Perceptual Tests on VBAP and Ambisonics Decoding Techniques for Multichannel Speakers System

Pinto, F.[3,4], Paulo, J.[1,2], Tavares, G.[3,4],

[1] ISEL-Instituto Superior de Engenharia de Lisboa, Portugal
[2] LAA-Lab. de Áudio e Acústica do ISEL, Portugal
[3] INESC-ID - Instituto de Engenharia de Sistemas e Computadores -ID Portugal
[4] IST-Instituto Superior Técnico
{franciscomcpinto@tecnico.ulisboa.pt, joel.paulo@isel.pt, goncalo.tavares@tecnico.ulisboa.pt}

ABSTRACT

Immersive sound is nowadays a hot topic on several fields, because it accurately builds the 3D acoustic environment. Using existing technology, it is possible to map sound anywhere in space, and decode it to stereo headphones or to any number of loudspeakers placed in a special geometry in a room. Vector-Base Amplitude Panning (VBAP) and Ambisonics are two common techniques used to achieve this.
VBAP places a virtual sound source in different locations by controlling each loudspeaker's output amplitude. Ambisonics is a recording and reproduction method containing a sound field representation, which may be decoded to any loudspeaker array.
This study investigates phantom sources localization, using binaural signals. This is accomplished by applying a set of virtual loudspeakers with Head Related Transfer Functions (HRTFs) to achieve auditory perception with headphones. To assess and compare each technique, a Bluetooth head-tracking system is used to map the subject's head relatively to the source position. The two techniques will be compared in terms of perception of virtual sound sources in space by ten subjects.
Future work will be held using loudspeakers distributed on the vertices of a real geodesic dome-shaped infrastructure.
The perceptual tests will take place at the anechoic chamber of Instituto Superior Técnico (IST), and at the Immersive Sound Room of the Audio and Acoustics Laboratory (LAA) of Instituto Superior de Engenharia de Lisboa (ISEL).

Keywords: 3D sound, VBAP, Ambisonics, immersive sound, multichannel loudspeaker array

RESUMO

O som imersvo é hoje em dia um tópico importante em diversos campos, visto que replica com precisão o ambiente acústico 3D. Usando tecnologia existente é possível colocar uma fonte sonora num determinado ponto do espaço e, posteriormente, descodificar para auscultadores estereo ou para altifalantes distribuídos numa geometria predefinida. *Vector-Base Amplitude Panning* (VBAP) e *Ambisonics* são duas abordagens usadas frequentemente para este efeito.
A técnica VBAP permite a colocação de fontes sonoras virtuais em diferentes localizações do campo sonoro, controlando a amplitude de saída de cada altifalante. *Ambisonics* é um método de reprodução e gravação que contém uma representação do campo sonoro, que será depois descodificado para um agregado de altifalantes.
Este estudo inestiga a localização de fontes fantasma usando sinais binaurais. Isto é alcançado aplicando Head Related Transfer Functions (HRTFs) a um conjunto de altifalantes virtuais, de modo a conseguir percepção auditiva com auscultadores. Para avaliar e comparar cada técnica, foi usado um sensor de Head-tracking Bluetooth para mapear os movimentos da cabeça do sujeito relativamente à posição da fonte sonora virtual. As duas técnicas foram comparadas por 10 sujeitos.
Em estudos futuros irão ser utilizados altifalantes distribuídos nos vértices de uma infraestrutura em formato de cúpula.
Os testes percetuais irão ter lugar na câmara anecóica do Instituto Superior Técnico e no Laboratório de Áudio e Acústica (LAA) do Instituto Superior de Engenharia de Lisboa (ISEL).

Palavras-chave: Som 3D, VBAP, Ambisonics, som imersivo, array de altifalantes multicanal

PACS: 43.60.-c

1. INTRODUCTION

The spatial properties of sound perceivable by humans are the directions and distances of sound sources in three dimensions, as well as the effect of the room. One of the main tasks in spatial audio is to position a sound source in a predefined direction in the virtual auditory space, replicating the spatial attributes of sound.
The perceived direction of a sound source heavily relies on two main localization cues: Interaural Time Difference (ITD) and Interaural Level Difference (ILD) [1]. The ITD corresponds to the difference in arrival time of a sound between the two ears or, simply put, the location of the object along the azimuth. The ILD is the difference in loudness and frequency distribution between the two ears (as sound travels further way, its intensity decreases). However, ILD and ITD cues could be ambiguous. Due to the symmetry of the head, sounds emitted from different directions could share the same ILD and ITD cues. This is the "cone of confusion", a particular volume where all sound source locations share the same ITD and ILD cues. To solve this ambiguity, two mechanisms have been proven to work: head movements and spectral filtering by the outer ears (pinnae) [1]. Head movements allow estimation of the sound sources lateral angle, thus determining the actual direction of the sound source, while pinnae collects sound, and performs spectral transformations to incoming sounds which enable the process of sound localization to take place [2]. Simply put, pinnae in auditory localization serve as linear spectra filters with direction-specific transfer functions, the head-related transfer functions (HRTFs) [1] [3].
Binaural reproduction techniques are often based on HRTFs. Fundamentally, these are functions that, for a certain angle of incidence, describe the sound transmission coming from a free field to a point in the ear canal of a human subject [4]. HRTFs are defined as multiple left/right pairs of head related impulse responses (HRIRs) measured at desired spatial increments. Thus in headphone reproduction, a convolution is computed between the input signal and the HRIR for each ear. The auditory object is perceived to be positioned to the direction from which the HRTFs have been measured.
This article investigates the localization of virtual sound sources created by VBAP and Ambisonics, both first order (FOA) and high order Ambisonics (HOA). The two spatialization techniques are compared over headphones by synthesising the position of virtual loudspeakers in 2D or 3D arrangements in specific geometries. This also provides some insight to the influence of the number of virtual loudspeakers and also their distribution in sound localisation and panning.

2. AMPLITUDE PANNING

Amplitude panning is the simplest method to create phantom sources in the horizontal plane or in a 3D sound field. Amplitude panning aims at steering the perceived direction of an individual auditory object by distributing its sound signal $s(t)$ to the l^{th} loudspeaker using frequency independent, real-valued, and positive gain g_l. To obtain stable loudness, the gains should be normalized as: $\sum_l g_l^2 = 1$. The output signal $x(t)$ of the lth loudspeaker is:

$$x(t) = g_l \cdot s(t) \tag{1}$$

The amplitude panning technique also takes into account the direction of the virtual sound source θ in auditory space. The amplitude panning algorithm calculates the gains as: $g_l = g_l(\theta_s)$. Directiont in auditory space can be defined as a unit direction vector $\theta = [\cos\varphi \ \cos\lambda, \sin\varphi \ \cos\lambda, \sin\varphi]$, which depends on the azimuth φ and elevation λ angles of the virtual sound source in the spherical coordinate system.

2.1. Vector-Base Amplitude Panning

VBAP [5] is one of the most commonly employed amplitude panning methods. This technique uses a triplet of speakers with gain weightings to pan a point source in a 3D speaker array. Three vectors are calculated for the origin of the a virtual sound source, defined by a vector (p) with components:

$$p_x = \cos\varphi\cos\lambda, \ p_y = \cos\varphi\cos\lambda, \ p_z = \sin\varphi, \tag{2}$$

The gain coefficients for the loudspeakers triplets is determined using equation 3:

$$g_{123} = p^T L_{123}^{-1} = [p_x \quad p_y \quad p_z] \begin{bmatrix} l_{1x} & l_{1y} & l_{1z} \\ l_{2x} & l_{2y} & l_{2z} \\ l_{3x} & l_{3y} & l_{3z} \end{bmatrix}^{-1} \quad (3)$$

where g_{123} corresponds to the gains for a triplet of loudspeakers, p^T is the transpose of the point source vector and L_{123}^{-1} is the inverse matrix of the same loudspeaker triplet. It is also important to point out that the loudspeaker gains are normalized. l_1, l_2 and l_3 are unit vectors that point in the x, y or z direction of loudspeaker 1, 2 and 3.

3. AMBISONICS

Ambisonics [6] panning techniques consider a continuous excitation of surrounding sources in terms of a finite-order expansion in spherical harmonics functions. Essentially, in this technique a sound source is split into directional components that represent the sound based on a given azimuth and elevation angle. This representation of the sound field is referred to as B Format. The B Format is then decoded to the listener's speaker setup.

Instead of discrete gains and vectors, Ambisonics is based on the expansion of the surround signal into spherical harmonics up to the order N that defines the angular resolution. A full periphonic (3D) Ambisonic signal set of order N results in $(N+1)^2$ channels or loudspeakers [7]. Spherical harmonics depend on the direction vector θ, and they have two integer indexes $0 \leq n \leq \infty$ and $-n \leq m \leq n$. An expansion into spherical harmonics of limited order n ≤ N, can represent any directional function $g(\theta)$ whose directional resolution is limited. Such an N^{th} order function with optional weights an expands the continuous pattern representing a virtual source at θ_s, and is given as:

$$g(\theta) = \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^m(\theta) \; a_n \, Y_n^m(\theta_s) \quad (4)$$

The encoding results may have several arbitrary degrees of accuracy, depending on the order of Ambisonics. Put briefly, the zeroth order Ambisonics corresponds to the mono signal and requires one channel. In first order Ambisonics (FOA), the components of the sound field in the directions x, y and z are encoded in three more channels [8]. Higher orders of Ambisonics (HOA) [9] also exist and in summary it adds information about higher order derivatives of the sound pressure field, which improves sound directionality.

3.1. Higher Order Ambisonics

Higher Order Ambisonics provides a rational and flexible way for spatial encoding, conveying and rendering of 3D sound fields.

It is known that one of the main drawback of the first order Ambisonics technology is the poor directionality. This translates in slightly blurry sources and a small usable listening area (or sweet spot). The main idea of HOA is to enlarge the sweet spot and improve the resolution by adding groups of more selective directional components to the lower orders of the B-format [9].

4. SOFTWARE PART

Encoding, processing and decoding spatial sound signals require support, configuration and routing for multichannel signals. Ambisonics encodes audio source into speaker independent representations of the sound field, called B-format [10] . Decoding is the process where an encoded sound field is translated to individual loudspeaker channel feeds, an advantage being that decoders can be designed for diferent speaker arrays. VBAP, on the other hand, makes possible the creation of two- or three-dimensional sound fields with any number of loudspeakers placed arbitrarily, equidistantly around the listener [5].

Sound reproduction and spatialization is achieved in realtime using an open-source VST (Virtual Studio Technology) audio plug-in suite, Spatial Audio Real-time Applications (SPARTA) [11].

SPARTA is a collection of flexible VST audio plug-ins for spatial audio production, reproduction and visualisation. All plug-ins are tested using REAPER, a very affordable and flexible Digital Audio Workstation (DAW), popular among many composers and sonic artists working with spatial sound.
Mono signals are encoded into Ambisonics using the AmbiENC [11], which are then decoded to a loudspeaker array with the AmbiDEC [11].The spatial encoding of the mono sound source can be listened also on headphones, using the Binauraliser plugin. A global scheme of this process is illustrated in Figure 1.
The same mono signals may instead also be directly panned to the same loudspeaker setup, using a VBAP-based plugin, the Panner [11]. A detailed description of the audio plug-ins as well as other tools that were employed, is presented in the following sections.

4.1. Encoders

The AmbiENC [11] is a relatively straightforward ambisonic encoder, that inputs multiple monophonic signals (up to 64 channels) and spatially encodes them into spherical harmonics signals at specified directions. These spherical harmonics signals describe a synthetic anechoic sound-field, where the spatial resolution of the encoding is determined by the transform order and, as previously stated, the higher order the greater the spatial resolution.
The Panner [11] is a frequency-dependant 2D and 3D panning plug-in based on the VBAP [5] method, and it also allows Multiple-Direction Amplitude Panning (MDAP) [12], given a certain configuration. The panning is frequency dependent, which allows more consistent loudness of sources, when they are panned in-between the loudspeaker directions, under different playback conditions [11].
In both encoders, presets for popular 2D and 3D loudspeaker setups are included for convenience.

4.2. Decoders

The AmbiDEC [11] is a frequency-dependent ambisonic decoder for loudspeaker playback. The loudspeaker directions can be user-specified for up to 64 channels, or alternatively, presets for popular 2D and 3D set-ups can be selected.

4.3. Binauralizer

The Binauraliser [11] plug-in convolves input audio (up to 64 channels) with interpolated HRTFs in the frequency domain.
The previously described AmbiDEC plug-in also features a built-in binauraliser. This allows the user to audition the loudspeaker output via convolution of the individual loudspeaker signals with interpolated HRTFs, which correspond to each loudspeaker direction.

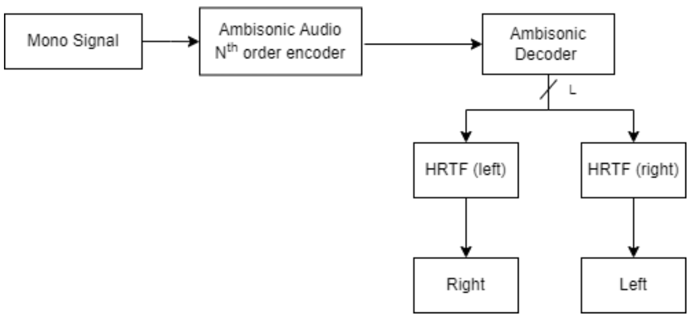

Figure 1 - Global scheme of a mono signal encoding to N^{th} order Ambisonics and decoding to binaural headphone reproduction. HRTFs with equal azimuth and elevation coordinates of the loudspeaker geometry are employed. L corresponds to the number of loudspeakers.

5. HEAD TRACKING

Since human auditory localization strongly relies on head movements, for plausible interpretation and assessment of virtual sound source localization, the incorporation of head movements is necessary.
The monitorization of head movements is achieved using a Bluetooth head-tracking sensor, the WitMotion WT901BLECL2. The sensor maps angles from the 3 Cartesian axis (x, y, z) as well as yaw, pitch and roll.

6. LISTENING TEST METHODS

The listening tests consisted of the comparison of two spatialization techniques, first order Ambisonics (FOA) and Vector Base Amplitude Panning (VBAP), in a virtual loudspeaker set-up made up of 8 loudspeakers. The virtual loudspeaker positions are presented in spherical coordinates in Table 1 and a visual representation is shown in figure 2.

Table 1 – Virtual loudspeaker position in spherical coordinates (azimuth and elevation) in degrees, relative to the subject's head.

Loudspeaker	Azimuth (°)	Elevation (°)
1	-90	-27.5
2	-18	-27.5
3	18	-27.5
4	90	-27.5
5	-36	24.8
6	36	24.8
7	-18	55.8
8	18	55.8

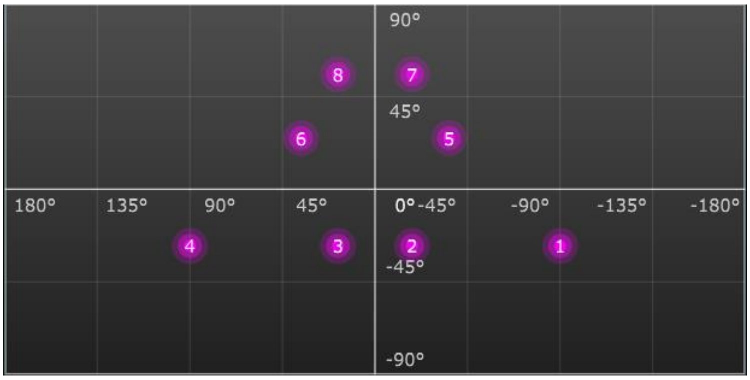

Figure 2 – Representation of the virtual loudspeakers setup in azimuth (x-axis) and elevation (y-axis)

For performing the test, five different stationary virtual source positions were chosen within the virtual sound environment (VSE), as presented in Table 1. To perform such evaluation, the subject performed a simple localization task, that consisted in reporting the perceived position of a spatialized virtual sound source and moving the head towards the perceived localization.

6.1. Subjects

The study group consisted of 10 subjects (4 women and 6 men) aged between 23 and 60. The subjects had no previous experience in psychoacoustic experiments. None of the subjects reported auditory loss or listening disabilities. All of them participated on a volunteer basis.

6.2. Stimuli

Two stimuli were used, one synthesised (pink noise) and one mono recording of a more familiar sound to the subjects (birds chirping).

6.3. Hardware

The AKG K512 MKII headphones were used in this study, as well as the Wit-Motion WT901BLEC head-tracking sensor, as mentioned. The sensor was attached to the top part of the headphones.

6.4. Procedure

Subjects sat with headphones on in front of a reference mark located at head height. Prior to the recording of head movements, the subjects were asked to stay still and stare at the reference mark allowing the sensor to be calibrated with the initial head position. Following the calibration, the sensor is set to record and the stimuli were presented. Figure 3 provides a visual representation of the listening tests.
Each stimulus lasted two seconds and came from the five different stationary sound source positions represented in Table 2. After hearing a stimulus, the subject would move its head in the direction of the perceived stimulus source location, while the sensor recorded the head movement.

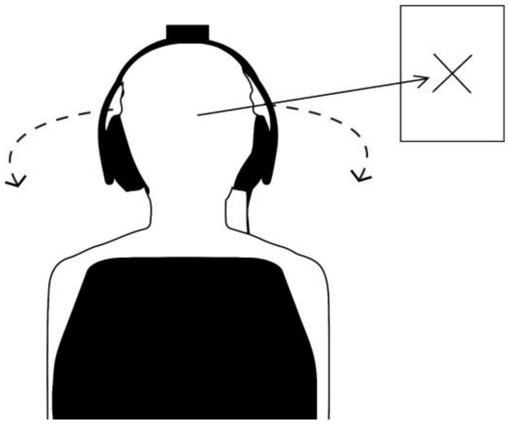

Figure 3 – Visual representation of the listening tests.

Table 2 – Localization coordinates for the stimuli.

	Azimuth (°)	Elevation (°)
Δ_1	-90	0
Δ_2	90	0
Δ_3	0	45
Δ_4	36	-24.8
Δ_5	-18	55.8

Subjects were seated with headphones on in front of a reference mark located at head height. Prior to the recording of head movements, the subjects were asked to stay still and stare at the reference mark allowing the sensor to be calibrated with the initial head position. Following the calibration, the sensor is set to record and the stimuli were presented.
Each stimulus lasted two seconds and came from the five different stationary sound source positions represented in Table 2. After hearing a stimulus, the subject would move its head in the direction of the perceived stimulus source location, while the sensor recorded the head movement.

For each spatialization method, the virtual sound source positions were presented in different sequences. This prevented the subject to form habits that could affect the results. For FOA the chosen sequence was Δ_1, Δ_2, Δ_3, Δ_4, Δ_5 and for VBAP Δ_1, Δ_3, Δ_5, Δ_2, Δ_4.
The total duration of each listening test was around 10 minutes.

7. RESULTS

The goal of the listening tests is to make a binaural assessment of two distinct, yet commonly used spatialization methods, first order Ambisonics and Vector Base Amplitude Panning, using the SPARTA plug-ins. Two different stimuli were used pink noise and a mono recording of birds chirping, to evaluate the influence of different types of sound material, synthesised and a mono recording, in spatialization.
Table 3 and 4 show the results for first order Ambisonics and VBAP, respectively. The result corresponds to the angle of azimuth rotation recorded by the head-tracking sensor when the stimuli were played to the subjects. The virtual sound source is placed in one of five different stationary positions (as presented in Table 2).
The results are a comparison between the virtual sound source position and the position perceived by the subjects during the listening tests, for the two spatialization techniques and both stimuli and are presented using a color map.
In this case, two colors are used: red and green. When the subjects perceive the virtual sound source on the opposite azimuth side from where it originated, the color red is used. The color green is used when the subject perceives the sound source on the correct azimuth side from where it originated.
Table 5 draws an overall comparison between the results obtained in both methods, in three parameters: error, most common coordinate mistake and stimuli with most errors. The first parameter, indicates the amount of times that the subjects incorrectly perceived the virtual sound source in space. The coordinate that caused most confusion to the subjects was Δ_3, in both VBAP and FOA, altough in VBAP the percentage was higher.
During the listening tests, all subjects were unable to detect the elevation component of the virtual sound sources. For this reason, the elevation results were inconclusive and therefore left out of this study. Future studies mostly in the binauralization process of the plugin will be conducted to improve this shortcoming.

7.1. FOA Results

The FOA results for the 10 subjects are shown in Table 3. It can be seen that generally the majority of the perception error occurred in Δ_3, situated directly in front of the subjects, as opposed to other virtual sound source positions. Beside Δ_3, perceptual errors in Δ_4 (2% overall) and Δ_5 (6% overall) coordinates were also frequent. This indicates that the subjects showed less effort to perceive the virtual sound source when it was positioned well to either the left or right side, closer to the ears then when it was centered or almost directly in front of the face. The difference in error between Δ_1 or Δ_2 and Δ_3 or even Δ_4 or Δ_5 illustrates such struggle.
The stimuli made no apparent difference in the perception of the virtual sound sources, since both stimuli had the same amount of misclassifications, which could lead to the conclusion that, for the FOA method, the stimuli used may not be as relevant to virtual sound source positioning as the localization coordinates. And the last parameter indicates the stimuli with most errors in the perception of the virtual sound sources in space.

Table 3 – FOA results.

	Subject 1		Subject 2		Subject 3		Subject 4		Subject 5	
Stimuli	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**
Δ 1	-37,0	-34,9	-40,1	-44,8	-46,8	-34,4	-27,0	-49,9	-57,8	-30,1
Δ 2	35,2	25,6	36,9	34,7	53,4	40,5	37,4	43,2	59,1	49,1
Δ 3	6,8	9,8	-42,8	34,7	-24,1	31,6	3,0	-34,2	-0,6	38,9
Δ 4	20,8	2,5	34,0	32,9	27,1	49,3	0,5	50,7	19,1	41,8
Δ 5	5,4	12,8	-35,3	33,7	-41,5	-14,8	-32,4	-43,3	-22,4	-27,0
	Subject 6		**Subject 7**		**Subject 8**		**Subject 9**		**Subject 10**	
Stimuli	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**
Δ 1	-41,7	-52,3	-37,1	-40,4	-6,2	-19,2	-47,2	-21,9	-37,3	-37,0
Δ 2	45,5	40,0	47,0	20,4	15,1	22,4	36,6	35,1	25,0	45,4
Δ 3	-0,7	-1,3	-12,3	-15,5	3,7	2,5	2,4	-17,2	34,8	-1,3
Δ 4	-0,8	33,2	23,8	43,5	14,4	16,4	36,2	23,4	36,2	20,0
Δ 5	-5,1	-1,5	-29,5	30,5	-9,1	-12,1	-28,2	-29,4	-27,2	-27,3

7.2. VBAP Results

The results obtained with the VBAP method are presented in Table 4. Great part of the errors occurred in Δ_3, as was the case in FOA, but in VBAP the amount of misclassification increased by 10%. In VBAP, the sound source positions with most errors were Δ_2 and Δ_4, in either stimuli.
In VBAP the stimuli played an important part in the positioning of the sound source, since most of the subjects had more difficulties with the pink noise, opposed to the bird mono recording, contrarily to what was verified in FOA.
Table V draws an overall comparison between the results obtained in both methods, in three parameters: error, most common coordinate mistake and stimuli with most errors.

Table 4 – VBAP results.

	Subject 1		Subject 2		Subject 3		Subject 4		Subject 5	
Stimuli	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**
Δ 1	-49,0	-37,8	-26,6	-35,2	-56,8	-49,0	-44,1	-40,6	-55,5	-59,1
Δ 3	-34,4	-15,6	11,4	-41,3	31,8	-33,6	-37,1	-37,8	-12,0	7,3
Δ 5	-32,6	-12,1	-22,4	-43,4	-3,4	-38,8	-42,9	2,6	-31,9	-4,2
Δ 2	30,7	15,5	31,8	41,3	30,4	41,5	-46,1	53,1	39,8	43,8
Δ 4	31,1	35,9	-33,4	38,6	34,9	38,6	-41,5	-49,2	37,7	43,8
	Subject 6		**Subject 7**		**Subject 8**		**Subject 9**		**Subject 10**	
Stimuli	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**	**Pink**	**Birds**
Δ 1	-35,6	-32,3	-48,2	-54,9	-55,6	-43,1	-56,4	-50,1	-27,1	-37,9
Δ 3	-47,1	30,4	48,0	-52,6	-0,1	2,2	-27,8	-23,1	2,2	-40,9
Δ 5	-30,9	-1,5	-35,5	-68,1	-24,3	-13,1	-47,0	-14,3	-47,1	15,4
Δ 2	-5,5	22,4	50,9	54,3	22,9	20,5	20,3	19,9	22,7	15,5
Δ 4	-6,0	27,3	-47,2	58,5	20,2	51,9	-1,5	43,6	-27,8	51,5

Table 5 – General results.

	VBAP	**FOA**
Error (%)	26%	22%
Most common coordinate mistake	Δ_3 (75%)	Δ_3 (65%)
Stimuli with most errors	Pink (61%)	Pink (50%) and Birds (50%)

Table 6 – Average results for VBAP and FOA

	VBAP Average (°)		FOA Average (°)	
Stimuli	**Pink**	**Birds**	**Pink**	**Birds**
Δ 1	-45,5	-44,0	-37,2	-36,5
Δ 3	-6,5	-20,5	39,1	35,6
Δ 5	-31,8	-17,7	-3,0	4,8
Δ 2	19,8	32,8	21,1	31,4
Δ 4	-3,4	34,1	-22,5	-7,8

8. CONCLUSIONS AND FUTURE WORK

The preliminary listening test results indicate that when placing sounds in a virtual sound environment, there is a difference in the subjects' perception of the localization of virtual sound sources. Out of the two techniques, FOA gives the best results. Such results show that the subjects have a better understanding of virtual sound sources when placed to either side of the subject's head than directly in the front, although the perceived angles and intended angles do not always coincide, even when the source is directly placed on the virtual loudspeaker. The elevation parameter was not perceived by any of the subjects, since it is quite difficult to replicate. This problem is under investigation to determine if the binaurelization plugin has issues with the elevation parameter or if the loudspeaker architecture is simply not adequate to provide subjects with a significant elevation perception.

Future work will focus mostly on using loudspeakers distributed on the vertices of a real geodesic dome-shaped infrastructure which is already build. A photo of the dome inside the anechoic chamber at IST is shown in Figure 4. This will allow a clearer perception of the sound environment and certainly an improvement in the elevation parameter.

Other tests will be petformed using moving sound sources, instead of static, as in this case, to evaluate the subject's percpetion to moving sound sources using the same two methods, VBAP and Ambisonics.

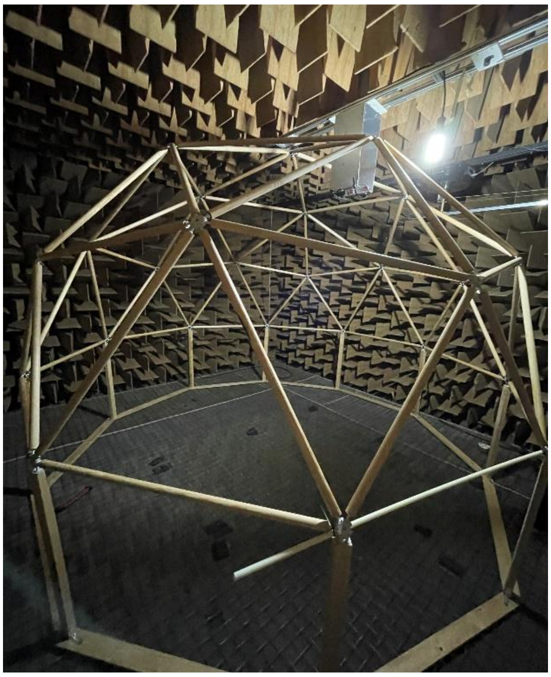

Figure 4 – Geodesic dome infrastructure to be used in future tests.

REFERENCES

[1] Blauert, J. "*Spatial Hearing. The Psychophysics of Human Sound Localization*", MIT Press, 1997.

[2] Shaw, E. A. G. "*Localization of Sound: Theory and Applications*" Ear and Hearing: May 1982

[3] Musicant AD, Butler RA. "*The influence of pinnae-based spectral cues on sound localization*". J Acoust. Soc. Am. 1984

[4] Møller, H., Sørensen, M. F., Hammershøi, D. and Jensen, C. B. "Head-related transfer functions of human subjects," J. Audio Eng. Soc., vol. 43, no. 5, pp. 300–321, May 1995.

[5] Pulkki, V., "*Virtual sound source positioning using Vector Base Amplitude Panning*", The Audio Engineering Society, vol. 45, no. 6, 1997.

[6] Gerzon, M. A. "*Periphony: With-Height Sound Reproduction,*" J. Audio Eng. Soc., vol. 21, no. 1, pp. 2-10, (1973 February.).

[7] Kronlachner, M. and Zotter, F. "*Spatial transformations for the enhancement of Ambisonic recordings*," in 2nd International Conference on Spatial Audio, Erlangen, 2014.

[8] Neukom, M. "*Ambisonic panning*", in Convention Paper 7297, New York, NY, USA, 2007.

[9] Daniel, J., Moreau, S. "*Further Study of Sound Field Coding with Higher Order Ambisonics,*" Paper 6017, (2004 May)

[10] Gerzon, M.A. "*Ambisonics in multichannel broadcasting and video*", Journal of the Audio Engineering Society vol. 3, pp. 859-871, Nov. 1985

[11] McCormack, L., Politis, A. "*SPARTA \& COMPASS: Real-Time Implementations of Linear and Parametric Spatial Audio Reproduction and Processing Methods*", in 2019 AES International Conference on Immersive and Interactive Audio, York, United Kingdom , 27/03/2019