5-2018

An Exploration and Analysis of 3D Audio

Bradley Peppmuller
California State University, Monterey Bay

Bradley Peppmuller
Lanier Sammons
MPA 475
May 14, 2018

An Exploration and Analysis of 3D Audio.

Alan Blumlein, an engineer for EMI's Central Research Laboratories, went to the cinema with his wife sometime in the late 1920s. He was very frustrated at the fact that one of the characters on screen could only be heard from one speaker on the other side of the room ("Alan Blumlein and the Invention of Stereo"). This made Blumlein to begin to think of new ways that audio could be conveyed. In the 1930s an amazing event occurred with respect to how sound was portrayed when Blumlein discovered that it was possible to achieve two signals at the same time on a record. This was facilitated by a complex cut on a gramophone record with two groves that could be read at the same time. Blumlein's discovery was the beginning of stereo sound ("Alan Blumlein and the Invention of Stereo"). Stereo sound was a revolutionary technology and is still the main way we listen to audio today. Due to this discovery we continue to build new ways the we interpret audio. This paper is a exploration of the new emerging technologies happening in the audio domain. Due to the growth in virtual reality, more research is occurring in creating audio environments that replicate 3 dimensions. This paper will look at how 3D audio is currently being implemented in a stereo domain using head related transfer functions. This paper will also go over new and innovative ways that may be the future domains used to help us perceive 3D audio.

First, I will talk about the basics of binaural audio, which “is the process of recording material in a way that simulates human hearing” (Fortin-Chabot 3). When recording binaurally the recording can put the listener right in the middle of the environment where the recording has been done (Fortin-Chabot 4). With a good binaural recording one should have a hard time distinguishing between the recorded audio and reality. Confusion surfaces when first learning about binaural audio: is binaural audio 3D audio? While 3D audio can be represented in binaural audio, they are not exactly the same thing.

Binaural audio can be best understood as a recorded medium. This means that either audio is either recorded binaurally or mixed prior to the recording to create binaural audio. Binaural audio can represent a 3D audio environment, but it is controlled by the creator as with audio for a movie. 3D audio is where the audio environment responds to user interaction. 3D audio is a much more immersive experience compared to binaural audio. This is because the audio can be controlled by the user. One can experience true 3D audio through a virtual reality system. Where this gets confusing is that virtual reality systems tend to use binaural rendering of 3D audio. This is because headphones, which are a binaural medium, are currently the most effective way to create a 3D environment. The way to remember the differences between binaural audio and 3D audio is that 3D audio is where the user can control the environment while binaural audio is controlled by the creator.

For understanding 3D and binaural audio, understanding spatial hearing and localization is important. Biologically, humans have two points of entrance for sound: the ear canals. To try and display how we localize sound we need to think of our head as the origin of a coordinate system. There are a few ways that we can give coordinates to a sound in space. We can use a directional vector pointing from one's head to the sound source to determine a position of a sound source in space. The location of the space is defined by three specific planes used in directional audio. These planes are the horizontal plane, the median plane, and the lateral plane.

Some common coordinate systems used to describe audio in space are spherical and interaural-polar. Figure 1.1 shown to the left is a graphical representation of a clockwise spherical coordinate system. The coordinates would be described as (r , θ , ϕ). r is the distance from the origin to the sound source. ϕ is the angle between the directional vector and the sound source where -90 degrees accounts for bottom, 90 degrees accounts top and 0 degrees accounts for the horizontal plane or the middle (Xie 2). Azimuth (θ) is described as the angle between the horizontal projection of the directional vector and the y-axis (Xie 2).

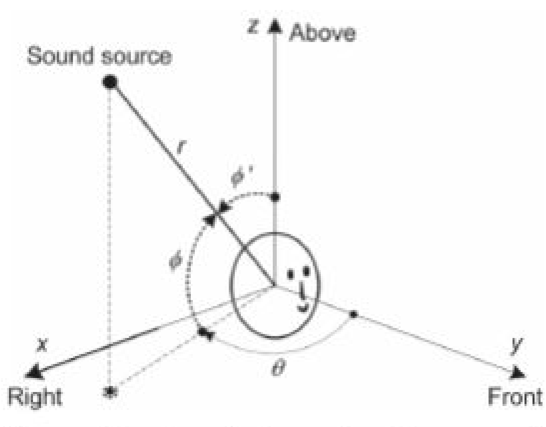

Figure 1.1(Xie 2)

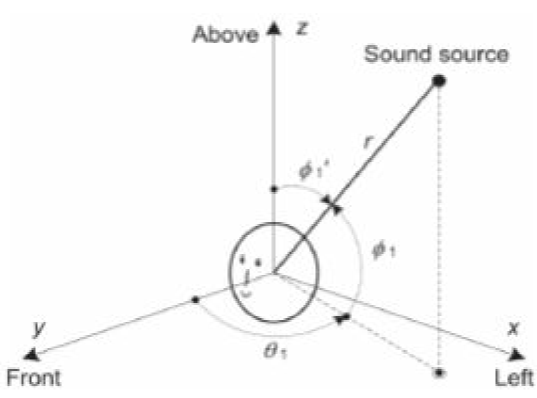

Figure 1.2(Xie 2)

Figure 1.2 is an example of the counterclockwise spherical coordinate system. Counterclockwise and clockwise are almost identical except for the fact that in a clockwise spherical coordinate system an angle of 90 degrees is represented by the right and an angle of 270 degrees is represented by the left while in a counterclockwise spherical coordinate system an angel of 90 degrees is represented by the left and an angle of 270 degrees is represented by the right (Xie 2).

Figure 1.3 is a graphical representation of a interaural polar coordinate system. The sound source location is determined by (r , Θ , Φ). Similar to spherical coordinate systems, r is defined as the distance the sound source is from the origin. The azimuth for interaural polar coordinate systems is represented by Θ and is defined as the angle between the direction vector of the sound source and the median plane (Xie 3). The interaural polar elevation is represented by Φ and is the angle between the projection of the directional vector of the sound source to the median plane and the y axis (Xie 3). For this system (Θ , Φ), (0°, 0°) is the front, (0°, 90°) is the top, (0°, 180°) is the back, (0°, 270°) is the bottom, (90°, 0°) is the right, and (−90°, 0°) is the left (Xie 3). The complementary angle of Θ in this graph is displayed as Γ.

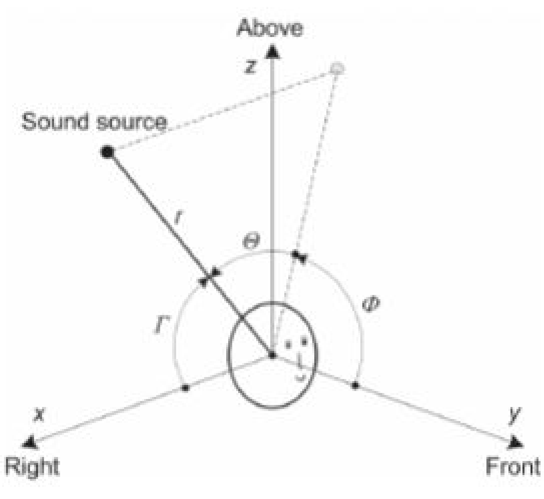

Figure 1.3(Xie 3)

When locating sound we have three cues, azimuth cues (Horizontal plane), elevation cues (Vertical Plane), and range cues (distance). First, let’s look at azimuth cues (Spherical: θ, Interaural polar: Θ). John Strutt over 100 years ago discovered the duplex theory ("Psychoacoustics of Spatial Hearing"). In this theory there are two cues that make up the azimuth cue: interaural time difference (ITD) and interaural level difference (ILD).

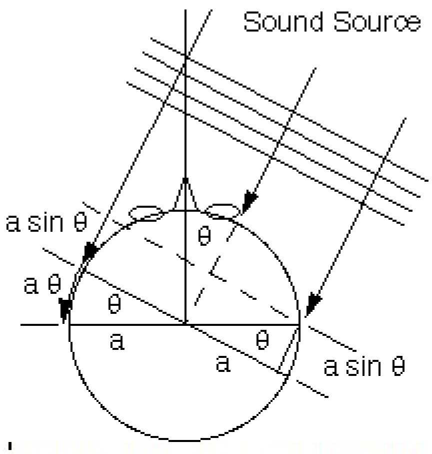

Figure 2 ("Psychoacoustics of Spatial Hearing")

A way to understand ITD would be to consider a sound wave propagating from a distant source. This sound wave would reach our head at radius *a* from a direction specified by the azimuth angle (θ) ("Psychoacoustics of Spatial Hearing"). As shown by this example the sound arrives at the right ear before the left; this is because the sound has more distance to travel to reach the left ear. This can be explained mathematically as a θ + a sin θ ("Psychoacoustics of Spatial Hearing"). We can develop an accurate formula for ITD. The formula would take radius *a* and divide it by the speed of sound ("Psychoacoustics of Spatial Hearing"). Given this we can use this formula to create an accurate understanding of interaural level difference: ITD = a/c(θ + sin θ) ("Psychoacoustics of Spatial Hearing"). To put this simply, ITD is the difference in time between when each of our ears hears a sound.

The next cue that helps make up the azimuth cue is Interaural Level Difference (ILD). ILD can be defined as the difference in sound pressure level reaching the two ears ("Psychoacoustics of Spatial Hearing"). ILD is highly frequency dependent. At low frequencies, where the wavelength of the sound is long, the sound pressure between the two ears has almost no difference. When looking at high frequencies, where the wavelength is much shorter, there could be a 20-dB or greater difference in sound pressure ("Psychoacoustics of Spatial Hearing"). This is called the head-shadow effect, where "the far ear is in the sound shadow of the head" ("Psychoacoustics of Spatial Hearing"). ILD can be simply explained as the the difference in perceived loudness in each ear.

The Duplex Theory developed by John Strutt states that the ILD and the ITD are complementary ("Psychoacoustics of Spatial Hearing"). When there are low frequencies below around 1.5 kHz there is not much ILD information because of the sound pressure that occurs within lower frequencies ("Psychoacoustics of Spatial Hearing"). The ITD makes up for this, though, to create easy location detection. When there are high frequencies above 1.5 kHz there is uncertainty in the ITD ("Psychoacoustics of Spatial Hearing"). ILD resolves this directional uncertainty. Thus, creating the azimuth cue provides localization information throughout the audible frequency range.

Now the the main cues for azimuth, left and right localization, are binaural, meaning two different listening points make up the localization; the cues for elevation are said to be monaural. Ofcourse when localizing a sound one would still hear with both of our ears when something is above or below us, but the way that is currently understood is through monaural cues. Yet we can estimate elevations there is still more studies being conducted so we can understand it better (Allen). The way we currently understand elevation cues is because of our ear and how sound travels inside it. The outer ear acts as an acoustic antenna. It has resonant cavities that boost some frequencies while the shape of the outer ear causes interference effects that attenuates other frequencies. This creates a frequency response that is directionally dependent ("Psychoacoustics of Spatial Hearing").

As shown in the figure, there are two different directions of arrival: above the ear and in front of the ear. In each instance we see two paths from the sound source to the ear. When low frequencies are present there is more energy from the wave so the sound stays in phase. When dealing with high frequencies we see the delayed signal being out of phase with the direct signal; this causes destructive interference ("Psychoacoustics of Spatial Hearing"). This is the most abundant when the delayed signal length is a half wavelength. This would look like f = c/2d where f is frequency, c is cycles and d is the length ("Psychoacoustics of Spatial Hearing"). This creates what is called the pinna notch, which is shown at 10 kHz in the figure above. When looking at typical values for the length you can see this notch occur from 6 kHz to 16 kHz ("Psychoacoustics of Spatial Hearing"). As seen in the figure the pinna reflects sounds coming from the front more effectively than for sounds coming from above. So, the pinna notch is more apparent for sources in front than for sources coming from above. This gives us our elevation cue because the path length difference

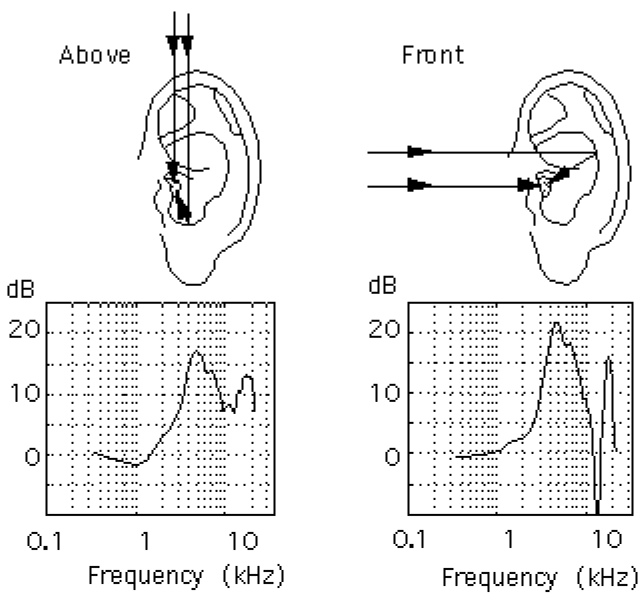

Figure 3 ("Psychoacoustics of Spatial Hearing") changes with elevation angle, which means the frequency of the notch moves with elevation.

Another aspect for spatial localization of sound is how we perceive distance or range of a sound source. The human auditory system is not as accurate at perceiving distance as it is perceiving direction. Even though this is the case, we can still form a basic perception of distance, though that perception is prone to errors: “Experiments have demonstrated that the auditory system tends to significantly underestimate distances for sound sources at a physical source distance farther than about 1.6 m, and typically overestimates distances for sound sources at a physical source distance closer than about 1.6 m” (Xie 19). Pavel Zahorik suggests to use a compressive power function, which then translates to a good approximation of a majority of distance psychophysical functions. The function is $r' = \kappa r \alpha$, where r' is an estimate of perceived distance, r is the physical source distance, and κ and α are fit parameters of the power function (Zahorik 410). Distance perception is based on multiple cues, including intensity, direct-to-reverberant energy ratio, spectrum, binaural cues, and dynamic cues. Even though the research for distance estimation has increased, there is still a lot that is not understood about the perception of distance and its estimation.

With an understanding of how one localizes sound, the head-related transfer function, or HRTF, makes more sense. According to Tilen Potisk, the

> [h]ead-related transfer function (HRTF) is a function used in acoustics that characterizes how a particular ear (left or right) receives sound from a point in

space. A pair of two transfer functions, one for each ear, is used for sound localization which is very important for humans. (Potisk 1)

What HRTF does is capture the transformations of a sound wave propagating from a sound source to our ear. The transformations looked at in HRTF are diffractions and reflections that occur from our bodies. Relevant body parts include our head, pinnae, shoulders and torso. With the function from each ear we are able to generate a illusion of spatially located sound (Potisk 1).

HRTF is created by a Fourier transform of a head-related impulse response (HRIR) (Potisk 2). This function must be defined for each ear, having the information about the magnitude and phase shift for each. The impulse responses for the left and right ear in the time domain are indicated as hL(t) and hR(t). For the frequency domain, they are indicated by using HR(ω) and HL(ω). x(t) describes the pressure of the sound source. xL(t) and xR(t) will be the pressure when reaching the left and right ear. Potisk states that “[i]n the time domain, the pressure at the ears can be written as a convolution of the sound signal and the HRIR of the corresponding ear:

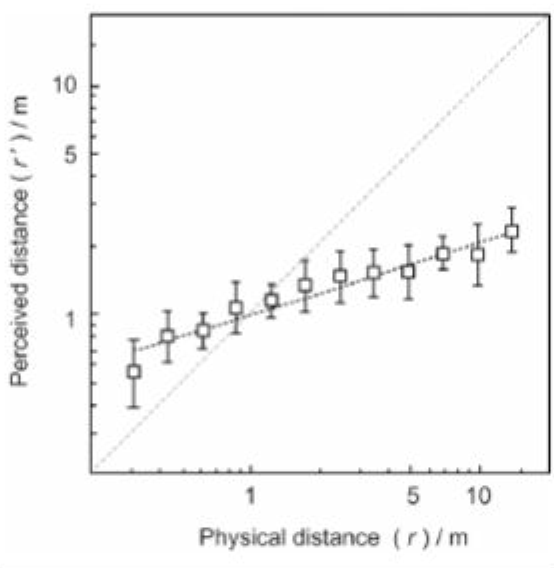

Figure 4 (Xie 19)

$$x_{L,R}(t) = h_{L,R}(t) * x(t) = \int_{-\infty}^{\infty} h_{L,R}(t-\tau)x(\tau)\mathrm{d}\tau.$$

In the frequency domain, convolution is transformed into multiplication:

$$X_{L,R}(\omega) = \mathcal{F}\left(h_{L,R}(t) * x(t)\right) = H_{L,R}(\omega)X(\omega).$$ ” (2).

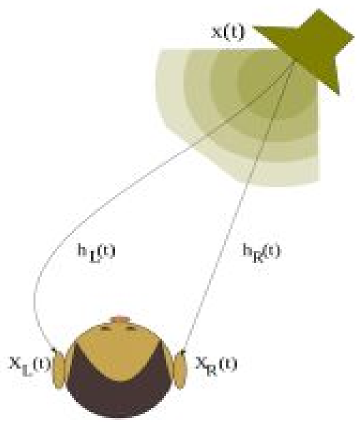

Figure 5 (Potisk 2)

Figure 5 is a diagram of a soundwave from a sound source moving to the ears with the conventional notations for filtering functions and pressure impulses. So looking at these functions and using the figure provided by Potisk, one can see how HRTF can be used to to give the illusion of spatially located sound.

With a understanding of how we naturally locate a sound source in space and how we can create an illusion of a sound source in space, we will talk about the current ways we use 3D and binaural audio. When working with binaural audio, a reliable method for creating binaural audio is by recording using a binaural microphone system. The earliest binaural microphone known was developed and displayed to the public was in 1933 when AT&T displayed their mechanical man, Oscar (Kall). Oscar had microphones for ears just as many modern binaural microphone systems place microphones in the ears of dummy heads. Oscar was presented to the Chicago World’s Fair, where people could put on headphones and listen in real time to what Oscar was hearing (Kall). This was very popular at the fair, yet no one really knew how this microphone system would be implemented. AT&T was ahead of their time because it was roughly another forty years when Neumann developed the dummy head that is used today for binaural recordings (Kall).

The first microphone created for binaural recording purposes was developed by Neumann in 1972. The microphone was a dummy head was called the KU80. The potential was quickly seen for use in the entertainment industry. Later in 1980, Neumann improved the KU80 and developed the KU81 which used diffuse-field equalization (Fortin-Chabot 8). Diffuse-field equalization a measurement which adjusts the signal so that the measured frequency response corresponds to that in the diffuse field. This helps when capturing the sound because prior to this technique the reverberations would be louder than the direct field causing a unnatural response (“What is Meant by Diffuse-Field Equalisation and Free-Field Equalisation”).

Today Neumann is still atop the best brands for any microphone, let alone binaural microphones. The latest dummy head is the KU 100, which is one of the top dummy heads on the market. This quality of microphone currently costs around 8,000 dollars (“KU 100 Dummy Head Microphone”). To achieve the best result for a binaural recording, a microphone such as the KU 100 might be the best option, but there are still many other microphones that are capable of creating binaural recordings.

When it comes to cheaper alternatives you can find solutions that cost about 100 dollars such as the Roland CS-10EM (“CS-10EM Binaural Microphones/Earphones”). The Roland CS-10EM is very interesting because the microphones are placed in one's ears. One fault that is apparent with this system is elevation cues would not be completely accurate. This inaccuracy occurs because how how sound hits the pinnae or the outer ear. When thinking about the pinnae notch discussed in the elevation cues, that notch is created from the delay that occurs from sound hitting different parts of the ear and causing phase cancellation in certain frequencies. Using these microphones would obscure the natural shape of the ear causing for possible complications. Despite the issue for elevational cues, one would still have great distance as well as location cues, and the recording would still be perceived as a 3D environment by listeners.

Another binaural microphone option that is somewhere between the Roland CS-10EM and the Neumann KU100 is from the company 3Dio. 3Dio has a variety of binaural microphones such ranging in price from 500 dollars to 2,000 dollars. For the purpose of this paper we will look at the 800-dollar Free Space XLR Binaural Microphone. This microphone comes mounted with two very realistic silicone ears. These ears are useful for replicating a realistic 3D environment ("Free Space XLR Binaural Microphone"). The one issue that stems from 3Dio binaural microphones is that they do not have a dummy head in between the two ears. This may cause the recording to differ from something such as the KU 100. If we look at how we establish HRTFs we account for head, pinnae, shoulders and torso. Without these we cannot get a true rendering of the space. This is because we would not be accounting for diffractions and reflections of sound caused by the body. With the KU 100 when recording you still achieve the diffractions and reflections from the head and pinnae, while the 3Dio binaural microphone only account for the pinnae. This would cause for a different response on how the sound reaches each ear making it not a completely accurate binaural recording. However, unless one is working on the research of spatial audio this microphone will get the job done.

Binaural audio is used in a bunch of different ways. It is used in music, video games, art installations, and video. In each way that it is used, binaural audio adds a new characteristic that transforms how the media is perceived. One way binaural audio had been used in music was in R.E.M’s second album, Reckoning, released in 1984. Don Dixon, the co-producer of the album, fell in love with binaural audio and wanted to implement it in this album (Chabot 14). When recording they did not track with a dummy head, instead they just used a box with a microphone on each side (Chabot 14). This technique was used as a type of room microphone to give a live feel to the album (Chabot 14). Even though more could have been done with the recording techniques, it is interesting to see how once people became aware of binaural audio they created their own techniques to make it work for themselves.

Another use for binaural 3D audio is shown in a variety of video games. In current video games, audio plays a huge role into the game play. While playing games such as *Fortnite*, where 100 online players are battling each other to be the last person standing, audio is used to give clues on where a opponent may be located. This uses 3D audio to give a representation of where this opponent may be even if they are not visible. When using headphones while playing *Fortnite,* the player may be able to distinguish whether a player is above or below. Most game audio is put into a binaural medium that way most people can use it. With games such as *Fortnite* the 3D environments created are not entirely realistic but give clues that are relevant to the game.

There are a few ways to listen to 3D environments. One way is through a binaural medium. This means that through any set of headphones one can experience 3D audio presented in a binaural state. Another listening environment for 3D audio would be virtual reality headsets. With these headsets one still uses headphones as a source of receiving audio. What is special about VR headsets is that they are able to track the user's head movements. This means that the user experience is very immersive and a very realistic 3D audio experience. With the head tracking capabilities the system is able to adjust the audio environment to wherever your head is in the given environment. Another listening environment is through a set loudspeaker format such as 9.1, 10.1, 11.1, or many others. One leader in consumer loudspeaker formats is a company called Auro-3D. The Auro-3D allows someone with a conventional 5.1 surround sound system to upgrade with at least 4 more speakers to give elevation cues to the environment ("What Is Auro-3D?"). The information about the height is captured by the recording or created in the mixing process. The mix is created in a standard 5.1 PCM stream. The Auro-Codec Decoder is able to take that 5.1 signal and decode it to the original Auro-3D mix (Doc). This then

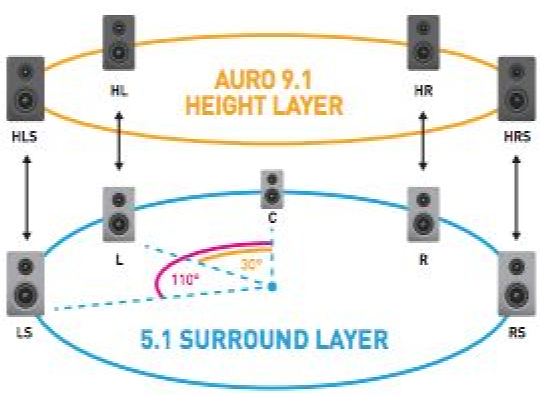

Figure 6 ("What is Aura 3D?")

can be used with the Auro speaker layout.. Through this system, one can experience a 3D audio environment.

The last listening environment is made possible with ambisonics. Ambisonics creates a spherical sound that can arrive from all directions. When using loudspeaker formats we are sending signals to each speaker to create a 3D environment illusion. It creates a decent representation of 3D audio but is not the most realistic display of a 3D environment. When using ambisonics we are able to create an arbitrary amount of points where a sound source can occur in a listener environment (Allen). It also gives us a arbitrary loudspeaker layouts. This means that we have multiple ways of achieving this spherical representation of 3D sound environments. When it comes to ambisonics the more outputs available, the more accurately we are able to localize sound. Systems such as the one seen in Figure 7 use ambisonics to create a truly realistic auditory display of 3D environments. Unfortunately spherical loudspeaker systems such as these cost roughly 10 million dollars (Allen). The great thing about ambisonics is that we are able to convert these signals into multiple different formats. We may lose the some spatial resolution with converting the signal down, but it gives the average consumer the ability to use ambisonics in a range of listening environments.

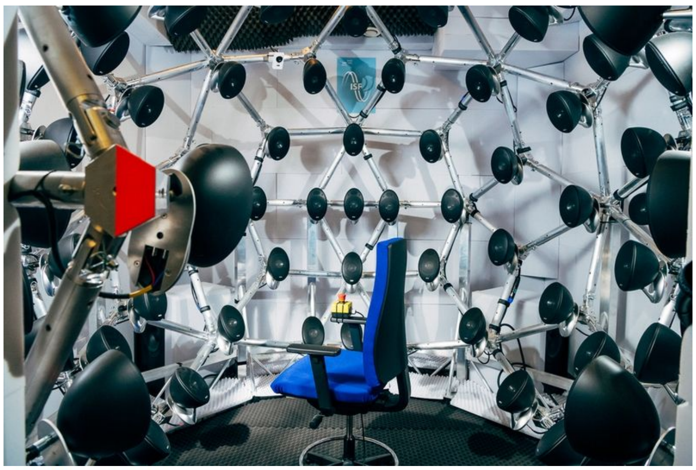

Figure 7 (Allen)

To achieve a result of a basic 3D audio environment in a binaural medium is as easy as recording that environment with a binaural microphone. Yet recording techniques for non-immersive 3D audio are fairly simple. There are interesting recording techniques that help us understand how to create immersive 3D audio environments. When creating a 3D environment, ambisonics is an effective method of achieving this:

> Ambisonics is a method for recording, mixing and playing back three-dimensional 360-degree audio…The basic approach of Ambisonics is to treat an audio scene as a full 360-degree sphere of sound coming from different directions around a center point ("Ambisonics Explained: A Guide for Sound Engineers").

To create an accurate binaural rendering of ambisonics one would need to use the idea of spherical loudspeaker arrays and HRTFs (Allen). The recording method to achieve this would be by placing a binaural microphone in the center of an ambisonics array and then generate sound from each loudspeaker of the ambisonics array independently to create an HRTF for each. This would help to create a set of virtual loudspeakers. With this information we could process the ambisonics with the newly created virtual speakers. This will create a binaural mix of an ambisonics signal (Allen). This process allows ambisonic channels to represent arbitrary loudspeaker layouts.

The future possibilities for 3D audio and the mediums to represent it can go in a few different directions. Currently we can experience two different mediums for 3D environments: ambisonic rendering of speaker layouts or binaural rendering for headphones. Currently spherical speaker layouts give us the best representation of 3D audio, but they are too expensive for the average consumer. The current rendering of ambisonics into a binaural medium does a decent job at creating realistic 3D environments. With these two mediums, would it be more practical to find cheaper solutions for accurate representation of 3D audio through speaker layouts? Or would enhancing the process of binaural rendering of ambisonics to a headphone system be the future of 3D audio?

When looking at speaker layouts to represent the best illusion of 3D audio we need to look at what goes into creating the system. When creating these speaker arrays 50 or more speakers are needed to create a great 3D environment. To create a accurate environment one would need high quality speakers. One may ask why wouldn’t normal speakers, such as a bunch of cheap car speakers, work for creating this spherical environment? The answer is that the basic premise would work, but to create the best environment great quality speakers are needed. The reason why quality speakers are important is because the speakers need to match almost perfectly in level and frequency response for the effect to work best (Allen). This is why, for now, the direction for advancement in 3D is to enhance what we can do binaurally.

The next question would be what needs to be done to enhance our current binaural rendering of 3D audio. One way to enhance binaural rendering would be by expanding higher order ambisonics (Allen). What higher order of ambisonics means is to give more channels to represent where sound comes from. Though this seems very simple, we have to understand that there may be psychoacoustic limit of abmisonic rendering for humans. Given this limit one may ask are we able to capture everything in HRTFs that is necessary to reach that limit and render it for humans (Allen).

To create the perfect HRTF for someone requires the exact measurements of that person's ear. This creates an issue for mass production of 3D audio in binaural rendering of ambisonics as it is impossible to get a individual rendering of HRTF for every single user. This means that the vertical confusion discussed earlier in the paper remain problematic without a completely accurate HRTF for the user. Currently with consumer products most companies aim for an ear shape that suits about 70% of the population. (Allen). Even though this ear shape is stated to suit 70% of the population, the probability of it actually being a perfect fit is much lower than 70%. To create a better HRTF, more understanding of timbral differences is assumed to help dissolve vertical confusion (Allen). Currently, studies are being done to understand these timbral differences, but nothing has been found that can be implemented into HRTFs. To create a perfect binaural rendering for 3D audio for the future more studies need to happen about the non-linear processing the brain does to recall and react spatially; until then we use our current HRTF to give the best result possible (Allen).

3D audio is a growing topic among audio enthusiasts. This growth of enthusiasm stems from the development of virtual reality and new forms of multimedia where 3D audio can be presented. Due to this growth of enthusiasm the implementation of 3D audio is becoming more commonplace. With 3D audio becoming more commonplace more studies will be conducted to try and enhance our experience of 3D audio. There is still a lot of work left to understanding how our brains localize sound, but the future in this research is looking very bright.

Works Cited

"Alan Blumlein and the Invention of Stereo." *EMI Archive Trust*, 16 Dec. 2013, www.emiarchivetrust.org/alan-blumlein-and-the-invention-of-stereo/.

Alexandre Fortin-Chabot. "Diversifying Binaural Recording Techniques: A Study of Recording Methods and Equipment for Binaural Audio." *SCAD Libraries* , *Savannah College of Art and Design*, May 2015, ecollections.scad.edu/iii/cpro/DigitalItemViewPage.external?lang=eng&sp=1002896&sp=T&sp=Pall%2CRu1000001%40%2CQdiversifying+binaural+recording+techniques%3A+a+study+of+recording+methods+and+equipment+for+binaural+audio&suite=def.

Allen, Andrew. "Current Practices In 3D Audio." 2 May 2018, Seaside, California State University, Monterey Bay.

"Ambisonics Explained: A Guide for Sound Engineers." *Waves.com*, www.waves.com/ambisonics-explained-guide-for-sound-engineers.

"CS-10EM Binaural Microphones/Earphones." *B&H Explora*, B&H Photo Video, 16 Oct. 2010, www.bhphotovideo.com/c/product/687077-REG/Roland_CS_10EM_CS_10EM_Binaural_Microphones_Earphones.html.

Fortnite. PS4, Epic Games, 2018.

"Free Space XLR Binaural Microphone." *3Dio*, 2018, 3diosound.com/products/free-space-xlr-binaural-microphone.

Kall, Alex. “A Brief History of Binaural Audio.” *Kall Binaural Audio*, 21 Aug. 2013, www.kallbinauralaudio.com/a-brief-history-of-binaural-audio/.

M. Kronlachner, W. Ritsch, “Towards usability of Higher Order Ambisonics in Digital Audio Workstations“, 2nd International Conference Music and Technologies, Kaunas, Lithuania, November 2012.

“KU 100 Dummy Head Microphone” *B&H Explora*, B&H Photo Video, 29 Nov. 2016, www.bhphotovideo.com/c/product/146368-REG/Neumann_KU_100_KU_100_Dummy_Head.html.

Zahorik, Pavel, et al. “Auditory Distance Perception in Humans: A Summary of Past and Present Research.” *Acta Acustica United with Acustica*, May 2005, pp. 409–421. *Research Gate*, www.researchgate.net/publication/229068125_Auditory_distance_perception_in_humans_A_summary_of_past_and_present_research.

“Psychoacoustics of Spatial Hearing.” *The CIPIC Interface Laboratory Home Page*, www.ece.ucdavis.edu/cipic/spatial-sound/tutorial/psychoacoustics-of-spatial-hearing/.

“What Is Auro-3D?” *Auro Technologies*, www.auro-3d.com.

“What Is Meant by Diffuse-Field Equalisation and Free-Field Equalisation?” *Beyerdynamic Help Center*, 9 Sept. 2015, support.beyerdynamic.com/hc/en-us/articles/201058811-What-is-meant-by-diffuse-field-equalisation-and-free-field-equalisation-.

Xie, Bosun. *Head-Related Transfer Function and Virtual Auditory Display*. J. Ross Publishing, 2013.