One strategy for an immersive recording using ambisonics

César Marino Villavicencio Grossmann
Universidade Federal de Uberlândia/Universidade de São Paulo
cevill@usp.br

Abstract: The process of setting-up an array of 10 microphones for recording a solo instrumental piece in a local church is described and the processing of the sound through plugins in 3rd order ambisonics for binaural rendering is explained. The results revealed that the multitrack recording in the acoustics of the church with binaural rendering produced depth and brilliance offering the possibility of managing the desired amount of reverberation without the necessity of using reflections available in the plugin.
Keywords: Recording techniques, Ambisonics, Multitrack recording, Binaural.

Uma estratégia para uma gravação imersiva usando ambisonics

Resumo: O processo de configuração de uma matriz de 10 microfones para gravação de uma peça instrumental solo em uma igreja local é descrito e o processamento do som através de plugins em ambisonics de 3ª ordem para renderização binaural é explicado. Os resultados revelaram que a gravação multipista na acústica da igreja com renderização binaural produziu profundidade e brilho oferecendo a possibilidade de gerenciar a quantidade de reverberação desejada sem a necessidade de utilizar reflexões disponíveis no plugin.
Palavras-chave: Técnicas de gravação, Ambisonics, Gravação multicanal, Binaural

Introduction

This project was idealized as an introductory investigation for the project entitled "Sound Reinforcement: A discrete system[1] for individual amplification" which aims to amplify acoustic instruments following the principles of sound reinforcement and the idea of stems through the use of digital sound capture, amplification and manipulation technology. In this investigation the main objective is to favor a performance of music with acoustic instruments in spaces where the acoustics are not suitable for the performance of musical presentations by means of acoustic instruments with low sound pressure.

In the late 1990s, during the period I dedicated myself to electronic music at the Institute of Sonology at the Royal Conservatory, The Hague, The Netherlands, I had the opportunity to develop recorder capable of manipulating live sounds, a practice known as live sound processing. With this instrument, that we named the *e-recorder*, I had the chance to participate in several concerts with important musicians from the electronic music scene. It was precisely with English composer Richard Barrett that the idea arose of using a separate amplification system for each performer and not routing the entire group's signals to a single PA. According to Barrett, this type of decision would allow listeners to identify the sound material with its source, integrating the visual and gestural factor of the performer into the sound. We believe that by means of such type of amplification setup the sonic result in the environment should become more organic, even facilitating the interaction between performers, since the sound sources are easily identifiable. This is the hypothesis that informs the current project, which is supported by The National Council for Scientific and Technological Development (CNPq) of Brazil. However, due to the restrictions derived from the COVID-19 pandemic that proscribed any social contact we had to change our plan of route slightly. So, we decided that we should first focus on the recording of solo acoustic instruments using a multitrack recording technology for further binaural rendering and, as the social contact restrictions fade-out, we will retake the original plan. In this manner, our first investigation was aimed to capture one low pressure acoustic instrument with an array of 10 microphones and process the sound through 3^{rd} order ambisonics. Our hope is that by developing advanced recording techniques we could profit by using this knowledge for registering the results of the original plan.

The Church

The main objective of the idea of multitrack recording was the creation of a cohesive sonic result that uses the 10 microphones to obtain an involving binaural rendering of a 3^{rd} order ambisonics. This idea comes from the assumption that "the higher the order of ambisonics, the greater the angular precision and the larger the audience listening area." (Barrett, 2012, p. 1). We chose the Parish Holy Spirit Church in the city of Uberlândia, Minas Gerais, Brazil, for recording a musician interpreting the first dance, Allemande, of Johann Sebastian Bach's Partita for flute (in a minor) in a Voice Flute[2].

Figure 1
The Parish Holy Spirit Church designed by architect Lina Bo Bardi.

Design by renowned architect Lina Bo Bardi in 1975, the building used almost all material made *in loco* for its construction. Clay bricks and wood structures from the region were used for building the round church which has 8 wooden pillars distributed equidistantly and perpendicularly. The floor consists of Portuguese stones and its ceiling has no encasement, exposing its wooden structure with terracotta tiles[3]. The latter reduces the amount of reverberance in its acoustics and also allows birds to pay the church a visit at times.

This building was declared historical heritage in 1997 by the *Instituto Estadual do Patrimônio Histórico Artístico* of the state of Minas Gerais, has an internal area of approximately 300m^2, a diameter of 19m[4] (Lazzarin, 2015, p. 46), less than 2 seconds of reverberation and, at some spots, a fairly fast echo due to reflection between the round structure. The roof has a slope of 40% and its highest point measures 8.3m and its walls are 4.30m high (Fig. 3).

Figure 2
Mockup of The Parish Holy Spirit Church (Lazzarin, 2015, p. 82).

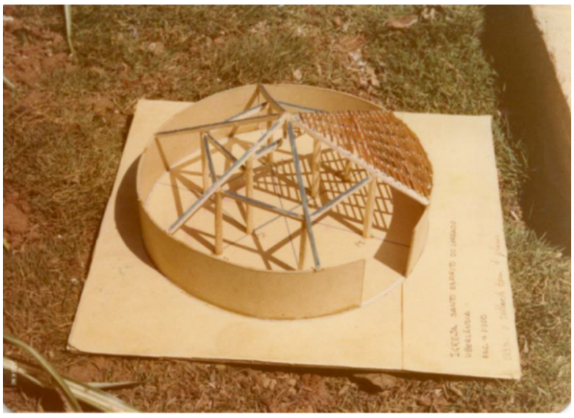
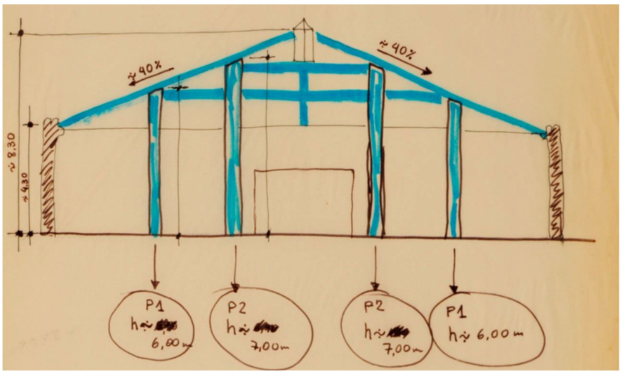


Figure 3
Measurements of The Parish Holy Spirit Church (Lazzarin, 2015, p. 83).



Another issue in the building is street noise coming mostly through the main door, which does come to interfere at times and for which we had to apply some technological solution, as it will be explained later.

Microphone setting

For this recording, we chose the M-Audio 2626 interface, the software Reaper and the following microphones: 4 cardioid KM184 (Neumann), 2 cardioid TLM103 (Neumann), 2 omnidirectional (DPA 4060) and 2 dual diaphragm condenser B2-PRO (Behringer). The signals were recorded in uncompressed WAV format at 24 bits of amplitude depth and the sample rate of 96 kHz.

The setting of the 10 microphones was based on the The Wide Cardioid Surround Array (WCSA) model, introduced by Mikell Nymand[5], which uses a mixture of cardioid and omnidirectional microphones to enhance the listening position from a sweet spot to a sweet area.

For our recording we chose the following distances between the three microphones from the front tree: Left and Right (LR) (Neumann KM 184) distanced by 75cm from the Centre (C), angled at 345° and 15° respectively. One Front microphone (F) (Neumann TLM 130) placed 20cm in front of C. Left Surround and Right Surround (LS, RS) (Neumann KM 184) distanced by 210cm from LR and 150cm from each other, pointing 30° upwards. However, we changed the angle of the back microphones to LS at 270° and RS 90° in order to avoid pointing them towards the main door, which would bring excessive noise from the street into the recording. All the microphones were placed at the height of 200cm. In order to have the possibility of a more direct sound, we used another Frontal microphone (F1) – a Neumann TLM103 – placed 70cm in front of the recorder player at 150cm high from the floor. The recorder player was placed 200cm away from F's TLM130 of the main front tree. Plus, we positioned 2 dual diaphragm condenser B2-PRO at 250cm L and R from C. We also placed 2 DPA 4060 omnidirectional microphones (240cm away from the player) at 37cm from each other and at 150cm high from the player's ground, since it was placed in the altar.

Figure 4
The Wide Cardioid Surround Array (WCSA).

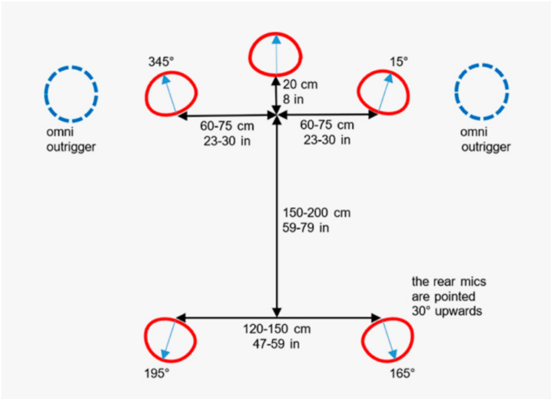
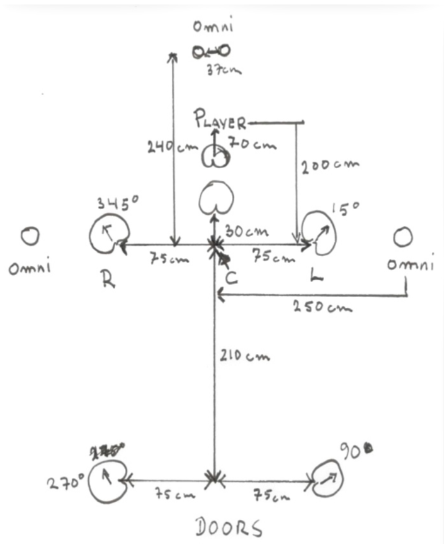


Figure 5
The Microphone Array used at The Parish Holy Spirit Church.



Ambisonics and processing the sound with plugins

As many already know, ambisonics is a full-sphere surround sound format that contain speaker-independent representation of a sound field called B-format, which can be decoded into the listener's speaker format, binaural included. As Natasha Barrett affirms: "For decades, ambisonic recordings have commonly be made with the Soundfield microphone. This microphone captures the complete sound-field using four cardioid response capsules located as close as possible in a tetrahedron geometry" (Barrett, 2012, p. 3). However, the use ambisonic microphones are questioned for obtaining high quality

3D recordings. In an equipment review from the Boom Library, we read: "Pretty much the same misunderstanding is widespread when it comes to Ambisonic. Ambisonic is a system developed in the 60s/70s. Today it is mostly a delivery format to be able to decode a specific location of a sound into a wide range of speaker setups, one being binaural for headphones. That means there is no need to have native Ambisonic recordings to deliver audio in Ambisonic format for VR or 360° videos to create a 3D environment." (Retrieved from The Boom Library: https://www.boomlibrary.com/blog/why-we-dont-use-ambisonic-microphones-even-for-vr/).

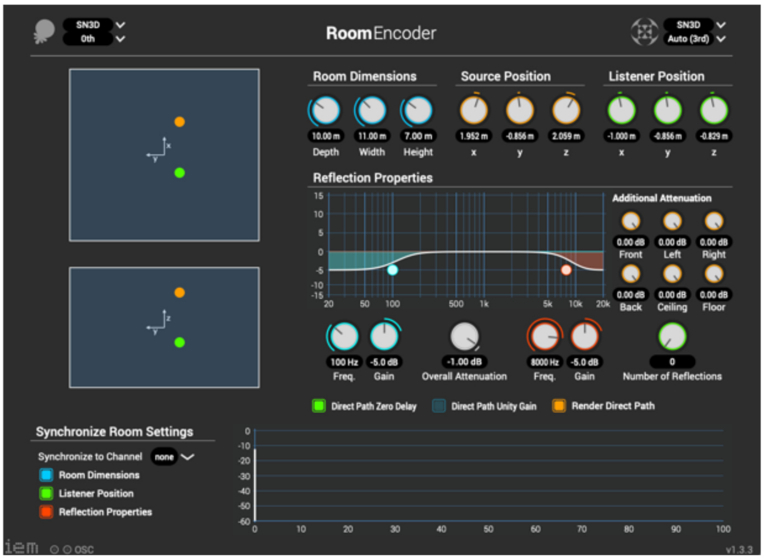
In our case, our objective in using a 10 microphone array for a recording was to create an immersive and engaging acoustic result. Judging the amount of unexpected noise leaking from the street into the church, it was necessary to apply a highpass filter at 260Hz first, taking into consideration that the lowest note of the voice flute is a D3 in A=415Hz, which is around 270Hz. Although not all the noise could be removed, especially the sounds coming from singing birds at the roof of the church and happy children outside the church's door, the results of the recording bring the idea of space very clearly.

The process we used consisted of applying IEM's RoomEncoder plugin into each of the 10 stereo tracks. This plugin allows you to choose the distances in 3D to artificially change the relative position between sound source and the listener. But first we needed to set each of the 10 tracks of our Reaper multitrack session to 16 channels because we were dealing with higher order ambisonics. Using the RoomEncoder in each of the 10 tracks allowed us to set the source position to the approximate distances that we used in the actual physical space. In this manner, we set the height of the labium of the recorder (1.497m) and the distance of the listener in relation to the source as the distance from the musician to the centre of the front (F) microphone tree (2.0m). Also, for keeping the distances used between the microphones during the recording it was necessary to apply the RoomEncoder plugin on each of the 10 tracks set to 16 channels before routing them all to the IEM's BinauralDecoder set to convert from 3rd order Ambisonics[6] to a stereo experience. Although the virtual shape of the plugin concerns the shape of a shoebox, differing from the actual shape of the physical building, the objective was not to recreate the original measurements of the church, but to choose the desired amount of natural reflections by changing the relative positions between the sound source and the listener. It is important to note that the setting for choosing the amount of reflections of the RooEncoder plugin was set to zero. Finally, all 10 tracks were put through the main bus containing the BinauralDecoder. The resulting sound is very involving due to a balance between the signals recorded by the main microphone array (L,R,F,LS,RS and F1) and the nicely reverberant components of the 4 omnidirectional microphones. Therefore, it depicts an interesting stereo image that is focused enough in the centre and spreads nicely to the sides, creating an immersive sonic spatiality.

Conclusion

We hope that this work in progress has exposed some of the possible manners of processing multitrack recordings taking advantage of the natural acoustical characteristics of the physical environment. Since the possibilities with the number of plugins available today are extremely extensive, there is an instigating line of research that can explore the use of different setups of microphone arrays in various places and their diverse types of acoustics. Also, we are looking forward to implement the ideas and knowledge derived from this investigation into the original focus of this research, namely registering the development of a system of sound reinforcement through the use of digital sound capture, amplification and manipulation technology.

Figure 6
Parametres used with IEM's RoomEncoder for the Front microphone (F).



Notes

1 In this project, the noun system is understood as the initiative to find the parameters to be able to understand the balance and sound quality in a pragmatic way based on hearing and not as the fixed configuration of hardware and software parameters.
2 Voice Flute is a recorder in D4, which is called by this name because of its proximity to the human voice.
3 For more information see Lazzarin, 2015, pp. 55 and 83.
4 Lazzarin, 2015, p. 44.
5 https://www.dpamicrophones.com/mic-university/immersive-sound-object-based-audio-and-microphones (accessed on April, 1st, 2022)
6 Follow this link to listen to it in SoundCloud: https://soundcloud.com/user-519752854/bach-roomencoder-iem-1?utm_source=clipboard&utm_medium=text&utm_campaign=social_sharing

References

Barrett, N. (2012). The perception, evaluation and creative application of high order ambisonics in contemporary music practice. Ircam composer in research report. Paris.

Lazzarin, A. L. (2015). A Igreja Divino Espírito Santo do Cerrado e suas Alternativas à Arquitetura Brasileira. Dissertação de Mestrado, Universidade de São Paulo, Instituto de Arquitetura e Urbanismo, São Carlos, SP.

Wilson S.; Harrison J. (2010), Rethinking the BEAST: Recent developments in multichannel composition at Birmingham Electroacoustic Sound Theatre. Organised Sound, vol. 15, issue 3, pp. 239-250.