3D Audio for Museum Exhibits

I. INTRODUCTION

The advancements in technology have allowed for individuals to have an immersive experienced through 3D audio. Some devices that include 3D audio capabilities can be found in video games, cell phones, speakers, and sound cards.

This paper describes how 3D audio can be utilized to help precisely guide guests through museum exhibits in Library East. To help guide people, beacons will be placed in front of every museum exhibit of Library East. Individuals will then walk through the simulated museum, while listening to the 3D audio streamed over the Bluetooth headphones. After pursuing the museum exhibits, an individual's quantitative feedback is recorded and analyzed.

This paper shows both how we can create 3D audio by retrieving a user's positional data, and how this 3D audio can be heard as a user changes location, rather than them being stationary.

II. WHAT IS 3D AUDIO?

A 3D audio system has the ability to position sounds all around a listener (Gardner,1999). The sounds are actually created by the loudspeakers (or headphones), but the listener's perception is that the sounds come from arbitrary points in space(Gardner, 1999). The position to create 3D audio is in terms of azimuth -- an angular measurement in a spherical coordinate system -- and elevation -- angle between the horizontal plane and the line of sight. The sound is generated at an azimuth and elevation relative to the user.

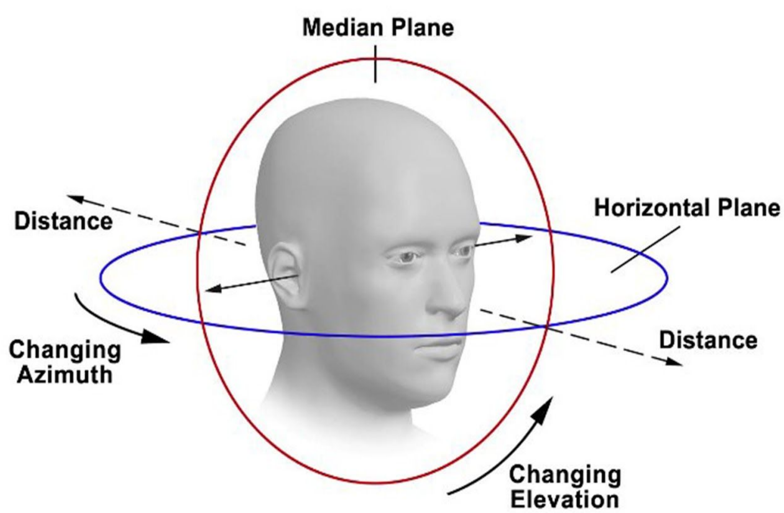

Figure 1. Understanding Azimuth and Elevation

The azimuth and elevation show where the sound originates from, however the sound travels to the left and right ear in different directions, thus a different sound is heard from each ear. The difference in arrival time of a sound between the left and right ear is known as interaural time difference (ITD). For example, the drop of a spoon on the ground might reach the left ear faster than the right ear if you left ear was closer to the spoon.

.

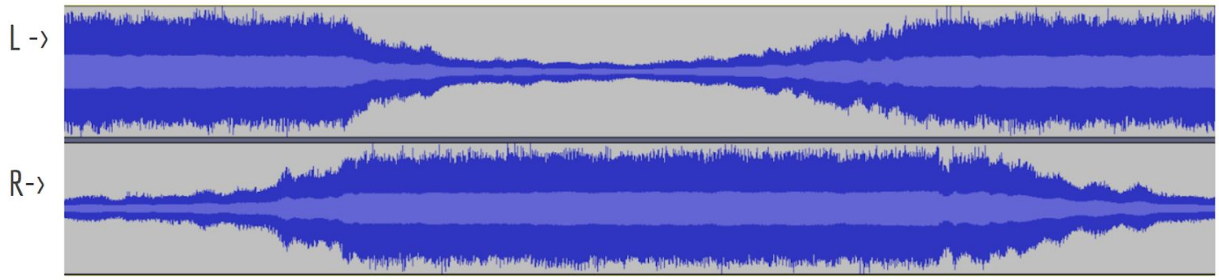

Figure 2. The same sound heard from the left and right ear (wav file)

Once the sound reaches the the ears, the incoming sound that goes into our head is now deformed due to how sound travels in different head sizes and ears as well as reflection from our shoulders. In other words, our head could be though as a filter than deforms the incoming sound. This is known as a head related impulse response (HRIR).

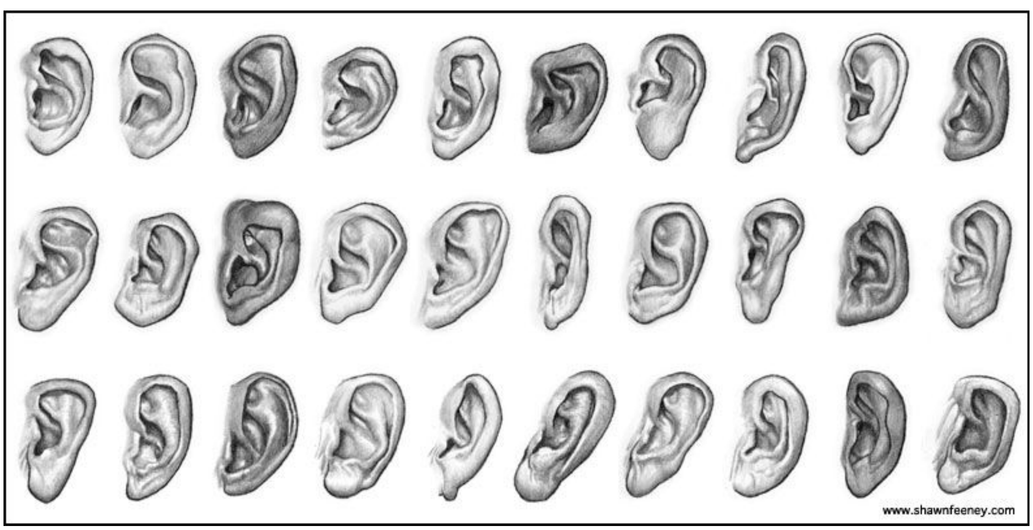

Figure 2. Different types of ears

The last component to creating 3D audio is getting the head related transfer function (HRTF). This is calculated by taking the Fourier transform and applying the convolution theorem. This is important because HRTF decomposes the incoming sound which is arriving from just one direction into its frequency components and helps to look at how each component is affected by the head.

III. CURRENT WORK IN CREATING 3D AUDIO

The National Public Radio (NPR) has experimented with bringing 360-degree video and spatial audio together. NPR audio engineers conducted an experiment with a technology called “Neural Surround Sound” in the mid-2000s enabled listeners with a surround sound setup to hear the mixes (Michael, 2017). Some other experiments where they combined both 360-degree audio and video include: Behind the Scenes at the Tiny Desk with Wilco, Standing at the Edge of Geologic Time, 360 scene from 2017 Inauguration 360 scene from the 2017 Women’s March and Wadada Leo Smith: Awakening Emmett Till.

IV. RETRIEVING THE USER’S POSITIONAL DATA

To retrieve the user’s positional data the SoundPad Lab used UCDavis’ CIPIC HRTF Database which is a high-spatial-resolution HRTF measurements for 45 different subjects. Their database includes 2,500 measurements of head-related impulse response for each subject. These measurements were recorded at 25 different azimuths (-80 through 80) and 50 different elevations (0-180).

To create a realistic audio through headphones, we would need the head measurements of each person who would be guided in the museum tour through the 3D audio, however this approach is not scalable. Instead the SoundPad Lab took candidate 30’s HRIRs and HRTFs from the CIPIC database (candidate 30 has an average head size). Candidate 30’s measurements were chosen because those measurements were meant to generalize what the average person’s head size participating in the study.

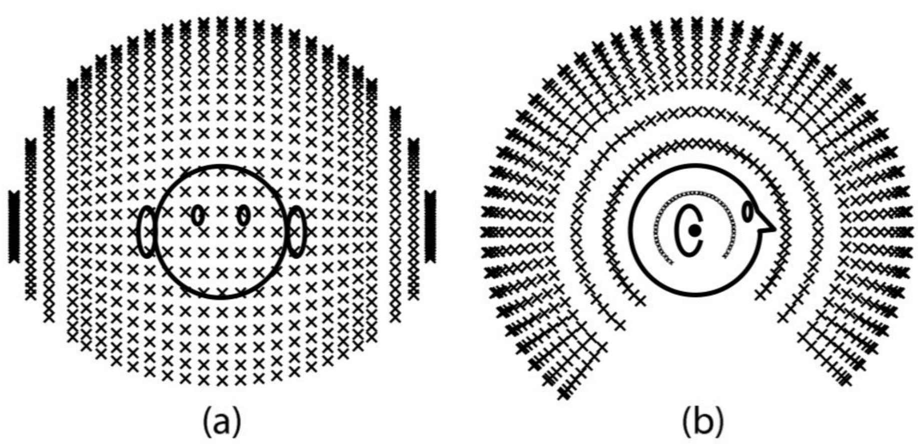

Figure 3. Locations of data points (a) front (b) side from CIPIC database

IV. 3D AUDIO AS A MUSEUM GUIDE

For the subjects participating in the study, they would need to carry a Raspberry Pi and gyroscope around their neck. The gyroscope produces the user's coordinates depending on where they are standing within the museum; through the positional data the programmers would know which audio clip to play and the sound to produce. This information from the gyroscope is sent to the Raspberry Pi and the Raspberry Pi contains the code to generate and render the new sound. The combination of these technologies allowed the user to walk around and hear different 3D audio, rather than them being stationary. Figure 4 shows a Raspberry Pi, which can be though as a mini computer. The SoundPad lab used Marvelmind Indoor Navigation System to retrieve the user's positional data through small beacon devices-- the gyroscope. The way we retrieved the data is through the software shown in Figure 5. Figure 6 shows the beacons (gyroscopes).

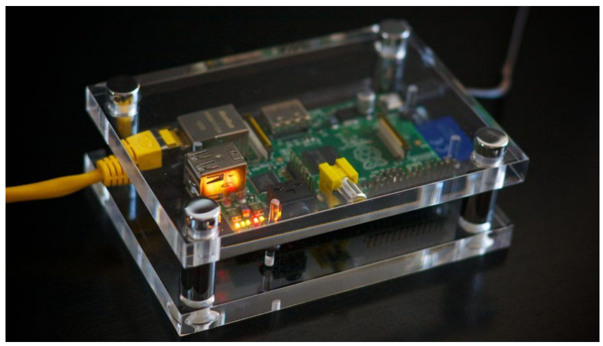

Figure 4. Raspberry Pi

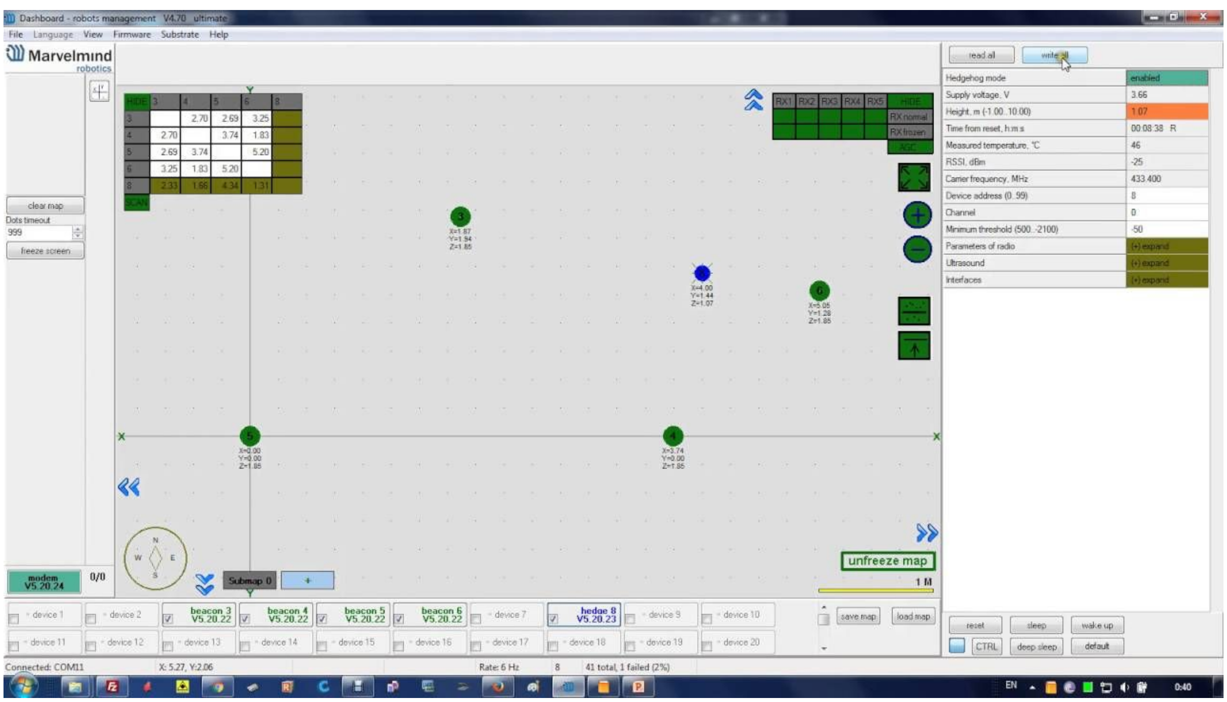

Figure 5. Tracking the user's position through a beacon device

Figure 6. Marvelmind Indoor Navigation System

V. Results

At the end of the summer research, the Sound Pab Lab team did not finish the experiment because there were issues in reading the data from the gyroscope. The positional data was not stabilized thus it kept changing rapidly due to the environmental noise. Additionally, the graphical user interface that one of my other teammates was working on to track where the user is located had not yet been integrated with the 3D audio I created.

However, during this research I successfully created 3D audio noise using a white noise audio clip (wav format) and programmed it such one could hear a white noise traveling around a person in 360 degrees. The output of the rendered clip was in a wav format so that next challenge was rendering this directly to the sound card so that the noise played in real-time which I was able to do using the museum audio clips. This new audio did not lag when played on my desktop. However when this played on the Raspberry Pi, there was about a 2 second lag between every azimuth change.

VI. Summary

The project has significant implications in terms of creating a more immersive experience. To push this project further, the graphical user interface and the audio will need to be integrated, the sound generated from the raspberry Pi will need to be rendered without a lag, and the gyroscope data will need to be stabilized. Some issues that this project might run into is that some users might not get an immersive experience if their head size deviates too much from average since the project uses an average head's HRTFs and HRIRs. Another issue could be if the user moves around too much, they might get dizzy from hearing different sounds that are not updated fast enough.

References

1. Gardner, W. G. (1999). 3D Audio and Acoustic Environment Modeling, Wave Arts, Arlington, MA

2. Algazi, V.R., Duda, R.O, Thompson, D.M (2001). The CIPIC HRTF Database, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, Davis, CA.

3. Michael, Nick. "A Beginner's Guide to Spatial Audio in 360-Degree Video." NPR, NPR, 17 Dec. 2017, training.npr.org/visual/a-beginners-guide-to-spatial-audio-in-360-degree-video/.