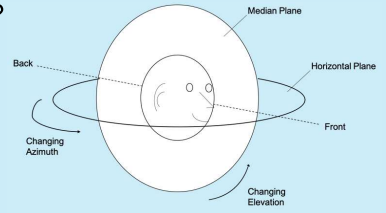
How Does the Positional Accuracy of Binaural Plugins Differ

Tom Aries

School of Media Arts and Technology

Research Question

How Does the Positional Accuracy of Binaural Plugins Differ and how does it determine its use?



Introduction

Spatial audio is a way of creating a 360° listening environment around and is used to create exciting and immersive content. Binaural is a very unique form of spatial audio as it is the only system that can deliver an immersive spatial mix over headphones.

In recent years there has been in an increase in the use of streaming services such as *Netflix* and *Amazon Prime,* a large amount of the content that appears on these services utilise spatial audio mixes. However, the popularity of these services has led to large percentage of users viewing the content with mobile devices and listening through headphones.

Currently these services do not support binaural audio, however, work is being done in done to create standards for broadcasting binaural audio.

Background

Binaural audio works by using complex filters and delays that simulate effects caused by the ear and the body. To achieve this, binaural plugins need to simulate three essential functions.

HRTF

The head related transfer function is a complex filter function that simulates the free field sound hitting the ear drum. Everyone has a unique HRTF as it is varied by factors such as the shape of the head. Because of this averaged HRTF's are used to create good general spatial accuracy for everyone (Bowman, Kruijff, LaViola Jr, & Poupyrev, 2004)

ITD&ILD

the interaural time difference (ITD), which is the difference in time between the wave front hitting the first ear and the second ear. And the interaural level difference (ILD), which is the difference in level between both ears (Rumsey, 2012) . Both of these functions use the delay caused by the human body, mainly the head, to locate sounds.

Limitations

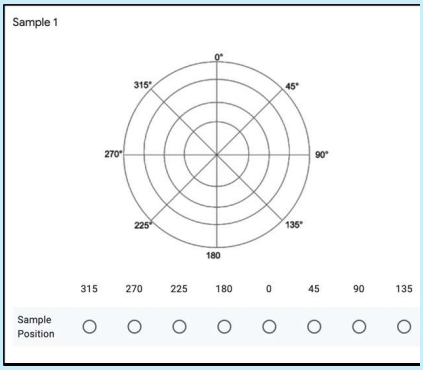
There are a range of limitations when using binaural audio. One of the main issues is caused by the use of average HRTF's. They can cause a high error rate and low performance in sound localisation (Shumaker, 2007).

Further more , it is common for binaural cues to be indistinguishable between the in front and behind the listener. This is known as the cone of confusion and has an extremely negative effect on the localisation quality of binaural renders (Mather, 2006)

Outcomes

The outcomes of this test may show the difference in the accuracy of binaural plugins ability to position sounds. This information can be used to get a better understanding of the different uses of binaural plugins, depending on their ability to accurately positions sounds. This can be devolved further to understand what is needed to broadcast binaural content.

Methodology



The test was distributed online and took roughly five minuets to complete. Due to the lack of control of the listening environment it was sent to forty individuals to achieve the required resolution (BS.2159-4, 2019) , however only sixteen responded, this was still enough to achieve a conclusion.

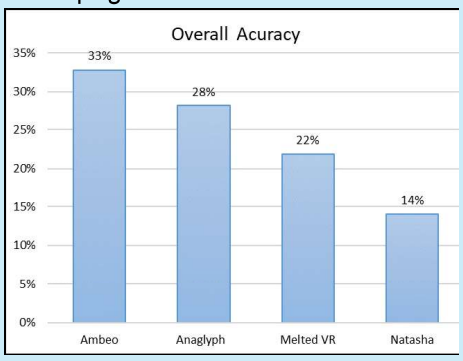
The test consisted of eight samples being played over sixteen binaural panned positions. It was the participants job to mark the angle they perceived the sound to come from.

The sounds were panned using the four plugins that are being compared; *Ambeo, Anaglyph, Melted VR* and *Natasha.* Each plugin was featured four times during the test.

To ensure each sound was tested in every position, four different variations of the test was created and distributed.

Results

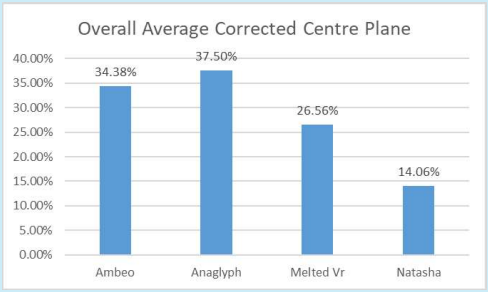
To form an initial conclusion, the exact accuracy of each plugin was found.



Whilst the accuracy of the plugins may seem low, similar results have been seen in other tests looking at positional accuracy (Reardon, Genovese, Zalles, Flanagan, & Roginska, 2018). A wide range of accuracy can be seen with *Ambeo* being the most accurate and *Natasha* being the least.

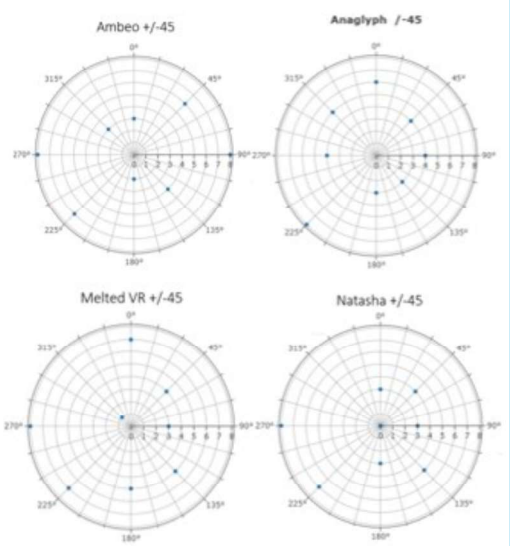
Results

Due to the common effect of reversal error caused by the cone of confusion, the results were also calculated irrespective of the reversal error.



This creates an increase in accuracy in all plugins, causing *Anaglyph* to be the most accurate plugin.

Finally, the results were also calculated, considering them correct in a 90° sector (±45°)



When considering the results correct across a wider accuracy it can be seen that *Melted VR* is the most accurate. It can also be seen that certain plugins are better at positioning sounds at different positions.

Conclusion

To conclude, the main aim of the project has been achieved, with the use of the listening test the differences between the accuracy of binaural content has been demonstrated. However, it has also highlighted that the difference ways in which plugins render binaural content causes them to be more effective at certain angles. The current data set does not provide enough information to determine which of these plugins is best at creating binaural content, to determine this, future research will need to be done.

References

Bowman, D., Kruijff, E., LaViola Jr, J. J., & Poupyrev, I. P. (2004). *3D User Interfaces: Theory and Practice, CourseSmart eTextbook.* Addison-Wesley.

Reardon, G., Genovese, A., Zalles, G., Flanagan, P., & Roginska, A. (2018). *Evaluation of Binaural Renderers: Localisation* . Milan : AES.

Rumsey, F. (2012). *Spatial Audio* . CRC Press.

humaker, R. (2007). *Virtual Reality: Second International Conference, ICVR 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007, Proceedings.* Springer.