Designing a System for Effective Use of Immersive Audio in Mixed Reality

Darin E. Hughes
Media Convergence Laboratory
Institute for Simulation and Training
University of Central Florida

Scott Vogelpohl, Charles E. Hughes
Media Convergence Laboratory
School of Computer Science
University of Central Florida

Keywords: Immersive audio, mixed reality, human-computer interface, sound design, audio engine.

ABSTRACT

Mixed reality experiences often demand technological solutions that are unnecessary or too complex for other applications such as desktop simulations. This is especially true of a mixed reality scenario that attempts to provide the user with an experience that is aurally as well as visually immersive. We discuss the shortcomings of technologies such as DirectSound and OpenAL, which are the typical solutions chosen for an application that must provide dynamic spatialized audio. We then present an overview of audio systems used in mixed reality platforms including our own previous efforts and describe the problems that one encounters in using such technologies in MR. We then describe the design and implementation of the audio system used in our next-generation mixed reality platform and discuss the ways in which it succeeds where other technologies fail.

Our audio library is designed to make it possible to compensate for speaker configurations which deviate from the ideal placement assumed by most 3D sound applications; this is of importance to mixed reality because a complex set may not permit ideal speaker placement. It also allows the sound designer to take advantage of speakers placed above and below the user to provide a better sense of spatialization for sounds in the vertical direction than is possible with speakers in a single plane. Further, it makes it possible for a sound system to contain point-source speakers, headphones, radios, and other unusual devices in additional to a surround sound system, and to run the full audio experience from a single machine, preventing a mixed reality exhibit from having to include several computers to run different aspects of the audio experience.

1. INTRODUCTION

Imagine this. You are a soldier embroiled in a tense hostage situation. The ambiance of an urban environment surrounds you, not permitting you to forget the lives of the civilians that will be disrupted and endangered if the situation is allowed to escalate into full urban combat. A small radio embedded in your helmet near your ear relays orders from command, though it requires concentration to make out the words over the distortion of the airwaves. You hear a Black Hawk helicopter circling somewhere overhead, ready to provide the air support you desperately hope you will have no need to call upon. From a window above you, one of the militants barks an instruction through a megaphone, his voice harsh from the amplification, and you hear a dull clink of metal on concrete from beside you as you gesture to your men to set down their arms. Suddenly, you notice the sound of footsteps emerging from the urban sounds behind you, and you whirl around, wary of an ambush on your disarmed troop, but you see only a civilian woman emerging from a door in the alley behind you. You open your mouth, about to advise her to stay indoors, but then you hear the sound of metal striking flesh above, and the cry of the hostage, and you whirl back to the occupied building, ready for anything….

When you take off your head-mounted display, of course, you realize that you are standing in a Hollywood-style backlot, a room with the flat fronts of buildings; that the characters you have been interacting with are computer generated models, except for the human who commanded the militants from somewhere behind the scenes; and that all the sounds that surrounded you were produced by a small speaker in your helmet, a megaphone mounted on one of the flats, and several more speakers hidden somewhere in the set. But for the minutes that you wore the HMD, it felt completely real. This is the objective of our military simulation, Mixed Reality Military Operations in Urban Terrain (MR MOUT).

Mixed reality is a class of interactive experience that incorporates both real elements (set pieces, other participants, traditional special effects) and virtual elements (computer-generated characters, synthetic audio). And while the visual aspects of such a scenario are of undeniable importance they would not have much impact without the aural environment that accompanied them.

Much work has been done on the unique challenges of seamlessly combining captured video with computer graphics in real time, and the significant issues that arise in doing so, such as occlusion of virtual objects by real objects. But for the most part, audio in mixed reality has been confined to traditional techniques—the prerecorded, premixed tracks used in theme park experiences, or the spatialized environments used in computer games through the use of personal headphones or 5.1 home entertainment systems. Mixed reality requires elements from both of these types of experiences, as well as unique elements such as live sounds being made by real objects in the environment or transmitted from a microphone; a mixed reality audio system must therefore be able to draw upon the capabilities of each of these setups without being subject to their limitations.

Imagine what would be required of traditional audio paradigms to realize the scenario presented above. One computer would be dedicated to run the spatialized audio for each vertical tier associated with the experience. At least one more would manage the various independent sources—more than one would typically be required unless fairly advanced software tools were used to output more than two independent channels of audio. But why is this complexity required? There are certainly sound cards with enough output channels to communicate with all the speakers involved; why is there no software that can run on a single computer and control all these components simultaneously?

To meet the unique challenges of mixed reality, we have designed and implemented a system for use in our MR experiences, drawing upon the shortcomings we have found in attempting to apply more traditional techniques for sound design to mixed reality. Because we were unable to effectively use the audio paradigm assumed by libraries such as DirectSound and OpenAL, and because no other library has significant support for real-time spatialization, our system is built on top of a custom audio rendering library engineered virtually from scratch, using only the low-level PortAudio library for interfacing with sound devices.

2. PAST WORK

The desire to create an easily configurable, powerful audio engine and high level interface came about over the course of designing audio for interactive experiences during the last several years. These experiences include exhibits at SIGGRAPH 2003, ISMAR 2003/2004, IITSEC 2002/2003, IAAPA 2002/2003 and the Orlando Science Center, as well as long-term installations at the US ARMY's Simulation Technology Training Center (STTC, Orlando). Standard media production tools such as Sonar, ProTools, Cubase, etc., while very useful for asset creation and multitracking, can not provide the kind of dynamic control necessary for interactive simulation. In addition to lacking any support for realtime spatialization, they do not have features to compensate for suboptimal speaker placement or expanded, multitiered surround systems. Both the previously mentioned features are essential since many interactive experiences must occur in environments where optimal speaker placement is not possible and where sounds along the vertical plane are essential both for immersion and for accuracy of training.

The shortcomings of using one of these media production tools in simulations are well documented in the Institute for Creative Technologies 2001 audio technical report [5]. Their system was based around ProTools. Sounds were generated dynamically by sending midi triggers to the applications. While this arrangement had some success with "tightly scripted and choreographed scenarios," it was entirely incapable of creating any dynamically created panned sounds. Additionally, their system was further hampered by an inability to trigger more than three premixed sound sets simultaneously.

We also purchased and investigated the AuSIM 3D Goldminer system which is an integrated hardware and software solution for audio simulation. While producing fairly realistic surround impressions, its main draw back for mixed reality is its use of headphone as its delivery method. It is very important to be able to hear real world sounds in an MR experience. This is especially true in a military training scenario where it is not only essential to hear the footsteps and movements of virtual characters but also of other human interactors. More information can be found about AuSIM 3D through their website, http://ausim3d.com.

For the specific demands of a highly dynamic and immersive mixed reality simulation, it became clear that a system must be built either on top of an existing API or from scratch. Some of our early attempts involved the use of the Java Media Framework (JMF) which, while providing dynamic cueing, did not support multichannel output, spatialization, and channel control among many other things [2].

An extensive review of the available technology was conducted including both proprietary and open source, hardware and software solutions. Due to the specific demands of the MR audio, EAX 2.0 was selected as the appropriate environment for building interactive audio experiences. However, this technology was also limited in its ability to address specific channels and provide control of multiple hardware arrangement through a single application. In addition, EAX does not provide the kind of low level support for creation of digital signal processing (DSP) effects—rather a static set of effects are provided. The frustration of working with these technologies due to the particular demands of interactive, immersive simulation led to the design and creation of a custom built audio engine and high-level interface.

3. THE DESIGN

3.1 The components

The delivery system for our mixed reality audio engine is a hybrid arrangement that may contain any combination of multi-channel surround sound systems, point source speakers, directed sound, 3D headphones, or haptic audio devices. This hybrid approach enables a wide range of uses and capabilities to deliver audio for a variety of environments with the potential for large numbers of participants. Sounds can be delivered globally to all participants, selectively to subgroups, or privately to individuals within an interactive experience [3]. The goal of this setup is to provide a dynamic hybrid range via the system's hardware and software that can adapt to a changing environment and to users within the environment, without giving up the ability to hear and mix real audio such as voice communication. In the following sections, the separate components of our hybrid approach to mixed reality audio are described.

3.1.1 Surround sound

Because the placement of speakers in a mixed reality experience is constrained by the design of the physical set, setting up a conventional 5.1 or 7.1 surround system with optimally placed speakers can be difficult. Furthermore, these systems can deliver audio only on a plane level with the listener's head, and provide little sense of the vertical placement of sounds. We avoid both of these limitations by allowing arbitrary placement of speakers, which is then compensated for by the controlling software. The software generates a sound field based upon the assigned speaker locations in relation to the listeners location (which can be dynamically altered based on tracking information) or assigned to a set position (or "sweet spot").

The multi-channel surround sound system can deliver dynamic audio cues across multiple locations in 3D space, including locations above, below and at head level. The addition of multiple vertical levels of speakers allows for a three dimensional audio experience. Vertical placement of sounds is particularly important in mixed reality scenarios where audio events may occur above and below the user, or in places where visual perception may be limited. Additionally, important cues such as footsteps and airborne vehicles can be more accurately modeled with a multi-level system. The goal of sound accuracy is complemented by the listener's enjoyment, as an effective spatial impression is important in obtaining a subjectively pleasing and realistic sound environment [6]. A recent study in our lab involving 39 participants investigated the discriminability between pink noise sounds originating from either head-level or above-head level speakers in a two-tiered configuration. Results indicated that participants were generally able to differentiate between these two locations along the vertical axis. These results suggest that this configuration is effective in delivering y-axis simulation.

3.1.2 Point source and intimate audio

A point source speaker is used when accuracy is critical and audio assets can be delivered at a predetermined position. This typically consists of a speaker device that is not part of a surround system, but is used to fill in the soundscape by providing audio that is either inappropriate or less effective when used as part of a surround system. Point source speakers are used for a variety of applications including personal audio (radio traffic), special effects (audio haptic vest), embedded sounds (haptic gun fire), and intimate audio (voices inside your head). Personal audio includes on-body speakers such as two-way radios, ear buds, and other such devices. These devices are important for delivering private audio including trainer commands, personal interface audio, and other audio associated with personal equipment (e.g., gun shots).

As a special effect, alternative speaker locations can be used to heighten tension and mood. One such example is the addition of a speaker on the participant's body for dramatic effect. Intimate audio refers to sounds that are very close to our personal space and may convey a strong emotional impact or even a sense of violation. One example of intimate audio is an effect played through a helmet embedded speaker behind the head. This intimate audio attempts to simulate extreme close proximity through the use of physically close speakers and creative sound design.

3.1.3 Directed and tracked audio display

Directed sound is a new technology sometimes referred to by the trademark names Hypersonic Sound or Ultrasound. Directed sound encodes audio into a high frequency beam (well above the range of human hearing), which "distorts" in a predictable manner when it collides with a surface and reproduces the encoded sound. The effect of directed sound is that the audio is perceived to be emanating from the source at which it is pointed, not at its source of origin. There are two distinct effects that can be created with directed sound: virtual sound sources, and private audio.

Virtual sound sources are sounds that appear to be transmitted by the object that the directed sound is pointed at and can produce the effect of speakers being embedded anywhere in an environment. Additionally, virtual sound sources can be moved so that a sound appears to travel along a surface or set of surfaces.

Private audio are sounds that are directed at a participant's head. These sounds are perceived by the listener as coming from inside their own head. They have the effect of being dramatically louder to the individual they are directed at than other participants who may be in the general area. Private audio can be used for delivering soft sounds such as whispers or unique effects such as buzzing flies. Additional applications could include simulating internal thoughts in the form of "thoughts-out-loud."

We have conducted informal experiments with both Hypersonic Sound and Ultrasound (Audio Spotlight). The Audio Spotlight, though significantly more expensive, had a greater clarity of sound with a lower noise floor. However, both units are limited in their ability to reproduce low frequency sounds (in fact, most anything below 400 Hz) and also sound best when used at quiet volumes – typically too quiet for many practical applications. For more information of Hypersonic Sound visit http://www.atcsd.com/. For more information on the Audio Spotlight visit http://www.holosonics.com/.

3.1.4 3D headphones

Headphone-based 3D systems are effective tools for creating complete immersion, as they provide the sound designer with complete control over an auditory environment [1]. The 3D headphone system uses an inertial tracking system to place sounds in space relative to the participant's head orientation. Headphone systems without tracking have the major problem of sounds moving as the user's head moves. In other words, if there is a sound source to the right of the listener, when they turn their head to look at it, the sound will still be to the right of them. In this situation, the participant would be constantly turning their head to try to find the sound but would never find the sound in front of them.

A 3D headphone system with tracking keeps the virtual sound in the appropriate location and, when the participant turns to look in the direction of a sound, the sound will be positioned in front of them. The main drawback to any headphone based system is that real world audio is blocked. A truly mixed reality headphone system is currently under development at the Media Convergence Laboratory (MCL) by merging the AuSIM 3D headphone system with external microphones and an intelligent "listening" system that will deliver to the listener the full spectrum of sound from real, to augmented, to virtual.

3.1.5 Haptic audio

Haptic audio refers to sound that are felt more than heard. Such devices include "bass shakers" and haptic vests. They can be used to increase the sense of realism and impact, but can also be used to provide informational cues.

In the case of haptic vests, pressure points can be used to alert participants to the direction of targets or potential threats. These feedback mechanisms can be used in coordination with detection devices. As a personalized tracked audio display within a haptic audio vest, it provides directional cues without cluttering up the already intense acoustic soundscape. With the use of speakers that vibrate more for feeling than for hearing, an intimate communication of stimulating points in the body provides the approximate orientation of events that may not be heard or seen. Thus, the user's attention may be directed to an event by a vibration or combination of vibrations. This information can give an immediate sense of direction and proximity (e.g., by making the vibration's intensity vary with the distance to the event). It works in essence like a tap on the shoulder to tell the user of a direction without adding to or distracting from the visual or acoustic noise levels. This message is transferred to an alternative sense and thus allows for this critical data to cut through the clutter of the audiovisual simulation. With targets outside of the line of sight, this approach can significantly reduce a user's response time.

3.2 The controlling application

Combined together, surround sound, point source, 3D headphones, and haptic audio provide a wide range of devices for displaying the auditory components of a MOUT simulation. With the many challenges of simulating this complex and diverse environment, a hybrid approach delivers the most options to the sound designer.

Designing sound for an interactive experience incorporating all of these elements, however, can be a tedious task that requires knowledge of audio programming and may also require familiarity with a scripting or story engine that drives the overall experience. Typically, a sound designer does not possess these skills and thus requires the assistance of programmers to achieve the desired design. The end result of this process is that changes are often difficult to make and aesthetic control is taken out of the sound designer's hands. In many instances, the client or trainer of a simulation may want to make minor modifications to certain sounds or even major changes to sounds in a particular scenario. This could be a very expensive task requiring additional time and funding for the programming team. What is needed is a high-level interface for total sound design and audio channel control that allows sound designers to modify the sound design and architecture of a sound system. This is the intended function of MR SoundDesigner (Figures 1 and 2).

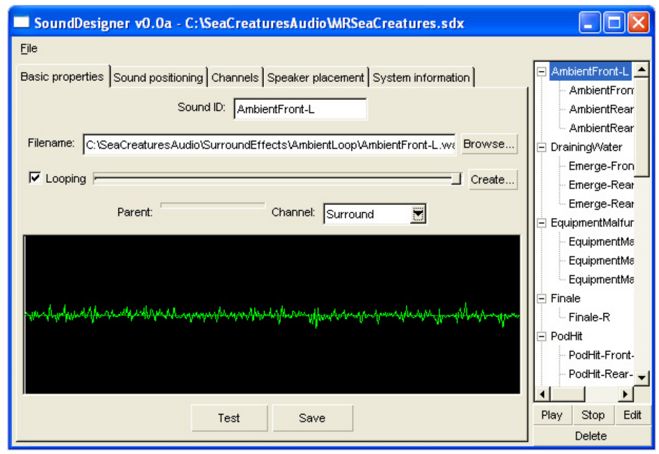

Figure 1. SoundDesigner main tab.

MR SoundDesigner was conceived as an application that allows the user to create or modify entire soundscapes, control all output channels of connected sound cards, designate sound states, and support a variety of delivery systems and classifications. The user of MR SoundDesigner can individually address audio channels and assign some to a surround system while leaving others open for use with devices such as point source speakers. To achieve this, MR SoundDesigner needed to be built on top of an API which allows low level control over audio hardware.

Most computer-generated military simulations lack the ability to be reconfigured easily and quickly. MR SoundDesigner allows non-programmers and audio novices to assert a high level of control over the soundscape and auditory structure of a scenario. This is particularly useful in simulations where various factors such as ambient noise, cueing, expectation, and other important variables can be modified for purposes of evaluation. This software also allows for easy configuration of new audio scenarios or the alteration of previous simulations without the need for reprogramming.

3.3 The paradigm

The implementation of a high-level interface to an advanced audio architecture such as MR SoundDesigner requires the definition of new abstractions to represent system components in an intuitive way. The MR SoundDesigner interface represents individual sound clips in terms of *buffers*, which represent the discrete samples of a sound and *sources*, which represent instances of a buffer currently being played. These are common audio concepts in other libraries such as OpenAL, but MR SoundDesigner is unique in providing an explicit representation of individual *speakers*, which it groups and addresses through the interface of *channels*. Each sound source is played on a specific channel, which is to say that the samples generated by that source are mixed, filtered, and output to the speakers bound to that channel. The two fundamental channel types are *point-source* channels (which simply mix and copy channels to all speakers), and *spatialized* channels (which use information about the position of sounds and speakers to perform a per-speaker attenuation on samples in order to associate each source with a specific spatial direction).

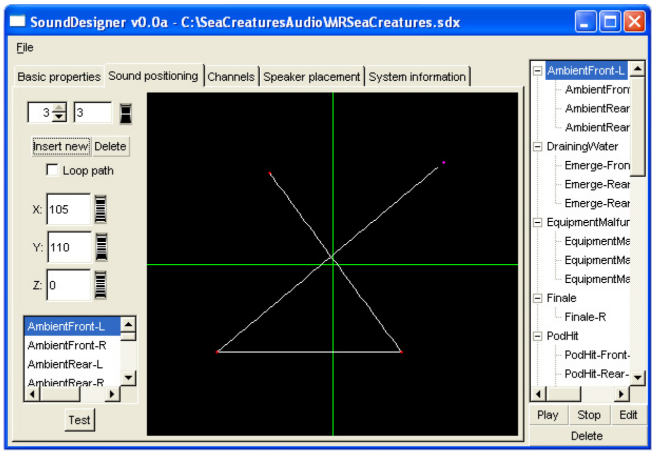

Figure 2. SoundDesigner sound positioning

4. THE IMPLEMENTATION

4.1 Low-level spatialization and rendering

As has been mentioned, there are no existing audio libraries with the capabilities necessary to implement the capabilities of MR SoundDesigner. In order to maintain a distinction between point-source and spatialized channels, the application requires direct control over individual samples in order to selectively apply spatialization algorithms only to those samples intended for spatialized channels and to simply pass through samples on point-source channels; this level of control is not provided by DirectSound or OpenAL.

To provide the necessary low-level support, we implemented a low-level audio rendering library, styled libsd, which provides an object-oriented framework for dealing with the elemental abstractions of the MR SoundDesigner paradigm. libsd is built on top of the PortAudio library, which uses ASIO drivers to provide a simple means for interacting with sound hardware. The programmer interfaces with PortAudio through a callback function that is invoked periodically by a separate thread. The callback function provides a buffer of raw data to be filled by the user's implementation. PortAudio is abstracted into the Stream class in libsd.

The main work of libsd is done in the Channel class. Each channel spawns a worker thread that is invoked periodically to do the work of mixing samples from individual sources and outputting them to speakers. Speakers are implemented as circular buffers in which samples are held until they are read by the PortAudio thread.

The most computationally and algorithmically complex part of the libsd rendering pipeline is the spatialization computation for spatialized channels, which are represented in the Channel3D class. The essence of the algorithm is the computation of a table of attenuation factors for each combination of source and speaker on that channel. This table is computed once per iteration of the channel thread, and then each sample in the block of samples read in that iteration is multiplied by the appropriate attenuation factor in the table. The computation of the attenuation factors is based on the techniques for spatialization described in [4]; the essence of it is that the dot product of the vector to the source and the vector to the speaker is used for attenuation, so that speakers in exactly the direction of the source play the sample at full volume, and the attenuation factor increases based on an increasing angle until at a difference of 90 degrees, the sample contributes nothing to that speaker. This also means that speakers at an angle of greater than 90 degrees—i.e., speakers in the opposite direction of the source, will not play those samples; this does lead to undesireable artifacts—for example, a sudden and jarring transition from the speakers on the left to those on the right when a source moves past the origin along the x-axis—but the simplicity of the technique makes it a desirable one to use despite the artifacts associated with it, and such effects can be avoided as long as the sound designer is careful to avoid having sources move very close to the origin. (OpenAL seems to exhibit the same artifacts, so it seems likely that our approach here is typical of spatialized audio applications.)

4.2 High-level API

Given the libsd library, we then implement a higher-level library to wrap its components in a more user-friendly manner. This is the MR SoundDesigner API, a library that is used both in the MR SoundDesigner application and as a component to be used in other applications that load and play sounds from soundscapes created by MR SoundDesigner.

The main interface point in the API is the Soundscape class. Most of the interaction of an application with the API consists of calls to the Soundscape to trigger sounds by name. The Soundscape maintains a mapping of names to buffers, so that every buffer in the system can be referenced by name; a similar mapping is maintained for channels. Each buffer can also have preconfigured data associated with it, including parameters such as volume, channel to be played on, and spatial position and motion. A call to trigger the buffer by name will play it using those predefined parameters. Triggers of this sort can also reference other "child" buffers, which should be triggered whenever the "parent" is. The Soundscape also provides support for "temporary" sources, which are one-time instances of a buffer created with dynamically assigned parameters and then discarded.

4.3 The MR SoundDesigner interface

The functionality of the API is then wrapped by an intuitive graphical interface, pictured above in Figures 1 and

2. The graphical interface is written using the cross-platform FOX windowing toolkit. The purpose of the application is to provide a simple way for a sound designer to sculpt a soundscape using advanced graphical tools and preview it interactively. Once the sound designer has produced a satisfactory soundscape, it can then be exported in an XML-formatted document, which can be read and referenced by other applications which link to the SoundDesigner API but do not necessarily provide complex interactive interfaces for editing the soundscape.

5. CONCLUSIONS

MR SoundDesigner has been used to create all the audio for the MR SeaCreatures exhibit (an educational and entertaining underwater experience created for the Orlando Science Center) and is currently in use on MR MOUT III. The system allowed for rapid design of the overall audio experience by our sound designer and greater artistic control. The auditory experience for both these simulations is richly layered with both prescripted surround events and dynamically triggered real-time 3D sounds.

Additionally, several sound designers for an interactive performance exhibit called StoryBox at the University of Central Florida have begun using MR SoundDesigner for their experiences. While they have suggested many potential improvements to provide more support for live sound applications, their overall experience has been very positive.

6. FUTURE WORK

MR SoundDesigner and the audio paradigm it represents seem to offer almost limitless possibilities in terms of the complexity of the dynamic soundscapes and audio output configurations that can be created, and the simplicity of creating and modifying them. The current list of features is very much a work in progress; there are already a number of expansions to the system's capabilities that we plan to investigate, and there are surely many more that remain to be envisioned by ourselves or by others in the field. A short list of features currently planned for development include:

- **Timeline triggers** controlled and managed internally by SoundDesigner. Not only do we intend to place triggers on a global timeline, but we will also permit *relative* timelines—i.e., to specify that sound B will be triggered 5 seconds after sound A, whenever sound A is triggered.
- **Multiple sound devices** for simultaneous output. We intend to investigate the possibility of allowing speakers across multiple devices to be placed into the same common pool, and to make the management of this completely transparent to the high-level API. This would allow sound designers an essentially unlimited number of channels to create, bounded only by the computer's capacity for sound devices and the limits of the processor.
- **Dynamic sound buffers**. We intend to redo the class structure of the libsd backend to allow for buffers of types other than sound files loaded from disk Simple examples would be waves computed on-the-fly with dynamically changeable frequency and amplitude, or feeds directly from microphones or other input sources—and perhaps the far future will bring even more interesting possibilities, such as on-the-fly voice synthesis.
- **Filters and environmental effects** such as reverberation and equalizer support. Truly complete soundscapes can hardly be engineered without some investigation into these areas. It remains to be seen how much of this can be done within the software rendering engine of MR SoundDesigner, and how much must be done in the setup of the physical components of the system through the use of dedicated DSP hardware (possibly with support for control of those components from within MR SoundDesigner through MIDI signals).

Beyond our immediate plans, it is our hope that the paradigm realized in MR SoundDesigner will find applications in projects other than our own. In mixed reality, as well as in other categories of experiences that require dynamic sounds in all three dimensions and on multiple channels, there is a need for tools that provide a sound designer with the level of control necessary to maintain a pleasing aural experience, as well as the degree of freedom necessary to make the experience fully dynamic. To our knowledge, there is no tool other than MR SoundDesigner currently available to provide these capabilities. We await the emergence of other tools of this type, to further refine the possibilities of dynamic audio. We await the development of hardware that implements the operations currently performed in software by libsd, so that the processor can be freed up for more complex operations such as DSP and synthesis. Most of all, however, we await the recognition of immersive audio as a part of interactive experiences as critical as stunning visuals and as worthy of the attention of developers. MR SoundDesigner is our contribution to that future.

ACKNOWLEDMENTS

This research was conducted within the Mixed Reality Audio research test bed at the University Central Florida's Media Convergence Laboratory, a partnership of the College of Arts and Sciences, School of Computer Science, and the Institute for Simulation and Training. The research is in participation with the Research in Augmented and Virtual Environments (RAVES) supported by the Naval Research Laboratory (NRL) VR LAB. It is also supported by the U.S. Army's Science and Technology Objective (STO) Embedded Training for Dismounted Soldier (ETDS) at the Research, Development and Engineering Command (RDECOM). Collaborative research and an ongoing dialogue are also taking place with the Army Research Institute (ARI) at the Science and Technology Testing Center (STTC) in Orlando, Florida. Finally, special thanks are due to Dr. Michael Moshell for his vision and guidance, and to the Mixed Reality Laboratory, Canon Inc., for their generous support and technical assistance.

REFERENCES

[1] Begault, D. R. 1999. "Auditory and Non-Auditory Factors That Potentially Influence Virtual Acoustic Imagery." In *AES 16th International conference on Spatial Sound Reproduction* (Rovaniemi, Finland, April 10-12), AES, New York, NY, 13-26.

[2] Hughes, C. E.; Stapleton, C. B.; Micikevicius, P.; Hughes, D. E.; Malo, S.; O'Connor, M. 2004. "Mixed Fantasy: An Integrated System for Delivering MR Experiences." In *Proc. of VR Usability Workshop* (Nottingham, England, Jan. 22-23). [Online]. Available: http://www.view.iao.fraunhofer.de/Proceedings/ [November 19, 2004].

[3] Muller-Tomfelde, C. 2002. "Hybrid Sound Reproduction in Audio Augmented Reality." In *AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio* (Espoo, Finland, June 15-17), AES, New York, NY, 58-63.

[4] Naef, M.; Staadt, O.; Gross, M. 2002. "Spatialized Audio Rendering for Immersive Virtual Environments." In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (Hong Kong, China, November 11-13), ACM Press, New York, NY, 65-72.

[5] Sadek, R. 2001. *3D Sound Design and Technology for the Sensory Environments Evaluations Project: Phase 1*. [Online]. Available: http://www.ict.usc.edu /publications/ICT-TR-01-2001.pdf [November 19, 2004].

[6] Soulodre, G. A.; Lavoie, M. C.; Norcross, S. G. 2003. "Objective Measures of Listener Envelopment in Multichannel Surround Systems." *Journal of the Audio Engineering Society, 51, no. 9*, (September) 826-840.