24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

DESIGN AND IMPLEMENTATION OF AN AMBISONICS SYSTEM CODEC

Andrés Felipe Espitia López and Diego Federico Barco Cruz

Universidad de San Buenaventura-Bogotá, Sound Engineering, Faculty of Engineering, Bogotá, Colombia.

Marcelo Herrera Martínez

University of Iceland, Faculty of Mechanical Eng. and Informatics, School of Engineering and Natural Sciences, Reykjavík, Iceland.
email: marcelo@hi.is

The general objective of the work is to develop an algorithm that is capable of encoding a signal captured with a microphone technique (Blumlein-omnidirectional) to be reproduced in an Ambisonics system. In that sense, the goals are presented, explained and analyzed one by one about the engineering development, including a study of acoustic parameters of the place of measurement, capture, and reproduction of the system.

Keywords: Ambisonics, 3D audio, Blumlein-omnidirectional capture, Matlab.

1. Introduction

Digital signal processing has encountered a lot of audio applications during the last decades [1], [2]. One of its more representative is spatial audio coding. Spatial audio coding can be done in two ways: (1) either with the use of an array of loudspeakers, or (2) headphones. Both these cases are aimed at the achievement of a more realistic immersion feeling in the listener. Ambisonics –originally conceived by Gerzon [3]- can be applied to audiovisual production, including musical productions, films, advertisement, among others. This paper focuses on the achievement of an algorithm which codifies various signals captured through the Blumlein-Omnidirectional (Ambisonics) microphone, so that later it may be played back through a four-speaker array, that are located in cross shape, they are situated at the same height and the same distance between them.

Ambisonics capabilities depend mainly on two aspects: (1) The number of used loudspeakers and (2) their order of placement. In this sense, the spherical harmonics have to be taken into account.

After the beginning of the new millennium, several contributions in the field have been made and published in AES Convention meetings. Gerzon criteria gave birth to new models which encapsulate concepts as ITD (Interaural Time Difference), and IID (Interaural Intensity Difference).

In Ambisonics, a recording/reproduction technique series are immersed, where their performance relies on mixing multichannel technology analog to the MIMO widespread technologies [4]-, and which may be applied to studio or live sound. There are required four (4) channels, in order to give a spherical sound field, and even 3 channels to give a horizontal distribution. One of the possible problems of the coder applications is the limitation to work with B-format.

Ambisonics is an audio recording and playback technique where at least four channels are used, one of which has the function of carrying amplitude information while the other three carry velocity information. Nevertheless, there is an important limitation and it is about the number of channels through the audio can be reproduced, taking into account that it is directly proportional to the order of approach sought to be implemented with the system. In this way, it is possible to generate a 3D sound field, which has the ability to move several sound sources simultaneously, using the decomposition of the sound field by the means of spherical harmonics.

2. Previous requirements for information acquisition

The first measurement will be an objective one, where an implementation of a pre-established protocol must be done, for system measurement. Since previously, there are no references about measurement for determining the procedure efficiency, then, it may be done in the following way:

1) A 3-dimensional grid may be implemented, with a wood cube. Inside the cube, with nylon, all the edges may be jointed, in such a way that several intersections may be configured throughout its trajectory, and then a reference point where height and depth values may be measured.
2) All over the grid, impulses will be generated, from bombs, voices and a guitar, taking into account the variety of audio nature for the test. At one side of the source, the SVANTEK sound level meter will be put.

The SLM measures the SPL that was generated by the source in order to compare it to the SPL that the system reproduces at a nominal level.

The SLM may measure as well the frequency response of the produced sound, which may give detailed information about the sound which was captured. In such a way, when this sound is reproduced through the system, then the frequency response may be compared to the one previously recorded.

After coding and decoding, the signal will be reproduced with a 4 speakers' system. At the reproduction moment, the recording grid is taking into account and the equipment will be installed following this order.

In the same manner, with an AKG 414 microphone, a re-recording is held, at the points previously calibrated. Including the centre, in order to perform a signal correlation for every position and compare the reproduction to the original audio.

The multidimensional acoustic field may be generated in a perfect manner with an unlimited number of loudspeakers creating a sphere and locating the sweet spot in the middle of this sphere.

Blumlein in the 20´s and the 30´s [5] implemented a formal base for the directional reproduction of sound. From that time, several improvements have been carried on, based on the same concept.

Ambisonics offers a series of advantages over other surround sound systems as:

1) Isotropy: Sound which arrives from all directions is treated in the same way.
2) All-the-speakers are used generally to localize a sound in any direction.
3) Stability and sound-field-imagery vary less with listener position, in comparison to other surround sound systems.
4) A minimum of 4 information channels is required for the distribution and storage of a complete sound field.
5) Other speakers do not need to be placed in rigid positions.
6) The Ambisonics signal is independent of the system reproduction.

If a B-format signal is manipulated (rotated or tilted) then it is necessary to scale the four channels of the signal with the correspondent coefficients.

2.1 Measurement of vibration modes

A grid is traced upon the measurement surface, with the aim to define the exact points of measurement. These points are shown in Figure 1.

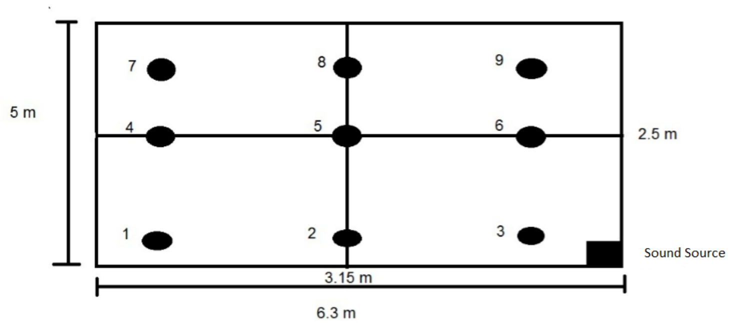

Figure 1: Grid for the vibration modes measurement

It is necessary to measure the dimensions of the studio room in order to obtain the room volume, taking into account that it is a necessary parameter for the critical distance calculation, as well as the minimum distance between the points in the grid. Afterwards the loudspeaker is located. This is connected to a pink noise generator, which plays back through an audio device connected to a console (therefore with gain control). In order to have an omnidirectional behaviour inside the room, the loudspeaker is located in one of the corners. The sound level meter is calibrated and configured so to take measurements by 1/3 octave, and with a time of 20 seconds per point. This measurement was performed in all of the points of the grid on the floor, taking the correspondent data for the analysis. An adjustment in the console´s gain was performed, so to have a pink noise of 94dB SPL at the first point of the grid. This point becomes the reference.

Complete tables for this measurement are omitted in the present text, but they may be found in [6].

In the case of the point No. 1 of the measurement, one vibration was found at 630Hz. In the case of the point No. 2 of the measurement, a vibration mode was found at 63 Hz. In the case of the measurement point No. 3, there was found a mode at 1 kHz. In the case of the measurement point No. 4, an anti mode vibration was found at 125 Hz. In the case of the measurement point No. 5, there were found no vibration modes. In the case of point No. 6 measurement point, there was found a vibration mode at 315 Hz, and an anti mode at 125 Hz. In the case of No. 7 measurement point, there was found a vibration mode located at 4kHz. In the case No. 8 measurement point, there were found no vibration modes. In the case of No. 9 measurement point, there was found one anti mode vibration point at 500 Hz.

2.2 Measurement of reverberation time (RT60)

A source was located in a corner of the room. One position was enough since the room dimensions are small enough. In order to compensate other source position, 3 measurements were taken for each sound level meter point. In this way, 6 points were the total number of points. Figure 2 shows the grid of the reverberation time measurement.

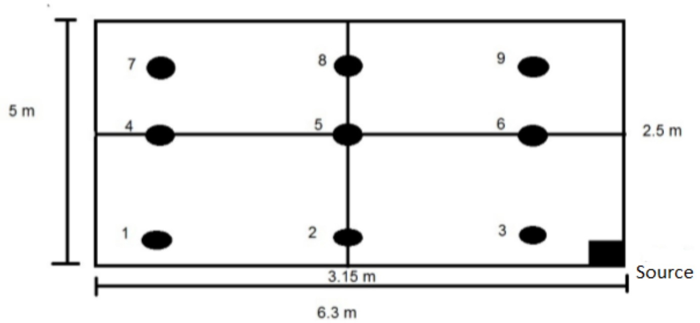

Figure 2: Grid surface for the measurement of reverberation time

For the measurement, the grid points have to be sufficiently separated from any surface. Therefore, the minimal distance was calculated, described in eq. 1, between the points and the source:

$$D_{min} = 2\sqrt{\frac{V}{ct}} \tag{1}$$

where,
V = Volume of the room, in m^3.
c = Sound propagation velocity in m/s
t = An estimation of the expected reverberation time, in seconds.

The values in the eq. 1 are, V = 113.4 m^3; c = 340 m/s; t = 0.4 s. The minimal distance between points is Dmin= 1.397 m. Taking into account this value, the grid points are allocated for the reverberation time. Within the measurement, the following acoustic parameters are taken: EDT, RT20, RT30 y RT60.

Taking into account the collected data, it is necessary to have the RT60 (reverberation time) of the room. In this order of ideas, Sabine´s equation will apply, as well as room absorption ABS. Taking into account that the room counts with different surfaces and materials –each one with different absorption coefficients- it is necessary to make the summation of the surfaces and their absorption coefficients.

Therefore,

$$RT_{60} = 0.161\left(\frac{V}{\sum_n S_1\alpha_1+S_2\alpha_2+\cdots+S_n\alpha_n}\right) \tag{2}$$

For the calculation, it is necessary to have exact dimensions of the wall surfaces, the absorbent material, the reflecting material (wood) and the glass, as well as their absorption properties (coefficients). With these acoustic parameters, the reverberation time may be found. RT60 for this room is 0.354 seconds.

For the construction of the grid, the calculation of the critical distance is performed, as it is defined in the eq. 1:

$$D_{crit} = 0.057\sqrt{\frac{QV}{Rt60}} \tag{3}$$

Where: Q = Directivity factor. V = Room volume. Rt60 = Reverberation Time.

Taking into account the omnidirectional sound source, then Q = 1. The room volume is V = 113.4 m^3, and the reverberation time is Rt60 = 0.3 seconds.

Replacing the resultant values, we have that, D_crit = 1.108m.

After finding the critical distance, it was determined that the distance to the source for exciting the source for recording has to be less than 1.108 m. Therefore, the grid was designed in an area of 2m x 2m, having marks each 10 cm. In order to calculate a distance less than 1.108 meters, a rectangular triangle was traced where the hypotenuse may be accommodated it to the required distance.

Replacing in the eq. 4:

$$h^2 = a^2 + b^2 \tag{4}$$

Where a=0.7m, b=0.7m, then the value of h=0.098m. After the analysis of the hybrid studio at the University of San Buenaventura, a comparison has to be performed, regarding advantages and disadvantages with the free field conditions, in order to define the most proper place for recording processes and Ambisonics playback. An ideal place for developing the project, including the measurement, recording and playback stages is an anechoic chamber (defined in the conceptual theoretical frame). Calculated measurements have to be taken into account for the system measurement grid, with the aim to avoid the influence of reflections on the capture and reproduction processes. After having the critical distance and after designing the measurement grid under the measured parameters of the room, it is concluded that the reverberation time of the place does not affect the necessary captures for testing the system.

3. Audio capture with the use of Blumlein-Omnidirectional technique

First consideration while recording is the room itself. A place with no acoustic problems was found at the moment when taking the pertinent captures, especially the eigen reflections of the room. This implies to be able to make a detailed analysis of the captured audios. The audios were captured according to the necessary array for the ambisonics system. This implies the use of Blumlein omnidirectional (Ambisonics) technique which includes the sound field microphone which possesses electret microphones in its capsule and it gives the necessary directionality for the system. These transducers are scarce and high cost. In order to emulate the transduction, 2 microphones AT4050 and 2 microphones AT4033 were used, which possess the same capsule. They were put into a stable acoustic centre.

Figure 3: Blumlein-omnidirectional (Ambisonics) microphone array

One of the AT 4050 were used in an 8-pattern figure in order to emulate one of the back parts. The two microphones AT 4033 were used in cardioid pattern configuration, one from each side, and another AT 4050 was located in the upper part in an omnidirectional manner, as it is shown in Fig. 3.

The array was implemented inside a conditioned grid for the playback, constrained by the acoustic parameters of the process. It is limited inside a 2 meter long, 2 meters wide and 3 height position cube, which varies according to the capture. This means, at 1 meter, 1.2 meters, and 1.5 meters.

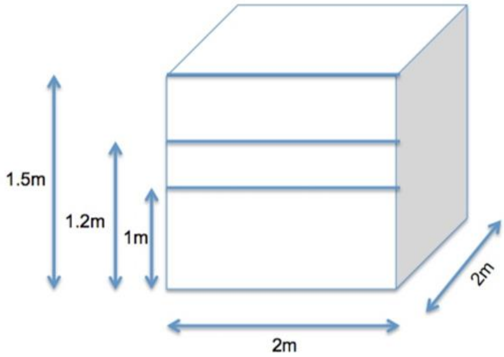

Figure 4: Implemented grid for the sounds capture and their further reproduction

Three different heights were taken, taking into account 1.2 meters as the standard value of any measurement protocol. It was given the option inside the grid to move the microphones each 20 cm over the horizontal axes of each one of the heights with a nylon system.

Regarding the sound level meter, it was located in each one of the points where the impulses were generated with bombs, voices and one guitar. An AKG 414 was located in the same position of the SLM, with the aim to make a correlation in the objective measurement and to compare the recorded signal to the recorded one.

After the signal is captured by the microphone array, they are preamplified in a SoundCraft Preamp. A/D conversion is followed, and after audio edition in Protools, they are played back through Active Monitors.

4. Coding Algorithm –counting with the acoustic parameters

The main objective of the present investigation is to implement an algorithm capable of coding the captured sounds through the Blumlein-omnidirectional technique (ambisonics). It was implemented in MATLAB.

4.1 Coding equations

The position of a sound within a three-dimensional sound field is encoded in the four signals that make up the B format, then,

$$X = \cos A \times \cos B \text{ (Front-Back)} \quad (5)$$

$$Y = \sin A \times \cos B \text{ (Left-Right)} \quad (6)$$

$$Z = \sin B \text{ (Up-down)} \quad (7)$$

$$W = 0.707 \text{ (Pressure signal)} \quad (8)$$

Where A is the horizontal angle (measured counterclockwise) from the center, and B is the angle of elevation. If the sound positions are limited to a sphere, taking into account that,

$$\mathrm{x}^2 + \mathrm{y}^2 + \mathrm{z}^2 \leq 1 \quad (9)$$

then, the equations may be written in a simpler manner as:

$$\mathrm{X} = \mathrm{x} \quad (10)$$

$$\mathrm{Y} = \mathrm{y} \quad (11)$$

$$Z = \mathrm{z} \quad (12)$$

4.1.1 Recorded signal edition

5 captured audio signals, with 3 velocity signals and 1 pressure signal.

All these signals were to be summed up, therefore they were edited in Protools. In this way, each audio acquired a 14 seconds length, considering that the longer had 13.2 seconds.

4.1.2 Coding of the recorded signals

In the recording process, the sounds were registered with: • X1-Microphone AT 4050 in the x-axis in the front position (figure of 8) • X2-Microphone AT 4050 x-axis (8-figure) • Y1-Microphone AT 4033, y-axis, left position • Y2-Microphone AT 4033 y-axis right position • Z-Microphone AT 4050 z-axis, omnidirectional.

Afterwards, signals were summed up:

$$X = x1 + x2 \quad (13)$$

$$Y = y1 + y2 \quad (14)$$

$$Z = z \quad (15)$$

$$W = x + y + z \quad (16)$$

In this way, 3 velocity vectors and 1 pressure vector become available. Therefore, the next step consists of locating it inside a sphere of radius =1.

This enables to put limits to the sounds within the unity-sphere:

$$x^2 + y^2 + z^2 <= 1 \quad (17)$$

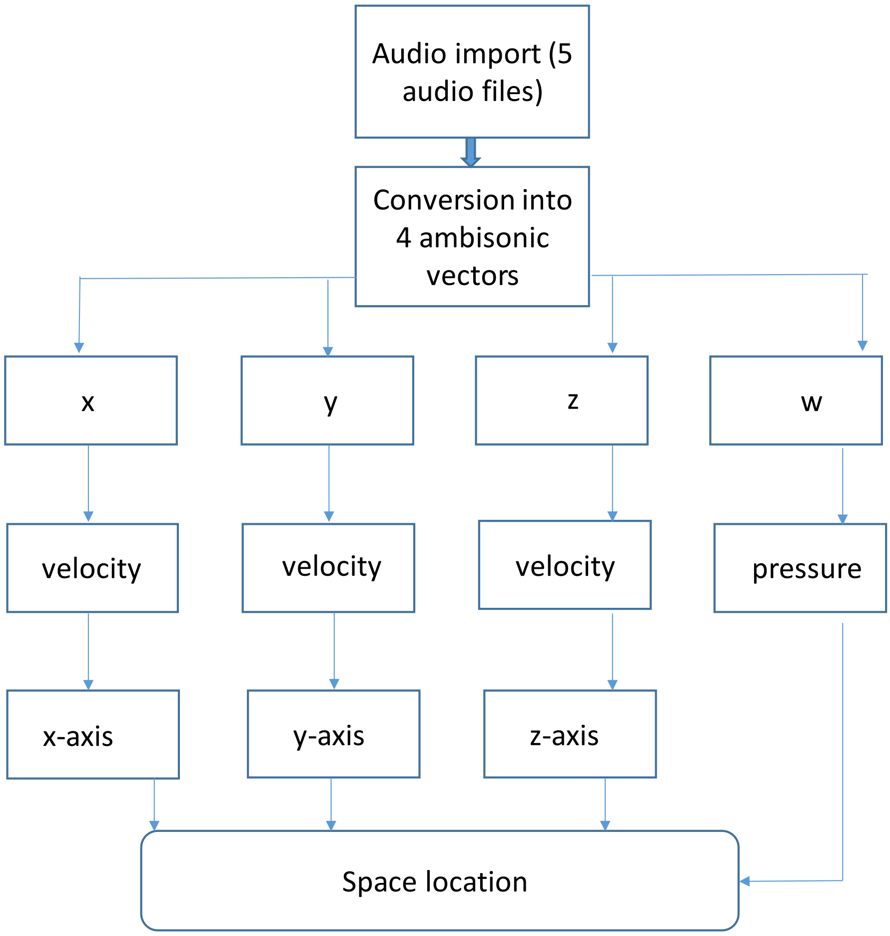

Figure 5: Implemented grid for the sounds capture and their further reproduction

5. Decoding

In the graphic interface of the algorithm, the user may select between 3 different signal types: Impulses, guitar, and voice.

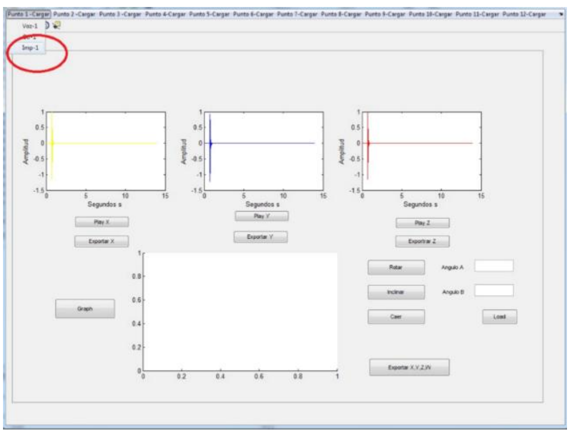

Figure 6: Implemented grid for the sounds capture and their further reproduction

For the recording process, 12 points are established in specific positions where guitar and voice are present. The generated impulses by the bomb blasts are only in the odd points.

All these signals are originally located through the coders in the point where they were recorded. Therefore, the algorithm has to have a decoder which locates them in the original point, and which gives them the possibility to facilitate to the user, their location in any position of the space. For this purpose, the work was done with three types of movement available with the ambisonics system.

The rotation of the sound event goes around the x-axis (taking into account that the positive angles orientate against clockwise). In second place, the inclination rotates the sound around the y-axis (taking into account that positive angles are orientated against clockwise). The last feature is the elevation, which makes the sound turn around the z-axis. With these 3 movements, there is the possibility to manipulate the sound in all the 3 dimensions in the ambisonics system.

5.1 Rotation

Eq. 5 and eq. 6 make possible this movement.

5.2 Inclination

$X' = x$ (18) $W' = w$ (19) $y\,! = y \times \cos B - z \times \sin B$ (20) $z\,! = y \times \sin B + z \times \cos B$i[14] (21)

5.3 Elevation

$x! = x \times \cos B - z \times \sin B$ (22) $y' = y$ (23)

Eq. 2 and eq. 3 are also part of the elevation process.

With these 3 types of movement, it can be said that the sound moves freely in the unity-sphere.

The user has as well the possibility to choose the angles. In the same manner, he will have the possibility to choose any of the 3 movements in an independent manner, making two at the same time, or even all.

After the signal multiplication with sines and cosines, the amplitude gets affected, therefore, after making any kind of movement, the signal is normalized in order to conserve the original amplitude.

Finally, after any process in the coder, the user has the option to export the 4 audios already modified, in order to get into the last stage of the process, which yields to the necessary routing for the system reproduction.

6. Conclusions

An ambisonics encoding algorithm developed in MATLAB software was developed. The rendering of spatial audio takes into account rotation, inclination and elevation of the sound source. The capturing process is also described in the present paper –with the Blumlein-omnidirectional technique.

REFERENCES

1 Herrera, M. Compresor perceptual basado en la transformada Wavelet-Daubechies Tipo IV, con dos niveles de descomposición, *Congreso Iberoamericano de Acústica, FIA*, Valdivia, Chile, 2014.

2 Kahrs, M., Brandenburg, K., *Applications of Digital Signal Processing to Audio and Acoustics*, Springer (1998).

3 Gerzon, M. A. Ambisonics in Multichannel Broadcasting and Video, *74th Audio Engineering Society Convention*, New York 8-12 October 1983

4 Herrera- Martinez, M.; Pérez-Pereira, M. Simulation of the OFDM technique for the transmission of digital radio, *TECCIENCIA*, **7** (4), 49–53, (2013).

5 Yang, D. T., Kyriakakis C., Kuo C. J., *High-Fidelity Multichannel Audio Coding*, Hindawi Publishing Corporation (2006).

6 Barco, F., Espitia, A. Desarrollo de un algoritmo para sistemas Ambisonics. Bch. Thesis, Graduate Program in sound engineering, Universidad de San Buenaventura, Bogotá, Colombia, (2013).