Project Sonology: An Experimental Project Exploring the Possibilities of Sound and Audio as the Primary Element of Interactive Entertainment

Shih-Han Chan, Dae Hong Kim, Eugene Massey, Katelyn Mueller, Fadzuli Said

Project Sonology: An Experimental Project Exploring the Possibilities of Sound and Audio as the Primary Element of Interactive Entertainment

Shih-Han Chan, Dae Hong Kim, Eugene Massey, Katelyn Mueller, Fadzuli Said,

Entertainment Technology Center, Carnegie Mellon University.
700 Technology Drive, Pittsburgh, PA 15219, USA
{Shihhan.Chan, Daehong75, Eugene.H.Massey,
Katelyn.Lisa.Mueller, Fadzuli.Said}@gmail.com

Abstract. The goal of the project is to show that audio can successfully be the primary element of interactive entertainment by delivering pure audio experiences that demonstrate both the creative potential and emotional power of an audio experience. We develop two proofs of concept with the technical foundation supported by prototypes. The core technology is a combination of a 3D game engine and an audio engine used to build sound environments. The interactions are based on 3D trackers and surround sound headphones.

Keywords: Interactive Audio, Virtual Environment, 3D Sound.

1 Introduction

Project Sonology is a student pitched project of the Entertainment Technology Center at Carnegie Mellon University in 2008 whose goal is to show that audio can play the major factor in interactive entertainment by creating pure audio experiences that are unique and present both the creative potential and emotional power of an audio experience. We create a flexible technical foundation supported by multiple prototypes demonstrating the technology that would help shape the experiences we would eventually deliver. With the technical foundation in place, we work on two proofs of concept as the final deliverables: The Story World and The Music World.

We combine a 3D game engine and an audio engine as the core technology to create 3D sound environments that can be navigated and interacted with intuitively. The interactions are based on 3D trackers and surround sound headphones. Guests perform actions and explore different parts of the environment.

2 Platforms

FMOD audio engine and OGRE 3D game engine are chosen as the fundamental software. FMOD gives sound designers the option to create complex audio event behaviors. OGRE comes with 3D math and game utilities that help to speed up

development. FMOD and OGRE allow us to divide up and modularize the programming work so that different team members can work on different aspects simultaneously. On the other hand, Polhemus 3D trackers and surround headphones are selected as core hardware platforms in the final proofs of concept. The surround headphones we use have speakers in the front and back for an added dimension.

3 Prototypes

In the early stage of the project, we developed a series of prototypes. From that we received valuable feedbacks directing us to revise and polish the work. The various prototypes that help define and demonstrate the technology foundation embrace:

- Head tracking to allow the guest to walk around and rotate his heads while listening to different aspects of a virtual 3D environment.
- Hand tracking to allow the guest to reach out sounds and then move them to new positions in the virtual 3D space.
- A fly catching game to integrate head and hand tracking into a system.
- Webcams with OpenCV to test out possible techniques for body tracking.
- DSP Plug-ins to simulate binaural effects.
- Speech recognition to implement auditory menus.
- Directional focused sound to assist the guest in locating a sound source.

4 Proofs of Concept

4.1 The Story World: Jack and the Beanstalk

Since audio is an ancient platform for story-telling, we want to deliver a story by using our interactive audio-only concept. The goal is to find a good balance between immersion and interactivity. An interactive version of "Jack and the Beanstalk" is made where the guest assumes the role of Jack. By using Jack's mom as the guide, the guest would feel that he is interacting with another character.

A variety of simple interactions with sound are applied in many places such as grabbing the beanstalk, catching the hen and chopping the beanstalk. We also utilize the voice recognition technique in which the guest is requested to sing and match the tone that the golden harp plays. Yet some guests have had a hard time finding objects, we add a "Hot & Cold" feature, by which our program would guide the guest to a certain object by providing information of the distance from the target to the guest.

4.2 The Music World: Whip It

Our other proof of concept is an experience based around listening to music. The guest would hear a song, in this case Devo's Whip It, and then the song would break into pieces at which point the guest would have to reconstruct the song to complete the experience.

The main interactions with this experience are finding sounds and then playing with them. We add a sonar feature that makes a beeping noise to increase in intensity as the guest gets closer to a sound that he needs to collect. Once the guest finds a sound, he can play with it. For instance, when the guest finds the guitar sound, guitar noises would be produced if he plays "air guitar." Most guests do make these kinds of motions without any prompting.

This experience is meant to be a tutorial for our technology and it works very well in this regard. It is also designed to be simple and fun to boost the guest's confidence.

5 Observation and lesson learned

People are happy with the simplest aspects of our technology. With a complex and new experience, very simple interactions that are intuitive and recognizable are more fun than more complicated combinations of the interactions.

Another lesson we learn is that it is very difficult to place a sound in 3D with enough information that a person can tell naturally where the sound is. When the sound is behind the guest, we apply a filter that blocks out all the high frequencies; when it is on the side, we pan the sound, attenuate volume and apply a Doppler effect for the distance, whereas in actual reality, sound is bouncing off walls and being absorbed by the material of the carpet. We also learn that the human hearing system is extremely complex. Our brain is automatically filtering out all the sounds that are not vital to allow you to listen to what is being said. We have attempted to solve this problem within the scope of the technology we have by adding story and context to help fill up gaps in information that our sound cannot convey. We moreover give other kinds of feedbacks like sonar beep noises to show the proximity of hands to objects and move sounds around in a certain pattern so that the guest can easily position the sounds. Though how to make the sound localization more clear is another course for us, our attempts meet with enough success that a guest can pick up quickly how to locate sounds in our world and proceed on to the rest of the experience.

6 Conclusion

As previously stated, our goal is to prove that audio can be the primary element of an interactive experience. We believe that our project is successful in demonstrating this, but that our proofs of concept have only scratched the surface of what is possible in this realm. For example, we only make "Jack and the Beanstalk" interactive, but this proves that the same can be done for any story; we only break apart "Whip It" and put it back together, but the same can be possible for any song in any genre. In addition, this technology can be adapted for use with mobile devices like cell phones, MP3 players, as an interactive museum installation, a musical instrument or even as an educational tool.

We hope that the lessons learned in our project and what we have achieved would motivate future audio-only projects to reveal more potentials of audio and sound with interactivities.