Full Body Acting Rehearsal in a Networked Virtual Environment—A Case Study

Abstract

In order to rehearse for a play or a scene from a movie, it is generally required that the actors are physically present at the same time in the same place. In this paper we present an example and experience of a full body motion shared virtual environment (SVE) for rehearsal. The system allows actors and directors to meet in an SVE in order to rehearse scenes for a play or a movie, that is, to perform some dialogue and blocking (positions, movements, and displacements of actors in the scene) rehearsal through a full body interactive virtual reality (VR) system. The system combines immersive VR rendering techniques as well as network capabilities together with full body tracking. Two actors and a director rehearsed from separate locations. One actor and the director were in London (located in separate rooms) while the second actor was in Barcelona. The Barcelona actor used a wide field-of-view head-tracked head-mounted display, and wore a body suit for real-time motion capture and display. The London actor was in a Cave system, with head and partial body tracking. Each actor was presented to the other as an avatar in the shared virtual environment, and the director could see the whole scenario on a desktop display, and intervene by voice commands. A video stream in a window displayed in the virtual environment also represented the director. The London participant was a professional actor, who afterward commented on the utility of the system for acting rehearsal. It was concluded that full body tracking and corresponding real-time display of all the actors' movements would be a critical requirement, and that blocking was possible down to the level of detail of gestures. Details of the implementation, actors, and director experiences are provided.

1 Introduction

A theatrical (or cinematographic) rehearsal consists of a series of preparatory events performed prior to the public performance, in which actors typically learn the dialogue, and solidify aspects of blocking (i.e., positioning, movements, and displacements of actors through time) and stage movement (Mitter, 2002; Mitchell, 2008). The most common method of rehearsal is to gather the director and the actors into a room where some elements of the scenery (or at least some representation of those elements) are available. The actors

Jean-Marie Normand*
EVENT Lab
Facultat de Psicologia
Universitat de Barcelona Campus de
Mundet, Edifici Teatre
08035 Barcelona, Spain
and
CERMA
École Centrale de Nantes
Nantes Cedex 3, France

Bernhard Spanlang
EVENT Lab
University of Barcelona
Barcelona, Spain

Franco Tecchia
Marcello Carrozzino
PERCRO
Scuola Superiore Sant'Anna
Pisa, Italy

David Swapp
Department of Computer Science
University College London
London, UK

Mel Slater
EVENT Lab
University of Barcelona
Barcelona, Spain
and
Department of Computer Science
University College London
London, UK
and
ICREA
Institució Catalana de Recerca i Estudis Avançats
Barcelona, Spain
*Correspondence to jean-marie.normand@ec-nantes.fr.

can then practice the dialogue and their blocking. The director can directly give the actors insights or guidelines in order to improve the performance. Technicians can also be present in order to work on camera angles (e.g., for television or cinema) and/or lighting, sound, and so on.

What if some of the actors, the director, or the sound or light directors cannot be in the same place at the same time? Probably, the rehearsal will have to be postponed, or some part of it would have to be cancelled. On the other hand, if it were possible to rehearse with some actors and/or the director away, significant savings, both in terms of time and money, could be made.

In this context, shared virtual environments (SVEs) seem to be an interesting medium to explore. Indeed, virtual environments allow easy creation and modification of scenery, costumes, and lighting conditions. In this paper, we explore the possibility that full body interactive shared virtual environments can be used for rehearsal. We describe a system whereby the actors and director can work together, and spectators can view the rehearsal, even though they are geographically dispersed. Rehearsals were held of a short segment of a film script, *The Maltese Falcon*, where one actor was located in Barcelona, another in London, and the director was also in London, although in a different physical location from the actor. We report on the evaluation of the process by the professional actor involved, the director, as well as the amateur actress.

2 Background

Virtual reality has typically been used in interactive theater where the focus has been on the role of virtual agents. In Cairco, Babu, Uliski, Zanbaka, and Hodges (2007), a participant could experience a scene of Shakespeare's *Much Ado About Nothing* by acting with a virtual co-actor displayed on a PowerWall. Geigel and Schweppe (2004) allowed participants to act by controlling virtual actors in the context of scenery elements, lighting, or audience in a virtual theatrical storytelling application. These interactions were achieved through specialized interfaces on normal desktop machines.

Recent film productions have relied more and more on motion capture systems; for example, movies such as *Beowulf*, *A Christmas Carol,* and *Avatar* made intensive use of motion capture systems in order to record real actors' motions before applying these to digital characters. Such virtual filmmaking has become more and more popular, especially with the availability of virtual camera tracking systems, for example, from Intersense,[1] Gamecaster,[2] Vicon,[3] and NaturalPoint.[4] These systems enable the combination of physical camera moves (performed by steadycam operators holding physical cameras) with software that articulates the camera movements in the virtual environments.

Mixed reality (MR) techniques for filmmaking applications have also been studied. In Tenmoku, Ichikari, Shibata, Kimura, and Tamura (2006) and in Ichikari et al. (2010), the authors propose a workflow to use computer graphics animation data (i.e., handmade animation in classical 3D animation software), motion capture data or 3D video data into specially designed software to be used for filmmaking. The positions and movements of the camera can be derived from the 3D animation and incorporated in a 3D model representing the physical location of the filming. By using 3D animation data and a 3D model of the scenery, the director is able to plan in advance the movements and positions of the camera for the shot to be filmed. In Ichikari et al., a real-time method for relighting the 3D scene was also added to the previous method.

Shared virtual environments, where people in remote places can collaborate together, already have a long history; see, for example, Normand et al. (1999) where a complete system is described. Such systems, where remotely located people carry out some task together such as puzzle solving, have been studied with respect to their co-presence inducing capabilities and power relationships among the participants and also their task performance (Tromp et al., 1998; Steed, Slater, Sadagic, Bullock, & Tromp, 1999; Slater, Sadagic, Usoh, &

1. http://www.intersense.com/categories/19/
2. http://gamecaster.com/news/GCS3-Revolutionary-Virtual-Camera-Control-
3. http://www.vicon.com/products/blade.html
4. http://www.naturalpoint.com/optitrack/products/insight-vcs/

Schroeder, 2000; Schroeder et al., 2001; Steed, Spante, Heldal, Axelsson, & Schroeder, 2003). A shared virtual environment system for acting rehearsal of a short play was described in Slater, Howell, Steed, Pertaub, and Garau (2000), where the problem was stated as evaluating "to what extent can actors and directors create sufficient acting performance so that a live performance can take place with hardly any live face-to-face rehearsal?" The actors and the director used a standard desktop system through which they could see avatars representing one another from a first-person perspective. They could move through the environment, make simple gestures, and change facial expression through keyboard presses and mouse strokes. Despite its simplicity, and the initial rejection of the system by some, the actors were able to use it for remote rehearsal and eventually acted together a short play in a live performance in front of an audience after only previously having acted together virtually without ever physically meeting one another. They found that the desktop rehearsal system was useful for blocking, spatial organization among actors, and allowed them to understand this before they ever met face to face.

3   Materials

3.1 The Virtual Reality Systems

There were two immersive virtual reality systems used by the actors, and one desktop system used by the director. Other spectators could also view the rehearsal on a desktop system. The actor in Barcelona used a stereo head-tracked head-mounted display (HMD) and wore a motion capture suit for full body tracking. The HMD was a Fakespace Labs[5] Wide5 with a field of view of $150° \times 88°$ and an estimated resolution of 1600×1200 displayed at 60 Hz. Head tracking was performed by a 6-DOF Intersense IS-900[6] device.

Motion tracking was achieved by a 12-camera Optitrack infrared system that tracks reflective markers attached to a tight-fitting motion capture suit worn by the actor (see

5. http://www.fakespacelabs.com/Wide5.html
6. http://www.intersense.com/categories/20/
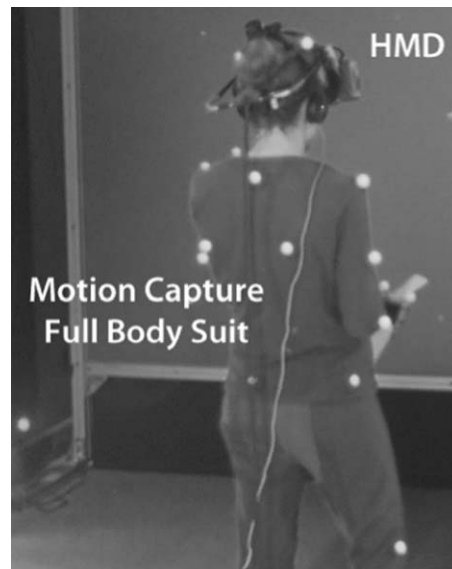
Figure 1. *Full body motion setup: the actor wears an HMD and a motion capture suit.*

Figure 1). The motion tracking volume was approximately 3 m in width and depth and 2.5 m in height.

The actor in London was located in a four-screen Trimension ReaCToR Cave-like system. (Henceforth, we refer to this using the generic name *Cave*.) The Cave was driven by a 5 PC cluster. The Cave system has three $3 m \times 2.2 m$ back-projected screens: front, left, and right, and a $3 m \times 3 m$ front projection surface on the floor. The computers in the cluster contained Intel Pentium 3.2 GHz processors with 1 GB of RAM and Nvidia Quadro FX 5600 graphics cards. The participants were fitted with shutter glasses that were synchronized with the projectors delivering active stereo at 45 Hz for each eye. Attached to the top of the glasses was an InterSense IS-900 tracking device to track the head of the participant.

There was partial body tracking of the actor in the Cave system (see Figure 2). The actor held an Intersense Wand in each hand and a simple inverse kinematics (IK) technique based on the law of cosines was used to infer the position of the arms and hands, and mapped to an avatar representing the actor. The head movements of the actor, obtained through the tracker mounted on the stereoscopic shutter glasses, were also mapped to the head movements of the avatar.
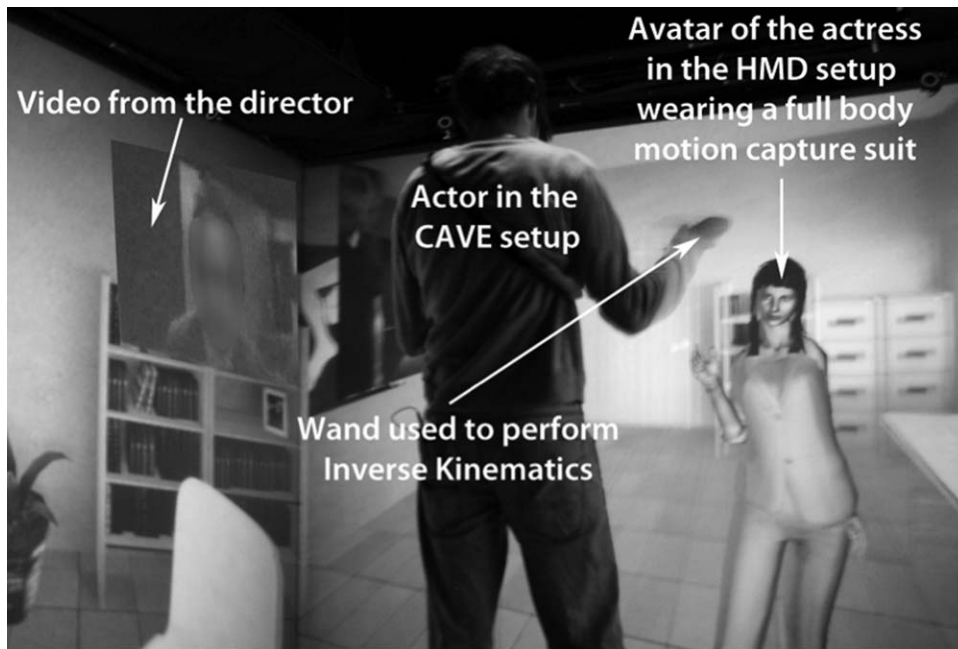
Figure 2. *Cave setup: the actor holds two wands used for inverse kinematics (IK). On the four-wall stereoscopic Cave, we can see the avatar representing the actress from the HMD setup and the video from the director's web camera.*
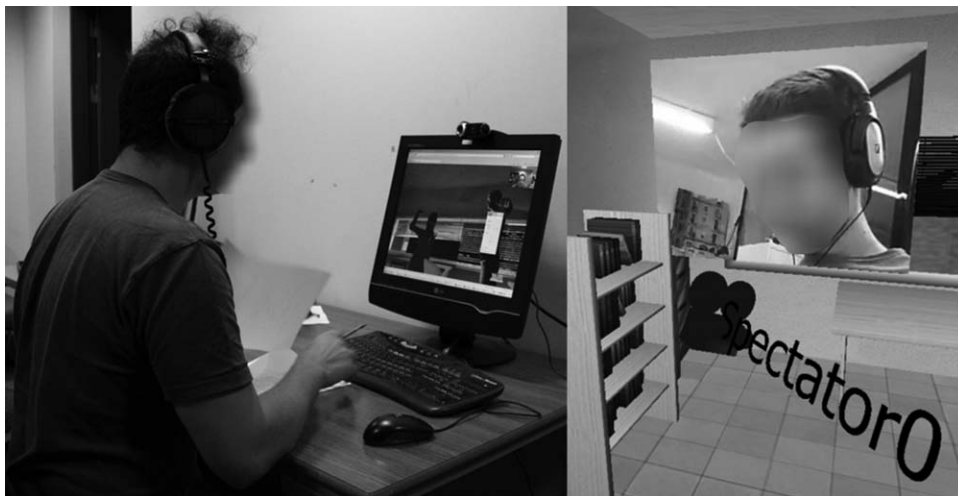
Figure 3. *Director/spectator setup. Left-hand side: A simple desktop/laptop computer is required. Right hand side: Representation of the director in the SVE. A 3D camera represents position and viewing direction, with live video from the web camera displayed on top.*

The director could view and intervene through a desktop system together with a webcam (see Figure 3). The director saw the scene with the two actors displayed on a desktop PC, and could navigate the 3D reconstruction of the scene using a standard keyboard and a mouse interface. The position and orientation of the director with respect to the scenario was displayed in the HMD and in the Cave as a representation of a 3D camera that

moved around the scene according to the view taken by the director. Additionally, the director's face was captured through the webcam, which was displayed in the HMD and Cave representations as a video window located on top of the virtual camera (Figure 3, right-hand side). The director could also speak to the actors via a microphone, his voice being streamed in real time over the network (see Section 3.3).

This 3D camera representation allowed the director to move freely in the virtual environment, in order to supervise the movements of the actors, and to be able to obtain an overview of the whole set, which might not be possible during a more typical rehearsal. This also allowed the possibility for the director to change the scenery, lighting conditions, and so on, thanks to observations made within the SVE.

Additionally, any number of spectators apart from the director could view the scene with the same setup. However, other spectators would not typically be represented or intervene in the virtual environment, with only the director having this capability.

3.2 The Software Framework

The full-body interactive SVE is built upon the XVR virtual reality software framework (Tecchia et al., 2010) and the Hardware Accelerated Library for Character Animation (HALCA; Gillies & Spanlang, 2010) for display and real-time animation of the virtual actors. XVR provides a framework to handle all the displays (HMD, Cave, desktop) and the networking activities related to streaming data, video, and voice over the network arrangement of the various connected peers.

The PC cluster of the Cave in London was also driven by the XVR system via a network-based graphics distribution strategy based on a master-slave paradigm, where a group of four slave PCs manages the rendering of each one of the walls. A "sort late" strategy is used to distribute the rendering task (Marino, Vercelli, Tecchia, Gasparello, & Bergamasco, 2007). In this approach, the application runs fully only on the master node, while the slave nodes receive just the OpenGL calls generated, to be used for rendering on the specific screen for which they are responsible. Compressed video streams

are sent to the master node only, where decoding takes place, and is then sent to the various nodes to be distributed on the local network in the form of OpenGL texture maps.

The HALCA library was used to animate and render the avatars. High-quality avatars from AXYZ Design[7] were used. HALCA provides functionality to retrieve and set the skeletal state of the virtual characters, which is important for the transmission of whole body motions over the network. The skeletal state of the virtual characters is stored as a root position and quaternion rotations for root orientation and all skeletal joint rotations. A typical avatar skeleton consists of around 70 skeletal joints, for each of which four float values describing a quaternion are used. Assuming 4 bytes per float, the required information adds up to approximately 1 KB per virtual character pose. The system transfers the whole skeletal state in order to make it independent of the tracking or IK methods used to control the virtual character.

The actor's motions acquired by the ARENA[8] software from NaturalPoint were sent to the HALCA library via NaturalPoint's NatNet protocol in order to animate the avatar representing the actor. Those motions were then transferred via the XVR network system to every other peer connected to the system. Additionally, the movements of the actor in the Cave were also transferred to all the other peers connected in order for everyone to be able to see the movements performed by the actor. Similarly, the position and orientation of the camera representing the director and each spectator is streamed to every other peer on the network.

3.3 The Network Peer Architecture

Each participant in the rehearsal was represented as a peer in the network architecture. Each actor, director, and spectator connected to the server, the purpose of which was to transmit all information (motions, voice, video from a web camera, etc.) to every connected peer.

The computing nodes in the system exchanged data over the internet using the IP protocols through a centralized server. This arrangement was selected over a

7. http://www.axyz-design.com/index.php
8. http://www.naturalpoint.com/optitrack/products/arena
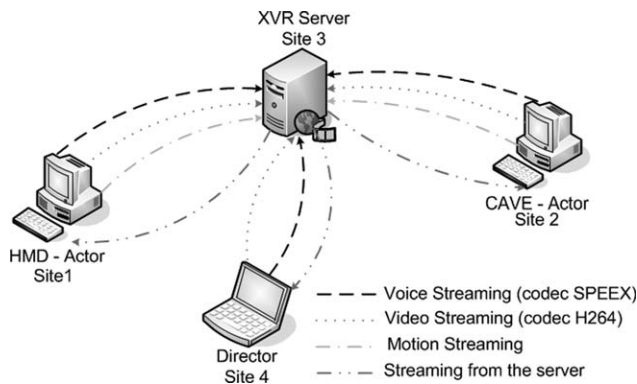
Figure 4. *The network configuration.*

peer-to-peer solution as it is more firewall-friendly and allows for a simple organization and ordering of the real-time streams sent around. A standard internet connection was used, with overall latency between the universities on the order of 15–20 ms.

Custom transmission protocols were developed on top of standard UDP calls, in order to have full control over critical aspects such as payload handling, latency control, and error correction. In particular, we avoided applying full-force error recovery for dropped UDP packets, as this would have caused 3D animation jitter, and opted instead for a time-bounded error recovery scheme (retransmissions are only allowed within a limited time window—80 ms in this case) and also the use of forward error correction by data redundancy. The resulting network infrastructure allowed the streaming of audio, video, and general data between computers, even though the endpoints were behind different private networks. The Google VP8[9] open-source codec was used for video compression/decompression, and the SPEEX[10] codec was used for audio streaming.

While CBR was used for SPEEX packets (resulting in 24 bytes of data for each 20 ms of audio), VBR for video transmission was adopted, resulting in highly variable frame data (from 300 bytes to 25 KB per frame, depending on frame type or content). In this system, data packets with a payload exceeding 450 bytes are fragmented in several subpackets by the protocols in order not to

9. http://www.webmproject.org/
10. http://www.speex.org/

exceed the typical MTU of the underlying network layers.

As shown in Figure 4, each peer connected to the application sent different streams to the centralized server. Video and audio streams were sent to the server from the director peer and the actors' peers. Moreover, the actors' peers had to send the motion of the avatars (either from the full body motion capture in the case of the HMD actor, or only from the IK for the Cave actor). These motions were represented by the positions and rotations of each bone of the skeletons of the avatars. The server received all the streams before sending them back to every peer connected to the network. Depending on the type of peer (actors, director, spectators) and the manner in which they should be handled, only a subset of the streams was used; for example, the video streams from the actor peers were not used. Due to the general low latency of data transmission and despite the large amount of data exchanged, participants in various locations did not perceive latency effects during their rehearsal, directing, or viewing experience.

3.4 Playback Mechanism

The system offered interesting possibilities for rehearsal; for example, a playback mechanism of the actors' motions, which could be used afterward by the director to give advice to the actors, for example, to correct movements or motions of the actors, and thus improve blocking and stage movements. One major advantage of VR over traditional rehearsing methods is that the acting can be viewed from any position and angle, to enable a more sophisticated means for the analysis and therefore possible enhancement of the performance.

A playback mechanism was implemented that enabled recording of the motions of the avatars while they were streamed to the network server. Each peer recorded the movements of its own avatar before sending the information across the network. Meanwhile, each peer also recorded the motions of the other avatars via the server. This allowed the movements of all the actors involved in the rehearsal to be stored on each peer. This mechanism would be useful for directors to replay the rehearsal, since it would allow the analysis of body language, movements, and

gestures of the actors. The director of photography could also use this feature to try different types of shots for the same scene, without having to ask the actors to play it again. This feature thus allows some extra savings for the stage play or movie shoot, both in terms of time and of money.

3.5 Implementation Details

Due to the flexibility of the XVR framework, the system's implementation was relatively straightforward. XVR offers a simple interface for video, voice, and data streaming over a standard network with a relatively low latency (see Section 3.3). The HALCA library used for avatar animation offers an interface to control skeleton-based 3D characters, which can be used within the XVR framework by means of a DLL. We developed a simple IK module on top of the HALCA library that allows control of an avatar's limbs based on a target position for each end-effector (hands or feet) of the avatar. The target positions are obtained via XVR by retrieving the positions of the Intersense wands inside the Cave. XVR offers a simple interface to obtain the positions and orientations of devices connected to the computer via the VRPN (Taylor et al., 2001) protocol. XVR is also designed so that the same program is able to transparently control the HMD, Cave setup, or desktop system without coding changes for the different systems.

The main difficulty resided in the implementation of a mapping interface of the full body motion capture data obtained via the NatNet protocol to the HALCA avatars since each software system has a different skeletal representation. This was implemented as a C++ DLL as described in Spanlang, Normand, Giannopoulos, and Slater (2010).

The IK method implemented was based on trigonometry only; the size of the avatar was modified according to the height of the actor in order to match the size of the avatar's limbs to those of the actor. The mapping of the actor's movements in the Cave to the avatar representing the Cave actor were limited since rotations of the bones were not taken into account; only positions of the wands were used while computing IK.

Finally, in the particular implementation in the Cave we did not allow the actor *M* to move around, since we did not track his leg movements, and therefore translation would have looked strange, with the avatar seeming to slide around. This was an accidental technical limitation of what equipment was available at the site.

4   The Scenario

We invited a professional actor *M* along with an amateur actress *K* and an amateur director *E* to participate in a virtual rehearsal. The chosen script was from the movie *The Maltese Falcon* and consisted of a 10-min dialogue between *M* playing the role of Sam Spade and *K* playing Brigid O'Shaughnessy. The story concerns a San Francisco private detective's dealings with three unscrupulous adventurers who compete to obtain a fabulous jewel-encrusted statuette of a falcon dating back to the Knights Templar.

M (Sam Spade) was located in the Cave at UCL in London. *K* (Brigid O'Shaughnessy) acted while wearing the full body motion capture suit 1138 km (707 mi) away, in Barcelona, Spain. *E* (the director) was in the same building as *M*, but not in the same room, using only a desktop machine with a web camera. During the performance, *M* could see *K*'s avatar moving in a virtual room, and being in a Cave his own body was visible. *K* was immersed with the HMD in a virtual environment where an avatar representing *M* was displayed. She could see a virtual body representation of her character from a first-person position; that is, if she looked down at herself she would see her virtual body, and due to the body tracking, the virtual body moved as she moved. The director *E*, using the desktop system, could change his viewpoint freely in the virtual environment in order to see the scene and be able to make comments or give instructions to both actors. Moreover, both actors could see the video from the director's webcam displayed in the SVE. Additionally, there was one spectator who was located in Sydney, Australia.

The rehearsal took place in July 2010 over several hours including preparatory work. In keeping with traditional rehearsal, the actors were familiar with the script but had not learned their lines beforehand, and hence frequently had to consult the printed version. The director took an active role in positioning the actors, suggesting postural

Figure 5. *Illustration of the rehearsal.* M *is in the Cave while we can see K's avatar displayed stereoscopically on the wall.*

and gestural movements, and taking care of overall blocking and timing considerations.

The proceedings were informally videorecorded by a number of people, and a short movie was made using this raw material.[11] The video especially illustrates the interaction between the director and the actors. Note that the actors were holding the scripts in this rehearsal, which can be seen with respect to the male actor directly, and the pose of one of the arms of the avatar representing the remote female actor.

Figures 5 through 7 illustrate various aspects of the rehearsal: physical setup in the Cave where some spectators attend the rehearsal (see Figure 5), and actors (both local and distant) being directed by the "virtual" director (see Figures 6 and 7).

5 Results

5.1 Network Performance

The network infrastructure performed well: with error correction disabled, the packet loss percentage was

11. http://www.youtube.com/watch?v=c9bLWQhbJz0

in the range of 2–3% with a point-to-point latency of 80 ms. When using error correction, the packet loss percentage went down to 0.5%, while latency increased to 200 ms (due to the time-bounded error correction and packet duplications), ensuring good audio-video quality and at the same time an acceptable latency for real-time applications. Application-level lag is obviously influenced by a variety of factors, not only network latency. Both for video and audio the overall perceived delay is computed as the sum of the capture delay, of the encoding and decoding processes as well as the rendering delay. Still, network latency (especially when it involves retransmissions) plays a fundamental role in the context of SVE. Latency was measured using the same PC as the source and the sink of packets. Packets were sent across the internet to servers of other institutions, each being assigned a unique ID; packets sent to the server were bounced back to the origin and the round trip time (in milliseconds) was measured.

5.2 Participants' Interviews

After the performance, both actors (*M* and *K*) and the director (*E*) were interviewed in order to have some feedback on the setup, and particularly on the impact of the full body motion interaction versus IK. The questions were open in order to let the participants express freely their feelings on the experience of the virtual rehearsal. The questions asked were:

- "How do you feel?"
- "Do you think it is important to know the location of the (other) actor(s) and to change your location according to the way the other actor changes his/her body?"
- "How did the whole experience feel?"

The actor *M* said that "*it was really interesting*" and that "*if you had two people in body suits . . . you could really experiment with the blocking of a scene for a film*" even if "*you had actors in different locations*" and the director need not even be present. He added that "*it would be such a useful way because you really get a feeling of the scene.*" He emphasized the fact that the

Figure 6. *Top: E is directing the actors; he asked K to turn around. Bottom left to right: the actress is turning around following the director. NB: These pictures are taken from the video available on YouTube and were taken during the rehearsal in the Cave system. As a consequence, the pictures seem blurred because the Cave system uses a double rendering needed for stereoscopy.*

whole body interaction is the key feature of the system. Indeed, he said that he first thought that "*if you do not have the actual person's* [facial] *expressions you are not going to be able to really play the scene for what it is worth, but I was really shocked and taken aback*" by the fact that "*you can* [play the scene for what it is worth] *and it is really an interesting technology to be developed.*"

For her part, the actress *K* said that "*it was a bit strange at the beginning*" to "*refer verbally and bodily to somebody who is virtually there while you are in a completely empty room, addressing no one.*" She continued, "*it seemed a little bit weird interacting with somebody lacking physical substance*" but that "*after a while you forget about this incongruence and you feel like being in the same but somehow different place as M.*" She said it was hard to express but that she "*strongly feel being in that room with the other actor*" (i.e., *M*). *K* also told us that "*acting with this setup was more*" like normal acting than "*practicing alone.*" She

Figure 7. *Top image: E is directing the actors. He tells K to put the back of her hand to her forehead in order to emphasize the dialogue. Bottom left to right: Movement of K's right hand in order to follow the director's request. As in Figure 6, the images appear blurred because of the Cave's stereoscopic rendering system.*

stressed the importance of the full body motion: "*the fact that when you move the virtual body moves correspondingly gives you the sensation of actually interacting in this virtual space.*" Moreover, "*seeing the virtual representation of M moving and interpreting the play not only verbally but also using gestures enhances the sensation of being there with him.*" She pointed out some problems related to the tracking of *M*'s IK movements "*there were moments when M's movements were somehow lost, this broke a little bit the whole atmosphere,*" but said that "*during most of the time I did have the sensation of being inside the same room.*" Finally, she emphasized the importance of knowing the location of *M* during the rehearsal: "*knowing his location was absolutely essential! Especially because the whole performance is based essentially on body movements, gestures,*

and postures. Knowing where M was directly impacted my movements. For example, when seeing M on my right, I was looking and gesturing towards my right. During the rehearsal I had to 'approach' M, I did so by reducing the distance between M's virtual body and my own virtual body even though I knew that M was never close to me."

The director *E* told us that "*the experience of directing a scene in virtual reality was very interesting.*" He emphasized the importance of the possibility of navigating freely in the 3D environment. "*The immersiveness that I felt by being able to move through the scene without obstructing the actors from their performance was an unprecedented advantage. This way I could move around the scene and choose the best possible camera positions independent of where the actors where*

standing. The immersiveness of the interaction with the camera was very impressive, making it possible for me to be anywhere I wanted within the scene and experiencing different camera positioning and shots. Of course in an actual set they would have to be set at specific positions according to the background environment, but I guess that this could be easily manipulated if I could have control of the background environment too (i.e., rotate it, change it, etc.)." He also described the interaction with the actors: "*the way I interacted with the actors was sufficient as even when I wanted to convey some postural information or movement, the actors could see me on an overlaid display within their views.*" Nevertheless, *E* pointed out some limitations due to the actors' motions (full body tracking or IK), "*What I missed a lot was the actors' avatars' motions as they were not very good and consistent with the real actors' motions (although the female actress' motions were superior to the male one's). Furthermore there were no facial expressions to the avatars, which is the most important drawback of the whole experience.*" Finally, he gave us his insight on the overall experience as an amateur director: "*Overall I think that it was a very interesting experience and that with some improvements it can be very useful from a director's point of view.*"

From their feedback, it is clear that the most interesting configurations of the system would be when most of the actors would wear body motion capture suits. Indeed, as *M* pointed out, "*body language is so crucial and there are specific directors that work very specifically*" [on body language]. "*In film and TV you have specific vision and image . . . it should be a perfect way to practice.*" He also pointed out that even if the director cannot take part in the virtual rehearsal, this system would be very interesting for actors who have a very busy schedule, because it would give them more flexibility to try to arrange rehearsals.

6   Discussion

The participants did not complain about any kind of lag in motions and/or voice/video transmission.

Nevertheless, further study is required regarding the performance of the network communication, in particular in terms of tolerance over the network's latency or the minimal ratio quality/latency required.

One disadvantage of the Cave system in this context is that virtual objects cannot occlude the actor's real body, that is, be between the eyes and hands. This could be problematic if close interactions were part of the rehearsal. Additionally, the Cave does not allow body substitution of the actor's real body by a virtual one, from his or her own perspective. Normally this would not be a problem, since the other remotely located actor would not see the Cave actor's real body but whatever avatar is representing that actor. However, if in the actual performance the actor would play a role with a visual appearance different from his own, then the use of a HMD would be preferable, if it is important that the actor sees himself as he is portrayed to the remotely located actor. Such body substitution systems have recently been explored, and can give the strong illusion of having a different body (e.g., Petkova & Ehrsson, 2008; Slater, Spanlang, Sanchez-Vives, & Blanke, 2010; Normand, Giannopoulos, Spanlang, & Slater, 2011).

Full body tracking of each actor is the preferred setup whenever possible. This gives the actors a much greater possibility of acting using their full body movements, and also gives the director more control over the rehearsal. For example, as can be seen on the video, the director interrupted the rehearsal whenever he felt necessary in order to correct postures and movements of the actress *K*, while his comments were minimal regarding *M*'s motions since he knew that only arm movements were taken into account by the system. Conversely, *M* was not convinced about the use of IK since arm movements did not seem to him to be enough to convey sufficient information while rehearsing. Our view is that IK based on limited tracking capability would be better than no body tracking, and it has been shown that quite good inference of body movement can be achieved with a minimal number of trackers (Badler, Hollick, & Granieri, 1993). This view of the importance of full body tracking was corroborated by the interviews that were conducted after the experiment. Both actors pointed out that having two actors in body suits would be much more

interesting in terms of rehearsal since not only dialogue but also the full blocking and movements of the scenes could be rehearsed.

One of the main limitations of the current implementation was the lack of the capability for physical interaction between the two remote actors. However, such a virtual rehearsal system could quite straightforwardly incorporate some degree of haptic feedback and be integrated into the avatar system (Spanlang et al., 2010). A collision detection mechanism could be integrated with the avatars and used for haptic rendering. An example of such an integration can be found in Spanlang et al., where exact collision detection is performed via a geometry shader and resulting collisions are mapped to vibro-tactile feedback provided through a reconfigurable vest. Other systems based on vibrators (Lindeman, Page, Yanagida, & Sibert, 2004; Bloomfield, & Badler, 2008), solenoids, actuators, or pneumatic cells (TNGames, 2008) could be used in order to give haptic feedback to the actors.

Although we wanted to explicitly demonstrate the flexibility of the system by using two levels of body tracking (full body tracking in the HMD compared to hand and head tracking in the Cave), the professional actor involved was particularly excited about the use of full body suits for motion tracking. Indeed, the use of full body motion would be the best choice to effectively rehearse with another actor. In order to offer the director the possibility of really rehearsing not only dialogue but blocking to a fine level of detail involving gestural components, the use of full body tracking is required.

We considered adding the possibility of supporting handheld devices such as mobile phones and tablets as the interface. Although this would be useful for the director, the feedback from *M* clearly pointed out that this would not be useful for the actors. He was quite thrilled about being able to use body language and movements during the rehearsal, which obviously forbids the use of mobile devices for the actors.

It would clearly be possible in this type of system to give the director the capability to dynamically control the virtual environment, for example, to be able to initiate visual effects (fire, smoke, introduction of extras), and play prerecorded sounds. Such features would add

significantly to the rehearsal and would give the director and actors a better understanding of the whole scenario. Of course, this would require some special development and user interfaces in order to be easily usable by the director during the rehearsal.

As mentioned previously, the system did not include haptic support; however, giving the actors the possibility to exchange or pick up theatrical props or objects would be useful. Some physical simulation combined with haptics rendering techniques could be used in order to allow such interaction.

Although it seems more difficult to handle face tracking, especially in the HMD, some work (Towner & Slater, 2007) has shown that it is possible to reconstruct facial expressions even if the top half of the face is occluded (which would be the case when wearing an HMD). An alternative solution would be to acquire simple emotional states of the actors from physiological measures (e.g., Palacios & Romano, 2010), and map them to prerecorded avatar facial expressions.

In the particular system described here, the actors could not move around much due to the limitations of tracking. However, in general, the physical size of the tracking area of the VR system need not be a great problem, since metaphors have been developed, such as walking-in-place and redirected walking, that allow virtual movement far greater than the corresponding physical real movement (Slater, Usoh, & Steed, 1995; Razzaque, Swapp, Slater, Whitton, & Steed, 2002; Wendt, Whitton, & Brooks, 2010). Other techniques such as scale translational gain (Williams et al., 2006), the seven league boots (Interrante, Ries, & Anderson, 2007) metaphor, or the use of distracters (Peck, Fuchs, & Whitton, 2009) also allow locomotion through large spaces. Another option would be to use an omnidirectional treadmill (Souman, Giordano, Frissen, Luca, & Ernst, 2010) that allows the participant to move naturally both in the physical and in the virtual space.

7 Conclusions

In this paper we have presented a flexible shared virtual environment for theatrical rehearsal: the system allows for various types of displays (ranging from immersive VR

to normal desktop displays) and different levels of partici-pant interaction (from full body motion capture to classical keyboard and mouse). The system allows actors and direc-tors to meet in an SVE in order to rehearse scenes for a play or a movie. This virtual rehearsal offers flexibility (e.g., when the actors and the directors cannot be present in the same place at the same time) and could reduce costs for the production of the play or the movie. Different interaction techniques were implemented to study their re-spective importance for virtual rehearsal. First, the actors can wear a full body motion capture suit that allows the system to stream his or her movements over the networks to the other actors, directors, or spectators, connected to the SVE. Second, we developed a setup where full body motion tracking is not available. In this case, a simple IK method was implemented to give some degree of tracked body movement. Finally, we provided a desktop version designed to be used by directors (or remote spectators) for whom no motion tracking is needed.

Moreover, there is a playback mechanism that records the movements of the avatars representing the actors. These replays can then be used by the director to work on gestures and movements of the actors in the scene. They could also be used by the director of photography for a movie to try to find interesting shots and viewing angles, without needing the actors to perform the scene another time. In this context, the system could also be used as a networked extension movie pre-visualization tool, allowing actors to record motions while not being at the production site.

According to the professional actor involved, the main feature of the system was to provide full body interac-tion, offering the possibility not only to rehearse the dia-logue (which may be the least important part in rehears-ing), but also, blocking. Indeed, even though facial expressions were not supported, the professional actor believed that body language by itself is very useful and offers a first step toward virtual rehearsal.

A possibility to enhance the current implementation would be to display the facial expressions and eye move-ments of the actors. Recent work (Steptoe, Steed, Rovira, & Rae, 2010) showed that eye movements are important for avatar-mediated communication. Eye tracking could be included in the Cave as well as in

HMD setups, since they can be mounted within stereo-scopic shutter glasses (Cave setup) or in an HMD. Data obtained from eye trackers could then be used to map actors' eye movements onto the avatars, in order to increase interaction. In *Avatar*, James Cameron and his team[12] used a specially developed technology based on head-mounted cameras that captured the facial expres-sions and the eye movements of the actors on stage. In addition, currently, the appearance of the virtual actor is different from the one of the actual actor; another improvement would be to digitize the actor's appearance in order to create faithful virtual actors in those situa-tions that required this.

We believe such a system could help directors and film crews to rehearse scenes when it is impossible for the crew and/or the actors to meet; for example, due to incompatibilities in principal actors' and/or director's calendars. Such a need was pointed out by the BBC, which organized a project that led to the development of the system presented in Slater et al. (2000). Immersive systems such as the one presented in this paper could greatly improve the rehearsal experience compared to existing systems or to videoconferencing software (e.g., Skype). Indeed, a simple videoconference that cannot give a first-person experience of spatial relationships would not support the actors and directors in rehearsal for blocking, that is, deciding the positioning of every actor at any time in the scene, determining the directions that actors should be facing, or which action he or she should be doing.

This work has laid the basis for and has been followed by another setup and study investigating acting in collab-orative multimodal mixed reality environments (Steptoe et al., 2011). The system described in the current paper directly led to the second iteration of our system for vir-tual acting. In the new system, the Cave was not used, and it was not a shared virtual reality system. Instead, while the actor in Barcelona still experienced the rehearsal from within an HMD-based virtual reality system, she was represented directly as a projected avatar in the physical space of the actor in London.

12. http://www.cgw.com/Publications/CGW/2009/Volume-32-Issue-12-Dec-2009-/CG-In-Another-World.aspx

Acknowledgments

References

Badler, N. I., Hollick, M. J., & Granieri, J. P. (1993). Real-time control of a virtual human using minimal sensors. *Presence: Teleoperators and Virtual Environments, 2*(1), 82–86.

Bloomfield, A., & Badler, N. I. (2008). Virtual training via vibrotactile arrays. *Presence: Teleoperators and Virtual Environments, 17*(2), 103–120.

Cairco, L., Babu, S., Ulinski, A., Zanbaka, C., & Hodges, L. F. (2007). Shakespearean karaoke. *VRST '07: Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology,* 239–240.

Geigel, J., & Schweppe, M. (2004). Theatrical storytelling in a virtual space. *SRMC '04: Proceedings of the 1st ACM Workshop on Story Representation, Mechanism and Context.* Retrieved from http://doi.acm.org/10.1145/1026633.1026642

Gillies, M., & Spanlang, B. (2010). Comparing and evaluating real-time character engines for virtual environments. *Presence: Teleoperators and Virtual Environments, 19*(2), 95–117.

Ichikari, R., Kikuchi, K., Toishita, W., Tenmoku, R., Shibata, F., & Tamura, H. (2010). On-site real-time 3D match move for MR-based previsualization with relighting. *ACM SIGGRAPH 2010 Talks.* Los Angeles, California.

Interrante, V., Ries, B., & Anderson, L. (2007). Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments. *3DUI '07: Proceedings of the 2007 IEEE Symposium on 3D User Interfaces.* Los Alamitos, California.

Lindeman, R. W., Page, R., Yanagida, Y., & Sibert, J. L. (2004). Towards full-body haptic feedback: The design and deployment of a spatialized vibrotactile feedback system. *VRST '04: Proceedings of the 2004 ACM Symposium on Virtual Reality Software and Technology,* 146–149.

Marino, G., Vercelli, D., Tecchia, F., Gasparello, P. S., & Bergamasco, M. (2007). Description and performance analysis of a distributed rendering architecture for virtual environments. *ICAT '07: Proceedings of the 17th International Conference on Artificial Reality and Telexistence,* 234–241.

Mitchell, K. (2008). *The director's craft: A handbook for the theatre.* Abingdon, UK: Routledge.

Mitter, S. (2002). *Systems of rehearsal: Stanislavsky, Brecht, Grotowski, and Brook.* Abingdon, UK: Routledge.

Normand, J.-M., Giannopoulos, E., Spanlang, B., & Slater, M. (2011). Multisensory stimulation can induce an illusion of larger belly size in immersive virtual reality. *PLoS ONE, 6*(1), e16128. doi:10.1371/journal.pone.0016128

Normand, V., Babski, C., Benford, S., Bullock, A., Carion, S., Farcet, N., et al. (1999). The COVEN project: Exploring applicative, technical and usage dimensions of collaborative virtual environments. *Presence: Teleoperators and Virtual Environments, 8*(2), 218–236.

Palacios, J. A., & Romano, D. M. (2010). A flexible bio-affective gaming interface. *Proceedings of the 3rd AIBS 2010 Symposium: AI & Game.* De Monfort University, Leicester, UK.

Peck, T. C., Fuchs, H., & Whitton, M. C. (2009). Evaluation of reorientation techniques and distractors for walking in large virtual environments. *IEEE Transactions on Visualization and Computer Graphics, 15,* 383–394.

Petkova, V. I., & Ehrsson, H. H. (2008). If I were you: Perceptual illusion of body swapping. *PLoS ONE, 3*(12): e3832. doi:10.1371/journal.pone.0003832

Razzaque, S., Swapp, D., Slater, M., Whitton, M. C., & Steed, A. (2002). Redirected walking in place. *Proceedings of the Workshop on Virtual Environments 2002,* 123–130.

Schroeder, R., Steed, A., Axelsson, A. S., Heldal, I., Abelin, A., Widestrom, J., et al. (2001). Collaborating in networked immersive spaces: As good as being there together? *Computers & Graphics-UK, 25*(5), 781–788.

Slater, M., Howell, A. J., Steed, A., Pertaub, D. P., & Garau, M. (2000). Acting in virtual reality. *Proceedings of the Third International Conference on Collaborative Virtual Environments (CVE '00),* 103–110. doi:10.1145/351006.351020

Slater, M., Sadagic, A., Usoh, M., & Schroeder, R. (2000). Small-group behavior in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments, 9*(1), 37–51.

Slater, M., Spanlang, B., Sanchez-Vives, M., & Blanke, O. (2010). First person experience of body transfer in virtual reality. *PLoS ONE, 5*(5), e10564, doi:10.1371/journal.pone.0010564

Slater, M., Usoh, M., & Steed, A. (1995). Taking steps: The influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interactions, 2*(3), 201–219.

Souman, J. L., Giordano, P. R., Frissen, I., Luca, A. D., & Ernst, M. O. (2010). Making virtual walking real: Perceptual evaluation of a new treadmill control algorithm. *ACM Transactions on Applied Perception (TAP), 7*(2), 1–14.

Spanlang, B., Normand, J.-M., Giannopoulos, E., & Slater, M. (2010). A first person avatar system with haptic feedback. *VRST 2010: Proceedings of the 2010 ACM Symposium on Virtual Reality Software and Technology,* 47–50.

Spanlang, B., Normand, J.-M., Giannopoulos, E., & Slater, M. (2010). GPU based detection and mapping of collisions for haptic rendering in immersive virtual reality. *IEEE International Symposium on Haptic Audio Visual Environments and Games (HAVE 2010),* 41–44.

Steed, A., Slater, M., Sadagic, A., Bullock, A., & Tromp, J. (1999). Leadership and collaboration in shared virtual environments. *VR '99: Proceedings of the 1999 IEEE Virtual Reality Conference,* 112–115.

Steed, A., Spante, M., Heldal, I., Axelsson, A.-S., & Schroeder, R. (2003). Strangers and friends in Caves: An exploratory study of collaboration in networked IPT systems for extended periods of time. *Proceedings of the 2003 Symposium on Interactive 3D Graphics,* 51–54.

Steptoe, W., Normand, J.-M., Oyekoya, O., Pece, F., Giannopoulos, E., Tecchia, F., et al. (forthcoming). Acting in collaborative multimodal mixed reality environments. Submitted to *Presence: Teleoperators and Virtual Environments, Special Section on RAVE/BEAMING 2011.*

Steptoe, W., Steed, A., Rovira, A., & Rae, J. (2010). Lie tracking: Social presence, truth and deception in avatar-mediated telecommunication. *CHI '10: Proceedings of the 28th International Conference on Human Factors in Computing Systems,* 1039–1048.

Taylor, R. M., II, Hudson, T. C., Seeger, A., Weber, H., Juliano, J., & Helser, A. T. (2001). VRPN: A device-independent, network-transparent VR peripheral system. *VRST '01: Proceedings of the ACM Symposium on Virtual Reality Software and Technology,* 55–61.

Tecchia, F., Carrozzino, M., Bacinelli, S., Rossi, F., Vercelli, D., Marino, G., et al. (2010). A flexible framework for wide-spectrum VR development. *Presence: Teleoperators and Virtual Environments, 19*(4), 302–312.

Tenmoku, R., Ichikari, R., Shibata, F., Kimura, A., & Tamura, H. (2006). Design and prototype implementation of MR previsualization workflow. *DVD-ROM Proceedings of the International Workshop on Mixed Reality Technology for Filmmaking,* 1–7.

TNGames, (2008). 3RD Space Vest. Retrieved from http://tngames.com/products/

Towner, H., & Slater, M. (2007). Reconstruction and recognition of occluded facial expressions using PCA. *ACII '07: Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction,* 36–47.

Tromp, J., Bullock, A., Steed, A., Sadagic, A., Slater, M., & Frecon, E. (1998). Small group behavior experiments in the COVEN project. *IEEE Computer Graphics and Applications, 18*(6), 53–63.

Wendt, J. D., Whitton, M. C., & Brooks, F. P. (2010). GUD WIP: Gait-understanding-driven walking-in-place. *VR 2010: Proceedings of the 2010 IEEE Virtual Reality Conference,* 51–58.

Williams, B., Narasimham, G., McNamara, T. P., Carr, T. H., Rieser, J. J., & Bodenheimer, B. (2006). Updating orientation in large virtual environments using scaled translational gain. *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization,* 21–28.