Mobile Audio Designs Monkey: A Tool for Audio Augmented Reality

Bruce N. Walker and Kevin Stamper

Sonification Lab, School of Psychology
Georgia Institute of Technology
654 Cherry Street, Atlanta, GA, USA 30332-0170
bruce.walker@psych.gatech.edu,
kstamper@cc.gatech.edu
http://sonify.psych.gatech.edu/index.html

Abstract. There are many audio design tools available that have filled some of a designers' needs, but these all are designed with context-ignorant audio in mind. To develop audio for a context-aware environment, at present a designer must have technical skills that are unrelated to the design process. Whereas the designer should be focusing on the creativity, design, and the logic of the Augmented Reality (AR), he or she is forced to spend time worrying about the details of the audio. To support this creative design process more effectively, and at the same time facilitate location-aware audio design, an audio AR designers' tool called Mobile Audio Design (MAD) Monkey was developed.

1 Introduction: The Need for Design Tools

In today's world of computing devices, there is a substantial amount of computer audio. There are many audio creation, editing, and manipulation tools currently available for composers and sound designers. These tools allow for complete control over the sound generation and manipulation. Unfortunately, though, there are no tools that provide appropriate functionality to designers of context-aware systems.

Currently, users carry around laptop computers, and are starting to use devices such as the Sidekick and Treo, which are, among other things, a cell phone, web browser, and email client. Users have also started wearing Bluetooth headsets that provide a constant audio link to their phones. With the proliferation of mobile devices comes the ability to provide a vast amount of data to the user. This provides them with constant connectivity, and a speaker with which to display auditory cues. These systems can provide an audio representation for any object in the user's location. With so many devices that can be used to provide contextual data, there is a clear need for software that provides tools for contextual audio design.

With computing moving away from the desktop and toward wearable computers, an audio interface can be less obtrusive, due to most users' heavy reliance on visual input. With a visual interface, the mobile user may have large portions their field of view obstructed, which is dangerous and distracting. An auditory interface can be engineered so that it does not interfere with critical auditory warnings in ways that a

visual interface cannot. These systems can present a significant amount of information from all directions, and because of the fact that humans associate a location with one type of data, can leverage the position of the sound to convey information, and thereby decrease cognitive load. [1]

In the current state of audio Augmented Reality (AR) design, there is a significant problem. In order to develop audio for an AR environment, a designer must have significant programming skills. Most designers do not have such skills, and those that do often find that the programming gets in the way of the audio design process. The design process is then relegated to a secondary position, and suffers because of that.
To allow designers to work with these context-aware systems, there must be a designer-friendly design interface. Without providing appropriate tools, the designer may miss significant problems in an audio design, and regardless of the technical quality of the other system components, the system will fail. Many, if not all current systems provide only very basic design tools. Experienced audio designers can overcome the limitations of current software, but the designers succeed in spite of the tool, not because of it.

2 Design Requirements

The design process for an audio AR is basically: (1) design audio objects (sources); (2) place those sounds (virtually) in the environment; and (3) evaluate the aesthetic and functional qualities of the audio sources as the users interact with them. Currently, there is only one tool that supports the design of visual objects in an AR environment: Designer's Augmented Reality Toolkit (DART). [2] The DART software focuses on designing for a visual AR display, and it supports the three design tasks fairly well for visual objects. DART also supports basic audio, but the audio manipulation is cumbersome. The tasks in an audio AR are basically the same as for a visual AR, but there is no need or ability to render a visual interface on the final AR. The visual interface is only necessary and allowable at design time.

To design for the state of the art sound systems now present even in handheld or wearable devices, there are several professional-grade tools in the marketplace. Unfortunately, despite their impressive capabilities, none is really suited for AR and mobile audio design. As an example, Digidesign's ProTools started as a tool to assist recording engineers mix music albums and soundtracks, but as audio technology got better, ProTools started incorporating surround sound design capabilities. This would seem to be an appropriate tool for developing audio AR experiences. Unfortunately, ProTools, and all the similar tools have been designed with the assumption that the listener is stationary or that the listener is not assuming that the sounds are tied to the real physical world. Thus it is clear that there remains a need for an AR design tool that supports audio in an effective way.

The requirements for the system are: (1) give the designer a sense of the environment for which they are designing; (2) provide a visual representation of the audio; (3) allow the designer to compare sounds based on a region of interest; (4) provide the ability to switch audio labels quickly; and (5) allow the designer to browse sounds quickly.

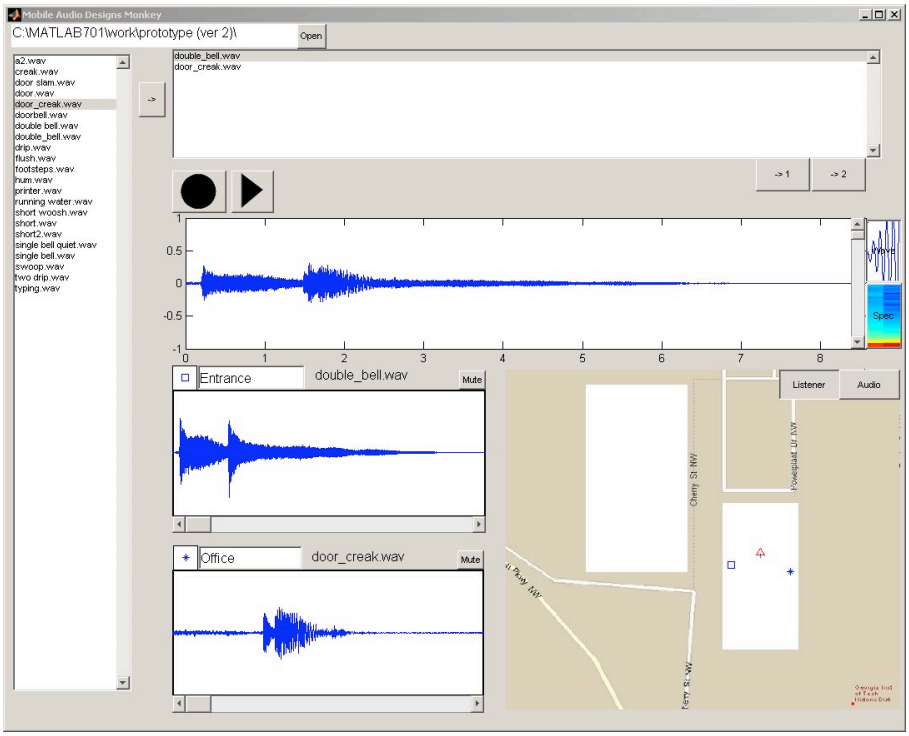

Fig. 1. Interactive Prototype – This software provides a visual representation of the location of the user and of the audio sources, as well as the characteristics of the sounds themselves. The designer can place a virtual user and hear the sounds spatialized as the user would

3 Mobile Audio Designs Monkey

In order to fulfill those requirements, an audio AR designers' tool called Mobile Audio Designs (MAD) Monkey was developed. The development of MAD Monkey was executed using the standard process for User Centered Design. The process consists of a survey of current software, and then a process of iterative design. This consists of a prototype being built, then evaluated by experts, refinement of the prototype and repetition of the process until time or money is depleted. [3]

3.1 MAD Monkey System Features

The system is built in MATLAB because of its sophisticated audio manipulation capabilities as well as its GUI builder. There is no other system that supports complex transformations of audio while still allowing for a familiar visual interface to be built relatively easily. Other programming platforms that do have significant audio capabilities lack standard GUI development features. The GUI prototyping tools, on the other hand, do not allow for the simple manipulation and representation of a wave-

form. MATLAB, by contrast, has a substantial number of tools that allow for the display of the waveform and also supports the use of ActiveX controls in its GUI development environment. These controls provide diverse and complex interface functionality, with a minimum of effort.

The system displays the audio as though it were a physical object in the environment. It also allows for the display of an overhead map of the environment, and a display of all the audio within a focus region. All of the audio in those regions can be played to give the designer the ability to determine conflicts or dissonance.
This system allows one or all audio files to be swapped out quickly. This is important in rapid prototyping and supports a fluid design process. MAD Monkey also provides access to simple manipulations of the audio, with a focus on those most relevant to augmented reality. Examples of these are delaying the onset of a sound, or manipulating the sound's attack.

3.2 MAD Monkey System Benefits

This system design has many benefits over any existing alternatives. In particular, it (1) gives the designer a sense of the environment for which they are designing; (2) provides a visual representation of the audio; (3) allows the designer to compare sounds based on a region of interest; (4) provides the ability to switch audio labels quickly; and (5) allows designers to browse sounds quickly.

MAD Monkey provides a representation of the environment in which the designer is placing the sounds, which allows the designer to leverage any knowledge of the space in the AR design. It provides a visual representation of the audio that will allow the designer to place audio that does not conflict with the audio currently in the system. By incorporating the overhead map of the space, MAD Monkey allows the designer to compare the sounds in their current focus region. The sound cues can be switched very quickly with a simple drag and drop interface. All of the designer's audio files are quickly and easily browsable, with easy access to those sounds for the current AR design.

4 Future Development

One of the most important features that will be implemented in the next version is a visual representation of a sound's extent. This allows the designer to see which sounds the user will interact with at a given place and time, and focus on those. Other planned features include integration with the System for Wearable Audio Navigation and DART, so that the designer can control the 3D engine in real-time. There are several features that help the designer create AR that is optimized for human perception, such as an automatic removal of, or a visual indication of, masked frequencies.

There is also a need for tools that generate audio that is specifically designed to be used in an AR environment. Currently, the designer must have significant experience with the audio design in order to design audio that attracts or repels the user. This software would include a library of those sounds, classified into the appropriate cate-

gories. In a related feature, instead of presenting a display of the sound's physical properties, the system could present a different representation of the audio. These alternative representations might be: amount of dissonance, emotional content, or other data that is more directly linked to the user's interpretation of the audio, rather than a physical representation.

The system needs to have the ability to turn on sets of audio. These may be all the sounds representing drinking fountains, or all the emergency exits. A real-time analysis of the audio that is currently being displayed is also planned for future development.

5 Evaluation

The process of User Centered Design that has been applied in the development of MAD Monkey largely consists of formative evaluations. Each step in the process incorporates evaluation after the development of each prototype. This approach reveals potential problems with the interface and interaction before substantial resources have been used to implement those inappropriate features. As a result of this process, the MAD Monkey interface has undergone extensive iterative redesigns to ensure effectiveness and usability.

The first interactive interface was evaluated with a think aloud protocol, which was possible and appropriate due to the relatively sparse nature of the auditory display. Additionally, the participants needed little, if any, guidance after the initial instructions. These sessions produced significant feedback about the interaction and layout of the interface, as well as the particular way that the system would play the audio.

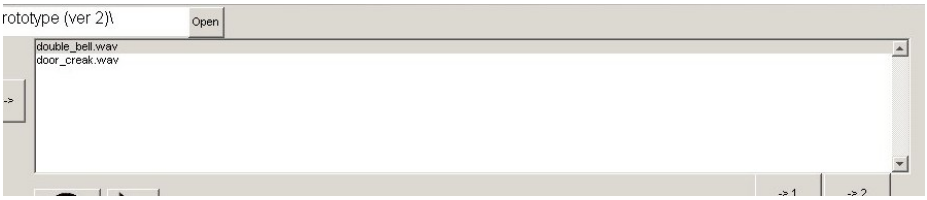

Fig. 2. Top portion of the first interactive prototype. This was moved and its functionality was modified due to expert feedback

None of the experts used the top section of the interface very much. That section is represented in Figure 2. This was partly because there were only two objects available in the prototype. After one session, an expert suggested that classifying the sounds into categories would be useful. Another expert confirmed this when he suggested a "layer" capability, similar to Adobe Photoshop®. As a result, this portion of the interface was moved to the left-hand side and will be represented as a hierarchical list. Figure 3 shows the redesigned interface incorporating all of the expert feedback.

The expert feedback about the sound playback did not change the actual functionality of the system, but brought up issues which the playback would not handle well.

Currently, the interface plays both “object” sounds when the listener is moved. If one sound is looping, and the other is not, then the playback does not give an appropriate representation of the audio. These issues need investigated further before a design decision can be made.

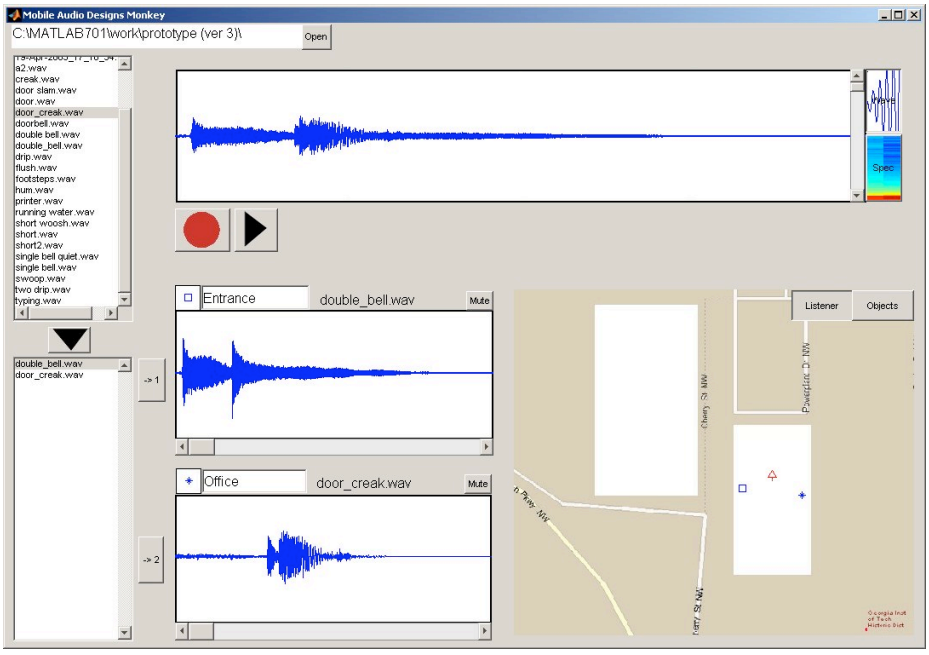

Fig. 3 The final interactive interface. This interface incorporates all of the expert feedback

6 Where to get MAD Monkey

Matlab files will be available at http://sonify.psych.gatech.edu/research/MADMonkey

References

1. E. C. Tolman, "Cognitive Maps in Rats and Men," *Psychological Review*, vol. 55, pp. 198-208, 1948.
2. B. MacIntyre, M. Gandy, S. Dow, and J. D. Bolter, "DART: A Toolkit for Rapid Design Exploration of Augmented Reality Experiences," presented at User Interface Software and Technology (UIST), Sante Fe, New Mexico, 2004.
3. A. J. Dix, J. Finlay, G. D. Abowd, and R. Beale, *Human-computer interaction*, 3rd ed. Harlow: Pearson, 2004.