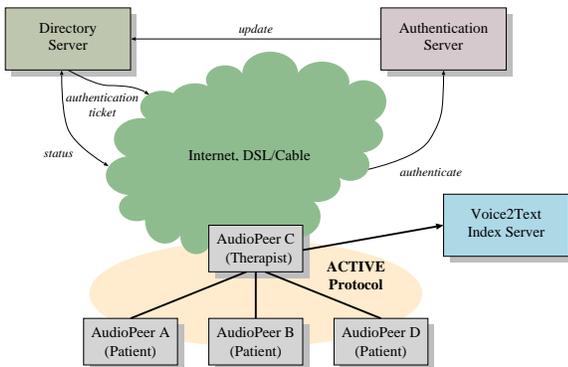
Figure 3. AudioPeer system architecture and components.

progress of the patient. If she feels that the patient needs additional help or instructions, she can conveniently provide the needed assistance through AudioPeer.

3.2 AudioPeer: Multi-Party Voice Conferencing

The AudioPeer system has been designed with a number of goals in mind [15]. Overall it aims to provide an efficient audio multi-party voice conferencing application with low audio latency. It is also designed to be easily integratable with complementary functions and applications. To be scalable, AudioPeer is built on a low bandwidth consuming peer-to-peer overlay network called ACTIVE [8]. Through its peer-to-peer (p2p) foundation, ACTIVE inherits all the desirable characteristics of p2p technology, namely (1) each node acts both as a server and client and hence resources scale with the number of participants, (2) the design is robust against node failures and other dynamic events in the topology, and (3) central resource usage is minimal.

The AudioPeer system design, based on the ACTIVE protocol, consists of four components as illustrated in Figure 3: a *Directory server*, an *Authentication server*, the *AudioPeer application* and an optional *Voice2Text index server*. We will briefly describe the functionality of each component.

The Directory server functions as the initial point-of-contact for an AudioPeer application to locate other AudioPeer users. It provides a convenient way to associate voice sessions with participants and applications (such as a therapy session). Once the relevant session information has been retrieved, the AudioPeer application runs independently of the Directory server in pure p2p mode.

The Authentication server is required in certain types of applications (e.g., tele-medicine) where the end-users must be authenticated to ensure that the privacy of the exchanged information is maintained, and that only registered users receive the services requested.

The AudioPeer application is at the heart of the voice communication system. It maintains the connection topology among multiple AudioPeer participants while dynamically optimizing the end-to-end audio latency based on speaker activity and maintaining the overlay multicast tree link structure. Each node performs audio mixing of the incoming streams (if required by the application) to reduce the network bandwidth. The AudioPeer design allows the support of a large number of simultaneous audio chat sessions.

A hands-free interface that can be controlled by speech commands is planned for the patient's system. The speech recognition provides real-time voice-to-text conversion for maintaining a transcript of the patient's dialog with the system. With some stroke patients, especially robust voice recognition techniques may be required. We are working with colleagues in speech science to develop techniques for real-time recognition and archiving of dysarthritic speech. This service will be implemented in the optional Voice2Text indexing component.

AudioPeer's capability to handle group communications can provide benefits for tele-rehabilitation. Whether multiple patients may be supervised together often depends on the functional ability of the individual. For example, people with very little movement often require more hands-on treatment while people with good movement and minimal compensatory patterns can often practice designated tasks with minimal supervision. Generally, one physical therapist will not supervise more than two stroke patients at once (again depending on the functional level; some patients need one-on-one treatment all the time). It is also becoming increasingly common to offer group classes for a fee to people who no longer qualify for traditional physical therapy (often due to insurance restrictions).

3.2.1 Experimental Validation

We conducted a preliminary experimental user evaluation during the Spring 2005 semester. Users filled out a web-

No.	Question	Rating (1 to 7)
1.	Once you had learned how to operate AudioPeer, did you find it easy to use?	5.06
2.	Did you find AudioPeer a useful way to communicate?	4.4
3.	Was AudioPeer helpful in your communications?	4.53
4.	Do you find AudioPeer an efficient way to communicate?	4.2
5.	Were you frustrated by any delays caused by communicating with AudioPeer?	5.2
6.	Did you like communicating using AudioPeer?	4.2
7.	Did you enjoy communicating using AudioPeer?	4.13
8.	Did you find AudioPeer fun to use?	4.13
9.	Did AudioPeer make you feel connected to other people?	4.4
10.	Did AudioPeer make you feel that people were at hand to answer your questions?	4.4

Table 1. User feedback on AudioPeer.

based questionnaire to rate AudioPeer's features on a scale from 1 to 7 (Table 1). Free-text comments were also collected. As is illustrated in the results, the average rating of the AudioPeer system is between 4 to 5. Many users found AudioPeer very easy to use. Issues that we are addressing in our ongoing work are the improvement of the audio quality and the reduction of the end-to-end latency.

4 Conclusions

We have presented our design of a distributed collaborative environment for tele-rehabilitation that integrates both haptic and aural modalities. Initial results are encouraging and we will continue on this exciting path and plan to pursue the following extensions.

We will continue initial patient testing of the environments to enhance usability and to iteratively evolve the interaction parameters. We are closely collaborating with the other members of the project to design the application interactions from a biokinesiology perspective to maximize therapeutic movement with cortical reorganization goals in mind. As well, operational interfaces will be designed that allow the rehabilitation professional to easily adjust stimulus parameters and interactional challenges "on the fly" based on individual patient need and progress. All systems will provide intuitive feedback as to the success of the individual actions of the patient and deliver performance data at the end of the trials for assessment of rehabilitative gains by the patient and rehabilitation professionals and for research purposes.

With AudioPeer there are several augments we are pursuing. First, we are continuing to refine our ACTIVE protocol and increase its robustness. We will improve its usability across firewalls and network address translation (NAT) devices (e.g., broadband routers). This will facilitate its practical deployment. Speech recognition is currently performed in a centralized location and this is in conflict with our peer-to-peer design. We will investigate the practicality of distributed speech recognition. Last but not least, we are currently using personal computers and we plan to extend the support of AudioPeer to PDAs and other handheld and mobile devices.

References

[1] S. Banerjee, B. Bhattacharjee, and C. Kommareddy. Scalable Application Layer Multicast. In *Proceedings of ACM SIGCOMM*, Pittsburgh, August 19-23, 2002.

[2] R. F. Boian, J. E. J. E. Deutsch, S. L. Chan, G. C. Burdea, and J. Lewis. Haptic effects for virtual reality-based post-stroke rehabilitation. In *11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 247-253 2003.

[3] J. Broeren, M. Georgsson, M. Rydmark, and K. S. Sunnerhagen. Virtual reality in stroke rehabilitation with the assistance of haptics and telemedicine. In *4th Intl Conference on Disability, Virtual Reality and Associated Technology*, pages 71–76, Veszprém, Hungary, 2002.

[4] M. Castro, P. Druschel, A. Kermarrec, and A. Rowstron. SCRIBE: A large-scale and decentralized application-level multicast infrastructure. *IEEE Journal on Selected Areas in Communications (JSAC)*, 2002.

[5] Y. Chu, S. G. Rao, S. Seshan, and H. Zhang. Enabling Conferencing Applications on the Internet Using an Overlay Multicast Architecture. In *Proceedings of the ACM SIGCOMM 2001*, San Diago, CA, August 27-31, 2001.

[6] B. B. Connor, A. M. Wing, G. W. Humphreys, R. M. Bracewell, and D. A. Harvey. Errorless learning using haptic guidance: Research in cognitive rehabilitiation following stroke. In *4th Intl Conference on Disability, Virtual Reality and Associated Technology*, pages 77–84, Veszprém, Hungary, 2002.

[7] P. Francis. Yoid: Extending the Internet Multicast Architecture. *Technical report, AT&T Center for Internet Research at ICSI (ACIRI)*, April 2000.

[8] L. S. Liu, R. Hampole, B. Seo, and R. Zimmermann. ACTIVE: A Low Latency P2P Live Streaming Architecture. In *SPIE Conference on Multimedia Computing and Networking (MMCN)*, San Jose, California, January 17-20, 2005.

[9] R. Loureiro, F. Amirabdollahian, S. Coote, E. Stokes, and W. Harwin. Using haptics technology to deliver motivational therapies in stroke patients: Concepts and initial pilot studies. In *1st European Conference on Haptics, Euro-Haptics*, 2001.

[10] M. L. McLaughlin, J. Hespanha, and G. Sukhatme. Haptic collaboration over the internet. In M. L. McLaughlin, J. Hespanha, and G. Sukhatme, editors, *Touch in virtual environments: Haptics and the design of interactive systems*. Prentice-Hall, 2002.

[11] M. L. McLaughlin, J. Hespanha, and G. Sukhatme. *Touch in Virtual Environments: Haptics and the Design of Interactive Systems*. Prentice-Hall, United Kingdom, 2002.

[12] M. L. McLaughlin, G. Sukhatme, W. Peng, W. Zhu, and J. Parks. Performance and co-presence in heterogeneous haptic collaboration. In *Haptics Symposium, IEEE VR 2003*, 2003.

[13] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker. Application-Level Multicast Using Content-Addressable Networks. In *Proceedings of the Third International Workshop on Networked Group Communication*, November 2001.

[14] A. Rizzo, M. McLaughlin, Y. Jung, W. Peng, S. Yeh, and W. Zhu. Virtual therapeutic environments with haptics: An interdisciplinary approach for developing post-stroke rehabilitation systems. In *International Conference on Computers for People with Special Needs (CPSN'05)*, 2005.

[15] R. Zimmermann, B. Seo, L. S. Liu, R. S. Hampole, and B. Nash. AudioPeer: A Collaborative Distributed Audio Chat System. In *10th International Conference on Distributed Multimedia Systems (DMS)*, San Francisco Bay, California, September 8-10, 2004.