The Fundamentals of Sound Field Reproduction Using a Higher Order Ambisonics System

Rok Prislan*
rok.prislan@innorenew.eu
InnoRenew CoE
Livade 6, SI-6310, Izola, Slovenia

ABSTRACT

Conventional sound recording methods are based on recording the sound pressure level with a microphone which is after some signal processing reproduced by loudspeakers. In spatial audio, more than one microphone and loudspeaker are required to provide the sound source location information to the listener. Several spatial audio formats have been developed and some have successfully entered our homes, such as the the multichannel 5.1 surround system. Among spatial audio formats, Ambisonics stands out due to its capability of capturing and reproducing the whole sound field and is not limited to predefined loudspeaker setups. In the paper, the InnoRenew CoE's Ambisonics system is introduced and some of its underlying principles are explained. Furthermore, practical examples of the use of Ambisonics, also in relation to Virtual reality applications, are presented.

KEYWORDS

higher order Ambisonics, sound field reproduction

1 INTRODUCTION

Michael Gerzon [1] invented Ambisonics in the 1970s, and since it has mainly been a research topic in acoustics. It's higher order version was developed twenty years later but only recently it has become a commercially available recording system [2]. Currently, more and more user applications of Ambisonics are emerging since Ambisonicsis is being positioned as the audio framework of choice for virtual reality [3, 4].

The acoustic laboratory of InnoRenew CoE has currently been equipped with a higher order Ambisonics system. The system is composed of a 32 channel microphone [2], a set of 64 full range loudspeakers, a dedicated low frequency loudspeaker, all the required AD/DA converters and accessories, such as stands and cables. The equipment in shown on Figure 1.

The system will be used for perceptual acoustic experiments, mainly by exposing test subjects to different acoustic conditions and investigating their response. In fact, room acoustic conditions are essential for a healthy and creative working environment – one of the important research topics at InnoRenew CoE. Another use of Ambisonics is in combination with virtual reality systems (e.g. [7]) that can provide a multi-sensoric immersion experience to users.

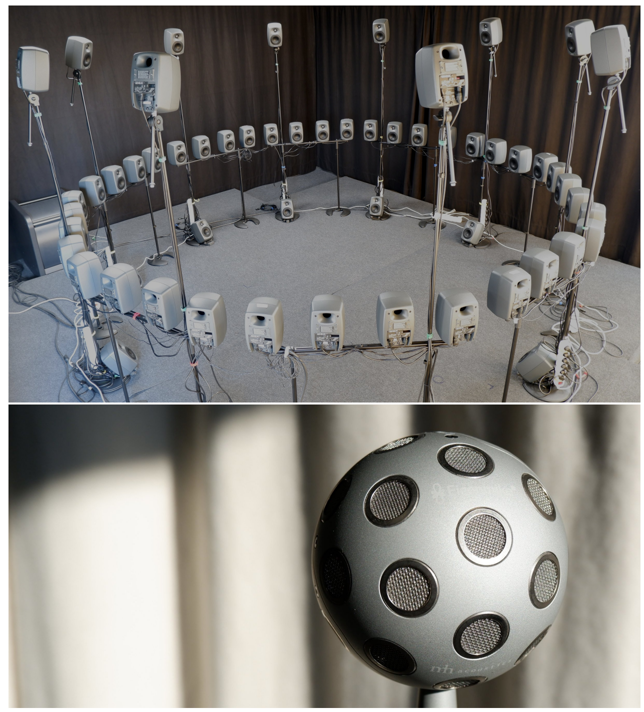

Figure 1: The higher order Ambisonics reproduction system with 64 loudspeakers (top) and the Ambisonics microphone [2] (bottom) which are part of the InnoRenew CoE's acoustic laboratory equipment.

2 RECORDING AND ENCODING

Ambisonics is a method of recording and reproducing a sound field and preserving its directional properties. The signal is coded, which is different in comparison with traditional multichannel audio formats (e.g., stereo, and 5.1 surround). In those, each channel contains the signal corresponding to a loudspeaker while in Ambisonics each channel contains derivatives of the pressure field. The encoded signals are known as B format.

In Ambisonics we record with several microphones spherically arranged on a (virtual) sphere. Summing properly weighted signals from each microphone is equivalent to recording with a microphone of a certain directional characteristic. Such processing is the basis of Ambisonics encoding [2], in which case the chosen directional patterns correspond to spherical harmonic functions (see figure 2).

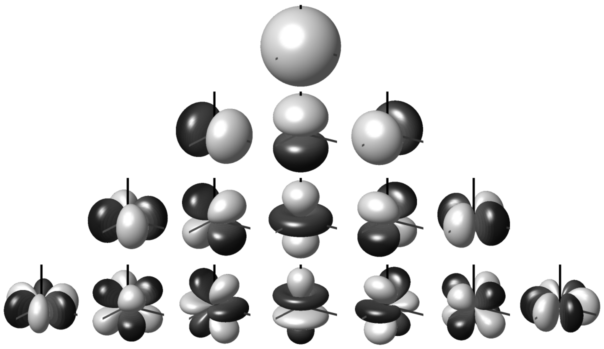

Figure 2: Polar patterns of spherical harmonics $Y_n^m(\theta,\varphi)$ of zero, first, second, third and fourth order (from top to bottom) (figure from [5]).

Spherical harmonic functions are grouped by their order number n and particular coefficient $m = -n, ...n$. Mathematically, each spherical harmonic corresponds to the angular portion of the solution of the wave equation. This way it is possible to capture the whole sound filed as it can be, in fact, decomposed into spherical harmonic functions

$$p(k, \mathbf{r}, \theta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} 4\pi i^n j_n(kr) A_{n,m} Y_n^m(\theta, \varphi) \qquad (1)$$

where φ and θ are the azimuth and elevation, $\mathbf{r}$ is the spatial coordinate and k is the wavenumber.

The general idea of a higher order Ambisonics encoding is to record sound with directionality patterns that correspond to polar patterns of spherical harmonics. As such, it is possible to encode the sound field in form of spherical harmonic decomposition factors instead of the sound pressure level at each microphone position.

The maximum order N at which we perform the expansion defines the order of the Ambisonic system. Each order contains $2N + 1$ channels, meaning that in total the ambisonics system of order N has $(N + 1)^2$ channels that have to be stored. Increasing the order to which the decomposition is done improves the directionality of the recording.

An important limiting factor for increasing the Ambisonic order is the number of microphones positioned on the sphere: the pressure is discretely sampled, which leads to artifacts, such as aliasing. Issues related to low frequency noise and several other technical limitations have been studied [3]. Generally, increasing the number of microphones is favored, although this obviously increases the cost of the system.

It is important to understand that the B format encoded signals can be as well manipulated with proper signal processing. For example, the sound field can be easily rotated for a certain angle, and it is also possible to focus to a certain direction of the sound field [6].

3 REPRODUCING THE SOUND FILED

The biggest advantage of Ambisonics over conventional multichannel spatial audio techniques (e.g. stereo, 5.1 and 7.1 surround) that consider fixed loudspeakers position is independence on the loudspeaker setup. In Ambisonics, the decoding from the B format takes into account the actual position of the available loudspeakers, which can be arbitrary chosen. Nevertheless, a high number of loudspeakers spatially distributed around the listener are required to provide a full and precise spatial impression.

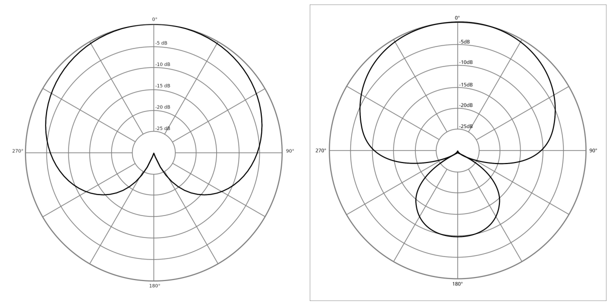

Figure 3: Example of a cardioid (left) and super-cardioid (right) microphone polar pattern (figure from [8]).

The number of loudspeakers required is as well dependent on the order of the system. The $N-$th order requires a minimum $(N+1)^2$ loudspeakers, meaning that 9 loudspeakers are required for the 2nd order, 16 for the 3rd and 25 for the 4th.

There are several strategies for decoding the B format to be reproduced on a setup of loudspeakers. The basic idea is to directionally filter the recorded signals by virtual microphones pointing in the direction of each loudspeaker.

Setting the proper directionality patterns (see Fig 3) is the important part of the decoding process. In a regular layout, the signal emitted by a loudspeaker is the same as it would be recorded by a supercardioid microphone pointing towards that direction [6]. This means almost all loudspeakers emit sound at the same time, and for a given sound source position, loudspeakers in the opposite direction emit in opposite phase.

4 THE AMBISONICS SYSTEM IN USE

Ambisonics systems are an useful research tool in acoustics, mainly because they enable to reproduce sound emitted by sources together with the acoustic environment in which they are located. An important example of such use are the investigations carried out by Tapio Lokki [9] with his group who have been investigating perceptually relevant acoustic properties of concert halls. In their research, listeners have been asked about their preferences about the acoustics of different concert halls in which the same orchestra was performing. As an individual's acoustic memory is strongly affected by the time that has passed since each concert experience, it is required for such research to migrate the listener and orchestra between concert halls immediately. This can be achieved by an Ambisonics system in which recordings can be switched by a push of a button.

Figure 4: Photo of a listener in the Ambisonics loudspeakers ring at the InnoRenew CoE's Acoustic lab. The control over the system and perceptual response is based on a tablet PC as an interface.

Currently at InnoRenew CoE, we are setting up the Ambisonics system for the listener to rate different acoustic environments. The research is not limited to a specific environment type, such as concert halls, but includes acoustic environments to which we are exposed on a daily basis (commonly referred to as soundscape [12]). The recording will be performed on several different locations that include noisy and pleasant environments, such as high-traffic roads, busy workspaces and nature.

The interaction of the user with the system can be designed in various ways. Firstly, we are relying on a tablet PC as shown in Fig. 4. Using the tablet, the playback is controlled and the response from individuals is gathered. The system can be upgraded with more advanced response tracking options, such as performing eye-tracking or tracking the electrodermal activity of the test subject.

Spatial sound can be incorporated into virtual reality (VR) interfaces, such as VR headsets. The most accessible approach is to use headphones for which the signals have to be processed based on Head-related transfer functions [10]. The main drawback in this case is that wearing headphones is not natural to users and can produce discomfort. It is well known [11] that the listener does not localize the sound source as being external, but rather positions it in between the ears. This phenomenon of using headphones is known as lateralization of sound sources [11].

Generally, the relative position/orientation of the sound source in relation to the listener's ears changes over time, meaning that Head-related transfer functions applied to process the audio content have to adopt accordingly. Therefore, when using headphones in VR head tracking and real time audio processing are required.

In this perspective, the use of Ambisonics advantageous as the full sound field is reproduced and the listener can freely rotate his head while localization clues are correctly perceived. Additionally, in Ambisonics the ears are free from wearable equipment, which is a more natural condition for the user.

A relevant use of Ambisonics in relation to VR is also recording the sound field using an Ambisonics microphone and reproducing it over headphones instead of an Ambisonics reproduction system composed of a high number of loudspeakers. In fact, the B format encoded signals can be processed for a binaural playback for any arbitrarily chosen head rotation. Recently, many commercial second order Ambisonics microphones containing four microphones have become available on the market together with dedicated digital audio work station plug-ins for binaural decoding.

5 ACKNOWLEDGMENTS

The author gratefully acknowledges the European Commission for funding the InnoRenew project (Grant Agreement #739574) under the Horizon2020 Widespread-Teaming program and the Republic of Slovenia (Investment funding from the Republic of Slovenia and the European Union's European Regional Development Funds).

REFERENCES

[1] Michael A Gerzon. *Periphony: With-height sound reproduction.* In: Journal of the Audio Engineering Society 21.1 (1973), pp. 2–10.
[2] Eigenmike — mh acoustics LLC, url: https://mhacoustics.com/products (visited on 16/09/2020).
[3] F. Zotter and M. Frank: *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, SpringerOpen, 2019
[4] S. Sherbourne et al. Ambisonics and VR/360 Audio in Pro Tools url: http://www.avidblogs.com/ambisonics-vr360-audio-pro-tools-hd/ (visited on 16/09/2020)
[5] Ambisonics — Wikipedia, The Free Encyclopedia. url: http://en.wikipedia.org/w/index.php?title=Ambisonics&oldid=656474391 (visited on 16/09/2020).
[6] D. Arteaga, *Introduction to Ambisonics.* url: https://www.researchgate.net/publication/280010078_Introduction_to_Ambisonics (visited on 16/09/2020)
[7] Virtual Acoustics — Institute of Technical Acoustics RWTH Aachen University url: http://virtualacoustics.org/ (visited on 16/09/2020)
[8] Microphone — Wikipedia, The Free Encyclopedia. url: https://commons.wikimedia.org/w/index.php?curid=50230608 (visited on 16/09/2020).
[9] T. Lokki, *Tasting music like wine: Sensory evaluation of concert halls,* Physics Today, 67(1), 27 - 32, 2014, doi: 10.1063/PT.3.2242
[10] Head-related transfer function — Wikipedia, The Free Encyclopedia. url: https://en.wikipedia.org/wiki/Head-related_transfer_function (visited on 24/09/2020).
[11] Yost W.A., Hafter E.R. *Lateralization.* In: Yost W.A., Gourevitch G. (eds) *Directional Hearing.* Proceedings in Life Sciences. Springer, New York, NY (1987).
[12] Soundscape — Wikipedia, The Free Encyclopedia. url: https://en.wikipedia.org/wiki/Soundscape (visited on 24/09/2020).