Acústica 2020
XI Congreso Ibérico de Acústica
51º Congreso Español de Acústica
TECNIACÚSTICA 2020

PERCEPTUAL COMPARATIVE TESTS BETWEEN THE MULTICHANNEL 3D CAPTURING SYSTEMS ARTIFICIAL EARS AND THE AMBISONIC CONCEPT

Paulo, J.[1,2], Banos, M.[1,2]

[1] ISEL-Instituto Superior de Engenharia de Lisboa, Portugal
[2] LAA-Lab. de Áudio e Acústica do ISEL, Portugal
{joel.paulo@isel.pt, a47024@alunos.isel.pt }

Resumo

A abordagem de realidade virtual e aumentada, VR/AR, veio realçar a necessidade de som imersivo em audição binaural. Assim, a captação de ambientes acústicos reais é realizada recorrendo a sistemas multicanal para som 3D, nomeadamente o conceito ambisonic e sistemas com vários pares de ouvidos artificiais. Neste trabalho usou-se um microfone ambisonic de primeira ordem e foi desenvolvido um sistema de cabeça com 4 pares de ouvidos em geometria de quadrado (OctoEar) no plano horizontal para obter o campo sonoro imersivo. Adicionalmente, foi desenvolvido um sistema *sound recorder* com armazenamento em cartão SD utilizando a plataforma Teensy para tornar a sua utilização versátil em situações de captação de música ao vivo e ambientes de natureza. A abordagem ambisonic de primeira ordem tem o problema de fraca resolução espacial. O sistema OctoEar tem a vantagem de permitir obter audição binaural pura para as quatro direções xx´, -xx´, yy´ e -yy´. Contudo, para as restantes direções de audição do utilizador aplicam-se técnicas de interpolação o que pode danificar o campo acústico conjuntamente com problemas de difração entre os ouvidos. Os testes de caracterização dos dois sistemas foram realizados em câmara anecoica para obter as suas respostas ao impulso para posterior utilização no processamento de sinal de conversão para sistema binaural. Os testes preliminares de audição binaural, realizados a um conjunto de utilizadores, revelam que o sistema OctoEar confere melhor realismo de som no plano horizontal. Contudo, é menos versátil, mais volumoso, não passa despercebido pelas pessoas e tem pouca informação sobre a direção de elevação. Este projeto encontra-se ainda em desenvolvimento prevendo a sua conclusão no final de Setembro onde teremos mais conclusões.

Palavras-chave: processamento de sinais, audição binaural, captação de som 3D, plataforma Teensy.

Abstract

The virtual and augmented reality approach, VR / AR, highlighted the need for immersive sound in binaural hearing. Thus, the capture of real acoustic environments is carried out using multichannel systems for 3D sound, namely the ambisonic concept and systems with several pairs of artificial ears. In this work ambisonic microphone of first order was used and a head system with 4 pairs of ears in square geometry in the horizontal plane (OctoEar) was developed to obtain the immersive sound field. In addition, a sound recorder system with SD card storage using the Teensy platform was developed to make its use versatile in situations of capturing live music and nature environments. The first-order ambisonic approach has the problem of poor spatial resolution. The OctoEar system has the advantage of allowing pure binaural hearing in the four directions xx´, -xx´, yy´ and -yy´. However, for the rest of

the user's listening directions, interpolation techniques are applied, which can damage the sound field together with diffraction problems between the ears. The characterization tests of the two systems were carried out in an anechoic chamber to obtain their impulse responses for later use in the conversion signal processing for binaural system. Preliminary tests of binaural hearing, performed on a group of users, reveal that the OctoEar system gives better sound realism in the horizontal plane. However, it is less versatile, more bulky, does not go unnoticed by people and has little information about the direction of elevation. This project is still under development and is expected to be concluded at the end of September, where we will have more conclusions.

Keywords: signal processing, 3D sound capturing –artificial ears and ambisonic, Teensy platform.

PACS no. 43.60.-c, 43.60.Dh, 43.60.Jn

1 Introduction

When we talk about realism in audio, most of the time we only take into account the raw quality of sound files, which is defined mainly by the accuracy (or lack) of the equipment used in recording, the sampling frequency and bit depth of the file, the presence of artifacts related to the compression algorithms (such as MP3)… but one of the characteristics of a sound that makes it realistic to our ears is the possibility of discerning where it comes from, how far away it is, how it reacts to the room that surrounds it… All of this aspects are studied when we talk about binaural hearing, a technique that applies different levels of signal processing to a base sound to make it sound like if it was coming from a specific point in space.

A. The human hearing system as a signal processing device

Our hearing system is an integral part of our connection to the world around us, the innate ability that we humans have to discern the position of a sound source just by hearing is as important as our sight to navigate our environment. To achieve this, our brain uses mainly two techniques, that complement each other and allows very high precision:

1) Interaural Time Difference

When a sound is produced by a source, the sound wave travels through the air and reaches our ears in different times, first it gets to the ear closer to the source, and then to the other. This delay is perceived and then used to pinpoint the average position of the source of said sound. This time difference is called Interaural Time Difference (ITD). The ITD can then be estimated by:

$$ITD = \frac{r(\theta + \sin\theta)}{c} \tag{1}$$

where r is the radius of the head (in meters), θ is the angle of incidence in radians, and c is the velocity of sound in the medium (for air, 340 meters per second), as shown in Figure 1.

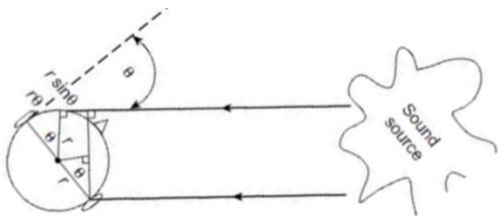

Figure 1 – Basic scheme for the Interaural Time Difference (ITD) cue.

2) Interaural Level Difference

When a sound comes from the right relative to the listener position, the sound arriving at the left ear is not only delayed, but it is also diminished in amplitude by the shadowing effect of the head getting in the way of the sound waves. This effect is known as Interaural Level Difference (ILD), and it is the second mechanism that our hearing uses to find the position of sounds. This effect is not constant in frequency, as the lower frequencies get less attenuated than the higher frequencies. This way, the shape of the filter can be approximated by the following equation:

$$ILD = 1 + \left(\frac{f}{1000}\right)^{0.8} \sin\theta \tag{2}$$

where f is the frequency of the sound (in Hz), and θ is the angle of incidence. As mentioned earlier, this formula gives less attenuation for the lower frequencies than for the higher ones.

3) Head Related Transfer Functions

This two phenomena (ILD and ITD), when applied to a certain monophonic sound, give a sensation of realism that is fairly limited, they only modify the sound with general and approximate formulas that don't really reflect reality that faithfully. Therefore, the approach of using the head related transfer functions (HRTF) become more interesting. The HRTF is a transfer function that defines how the sound is modified when arriving with a certain direction to the listener, this means that there are as much HRTF as resolution we want to have in our system. The typical way to create this HRTF is using a dummy head, that simulates an average human head in size and shape, making the sound recorded by the two microphones inside (one for each ear) much more realistic than that you can get from a microphone pointed directly at the sound.

Using this HRTF as a filter by any other sound, the output will be a simulation of the sound as if it was recorded using the setup with the dummy head, and thus fairly similar to what a human would hear.

B. Approaches to tridimensional sound fields

A tridimensional (3D) sound field is a sound environment that allows the listener to hear sound as if it was coming from any position in the 3D space around his head. Therefore, the techniques, signal processing algorithms and devices used to create this kind of sound fields are studied under the concept of 3D sound synthesis. This field of sound processing is one of the less developed because the need for 3D sound synthesis is recent necessity related to the increase of virtual reality (VR) devices and applications. The first instances of 3D audio can be found in some music recordings from the 80's and 90's where the musicians would play in a studio room, but instead of using the usual microphone setup, they would record everything using a dummy head or similar approaches, this way, when listened with headphones, the music would sound like if each instrument was coming from a different point in the room. The first example of what was called "binaural recording" was the album Street Hassle, by Lou Reed, released in 1978. This first approach was limited to the constraints of the processing power of the time, it didn't take into account the listener turning and moving his head around, and was overall a novelty, more of a curiosity than an actual useful recording technique.

In recent years the surge of VR and the exponential increase of computing power has put 3D sound on the spotlight, now that we have immersive visual environments, we need to take it a step further.

1) Google Omnitone

This is Google's attempt on 3D sound synthesis, and as almost everything Google does, it is oriented for online use. The Omnitone is a tool based on Ambisonics, that allows online reproduction of 3D sound and 360º video and its compatible with Google's own Google Cardboard VR headset, Oculus Rift and many other VR sets. Its main characteristic it's the capability of being integrated on a webpage, and still working in real time.

2) 3D Audio Spatialization by Oculus

Oculus is one of the biggest companies investing in VR techniques and 3D sound development, being the creators of the Oculus Rift headset, they want all videogame studios and developers to have the capability of creating a 3D soundscape realist enough to boost sales of his product. This approach of 3D sound is fairly simple, they use the head tracking ability included in their VR headset in combination with a big HRTF database to simulate the orientation of a sound, and then they apply algorithms that take into account the attenuation of sound, the doppler effect and the reverberation… in order to make the sound as close to reality as possible.

3) Vector-Based Amplitude Panning

The vector-based amplitude panning (VBAP) is a case of amplitude panning that imitates the sensation of spatialization of sound by creating an array of virtual loudspeakers around the listeners head, each one having a different amplitude gain calculated by its position relative to the listener [1-4]. Consider a triangle in front of the listener, with l_m, l_n and l_k being the position of 3 loudspeakers, and g_m, g_n and g_k being the amplitude gain for each loudspeaker. By applying a certain value of gain to each speaker, we can simulate the sound coming from any point inside the triangle. Increasing the number of loudspeakers a bigger area around the listeners head can be simulated. Figure 2 shows this approach.

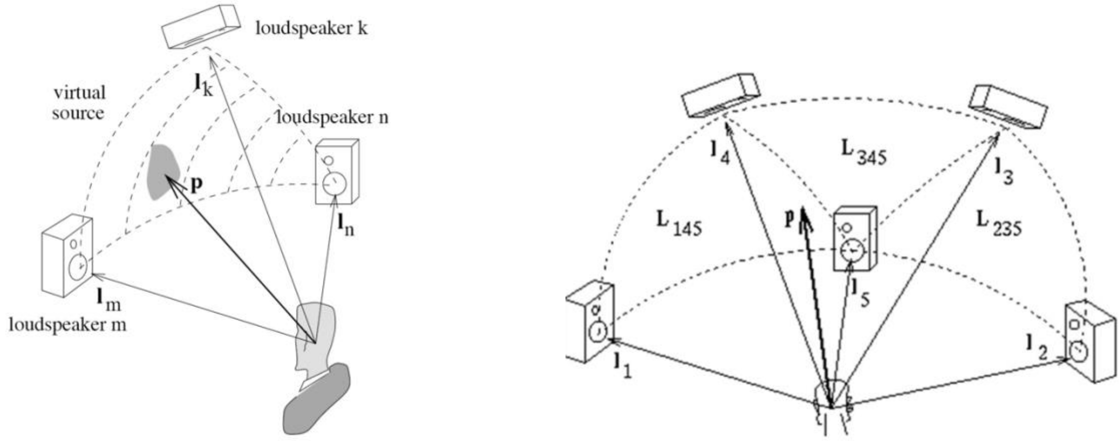

Figure 2 – Vector-based amplitude panning (VBAP) technique.

This technique is limited to giving the sensation of a sound coming from a specific point in space, but it is uncapable of giving information of the distance, and also doesn't take into account the effects that the physical elements of the hearing system (the head, the pinnae…) has on the sound. Nonetheless, the results are fairly accurate, and it has the advantage of being easily extrapolated for real-life loudspeaker systems [5].

4) Interpolation of HRTFs using VBAP

The gains obtained by the VBAP approach have proven to be useful for application in binaural audio. If we take the signal generated applying three HRTFs (h_m, h_n and h_k) related to the position of the loudspeakers from the example before (l_m, l_n and l_k) and its associated gain values (g_m, g_n and g_k) and we apply them to the signals generated before, we would get an output like this:

$$y(t) = (g_n h_n + g_m h_m + g_k h_k) * x(t) \tag{3}$$

With this approach, we are doing a sort of linear interpolation of the HRTFs.

$$\hat{h} = g_n h_n + g_m h_m + g_k h_k \tag{4}$$

5) Spectral Interpolation

If the technique of VBAP interpolation works by making a linear combination of waveforms, the spectral interpolation works by shaping the spectra of the interpolated HRTF. This technique is better explained in graphic form:

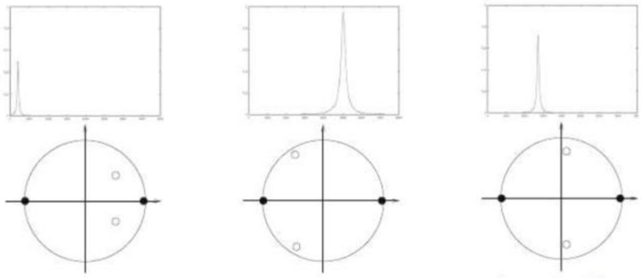

Figure 3 – Pole-zero diagrams and spetral interpolated filter.

To use this method, we would need to synthesize the HRTF as low order IIR filters, which is in itself a complicated process [1]. Then, we would need to analyze this HRTF filters in polar coordinates, given that the position of the poles and zeros in the interpolated filter would be on median positions between the coordinates of the poles and zeros in the A and B filters, both in angle and magnitude.

C. First conclusions

Taking into account the ease of use, computational power needed, and the final objective of this project, we will settle in using a modified algorithm to the one explained before, Interpolation of HRTFs using VBAP. For our particular case, we will only simulate information regarding the azimuth of the sound, but it would be easily modified to regard the elevation, more on that later.

2 SETUP AND DEVELOPMENT

A. Measurement setup

The setup consists of a sound source and the recording equipment. The recording equipment consists in 4 pairs of microphones with their respective artificial ears, recording in each one of the directions drawn as dotted lines: red for the front (0º), yellow for the right (90º), magenta for the back (180º) and blue for the left (-90º), as shown in Figure 4. Thus, we will obtain 4 sound files, one for each pair, plus the original sound that was emitted by the sound source.

B. Treatment of the recorded material

As stated before, we will obtain 4 sound files from our setup, one for each pair of ears, each one associated with one direction (front, back, left, right). The first step would be to apply the FFT function to each sound file, including the original one. Now, to calculate the HRTF associated with each direction we just need to use the equation:

$$HRTF = \frac{FFT(Recorded\ sound)}{FFT(Original\ sound)} \tag{5}$$

Therefore, we will get 4 HRTFs: HRTF_F for the front, HRTF_B for the back, HRTF_L for the left and HRTF_R for the right. Now, we have all the information we need from our measurements.

C. First considerations

1) Planning ahead

Before we can start coding, we need to plan what we need our code to do, this is easier to see using a simple scheme (Figure 2). Our virtual environment would be formed mainly by HRTF_F, HRTF_L, HRTF_B, HRTF_R and finally, the listener.

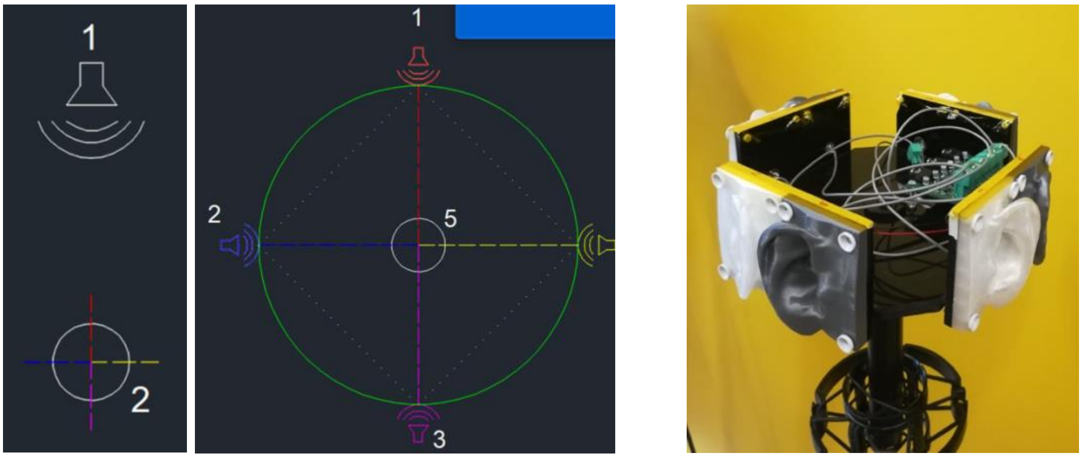

Figure 3 - System basic scheme and a photograph of the OctoEar, apparatus with 4 pairs of ears.

Using the VBAP approach, we can interpolate any HRTF in the green circle, this is, any HRTF around the listener. To do this, we will apply linear interpolation between pairs of the HRTFs:

- HRTF_F and HRTF_R for $0 \geq \theta \geq 90$
- HRTF_R and HRTF_B for $90 > \theta \geq 180$
- HRTF_F and HRTF_L for $-90 > \theta \geq 0$
- HRTF_L and HRTF_B for $-180 \geq \theta > -90$

We will also try to use another VBAP approach, which would be a cubic interpolation, using the four HRTF for each calculation:

- HRTF_F, HRTF_R, HRTF_B and HRTF_L for $0 \geq \theta \geq 270$
- HRTF_R, HRTF_B, HRTF_L and HRTF_F for $270 \geq \theta \geq 360$

Even though we use the four HRTF for each calculation, in this case the order of the factors does in fact alter the results, more on this on the calculations and formulas part. We will try these two approaches (linear and cubic interpolation) and test the results on a subjective way, as in this project the subjective perception is more important than any other factor.

2) Formulas and calculations

In this part we will explain the different mathematical expressions in which we based our algorithms.

a) Linear interpolation

This approach is fairly simple, we will interpolate any HRTF in between two known HRTF in intervals of 90 degrees. We will use the following formula:

$$Interpolated\ HRTF = a * HRTF_{n*90} + b * HRTF_{(n+1)*90} \quad n = 1, 2, 3 \tag{6}$$

This way, any HRTF can be found as the sum of two HRTF weighted by two coefficients, a and b. These coefficients can be calculated as:

$$a = 1 - b;\ b = \frac{\theta_{int} - \theta_{fst}}{\theta_{last} - \theta_{fst}}; \tag{7}$$

where θ_{int} is the angle associated with the HRTF we are interpolating, θ_{fst} is the angle of the first HRTF and θ_{last} is the angle of the last HRTF. This is the simplest approach; we will see if the results are good enough.

b) Cubic interpolation

In this case, all the HRTFs will be used during the calculations:

$$Interpolated\ HRTF = \\ = \begin{cases} a * HRTF_F + b * HRTF_R + c * HRTF_B + d * HRTF_L \ (0 \leq \theta_{int} \leq 270) \\ a * HRTF_R + b * HRTF_B + c * HRTF_L + d * HRTF_F \ (270 \leq \theta_{int} \leq 360) \end{cases} \quad (8)$$

where θ_{int} is the angle associated with the interpolated HRTF. Even though the two equations seem similar, the difference is on what angle we associate with the HRTF_F:

- For the first case ($0 \leq \theta_{int} \leq 270$) HRTF_F is the HRTF for a 0 degree angle.
- For the second case ($270 \leq \theta_{int} \leq 360$) HRTF_F is the 360 degree angle.

these two angles are the same in a mathematical sense, but we need to make the distinction in order to use the weighting values correctly:

$$a = \frac{(\theta_{int} - \theta_2)(\theta_{int} - \theta_3)(\theta_{int} - \theta_4)}{(\theta_1 - \theta_2)(\theta_1 - \theta_3)(\theta_1 - \theta_4)};$$

$$b = \frac{(\theta_{int} - \theta_1)(\theta_{int} - \theta_3)(\theta_{int} - \theta_4)}{(\theta_2 - \theta_1)(\theta_2 - \theta_3)(\theta_2 - \theta_4)};$$

$$c = \frac{(\theta_{int} - \theta_1)(\theta_{int} - \theta_2)(\theta_{int} - \theta_4)}{(\theta_3 - \theta_1)(\theta_3 - \theta_2)(\theta_3 - \theta_4)};$$

$$d = \frac{(\theta_{int} - \theta_1)(\theta_{int} - \theta_2)(\theta_{int} - \theta_3)}{(\theta_4 - \theta_1)(\theta_4 - \theta_2)(\theta_4 - \theta_3)}; \quad (9)$$

where:

	$\boldsymbol{\theta_{int}}$	$\boldsymbol{\theta_1}$	$\boldsymbol{\theta_2}$	$\boldsymbol{\theta_3}$	$\boldsymbol{\theta_4}$
First case	$0 \leq \theta_{int} \leq 270$	0	90	180	270
Second case	$270 \leq \theta_{int} \leq 360$	90	180	270	360

These calculations are a bit more complex than in the linear case, we will see if this increase in complexity translates into a better subjective result.

D. Data acquisition

In this project, we recorded all of the signals in the Anechoic Chamber located at the Instituto Superior Técnico, in Lisbon. This is to get a sound as clean as possible without any reflections. We will use two sets of microphones:

- The first one would be a set of two microphones with artificial ears inside an artificial head to simulate the listener's head, binaural listening. With this approach we will need to record 4 times to get signals for each direction in our scheme (Figure 3).
- The second one would be 4 pairs of microphones at the same level using the OctoEar, each with its own artificial ears, but without any artificial material to simulate the head.
- On both setups we will use a loudspeaker with a known frequency response, so we can compensate it later.

The ITA-toolbox for MATLAB was used as it gives us the option to directly obtain the transfer function. It uses a sweep function to calculate the frequency response of the system, and gives very accurate results. The results of these measurements are as follows:

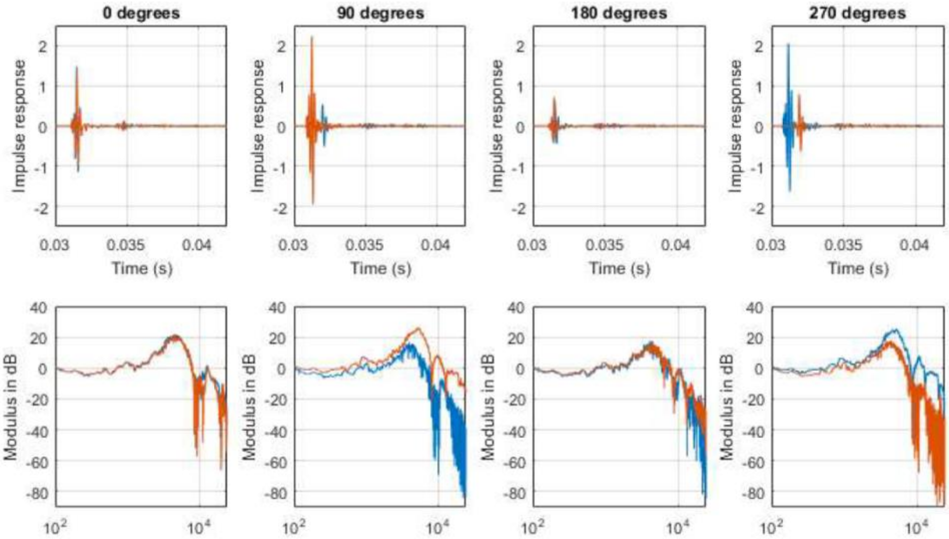

Figure 4 – HRTF for the 4 pairs of ears measured with the OctoEar apparatus.

We have both the response in frequency and in time. We can see that around the 0.034 seconds there appears an unexpected reflection on the impulse response (it is easily seen on the 0 degrees impulse response). This reflection should be the sound coming from the loudspeaker bouncing off of some part of our measurement setup. This corresponds to a time delay of about 2 ms which is an extra path of about 0.7 m.

E. Signal conditioning

These raw signals could be used directly, but since our objective is to get a realistic sound, we need to eliminate any other conditions that may affect this output, such as, the:

- Unwanted reflections - The first way to eliminate this unwanted reflection is to make the Impulse response zero from just before the reflection appears, but this approach is not good enough as we will lose many low frequency information.
- Loudspeaker compensation - The loudspeaker we used on our calculations does not have a perfectly flat response but, on the other hand, we can measure its response or use the manufacturer's specifications. Given this impulse response, we can develop a filter that inverts the frequency coloring created by the speaker. As expected, the flat line indicates that the response is now flat. We will now apply this to the HRTFs to compensate them (Figure 5).

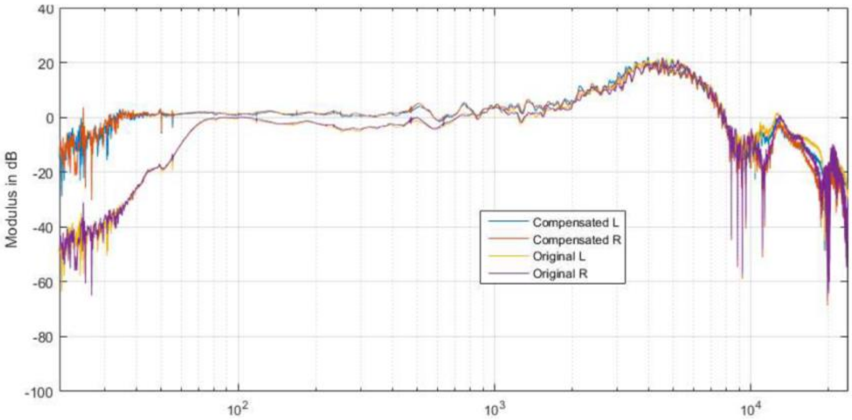

Figure 5 - Compensated vs. Original frequency response for 0 degrees.

F. *Algorithm development*

In this part, the main function we have to develop is the weighting coefficients for each kind of interpolation, as it is the basis of the VBAP method. We will be using mainly MATLAB.

1) Linear interpolation

The function developed just needs to be given the angle of incidence and the four known HRTF (Front, Back, Left and Right), and it outputs the interpolated HRTF. To calculate it, it uses the mathematical equation 6 and calculates the weighting the values using the formula from the same section. The results in frequency are depicted in Figure 6.

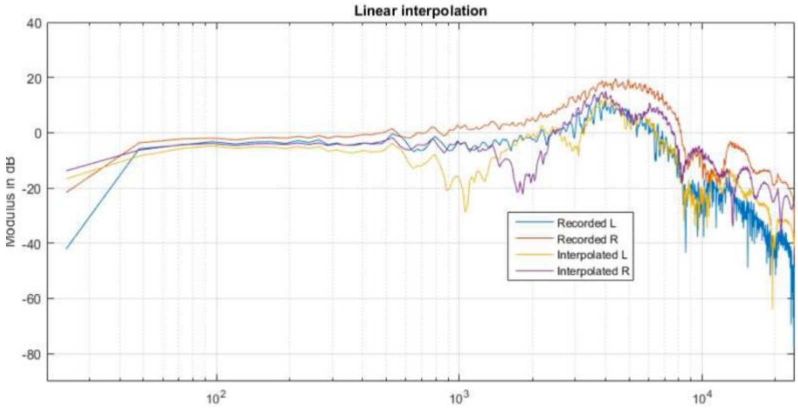

Figure 6 - Interpolated HRTF for 45 degrees using Linear Interpolation approach.

To compare, we used the HRTF associated to 45 degrees that we recorded on the Anechoic chamber, also compensated by the loudspeaker response. As we can see, the frequency response is not very similar, but the important part is the subjective perception of the listener.

2) Cubic interpolation

The process is analogue to the one we did for the linear interpolation, using in this case the equation 8. The results in frequency are depicted in Figure 7.

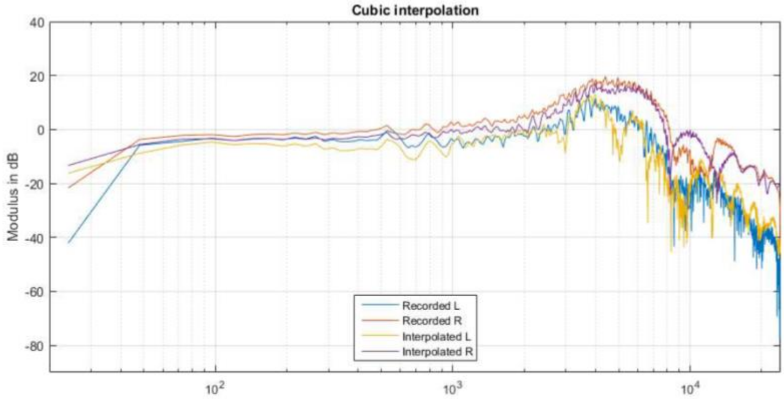

Figure 7 - Interpolated HRTF for 45 degrees using Cubic Interpolation approach.

Similarly, the interpolated response is not equal to the original, but we need to compare on the perceptual listening experience, not on the mathematical likeness.

With a few simple tests, the cubic interpolation has a better response than the linear one, and in terms of complexity the difference is not that big. We can now start to develop a real-time solution to put these conclusions to use.

3 Measurements for perceptual tests

The other interesting approach is to use the OctoEar system for recording musical events and to be able to use this material for VR devices. Based on the information of the 4 recording directions, any other direction can be obtained using the interpolation techniques previously mentioned.
However, the OctoEar concept has a limitation in terms of sound reproduction. It can only be used conveniently with headphones, that is, in binaural hearing, since the recorded sound material itself already has the influence of the shape of the ears and the elevation parameter cannot be used.
The comparative tests carried out used the OctoEar system, with a spherical shape to simulate the human head, and the AMBI1 microphone of order 1, that follows the ambisonic concept. This last approach has the advantage of having a very small volume and allowing to encode the sound in an arbitrary number of speakers for surround systems. Therefore, it is nowadays widely used for creating immersive environments and / or with an artistic perspective. Figure 8 shows the two capture devices. The tests were carried out in an anechoic chamber of the Instituto Superior Técnico and in the Audio and Acoustics Laboratory of ISEL.

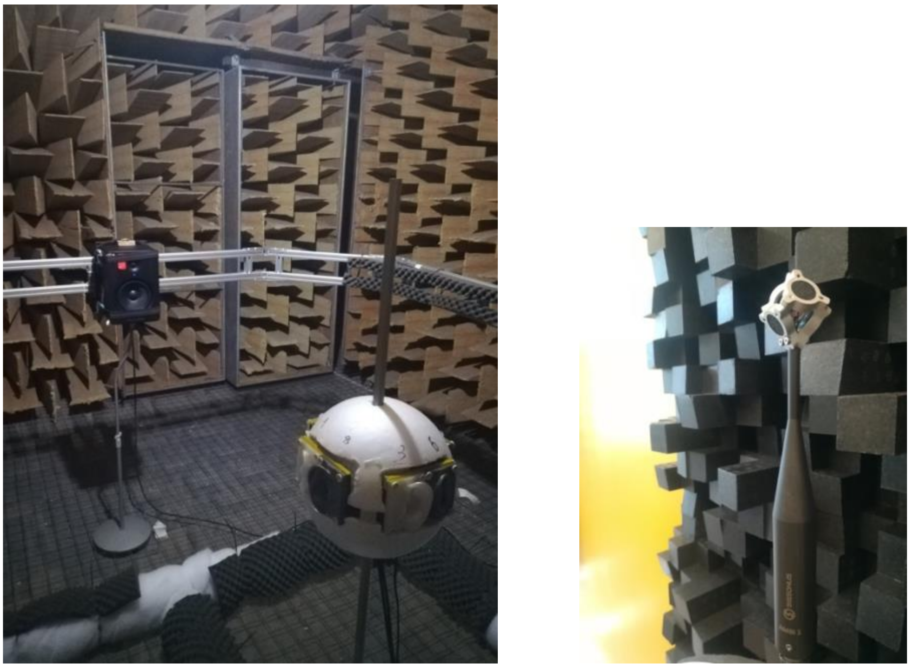

Figure 8 – Microphone systems, OctoEar and AMBI1 used to capture 3D sound in an anechoic chamber.

Several musical excerpts were recorded with both systems and the sound was then encoded in a binaural format for listening with headphones.
The results obtained show that the OctoEar system surpasses the AMBI1 system in all metrics applied. The following parameters were considered for comparison: 1) location/image of the sound source in azimuth (panoramic), 2) sharpness of the recorded material and 3) background noise levels. These results are mainly due to the fact that the ambisonic system is of order 1 which does not provide a high spatial resolution and because it requires more signal processing. However, we are still improving the microphone calibration and compensation algorithms, which could greatly improve the results.

G. Other developments

After achieving a clear idea of how to interpolate these HRTFs we can advance to produce real-time response, as with our MATLAB code we need to have the whole signal beforehand. The first idea that comes to mind is to use the PureData software (the predecessor of MaxMSP software), as it is focused on real-time audio processing, and allows us to both develop our own "patches" (like MATLAB functions, patches have input and output variables and can compute any operation we create) as well as having a vibrant community with many users sharing their own patches for different uses [6]. In this part we can divide in two different approaches:

- Using the (mono or stereo) input from a microphone and treating it with our known HRTF using a similar algorithm that we developed in MATLAB.
- Using the 8-mic setup from earlier directly and using our known weighting coefficient calculations to weight the gain of each input.

1) HRTF approach

In this case, we will use a convolution and a multiplexer patch we downloaded from the PureData forum. In our first approach, we will use a slider to determine an angle from 0 to 360 degrees.

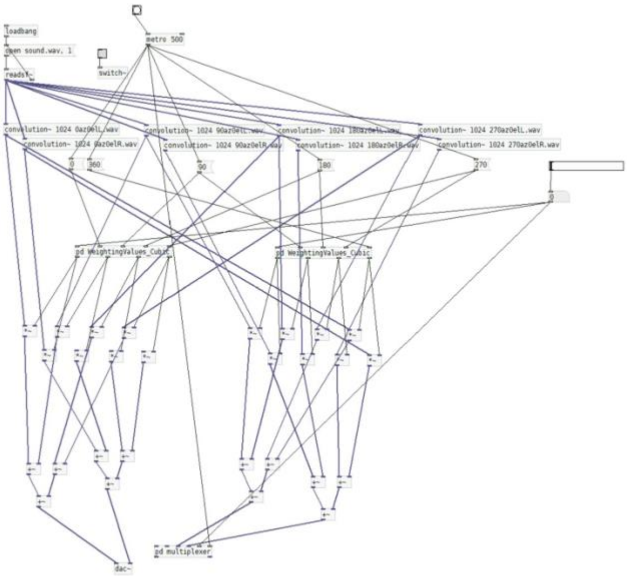

Figure 9 – PureData patch used to implement the HRTF filtering depending on the desired angle.

In this patch, we were using a sound file, but it could be done also using an ADC patch to obtain the sound from the PC's soundcard.

2) Audio approach – live recordings

For this case, the approach is very similar, the only difference is that we don't need a convolution patch. We could use an ADC patch with 8 outputs to read all of the microphone signals instead of reading 8 different recordings, but for testing purposes we opted for this.

These two approaches have a patch in common, the one used to calculate the weighting values: This patch works in the same way that the MATLAB function, given the angle of incidence and the angles used to interpolate, it gives the four weighting values needed on the output.

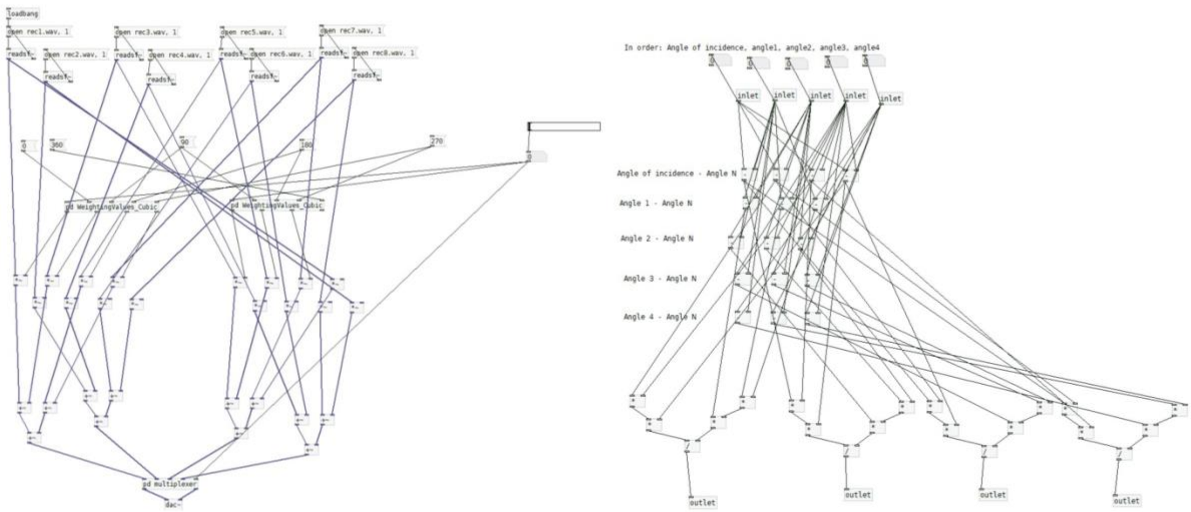

Figure 10 - PureData patches used for live recordings (left) and the sub-patch for weighting values (right).

4 CONCLUSIONS

A. Regarding the research

The two final solutions we achieved we can say that each of them have a different application:

- The HRTF approach can be used we only have a stereo or mono recorded sound and we want to give it an auralization effect, or when the realtime input is a single mono or stereo microphone and we want to develop auralization in real time.
- The audio approach can be used when we have a recording made with the an 4 pairs of ears setup (or similar), or when the audio input in real time is being recorded with this setup.

This way, we can determine that the HRTF approach works better to develop auralization to any sound environment, and the audio approach needs a setup design for that specific purpose.

A set of comparative tests were carried out between the multichannel capture systems, OctoEar and AMBI1 (ambisonic concept) to understand the advantages and disadvantages of each one for recording 3D sound. The results obtained show that the OctoEar system surpasses the AMBI1 system in all metrics. The following parameters were considered: 1) location of the sound source in azimuth (panoramic), 2) sharpness / quality of the recorded and encoded material and 3) background noise levels. In part, these results are due to the fact that the ambisonic system is of order 1 which does not provide a high spatial resolution and because it requires more signal processing, which adds some processing noise. However, we are still improving the microphone calibration and compensation algorithms, which could greatly improve the results. However, the OctoEar concept has a limitation in terms of sound reproduction. It can only be used conveniently in binaural hearing, with headphones, and the elevation parameter cannot be used.

B. Future developments

The problems we found while using PureData kept us from comparing one on one the solutions we developed, so in the future more work could be used to fix this problems. Other option could be to develop some kind of User Interface in MATLAB to use in a similar way to the PureData patch, with a slider to select the angle of incidence. The main objective would be to integrate the PureData patch with a VR system and to use the position and rotation information from the headset to calculate the angle of incidence. The immersion can be even bigger if we pair our recording equipment with a 3D camera, as PureData allows us to treat video at the same time as audio.

Acknowledgment

This project was partially supported by Laboratório de Áudio e Acústica do ISEL.
https://acusticaudiolab.isel.pt
and by Anechoic Chamber of CAPS – IST for the sound tests.

References

[1] V. Pulkki "Spatial Sound Generation and Perception by Amplitude Panning Techniques". http://lib.tkk.fi/Diss/2001/isbn9512255324/

[2] An Integrated Approach of 3D Sound Rendering Techniques for Sound Externalization - Yong Guk Kim, Chan Jun Chun, Hong Kook Kim, Yong Ju Lee, Dae Young Jang, Kyeongok Kang.

[3] A Study of 3D Audio Rendering by Headphones - Eran Hadad, Dmitry Fishman, Elior Hadad, Sharon Gannot

[4] V. Pulkki "Compensating displacement of amplitude-panned virtual sources." Audio Engineering Society 22th Int. Conf. on Virtual, Synthetic and Entertainment Audio pp. 186-195. 2002 Espoo, Finland

[5] Two Approaches for HRTF Interpolation - Gustavo H. M. de Sousa, Marcelo Queiroz

[6] Recording and sound processing using Pure Data, Pablo Cetta