INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2026, VOL. 72, NO. 1, PP. 1-6
Manuscript received January 19, 2026; revised January, 2026. doi: 10.24425-ijet.2025.155466

Immersive recording of vocal ensemble on location – practical considerations and case study

Kaja Kosmenda, and Witold Mickiewicz

***Abstract*—The dynamic development of spatial sound technology has led to a rapid movement from classic stereo and surround formats to immersive systems, including 7.1.4 and other Dolby Atmos configurations as a primary production format. As part of the project, a custom 3D microphone array was developed and built, inspired by contemporary research on techniques for recording three-dimensional acoustic space. The aim of the work was to obtain a realistic representation of space in the form of a true multi-channel recording while maintaining configuration flexibility and low construction costs. The recorded material was used to study various methods of reference recordings mixing for immersive loudspeaker configurations. Two approaches were analyzed: direct assignment of signals to 7.1.4 system channels and object-based rendering using Dolby Atmos Renderer within Pro Tools, allowing for dynamic positioning of sources in space. The results indicate that the channel-based approach regarding immersive mixing provides the best sound quality , while object-based rendering offers greater creative flexibility and the possibility of reducing the number of microphones. Further research focuses on the development of binaural immersive listening methods using obtained audio and the creation of authentic 3D recordings for streaming platforms.**

***Keywords*—immersive sound; immersive recording; microphone techniques; 3D microphone arrays; recording techniques**

I. Introduction

THE development of sound technology in recent decades has brought a significant evolution in the way we record, process, and perceive sound. From the first monophonic systems, through stereo and multi-channel formats, to immersive solutions. Not only has the quality of reproduction changed, but also the perception of the spatiality and realism of the sound scene [1]. With growing technological capabilities and the development of immersive rendering methods, sound is no longer just an addition to a picture but it is becoming a full-fledged medium for three-dimensional experiences, engaging the listener in a way that is similar to natural acoustic perception.

In response to these trends, our department has constructed and equipped an immersive laboratory shown in Fig. 1, featured with a 7.1.4 speaker system, enabling the creation, analysis and playback of various immersive sound compositions. This system allows for research into the subjective perception of three-dimensional sound, testing of mixings and rendering algorithms, and the development of tools used for the production of immersive audio. The laboratory also serves as a platform for interdisciplinary collaboration. Those include musicians, composers, and most importantly choirs, since the laboratory enables realistic playback of created immersive works in a set acoustic environment in different loudspeaker setup scenarios.

Fig. 1. Immersive sound laboratory in Electrical Engineering Faculty of West Pomeranian University of Technology in Szczecin

The main goal of our research is to develop an immersive listening system for binaural playback that will allow the immersive experience to be transferred from the speaker laboratory to individual usage [2]. The aim is to develop methods of sound recording and processing that enable faithful reproduction of the acoustic space in a headphone environment. A key element underlying the work is the acquisition or creation of high-quality, multi-channel reference material that will allow the accuracy and natural feeling of the generated immersive experiences to be assessed. In this context, however, there are significant challenges, among other things the acquisition of any reference recordings of sufficient quality and number of channels.

II. Objectives

As mentioned, one of the key challenges in conducting research on immersive sound is obtaining high-quality reference

Witold Mickiewicz and Kaja Kosmenda are with West Pomeranian University of Technology in Szczecin, Poland (e-mail: witold.mickiewicz@zut.edu.pl. kaja_kosmenda@zut.edu.pl).

multichannel audio sample. Despite many audio tracks available online in various databases, the high requirements for this specific research are far from what is publicly accessible. Currently available recordings are often of dubious technical quality, resulting from both limitations in the recording and mixing process and also secondary processing during publication. An additional problem is the limited availability of material in multichannel formats, especially those offering full immersive configurations such as 7.1.4 which need 12 separate channels for best post-processing results.

On the other hand the common music streaming platforms nowadays declare support for a whole range of immersive formats in cooperation with technology companies developing relevant software [3]. However despite the claims made by providers of these recordings and authors of said technologies, the resulting listening experience remains imperfect, with the quality varying significantly across both the platforms and individual listeners. Because of this and a lack of information about the specific encoding technologies used, the overall quality of the audio is thought of to be too inconsistent for proper research. The most noticeable cases are duplicate channels, e.g., rear or side channels, especially at the added immersion height, or minor errors in spatial mapping, which lead to distortion of the sound scene. All pointed above issues are significant and prevent the creation of a consistent reference set for research on the subjective perception of immersive sound. For these reasons, the main goal for the beginning of our research is to develop and acquire our own controlled multichannel material that can serve as a reference point in analyses and listening tests in both immersive loudspeaker environments and binaural systems.

The development of sound recording techniques reflects the evolution of how acoustic space is perceived, from simple stereo systems to complex three-dimensional (3D) systems. The first commonly used stereo configurations, such as AB, were based on the principle of phase differences between omnidirectional microphones. Despite its wide spatial image, this system required relatively large distances between microphones, which led to phase coherence problems and limited its practical application in studio conditions [4]. In response, systems based on directional microphones were developed, such as OCT (Optimized Cardioid Triangle), proposed by Theile and Wittek [5]. The OCT system used cardioid microphones arranged in a triangle with different angles and shorter distances, which improved source and phase separation as well as localization accuracy. These types of configurations were the starting point for further experiments with surround microphones, leading to the expansion of systems from flat 2D configurations to spatial 3D arrays.

The next step in the evolution of spatial recording techniques was the emergence of coincident and quasi-coincident microphones, such as Soundfield [6] and first-order Ambisonic systems. These microphones, usually based on tetrahedral capsule geometry, four cardioid or supercardioid microphones arranged at the tips of a tetrahedron, enabled the encoding of the acoustic field in the so-called B-format domain (W, X, Y, Z channels). The tetrahedral geometry allowed for 3D sound to be recorded, but it wasn't until years later that systems with playback for such audio were created. Ambisonics as a representation of sound directionality in three dimensions became the foundation for later higher-order systems (Higher Order Ambisonics) [7],[8].

With the development of digital technologies and increased computing power, it has become possible to expand microphone arrays with a larger number of capsules and more accurate reproduction of the acoustic field. Commercial solutions such as Eigenmike (mh Acoustics) and Zylia ZM-1 have emerged, integrating more microphones in a single body, enabling three-dimensional reconstruction of the sound field in real time. [9].

Our work is part of this research trend. Based on an analysis of existing solutions and literature on the subject, we decided to build our own three-dimensional microphone array, inspired by the research of Lee H., Johnson D. and Gribben C. [10],[11], who demonstrated that the appropriate configuration of 3D microphones significantly affects the subjective sense of immersion and the precision of sound localization. The authors conducted a series of experiments comparing microphone arrays of different sizes, configurations and capsule types, analyzing the impact of these parameters on the perception of spatiality and source localization. The results of their work showed that even slight changes in microphone spacing or directional characteristics can significantly affect the perception of immersion. The aim of this work is to create high-quality multichannel material that will serve as a basis for further research on immersive listening, especially in 7.1.4 system.

In the context of immersive sound recording, one of the key factors determining the quality of the subjective perception of spatiality is the size and geometry of the microphone array used. Greater distances between microphones allow for better reproduction of time and directional differences, which are responsible for the sense of depth and width of the acoustic scene. On the other hand, large arrays generate practical challenges: they require stable mechanical construction and, most importantly, increase equipment costs. In our research, we therefore adopted a compromise solution, using mid-range microphones that provide sufficient signal quality while keeping the overall system costs reasonable.

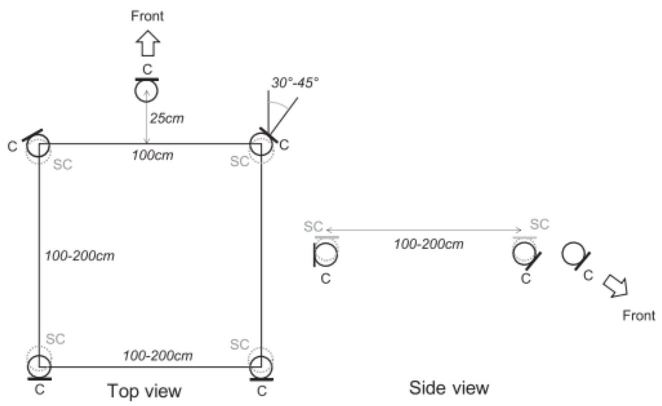

Fig. 2. Top and side views of PCMA-3D. The solid black and dotted grey circles represent the middle and upper layer microphones, respectively. C = cardioid, SC = supercardioid. [10]

When analyzing the sample recordings from various configurations provided by Lee, we noticed their exceptionally realistic spatial image and soundstage coherence. This experience prompted us to develop our own recording system based on similar principles. Our goal was therefore to investigate the practical aspects of the construction and use of a three-dimensional microphone array in real acoustic conditions and to present the results in the context of research on the practical approach to immersive recording on site.

Fig. 3. Church interior during immersive recording session

A special opportunity to test the designed system happened with a curtesy of a friendly choir, within a recording session organized in a sacred space with rich reverberation acoustics (the interior of the Church of St. Peter and St. Paul the Apostles in Police-Jasienica shown in Fig. 3). The choir, planning a professional recording, invited us to participate in the project, which allowed us to combine artistic needs with research goals. During the session, our experimental microphone array, tentatively referred to as a "spider," modeled and lightly based on the PCMA-3D shown on Fig. 2 was used. Its configuration is described in detail in the literature [10].

Due to limited equipment availability, only cardioid microphones were used in our implementation, whereas the original PCMA-3D system also provides for the use of supercardioids in certain positions. Despite these limitations, the results obtained proved satisfactory, and the recordings were characterized by a clear sense of depth and spaciousness. The resulting multichannel material was a foundation for further research and most importantly as a base for true immersive playback in our 7.1.4 listening environment.

The whole study had two main objectives. Firstly, to assess the feasibility of building a low-cost 3D microphone array capable of producing high-quality reference recordings in 7.1.4 format. Second one was to conduct an exploratory comparison of two methods of creating an immersive playback based on the obtained material: direct signal mapping to the channels of the listening system and object-based rendering using the Dolby Atmos engine. The ultimate goal was to determine the practical advantages, limitations and potential of both approaches in the context of creating accessible, high-quality immersive reference recordings.

III. Methodology and Recording Session

As mentioned earlier for the purposes of research on immersive sound recording, a proprietary 3D microphone array was developed and constructed. The design involves placing eight microphones in the corners of a cube with sides measuring 1 m, which allows for spatial mapping of the acoustic field with a clear distinction between vertical and horizontal layers. In addition, a central microphone was used and placed approximately 50 cm in front of the center of the front vertical plane of the cube, served as a reference point and central channel in the subsequent loudspeaker system positioning.

All microphones used were RØDE NT-5 cardioid microphones, characterized by are characterised by a relatively low level of inherent noise (16 dBA) and a balanced frequency response in the range of 20 Hz–20 kHz, as stated by the producer which helps to obtain reference material with high colour transparency. Cardioid characteristics provide moderate directional discrimination, but at higher frequencies they offer limited vertical separation and more noticeable off-axis colouration compared to supercardioids. This has a direct impact on the stability and precision of height channels. Immersive systems using upward-facing supercardioid microphones provide a clearer representation of direct sound and a clear reflection pattern consistent with the room acoustics. In churches, for example, a blurred characteristic may often be more desirable, but the lack of precise depth and directionality reproduction was an important aspect for us in the subsequent creation of reference recordings. The RØDE NT-5 model was chosen as a compromise between sound quality, availability and cost, while maintaining consistency across all channels.

Fig. 4. 3D microphone array build by the authors for the purpose of immersive recording session

The directions of the individual microphones were selected in such a way as to obtain the most natural reproduction of the sound stage and differentiation of signals between channels and height levels:

- The lower front microphones were directed slightly downward, toward the performers, and spread at an angle of approximately 30° from the central axis, which allowed for the capture of direct sound while maintaining the width of the soundstage.

- The lower rear microphones were positioned horizontally, facing backwards, which allowed for the recording of the reverberation field and acoustic reflections.
- The upper front microphones were positioned perpendicular to the top in order to record vertical reflections and height information.
- The upper rear microphones were directed slightly upwards and backwards, along the axes connecting the center of the performers with the rear vertices of the array, which allowed for the capture of reverberation and natural spatial dispersion.

The configuration was designed to increase the separation of direct sound from reverberant reflections, especially in the case of the four top microphones and the two rear bottom microphones. This arrangement allows for a clear three-dimensional character of the recording, with a well-defined vertical layer and a natural sense of depth.

Professional solutions designed for building and mounting multi-channel microphone arrays are characterized by high precision, but also very high costs. For example, an aluminum beam with holders for five microphones, without a stand, can cost as much as $1000, making this type of equipment inaccessible to many researchers or educational institutions. An additional challenge in designing such structures on one's own is the need to make precise inch threads and holders with adjustable geometry, allowing for free change of the spacing and orientation of microphones.

Therefore, as part of the project, a custom solution based on commonly available components was created. It eliminated the need for a professional metalworking workshop while maintaining the functionality and adequate stability of the structure. The support system shown in Fig. 4 was built from the following components:

- Lighting stand with a T-bar crossbar (purchased at a music store), which is the main supporting element and allows for stable suspension of the entire structure.
- Structural elements made of aluminum sections (20×10 mm profile and 16 mm diameter tube), purchased from an aluminum wholesaler, forming the cubic frame of the matrix.
- Connectors and brackets for attaching balusters to steel railings, purchased from a specialist online store, which provide three degrees of freedom of adjustment – movement along the axis and rotation in two planes. This allows for precise adjustment of the direction and angle of each microphone, are shown in Fig. 5b.
- Inch thread screws (suitable for standard microphone holders) and metric screws with wing nuts for quick adjustment of the position of the elements, purchased from a hardware store, are shown in Fig. 5c.
- Photographic brackets, adapters, and adapters with inch threads, purchased from online stores, are shown in Fig. 5a.
- The matrix cabling was made using a ready-made, pre-assembled 8-channel multicore cable purchased from an online store, which enabled a safe cable routing and aesthetic, as well as a durable connection of the microphones to the recording system.

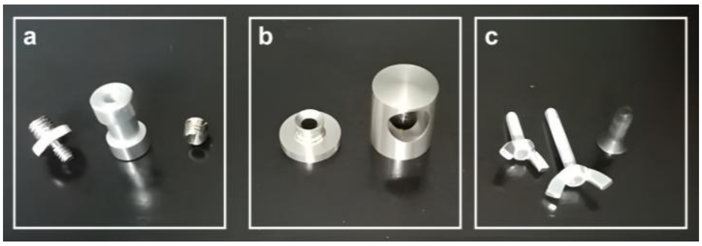

Fig. 5. Components used in inch threads and handles for the microphones on the aluminum pipes

As shown in Fig. 6 the microphone holder created provides three degrees of freedom in case of movement: sliding along the beam, rotation around its own axis, and vertical position adjustment. The design is based on articulated elements that allow full adjustment of the microphone's position and orientation in space without the need for tools.

By using inexpensive, widely available components, we were able to create a design that allows for full adjustment of the microphones' position in space, while maintaining adequate stability and repeatability of settings. This solution proved to be not only economical, but also flexible as it allows for quick reconfiguration of the array geometry and adaptation to various immersive sound recording scenarios, when more or less channels are needed.

Fig. 6. Microphone holder positioned on the aluminum section

The choice of microphones used in the custom array was dictated by availability and the need to maintain uniform directional characteristics. We had a sufficient number of RØDE NT-5 cardioid microphones, which made it possible to forego a mixed capsule arrangement (e.g., supercardioid or omnidirectional) and maintain tonal consistency across all channels.

As part of the comparative session, parallel recordings were made using two configurations:

1. The classic Decca Tree array, which is the reference point for spatial sound recording in a 2D configuration,
2. An experimental 3D microphone array based on the described above design.

Both configurations worked simultaneously, which allowed for a direct comparison of surround material with immersive recordings.

Two independent devices were used for recording:

1. Millennia Media HV-3D – a high-end preamplifier ensuring very low noise and high bandwidth linearity,
2. Focusrite Clarett+ 8Pre – an audio interface with mid-range preamplifiers, used as a subsidiary channel.

Analog-to-digital conversion was performed using RME ADI, and the entire system was synchronized and connected to a computer via the RME Digiface USB interface. Avid Pro Tools 12 was used as the recording software.

The recordings were made at 24-bit resolution and 44.1 kHz sampling rate, using a laptop as the recording unit. In addition, in parallel with the 3D matrix, material was recorded using a Zylia ZM-1 microphone, loaned to us with the courtesy of another professor from our faculty, as seen in Fig. 4, allowing comparison of the custom solution with a commercial third-order ambisonics system. The obtained result was not satisfactory due to some technical difficulties but the differences between custom made 3D array and ambisonics array are considered for future assessment and comparisons.

IV. Results and Postprocessing

The recorded material was used for preliminary reconstruction and spatial analysis tests on a 7.1.4 listening system installed in an immersive sound laboratory. The aim of the experiments was to compare two methods of acoustic space reconstruction for immersive loudspeaker system: direct channel positioning and immersive rendering using Dolby Atmos renderer integrated withing Pro Tools.

1. **First approach** - direct channel mapping

In the first variant, the method of positioning microphone signals directly to the channels of the 7.1.4 system was adopted, without additional spatial processing. Each microphone from the 3D array was assigned to a specific speaker in the listening configuration in a manner corresponding to its actual position in the recording space. Signals from top channels corresponded to the top loudspeakers, front bottom and central signals were used in the three front loudspeakers, and the bottom rear channels were doubled in the side and rear loudspeakers.

This approach allowed for faithful reproduction of the original acoustic scene, maintaining the proportions of the levels, angles and suspension heights of the individual microphones. In addition, it allowed for experimental changes to the array configuration – by modifying the channel assignments, it was possible to simulate different variations in the placement of sources and microphones, analysing their impact on the impression of immersion of the recorded space.

2. **Second approach** - immersive rendering using Dolby Atmos

The second approach was based on the use of the Dolby Atmos Renderer environment in Pro Tools (Fig. 8), which allowed for the virtual placement of individual sound sources in a three-dimensional acoustic field. In this variant, microphone signals were treated as audio objects, which were assigned positions in the listening space in three different variants.

This solution enabled dynamic shaping of the sound space, including simulation of dynamic changes in the position of sources over time. Since flexible adjustment of the stage layout was possible, proposed positioning was as followed:

- Sources corresponding to the positioning of microphones in the custom built array (Fig. 8a)
- Further moving upwards the top microphones for better height separation in playback (Fig. 8b)
- Positioning sources to correspond to the layout of loudspeakers in the laboratory (Fig. 8c)

By using real, immersive multi-channel recordings, this approach provided a safe and controlled method of creating spatial object-based audio without the need for artificial signal generation or virtual reverberations. Comparing the two methods allowed us to assess the practical potential of using authentic 3D recordings in Dolby Atmos production environments (Fig. 7).

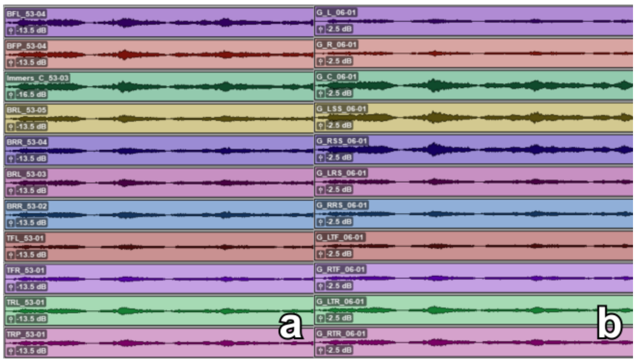

Fig. 7. Results from two different positioning approaches

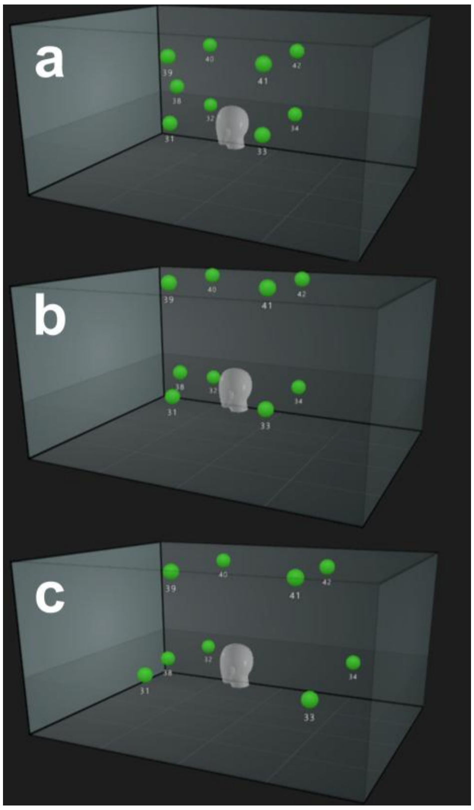

Fig. 8. Three different Dolby Atmos renders using sound sources as objects

V. Conclusion and Future Works

The experiments conducted have shown that the direct placement of microphone signals in the channels of the listening system allows for the preservation of very high sound quality and faithful reproduction of the acoustic scene of the recorded event. The authors compared all presented approaches by conducting comparative listening tests in the immersive laboratory with the 7.1.4 loudspeaker setup shown partially in Fig. 1. The loudspeaker setup is calibrated for a specific spot by the desk in the centre of the system which is the ideal place for listening to any mixed multichannel recordings and that is also the place from where the authors have listened to prepared audio. Although the findings were based on a critical listening tests there are few conclusions that can be drawn from it.

Comparing the three rendering variants, the first configuration, which reproduced the original physical layout of the microphone array as closely as possible offered the most stable localisation of the choir and preserved the spatial relationships captured during recording. The second variant, designed to emphasise the separation between the vertical planes, strengthened the perceptual impression of the church acoustic by enhancing the contribution of both the upper and lower layers of the array, resulting in a clearer rendering of the reverberant field. The third, fully object-based variant allowed the signals to be positioned and distributed more freely across the loudspeaker system, providing the strongest sense of envelopment and immersion. However, this increased spatial diffuseness came at the cost of slightly reduced localisation stability of the choir. Despite its simplicity, this approach allows for a detailed analysis of the spatial relationships between microphones and sound sources, which is a valuable reference point for further research on immersive playback.

At the same time, there has been a shift in the expectations of listeners and sound producers as a result of the popularization of immersive technologies. Consequently, standards such as Dolby Atmos and binaural listening systems are becoming increasingly important, as they allow for more flexible spatial experiences while maintaining compatibility with different listening environments.

The use of Dolby Atmos Renderer in Pro Tools has opened up new possibilities for processing and presenting material. In particular, it has made it possible to reduce the number of microphones while maintaining the immersive nature of the sound. This points to the potential of hybrid methods combining multichannel recording with object-based rendering techniques for added virtual reverberation especially in places where recording with a whole 3D microphone array is not possible.

In future, the research is planned to be developed towards binaural forms of immersive listening, intended mainly for playback on headphones. To this end, work is underway on cooperation with choral ensembles, which will result in authentic, multi-channel immersive recordings prepared specifically for streaming platforms.

The recording session allowed us to formulate several practical conclusions regarding the use of cube geometry in a church with strong reverberation. Regarding the array itself, listening tests and reference recordings showed that a more complex array with a different shape, perhaps similar to the geometry of spatial listening systems, would be more appropriate. The array itself can also be further modernized. The main problem is the weight of the structure, which can be changed for example by replacing steel elements with aluminum or carbon fiber ones. Although these solutions would increase production costs, their implementation could significantly improve the ergonomics and mobility of the system. The array is maybe not an ideal solution for manufacturing but we would like to highlight its simplicity for any interested researcher especially in the terms of building and setting it up. Given the assumption of maximum accessibility of the solution for other researchers and practitioners, the compromise between quality, price, and functionality remains a key aspect in the future for improving the custom 3D microphone array.

Acknowledgement

The authors would like to thank the choir Sonus Cantus Chamber Choir for participating in the research [13]. We highly encourage readers to listen to their recent works as well as the recording mixed and prepared for streaming platforms by one of the authors [14].

References

[1] A. Rogozinska, P. Geluso, “Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio”, Routledge, Taylor & Francis, 2018.

[2] W. Mickiewicz, K. Kosmenda, “Creation of immersive experience for headphone listening using directional room impulse responses for BRIR synthesis”, International Journal of Electronics and Telecommunications, vol. 71, no. 1, pp. 1–6, 2025. https://doi.org/10.24425/ijet.2025.155462

[3] O. Nosenko, “Immersive technologies in the formation of musical audio space”, Journal Homepage, vol. 24, pp. 23–32, 2024.

[4] J. Eargle, “The Microphone Book: From Mono to Stereo to Surround — A Guide to Microphone Design and Application”, 2nd ed., Routledge, 2005. https://doi.org/10.4324/9780080473468

[5] G. Theile, H. Wittek, “Principles in Surround Recordings with Height”, Journal of the Audio Engineering Society, vol. 59, no. 5, pp. 296–303, 2011.

[6] M.A. Gerzon, “Periphony: With-height sound reproduction”, Journal of the Audio Engineering Society, vol. 21, no. 1, pp. 2–10, 1973.

[7] J. Daniel, J. Rault, J.-D. Polack, “Ambisonics Encoding of Other Audio Formats for Multiple Listening Conditions”, 105th Audio Engineering Society Convention, San Francisco, USA, 1998.

[8] F. Zotter, M. Frank, “Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality”, Springer, 2019. https://doi.org/10.1007/978-3-030-17207-7

[9] J.G. Tylka, E. Choueiri, “Soundfield Navigation using an Array of Higher-Order Ambisonics Microphones”, Princeton University 3D Audio and Applied Acoustics Laboratory Report, 2016.

[10] H. Lee, “Multichannel 3D microphone arrays: A review”, Journal of the Audio Engineering Society, vol. 69, no. 1/2, pp. 5–26, 2021.

[11] H. Lee, D. Johnson, “3D Microphone Array Recording Comparison (3D-MARCo): Objective Measurements”, Journal of the Audio Engineering Society, vol. 69, no. 10, pp. 732–747, 2021.

[12] H. Lee, C. Gribben, “Effect of Vertical Microphone Layer Spacing for a 3D Microphone Array”, Journal of the Audio Engineering Society, vol. 62, no. 12, pp. 870–884, 2014.

[13] https://sonuscantus.azurewebsites.net/

[14] https://push.fm/fl/uc5u5a9q