Experiments in Mixed Reality

David M. Krum *[a], Ramy Sadek[a], Luv Kohli[b], Logan Olson[ac], and Mark Bolas[ac]

[a]USC Institute for Creative Technologies, 13274 Fiji Way, Marina del Rey, CA, USA 90292;
[b]Department of Computer Science, UNC-Chapel Hill, Chapel Hill, NC, USA 27599;
[c]USC School of Cinematic Arts, 900 West 34th Street, Los Angeles, CA, USA 90089

ABSTRACT

As part of the Institute for Creative Technologies and the School of Cinematic Arts at the University of Southern California, the Mixed Reality lab develops technologies and techniques for presenting realistic immersive training experiences. Such experiences typically place users within a complex ecology of social actors, physical objects, and collections of intents, motivations, relationships, and other psychological constructs. Currently, it remains infeasible to completely synthesize the interactivity and sensory signatures of such ecologies. For this reason, the lab advocates mixed reality methods for training and conducts experiments exploring such methods. Currently, the lab focuses on understanding and exploiting the elasticity of human perception with respect to representational differences between real and virtual environments. This paper presents an overview of three projects: techniques for redirected walking, displays for the representation of virtual humans, and audio processing to increase stress.

Keywords: Virtual reality, mixed reality, redirected walking, head mounted display, head mounted projector, virtual humans, audio.

1. INTRODUCTION

The Mixed Reality lab develops technologies and techniques for presenting realistic immersive training experiences. Such experiences must represent complex ecologies of social actors, physical objects, and collections of intents, motivations, relationships, and other psychological constructs. It is difficult to completely synthesize the interactivity and sensory signatures of these ecologies. Therefore, the lab advocates mixed reality for training which allows a mixture of real world elements for setting the stage, and virtual elements that can be easily scripted and controlled. The lab currently focuses on understanding and exploiting the elasticity of human perception with respect to representational differences between real and virtual environments. An example of such elasticity is the technique of redirected walking. This approach provides the illusion of representing a large virtual environment in a physically small motion-tracked space by virtually rotating users away from the constraints of the tracked space – typically physical walls.

The lab is beginning a course of research to better understand the interplay between the perception of real and virtual environments with an emphasis on learning where the "seams" or inconsistent artifacts lie. This paper outlines three areas of such research:

- *Stretching Space* is an investigation of redirected walking, which should enable the presentation of virtual spaces that are larger than the physically available space while still supporting physical locomotion.

- *Sharing Space* is an effort to characterize emerging displays that enhance the perception that a virtual human is in the same space as the user. Ideally, users will feel that virtual humans co-inhabit their *volumetric space* due to the social and emotional implications of physical proximity.

- *Stressing Sound* is an audio research project investigating how noise can be subtly injected into a training environment as a stressor, without breaking immersion.

*krum@ict.usc.edu

2. STRETCHING SPACE

Physical locomotion (walking, running, etc.) is essential for individual combatant simulations in which the soldier interacts directly with the surrounding environment. Spatial cognition and other processes of understanding an environment are weakened when only virtual locomotion is provided (i.e. through a joystick or button press) [Usoh1999]. Furthermore, virtual locomotion could result in negative training in scenarios involving dismounted soldiers, like urban combat, where precise physical locomotion and coordination are important. Unfortunately, most virtual environments are limited in physical size and thus cannot allow unbounded physical locomotion.

A variety of hardware has been devised to allow free locomotion while confining the user to a small area. These include omni-directional treadmills, unicycles, and large "hamster balls" within which users can walk or run [Darken1997, VSD2006, Cyberwalk2005, Ackerman2006, Templeman1999, Kaufman2007]. However, there are some differences between normal locomotion and device borne locomotion. Omni-directional treadmills detect the location of the user and apply appropriate translation to bring the user back to the center of the treadmill. Thus, omni-directional treadmill experiences present some time lag related to determining user position and direction, as well as the actuation of a large mechanism. "Hamster ball" style devices induce lag due to the rotational inertia of the large sphere. Both forms of lag can become apparent under sudden changes in speed and direction, requiring changes in locomotion on the part of the user, additional training, and perhaps safety harnesses or other protective gear.

Figure 1: Wide5 Head Mounted Display with a 150 degree field of view.

Redirected walking may provide an alternative to locomotion hardware for limiting the position of the user. Redirected walking is a set of techniques that allow reorientation and repositioning of a freely walking user within the tracking area of a virtual environment.

The Mixed Reality lab employs a head mounted display (HMD) with a 150 degree horizontal field of view within a wide area motion capture system (approximately 35x40 feet) to allow users to traverse large virtual scenes. While this motion capture space is large, it is still limited in size. Therefore, it is easy to devise virtual environments far larger than this space, for example, a large meadow, an office building, or even a scaled-up molecule. As such, the lab is investigating a variety of experimental techniques to compress the virtual representation with respect to the real, tracked space. Emerging ideas and preliminary findings include:

- Applying a scaling factor to all three axes (X, Y, Z) of a virtual space, which uniformly shrinks (compresses) the virtual scene.

- Applying a high scaling factor to the X and Y axes of that space allows users to lean, without taking a step, to traverse the virtual scene.

- Applying the scale factor along a single axis of the space is disconcerting since the scale depends on the direction the user is moving. In this configuration, walking in a circle is difficult, if not disorienting.

- If the scaling is applied only to the forward axis that is aligned with the user's head, then the user can look at an object, and extend her head forward and back to zoom in for more detail. The user can also walk forward while looking up to gain altitude, or move forward while looking down to lose altitude. This scheme functions much like a set of virtual stairs.

- If the scaling is applied to two axes, X and Y, which are aligned with the user's head, the previous advantages exist, with the additional property of also allowing exaggerated side to side movement, which can help in

examining an object in detail. This method allows users to more easily invoke exaggerated zooming as well as more widely spaced horizontal viewpoints.

- Application of scaling along a single axis that is in line with the user's head, but kept horizontal, can provide space compression that does not affect altitude. This type of compression may be less noticeable to the user, which would be useful when redirection should not interfere with immersion in a realistic training scenario.

This work led us to consider techniques to trigger space compression. While a key press or voice command could be used, these actions may not blend well with the task (navigation, exploration, examination, etc.) that is being performed. System configurations considered include:

- A hand-held, squeezable control, used to indicate a continuum of compression factors, enabling rapid traversal, or standard locomotion through a virtual scene.

- Trackers could be placed on the hands. The height of the hands above some predefined threshold, such as the shoulder height, could be used to indicate the level of compression.

- Space between hands could also be used to indicate the compression factor. Shoulder width could represent the default: no compression. Larger distances between the hands would apply larger scale factors, much like the allegorical fisherman exaggerating the size of a caught fish.

- A real time gait analysis algorithm could determine the type of steps that a user is making. Very large, exaggerated steps could invoke high scale factors, allowing users to easily specify when they wish to quickly traverse a space [Interrante2007].

- A tracked, physical turnstile was built to detect when a user was passing through. During the rotation of the turnstile, users experience an additional scale factor on the rotation, providing additional redirection to the user's travel. The interesting characteristic of the turnstile is that the scaled rotation does not occur on axis of the user's head, as in most redirection algorithms, but is instead centered on the turnstile axis.

- A heavy wheeled cart was constructed that detects the pushing force applied by a user. Within a hospital gurney scenario, the user would push the pushcart/gurney through a building. The amount of space compression was dependent on the force applied by the user. Preliminary results point to this being a promising technique. Compression appears to be less noticeable to the user, since the user is either distracted by pushing the cart, or psychologically prepared for higher levels of compression due to the effort exerted on the cart. Effort exerted into the environment may affect how individuals perceive the environment [Bhalla1999].

Figure 2: Force sensitive cart compresses space in this gurney scenario.

3. SHARING SPACE

Virtual characters are often presented as two-dimensional characters on a large screen. Accordingly, they may be perceived as being behind a glass window, as a distant (either perceived or real) character. This framing effectively confines the virtual human to the space behind the screen.

Techniques which enhance the perception that a virtual human is in the same space as the user –ideally such that users feel that virtual humans share the same volumetric space– may enhance the illusion that a virtual character is a sentient entity with whom the user can relate socially. Humans have a strong drive to relate socially with items that display only a glimmer of what can be described as personality. In fact, while many humans may not consciously perceive that they are interacting with unintelligent objects in a social fashion, they often exhibit a tendency to ascribe personalities and emotions to animals, computers, cars, and other machinery. Removing barriers to this tendency can elicit more realistic responses to virtual human characters, making them more effective in supporting roles in training scenarios.

Employing wide field of view displays like one of the new generation of wide HMDs (providing up to 150 degrees), the lab is exploring the ability to create uniquely compelling experiences with virtual humans. Peripheral vision cues, previously unavailable using standard HMDs, engender a variety of spatial cognition and pre-attentive behaviors that appear to be important in engaging users into a situation. Some photographers anecdotally mention that they feel as if they are detached observers of real world events. While some of that disengagement comes from their journalistic role, perhaps the limited field of view of a camera lens may play a part by lessening the immersion that photographers feel with the real world. This line of thought leads to questioning if a narrow field of view HMD could limit the sense of presence of virtual characters.

**Figure 3: SGT Star, a virtual human presented on a trans-screen**

Emerging display techniques employed by the lab to present such characters and early results include:

- Semi-transparent screens, consisting of plastic film, or plastic sheets, coated with a thin film of Mylar™ or similar material.

- Trans-screen material illuminated with stereoscopic images.

- Feathering the edge around projection screen with occlusive props.

- Users flinching when presented with an HMD view of aggressive virtual human moving to occupy the user's space, yet not flinching when presented with similar imagery on a flat screen display.

- Overlaid image of virtual human in a user's personal camera view in a teleconferencing setting leading to the user turning around to seek the virtual human in the surrounding real room.

4. STRESSING SOUND

Immersive training scenarios are often used to expose trainees to situations that are too dangerous to simulate physically. In many such scenarios, interpersonal communication and situational awareness are crucial to success or survival (e.g. in emergency medical, police, or military operations). The addition of audio noise into an environment can serve as a significant source of stress, fatigue, and disruption of communication. Training to function successfully with inhibited communication and increased stress may help trainees better cope with the effects of stress and hearing impairment in the field.
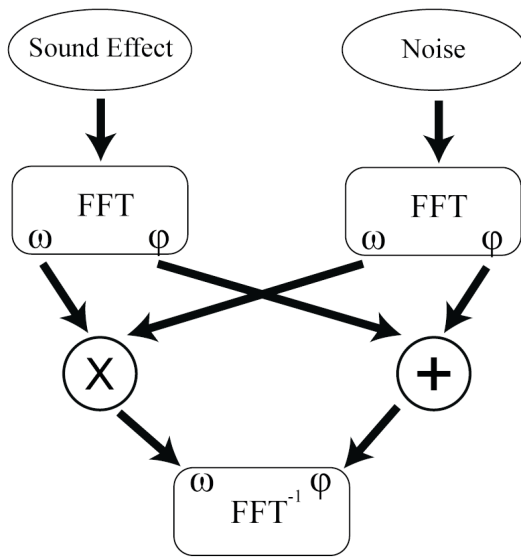
Figure 4: Cross Synthesis for Contouring White Noise as a Masking Sound.

A broadband white noise signal can disrupt communication, since it contains energy at frequencies throughout the human hearing range. However, such a sound is easy to notice over typical ambient noises in an immersive training environment. The lab's goal is to create inconspicuous maskers, since the presence of an obvious noise signal (e.g. audible rumble or hiss) could undermine the realism and believability of the scenario. Thus, the focus of this research effort is to camouflage masking signals so that they remain undetected by trainees.

This approach uses loudspeakers placed in the immersive training environment to add contoured noise, which can mask or "block" content sounds such as speech and environmental audio cues. Contoured noise is noise that has been spectrally shaped. This shaping allows the noise to hide in ambient sounds that share context with the immersive training scenario. Contouring and hiding noise within ambient sound effects, such as an idling helicopter, can introduce significant disruption in the speech band, without increasing the sound pressure to dangerous levels or changing the perceived nature of the sound effect.

To create a contoured noise signal, this approach leverages a technique common in computer music called cross-synthesis. Cross-synthesis imposes the spectral envelope of one signal (the modulator) onto another (the carrier). This approach is the basis for several different synthesis schemes (e.g. different types of vocoders), which employ varying schemes for spectral envelope estimation, modulation, and synthesis.

Preliminary results include:

- Matching the ambient noise spectrum in the manner similar to maskers used in office environments [Chanaud2008] is the first step to hiding noise. This approach allows for a modest baseline level of masking. In order to allow additional masking energy by hiding maskers in virtual sound effects, the masker must share the same spectral envelope as the sound effect. This criterion requires time-varying filtering which we achieved via cross-synthesis. Matching time variance appears to be more important than smooth transitions in these hidden maskers.

- The noise signal must be diffuse; otherwise it is detectable as an isolated sound. Using multichannel audio systems (e.g. 5.1 and 7.1), we implemented separate noise generators per channel to avoid correlation and phase effects, thus also preventing detection via localization.

- Phase vocoding appears to work poorly: the modulated noise signal was easily detectable and its timbre was unacceptable, failing to blend with the virtual scene. Instead, the simple approach of replacing the frequency

component of the carrier signal with that of the modulator, while preserving the carrier's phase produced acceptable results.

5. CONCLUSION AND FUTURE WORK

For the Stretching Space research, several different ways to invoke and use space compression have been enumerated. Most interesting perhaps are interfaces involving mapped exertion. Next steps are to develop further understanding of to encourage users to exert varying levels of force into the environment and how these forces affect perception of locomotion. Furthermore, incorporation of studies showing field-of-view's influence on vection may help guide development. Finally, the influence of wide field of view HMD on redirection is of interest.

Future goals related to Sharing Space using virtual human displays are to closely examine and compare the level of physical and social reactions induced by virtual humans when presented on various display types. Another goal is to determine what factors influence human users to establish particular distances from a virtual human character. Questions we hope to answer include: Is distance chosen to establish the proper social distance to another entity? Do users place themselves to establish proper stereo fusion? Does the need to frame the head, head and torso, or whole body of the virtual human within the user's field of view play a role?

Finally, the Stressing Sound research will next develop masking algorithms which contour noise in real time based on the live sound that in the training environment (e.g. shouts of participants or shots from simulated firearms). Additionally, it will implement a system for selecting and playing relevant sound effects when the ambient noise in the scenario is not loud enough to mask the added noise.

REFERENCES

[1] Usoh, M., K. Arthur, M. Whitton, R, Bastos, A. Steed, M. Slater, and F. Brooks. "Walking > Walking-in-Place > Flying, in Virtual Environments." SIGGRAPH 99, 1999.
[2] Darken, R., W. Cockayne, and D. Carmein. "The Omni-Directional Treadmill: a Locomotion Device for Virtual Worlds." User Interface Software and Technology (UIST), 1997.
[3] "Present." Virtual Space Devices Inc., 2006. http://www.vsd.bz.
[4] Ackerman, RK. "Navy Researchers Target, Virtually." Signal Online July 2006. http://www.afcea.org.
[5] Cyberwalk. 2005. http://www.cyberwalk-project.org.
[6] Templeman, JN. "Virtual Locomotion: Walking in Place Through Virtual Environments." Presence 8 (1999): 598-617. MIT Press.
[7] Kaufman, RE. "A Family of New Ergonomic Harness Mechanisms for Full-Body Constrained Motions in Virtual Environments." IEEE Symposium on 3D User Interfaces, March 2007.
[8] Chanaud, R., *Sound Masking Done Right*, Magnum Publishing, 2008.
[9] Bhalla, M., & Proffitt, D. R. "Visual-Motor recalibration in geographical slant perception." Journal of Experimental Psychology: Human Perception & Performance. 25(4), 1999.
[10] Interrante, V., Ries, B., Anderson, L., "Seven League Boots: A New Metaphor for Augmented Locomotion through Moderately Large Scale Immersive Virtual Environments", IEEE Symposium on 3D User Interfaces, 2007.