Ear Hands

Applications of Spatial Sound in Video Games in VR using Unity

CHRISTIE ZHAO and THOMAS HSU, University of Washington

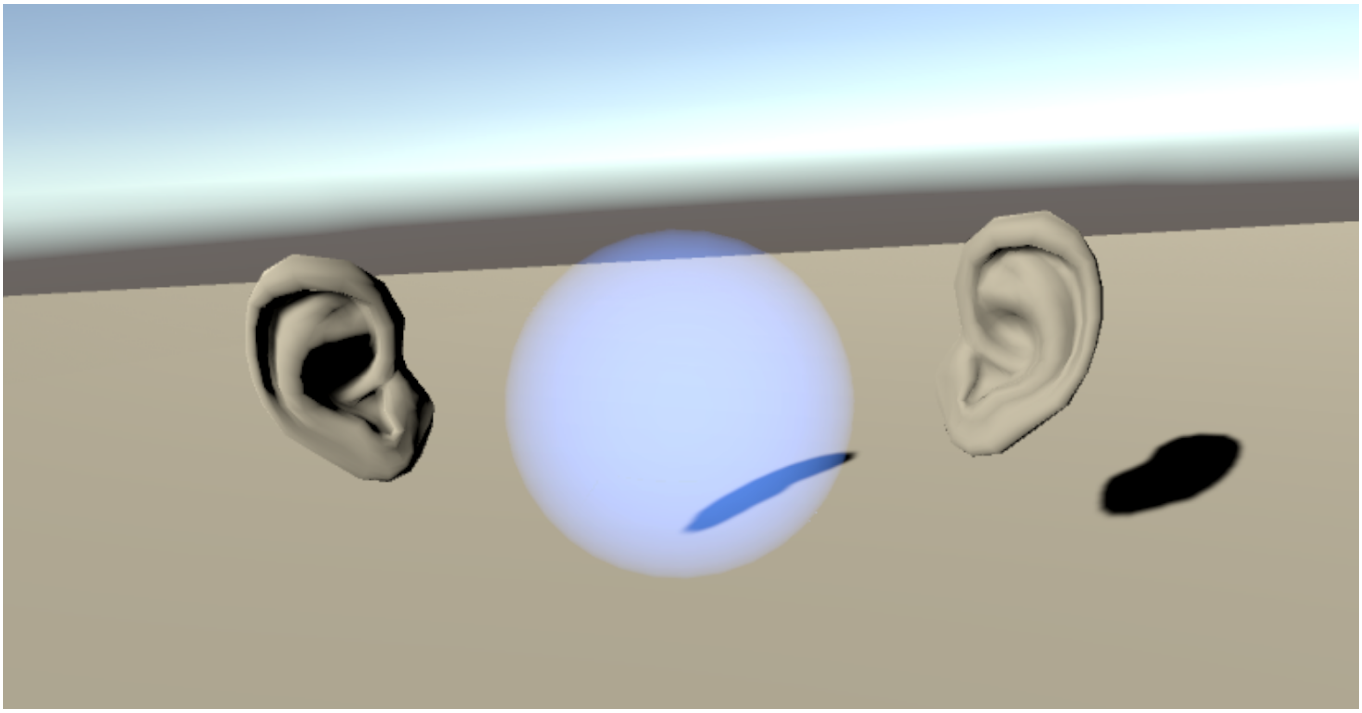

Fig. 1. With ear hands, players will get to experience what it would be like to hear out of their hands instead of their head.

Spatial audio in VR allows for new, immersive ways to interact with the environment. In our game, we developed a system where your hands are ears, and used this system to develop an audio focused game that tests the player's spatial audio awareness skills.

1 INTRODUCTION

Most audio focused games are rhythm games. However, this is a very narrow part of the possibilities of an immersive audio game. In this project, we made a game that focuses heavily not only on audio cues, but also on spatial audio. In this game, the player's hands function as their ears which they need to use to navigate and survive in the environment. The player will be able to hear the invisible, audio based enemy "hear-onavirus" coming from a clear direction where they need to destroy the enemy to survive.

Here, we also highlight how earhands are implemented, which is by simulating each source of audio twice, one for the left and right audio channel. This is done to bypass the constraints of Unity and to realistically simulate the different audio space reference frames that occur when having ears for hands.

Authors' address: Christie Zhao, crystor9@cs.washington.edu; Thomas Hsu, wassaman@cs.washington.edu, University of Washington.

This project showcases how spatial audio can be used as the central focus in a VR experience, not just serving a complimentary role like in most games. Concepts such as having hands for ears shows how spatial audio in VR can craft unique experiences, concepts which can be further used for games.

1.1 Contributions

- We created a full game play loop for a game that focuses on spatial audio.
- We introduced the idea of ear hands and successfully implemented this idea despite Unity constraints.

2 RELATED WORK

2.1 Existing Audio Games

The visual and audio components of VR headsets are improving throughout the years. A lot of current audio games do a great job on giving the players an immersive audio environment on music, especially rhythm games. Rhythm games or rhythm action games is a genre of music-themed action video game that challenges a player's sense of rhythm. Games like Beat Saber, Audica, and Synth

Riders generally fall into this category. It is not that easy to find an immersive VR game that focuses on the spatial audio component.

2.2 Approaches to Multiple Listeners in Unity

In Unity, AudioListeners are the ears - they take in audio for the players to hear, and the locations between listeners and audio sources are what determines the spatial audio effects that gets applied to the sound. Unfortunately, Unity only allows for one listener, which means we need to find workarounds to get earhands to work using Unity. Similar problems have occurred to developers working on split screen multiplayer games, and their solution was to place the listener in the middle of the two cameras. In our case, this means to place the listener between the two earhands. Even though this works for some cases, this doesn't work for cases when one earhand doesn't move while the other hand is moving. Because the listener is placed in the middle of the two hands, the earhand that remains still hears a change of audio due to the listener moving further away. However, the ear that remains still should not have any changing sound whatsoever, making this approach flawed if we want realistic spatial audio when dealing with multiple reference frames.

3 METHOD

To simulate having two ears as hands with only one reference frame (listener), we need to capture the separate reference frames of the ears, then convert them into one. Our solution was to instantiate each sound twice and place them relative to where they would be in the reference frame with one listener.

Fig. 2. In the actual game, this is what it looks for the player.

Fig. 3. This is how the sounds are actually simulated in game.

This method accurately captures translation of the earhands, since relative position between ears and sound objects are used. By moving the left earhand away from an audio source for example, the audio source in the reference frame where there is only one listener moves away from the listener's left ear. In other words, when we move the earhands, what actually happens is that the sounds move instead, reacting as though the person was standing still.

However, rotation of the earhands works differently. Say we had a earhand facing downward, above a sound object, so that it is facing toward it (Fig. 4.). In the reference frame where there is only

Fig. 4. An ear resting above a sound object facing towards it.

one listener, the sound should be directly in front of the left ear. However, if the earhand rotates upwards, the sound should sound like it's coming from directly below the left ear. This means that for rotations, we need to keep track of two vectors: a vector that points to the direction of the sound object from the ear and a vector that points in front of the ear (imagine a arrow shooting out of the ear canal).

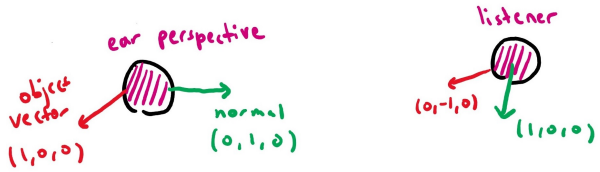

Fig. 5. These two vectors allow us to get from one reference frame to another.

With these two vectors and an additional vector that points in front of the associated ear in the reference frame with a single listener, mapping from one reference frame to the other is a matter of rotating from one ear front vector to the other. Because this method captures relative angles between ears and sound objects, this allows for accurate spatial sound when rotating the earhands when audio is perceivable.

Lastly, a rudimentary audio occlusion system is added by shooting a set number of rays from each ear to each sound object. Each ray that collides with a piece of the environment lowers the volume of the sound and applies a low-pass filter.

These audio systems are designed for our game's combat system, as our enemies are invisible and emit a sound that the player has to detect using their hands. Picking up weapons against these enemies occludes and blocks the sound from the earhands, providing a risk-reward of being able to fight back, but not being able to detect the

Fig. 6. The left hand is listening the environment and the right hand is ready to combat

enemy. In the game play, the earhands are used to navigate the player and help them detect the direction and the position of the Hearonavirus. To increase the engagement of the game, the enemies sense the location of the player and automatically traverse to the player location in different speed and size. They will be destroyed once they collide with the bullets from the disinfect spray or they collide with the player. The second condition also costs player's health, which speeds up the ending of the game.

4 IMPLEMENTATION DETAILS

We used Unity 2019.1.10f, Blender, and an Acer Mixed Reality headset. All software done with default Unity environment.

We built two models with Blender. The disinfect spray that serves as the major weapon and the virus that serves as the major enemy.

Fig. 7. The disinfect spray that we modeled from scratch

In each scene, the real and the fake audio sources are separated, the real audio sources having two audio sources for each object in the fake audio source. Each of the two real audio sources represent the left and right audio source and are synchronized so they play at the same time. The fake audio source is where the player feels like the audio is coming from, and it is these fake audio sources that raycast towards the player to generate the occlusion, which effects the audio from the real audio sources.

The separate reference frame with the actual listener should be as separate from the actual player as possible to prevent unintended audio occlusion between the walls and floor.

5 EVALUATION OF RESULTS

5.1 Benefits

We successfully implemented the earhands concept as players can hear the difference between audio channels from the from left hand and right hand very clearly. Although the effect is subtle, the auditory shift between hand rotation is also noticeably in effect, as when each hand rotates upwards, the audio shifts downwards below the hand. Hands truly start feeling like ears after a few moments of interaction with our environment.

Due to our implementation, adding multiple audio objects for the earhands to interact with is easy, as it's only a matter of copy and paste to add a separate audio object to our scene.

We also achieved partial game play. Players can explore the scene and interact with assets in the scene. They can also use the disinfect spray and destroy enemies by following the spatial sound of the enemies.

5.2 Limitations and Failures

5.2.1 Gameplay. The game play loop is not fully implemented yet due to the time and skill constraints. We decided to focus more on the intermediate game play to provide users an audio experience. Thus, the introduction, ending and the health system are still incomplete. We also decided to use walking as the main movement mode instead of teleportation because we wanted a more realistic effect. However, the walking movement mode does not simulate the walking in real life and it can cause some slight motion sickness after a while. Lastly, we are also facing the trade off between the visual cues and audio cues. According to one of the players who tested our demo, the fast speed of the game play sometimes distracted them from using the earhands to navigate.

5.2.2 Technical Issues. The technical issue we encountered with the software is a big limitation. Since we are very new to Steam VR, Unity and Blender, we did not achieve the ideal animation effects for the explosion, disinfect spray and the earhands. Other than this, we can also optimize the generation of hearonavirus(enemy) by randomizing their amount and position, which will add up replay value for the players.

5.2.3 Sound Effects Management. Although we can clearly hear the difference from left and right hands, managing the sound effects system is difficult for the developers. There are two instances of the audio sources for left and right channels and it takes some redundant work to input sound effects. This is especially difficult for the multiple hearonavirus sound effect generation. Currently, this current audio system doesn't easily allow for audio sources that are quickly instantiated and then deleted.

6 FUTURE WORK

In the future, we would like to complete the game loop and make it more user friendly by adding better UI and having an intuitive introduction that allows the players to better know what to do. Invisible enemies are an intimidating concept, so we would like to find ways to better prepare our players to fight a purely auditory challenge. We also want to add better replay value into the game by adding different difficulties and interesting plot developments.

In terms of the audio systems, we would like to improve our current audio occlusion system to make sure the occluded sound is spatialized as well, as the current occlusion system only lowers the volume and applies a low-pass filter based on the amount of raycasts that collide between the environment. We would also like to conduct a user study on the application of ear hands. We want to

see how well people can navigate themselves through spatial audio and see how different users react to audio sources coming out from two different places.

The community should follow our lead and create interesting applications that utilize sound in VR in unique ways. For example, perhaps our earhands idea could be extended into a spy game, where the player would need to extend their hand in otherwise unreachable places to find secrets to reach their goal.

7 CONCLUSION

Spatialized audio in VR should be focused on more due to the immersive nature of VR applications. By experimenting with novel ideas such as having ears for hands, players can be provided with unique experiences they could not otherwise experience elsewhere. With the rise of VR, developers should use explore the avenues of sound to make virtual reality feel more like reality.

ACKNOWLEDGMENTS

A special thank you to Cameron Lee Astor for modeling the hearonavirus and being our sole user tester for this game. Another special thank you to Douglas Lanman for being a constant source of inspiration.