VPS VIBRATIONS IN PHYSICAL SYSTEMS

DOI: 10.21008/j.0860-6897.2024.2.11
Vibrations in Physical Systems, 2024, 35(2), 2024211

First-order ambisonic microphones with MEMS and condenser capsules

Marcin LEWANDOWSKI[1], **Aleksander AUGUSTYNIAK**[2]

[1] Warsaw University of Technology, Institute of Radioelectronics and Multimedia Technology, Faculty of Electronics and Information Technology, Nowowiejska 15/19, 00-665 Warsaw
[2] Warsaw University of Technology, Institute of Electronic System Design, Faculty of Electronics and Information Technology, Nowowiejska 15/19, 00-665 Warsaw

Corresponding author: Marcin LEWANDOWSKI, email: marcin.lewandowski@pw.edu.pl

Abstract Technologies that produce and deliver immersive VR content are still growing. Ambisonic microphones are available in various types, from the FOA (first-order ambisonics, e.g., Sennheiser AMBEO VR Mic) to HOA (up to fourth-order ambisonics, e.g., MH Acoustics Eigenmike). They are designed with different capsules, from low-cost MEMS transducers (e.g., Zylia ZM-1) to high-grade condenser capsules (e.g., NEVATON VR). Each type of microphone has different parameters, design, spatial quality, and sound quality. This paper presents two designs of a low-cost FOA microphone based on low-noise MEMS (Infineon IM69D130) and electret capsules (JLI-2590A) in an identical tetrahedral arrangement. Details of the design, fabrication, and testing of two FOA microphones, including their frequency and directivity responses and subjective audio quality evaluation compared to commercially available solutions, are presented. This study aims to indicate whether the microphone capsule's type affects the recorded sound field's quality and to what extent.

Keywords: ambisonic microphone, sound recording, MEMS microphone, electret microphone.

1. Introduction

Ambisonics is a sound system that allows spatial encoding and reproduction of sound to give the impression of a sound environment similar to the real acoustic space. It was developed in the 1970s by British sound engineer Michael Gerzon, who created a sound coding system based on mathematical models of acoustic spaces [1].

Its main advantage is that once a recording is created in ambisonics format, it is possible to reproduce the sound field on any audio system, regardless of the number of speakers.

Ambisonic microphones are sound recording tools that allow precise recording of sounds in a three-dimensional space. They are beneficial in areas such as music and film production, video games, and virtual reality, where high sound quality is crucial to the viewer's experience. In the music industry, ambisonic microphones are used to create virtual concerts and other sound experiences that produce a spatial sound similar to a live performance.

The aim of this work is to compare FOA microphones built with MEMS (Micro Electro Mechanical Systems) and electret capsules. Both microphones are measured in an anechoic chamber, and the quality of various 3D sound field recordings is evaluated in listening tests.

2. Design and construction

2.1. Electret microphone capsule

The electret capsule model is the JLI-2590A [2]. It features good acoustic performance, such as a signal-to-noise ratio of 70 dB and a sensitivity of -44 dB. As shown in Fig. 1, the frequency response is flat for the angle of 0º (3 dB drop between 80 Hz and 10 kHz), and the microphone's directivity is cardioid.

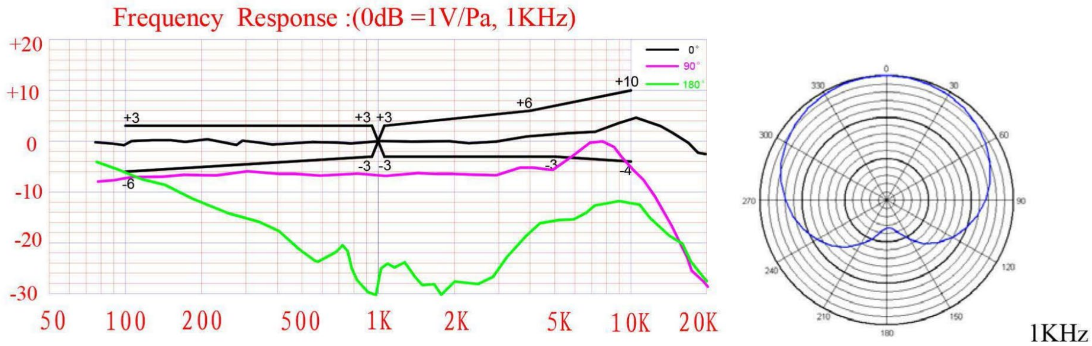

Figure 1. Frequency response (left) and directivity diagram (right) of the electret microphone capsule [2].

2.2. MEMS microphone capsule

The MEMS capsule model is the Infineon IM69D130 [3]. Its signal-to-noise ratio is 69 dB, and its sensitivity is -36 dB. The frequency response is flat between 100 Hz – 10 kHz (with 1 dB drop), as seen in Fig. 2. The microphone's directivity is omnidirectional. The capsules additionally contain a built-in ADC, which transmits the digital audio signal in the PDM (Pulse Density Modulation) standard. A significant advantage of these capsules is the availability of commercial evaluation modules designed on miniature PCBs, which are much easier to mount in a microphone housing.

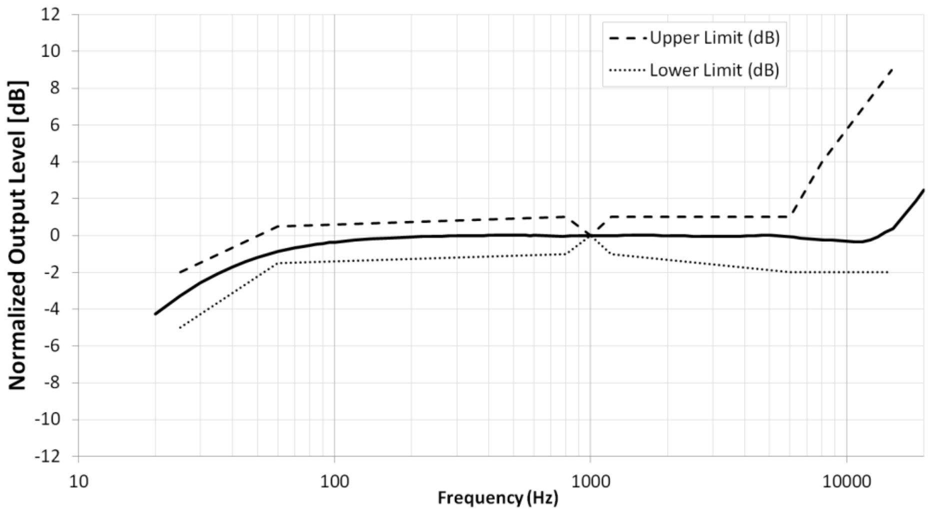

Figure 2. Frequency response of the MEMS microphone capsule [3].

2.3. Housing

MEMS Infineon IM69D130 and electret JLI-2590A microphone capsules have different dimensions and wiring. It was decided to design the microphone housings from scratch to ensure the same measurement conditions. Microphone capsules were placed in a tetrahedral orientation, and a 3D design was prepared in Autodesk Fusion 360 software [4]. Both microphone designs have the same shape and dimensions, and only the capsule mounts differ. In Fig. 3, the 3D models of the central part of the housing, which holds the microphone capsules, are presented.

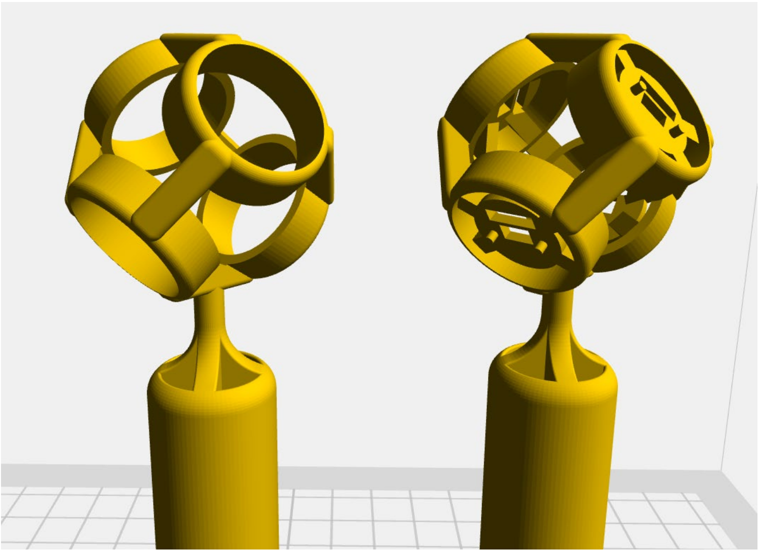

Figure 3. 3D models of the microphone housings for electret (left) and MEMS (right) microphone capsules.

Housings were printed using 3D SLA technology, which enables high-resolution prints with spatial complexity. The printer model used was Formlabs Forms 3+, and Clear V4 was used as a resin. It is characterized by high print quality and partial transparency.

3. Measurements

Sound can be measured by many different parameters, but the final impression of sound quality depends on the subjective conditions of the individual listener. Because of that reason, measurements have been divided into objective and subjective parts.

3.1. Objective measurements

The measurements were conducted in the anechoic chamber at the Warsaw University of Technology. Microphones were placed on a turntable at a 1.5 m distance from the arc-shaped arm with 10 speakers attached to it at every 15º in the elevation plane, covering a range of angles from -45º to 90º. The turntable was rotating in the range of 0º – 360º in a horizontal plane with an interval of 15º. The system setup is presented in Fig. 4. Reflections from a turntable were minimized using sound-absorbing materials placed at the table.

A personal computer with a Focusrite Scarlett 18i20 3rd Gen external audio interface and a B&K Type 2735 amplifier was used to generate the needed audio signals. The analog microphones (with electret capsules) were connected directly to the audio interface via XLR connectors, while the digital microphones (with MEMS capsules) were additionally connected via a miniDSP MCHStreamer [5] interface supporting the PDM digital standard.

A sweep signal was played individually from each speaker for each azimuth angle, giving 240 measurement points spread over the sphere surrounding the microphone. Two custom-made ambisonic microphones and one reference microphone (Sennheiser AMBEO VR Mic) were measured. The recorded signals were then processed to obtain impulse responses for each capsule in each microphone using the AFMG Easera software [6]. Then, all files were converted from ambisonics A format to ambisonics B format using the Reaper digital audio workstation [7]. The conversion was made using the Sparta Array2sh plug-in [8]. Further processing was held using the Matlab software, i.e., impulse responses were converted to the frequency domain using the pwelch algorithm, and the directional characteristics for the W, Y, Z, and X channels were plotted for each frequency. All results are presented and analyzed in Section 4.

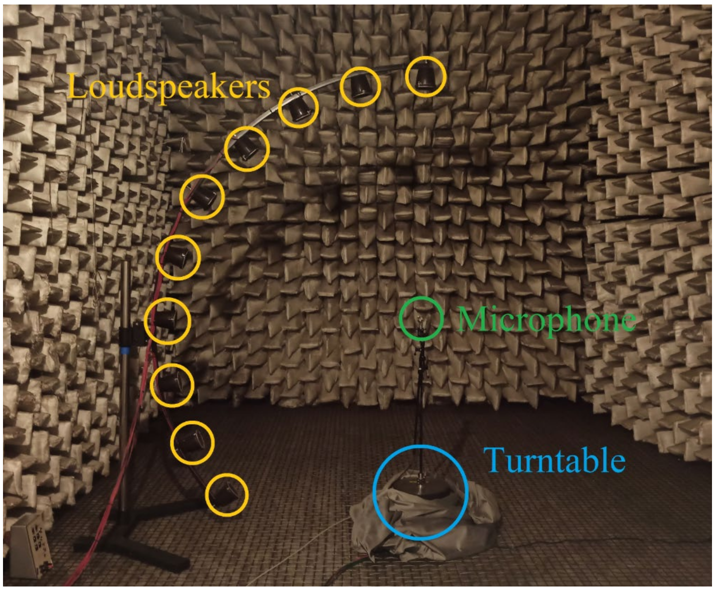

Figure 4. Measuring system setup in the anechoic chamber.

3.2. Subjective measurements

Listening tests were designed to compare recordings made with MEMS and electret ambisonic microphones. Two sets of recordings were prepared in diffuse field conditions using custom-made microphones (MEMS and electret) and a reference microphone (Sennheiser AMBEO VR Mic). The first group of recordings was made in real-world scenarios shown in Fig. 5: a location near the noisy street Fig. 5a, a city fountain Fig. 5b, a tunnel Fig. 5c, and the Main Hall of the Warsaw University of Technology Fig. 5d.

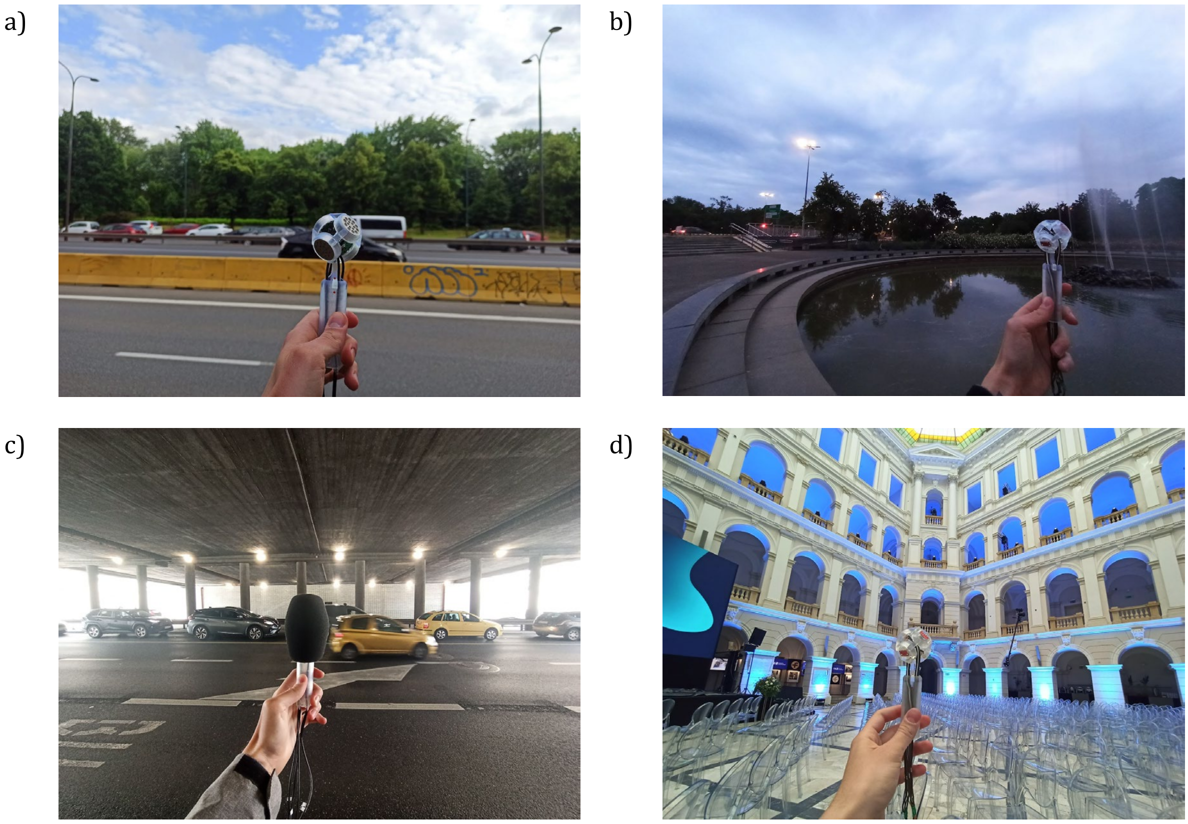

Figure 5. Real-world scenarios recorded: a) near a noisy street, b) near a city fountain, c) in a tunnel, d) in the Main Hall of Warsaw University of Technology.

The second group of recordings (music played on the piano and human speech recorded around microphones) were made in the sound studio shown in Fig. 6. Samples were recorded with Tascam Portacapture x8 portable audio recorder and miniDSP MCHStreamer audio interface supporting the PDM audio standard.

Figure 6. Sound studio recording setup.

4. Results and discussion

4.1. Objective results

The first step of the analysis was to compare the frequency response of single microphone capsules mounted in three ambisonic microphones. In Fig. 7, it can be seen that the characteristics of all 3 microphones are relatively flat in the 100Hz – 10kHz range. These characteristics are achieved from the front-left-upper capsule and were obtained for 45 degrees of azimuth angle and 30 degrees of elevation angle. This setting is one of several that allow the speaker to be positioned almost perfectly in front of the microphone capsule. The fixed angular distances between the loudspeakers cause an inaccuracy of about 5 degrees in the elevation plane. However, given the cardioid and omnidirectional characteristics of the capsules, this does not significantly affect the shape of the frequency response. There is some ripples in the frequency responses, which could be related to the capsule's mounting and resonance from an arc-shaped arm with loudspeakers. It can be seen that MEMS capsules have a slightly more comprehensive frequency range near 10kHz and that the gain was lower because of different system setups.

The second part of the analysis examined the microphones' directivity. The directional characteristics of the B-format ambisonics channels were analyzed in the horizontal and vertical planes. The analysis shows that the MEMS microphone's characteristics are significantly closer to the ideal characteristics than those of the other two microphones (with electret capsules). This can be easily observed in Fig. 8, where the characteristics of all 4 channels, W, Y, Z, and X, are presented. The W, Y, and X channels are presented in the horizontal plane in Figs. 8a, 8b, 8d, while the Z channel is presented in the vertical plane in Fig. 8c. The speaker arm (shown in Fig. 4) did not cover a whole range of elevation angles, so the Z channel graph is shown entirely from the top and half from the bottom. The Ambeo and Electret custom-made ambisonic microphones still perform satisfactorily, and one can quickly identify the corresponding B-format ambisonics channels by the shape of their directional characteristics. Deviations from the ideal shape may be due to reflections and resonances created during measurements (speaker metal arm, the rotary table microphone holder, and part of the measurement apparatus located in the corner of the anechoic chamber). All frequency responses were equalized with the speaker's characteristics measured with B&K 4189 type reference microphone.

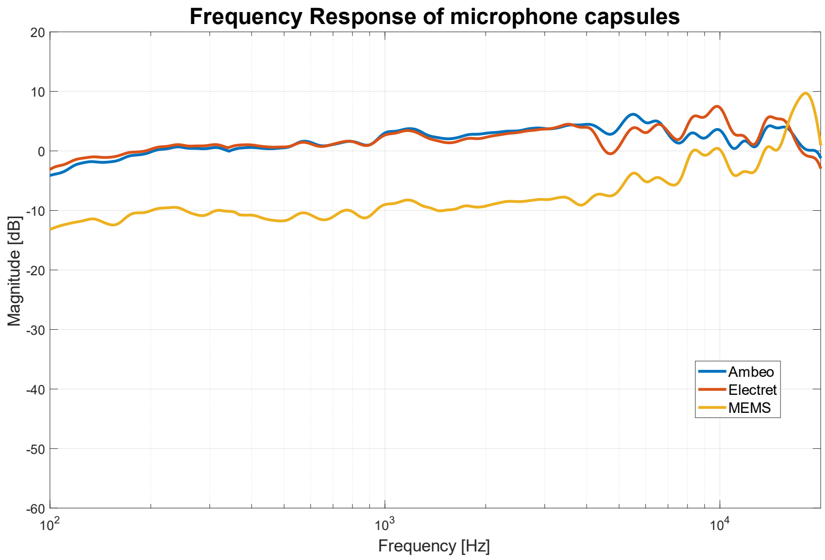

Figure 7. Frequency response of single capsules mounted in ambisonic microphones.

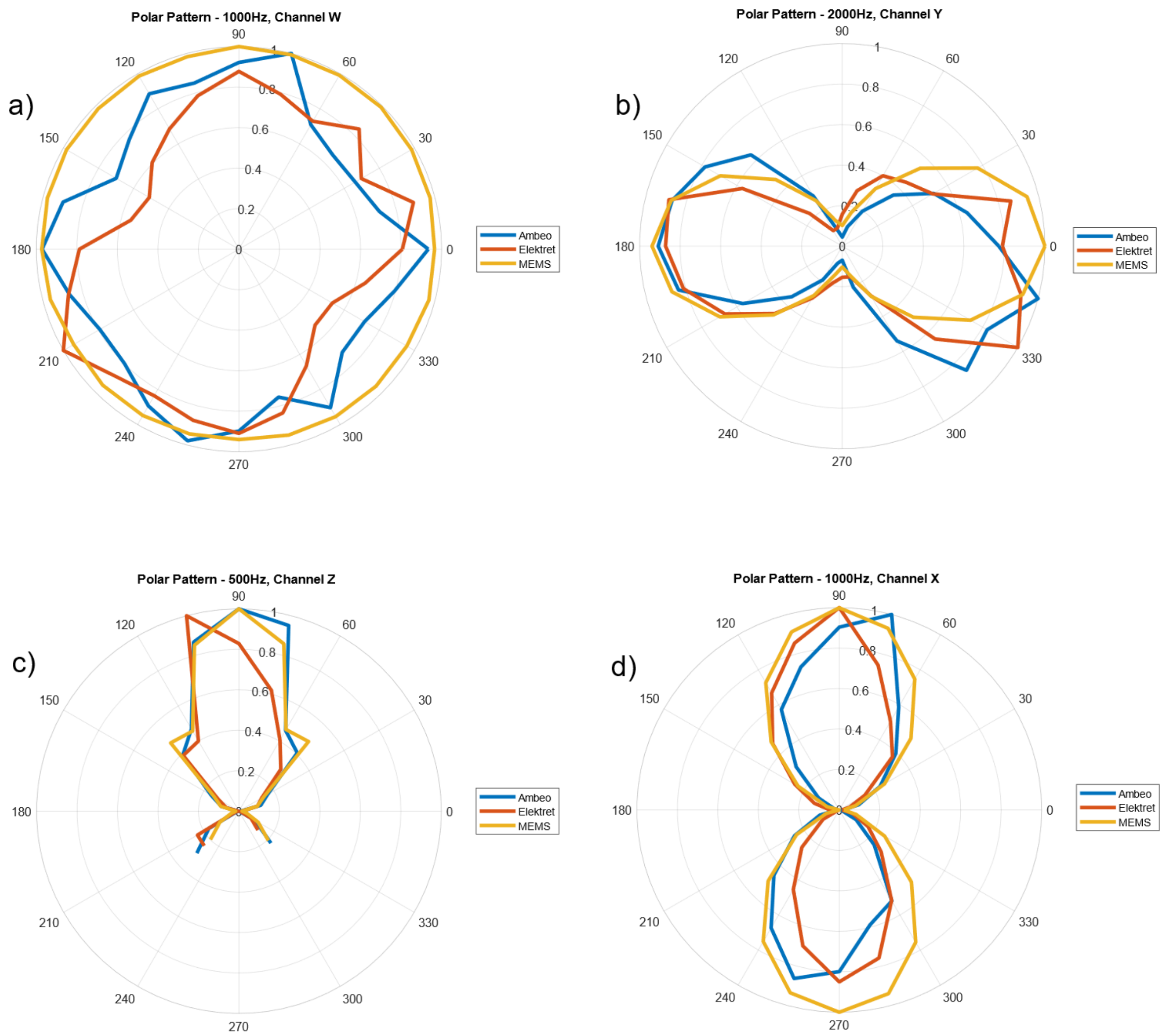

Figure 8. Directivity patterns of ambisonics B-format channels for three ambisonic microphones.

4.2. Subjective results

To investigate the subjective timbral sound quality and spatial properties of recordings described in section 3.2, a modified MUSHRA test and preference test were conducted. MUSHRA listening test were prepared as a web application based on webMUSHRA [9]. Participants didn't have any previous experience in listening tests, and all of them used different pairs of headphones, so preference tests were preceded by the training phase and headphones screening proposed by Milne et al. in [10]. Test stimuli for the MUSHRA test consisted of binaural B-format audio samples of different recordings described in section 3.2 decoded with SPARTA suite [11], reference file and two anchors: lowpass filtered reference file at 3.5 kHz (named anchor35) and W channel decoded from the reference microphone (named ambeo_anchor). A listening test was conducted with 42 participants, all of whom self-identified as having normal hearing. Results of the MUSHRA test presented in Fig. 9 for all subjects show that most of them had problems correctly identifying the W channel anchor and that the custom-made electret microphone was rated high compared to the reference. These results allowed to exclude from the aggregated responses subjects who rated reference samples (recordings made with Sennheiser Ambeo VR microphone) for > 15% of the test items lower than a score of 90 and who rated anchors for more than 15% of the test items higher than a score of 90 [12].

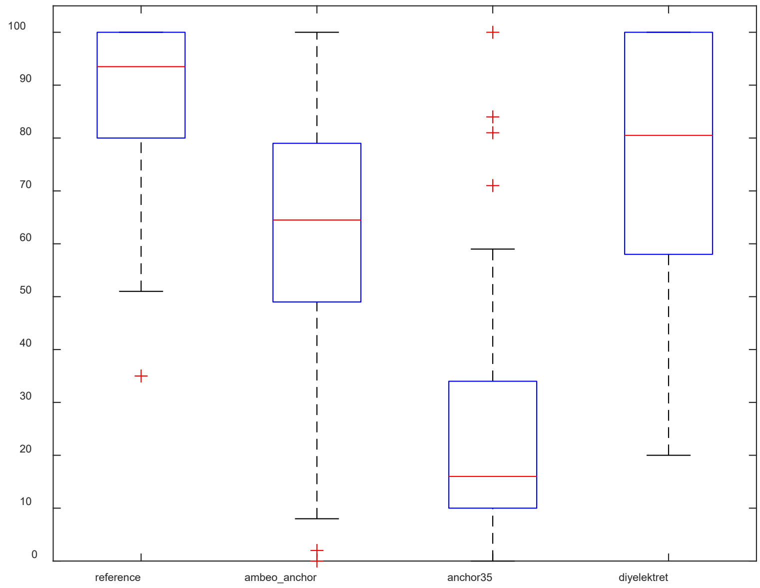

Figure 9. Results of the MUSHRA listening test for all subjects.

The post-screening process excluded 29 of 42 subjects, so the main preference test results were analyzed with 13 participants. Test stimuli consisted of the same binaural B-format audio samples of different recordings described in section 3.2 used in the previous MUSHRA test. In each test, two audio samples were presented (recorded with a custom-made electret and custom-made mems microphone). Every sample pair was presented twice randomly to exclude subjects who rated samples differently. Results of the preference test are shown in Fig. 10. Electret type microphone recordings were chosen 72 times out of 120 total answers. The calculated p-value from the binomial test was 0.00662 (lower than the critical level of 5%). The probability that listeners were guessing was below 1%, meaning there was a significant difference in sound quality between these two types of microphones.

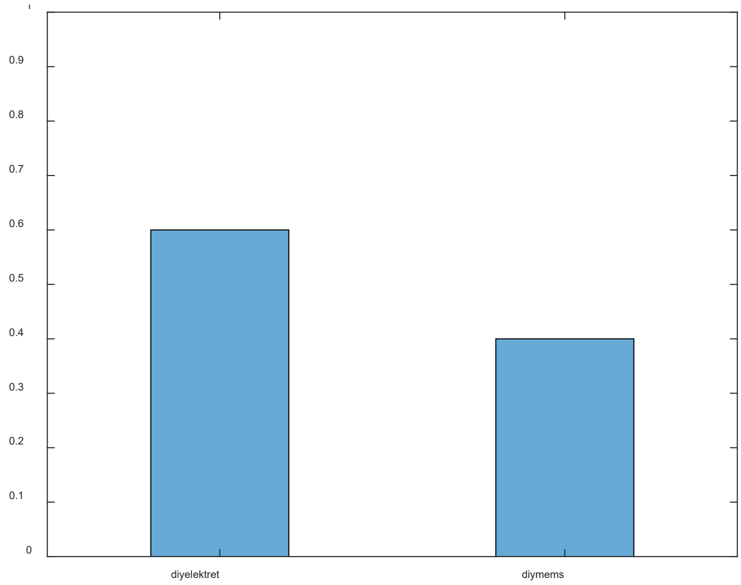

Figure 10. Results of the preference listening test for 13 subjects.

5. Conclusions

The paper compares two FOA microphones designed with the same housing but different types of microphone capsules – electret JLI-2590A and MEMS Infineon IM69D130. The frequency response measured for a single microphone capsules in an anechoic chamber was close to those shown in the datasheets. Differences were caused probably by housing design, materials and proximity to other capsules mounted in the FOA microphone. Measurement results of directivity patterns for W, Y, Z, and X channels for selected frequencies shown in Fig. 8 are slightly distorted probably by resonances and reflections present near capsules. Frequency responses of all types of microphone capsules were equalized with reference measurement made with B&K 4189 microphone. Figure 8 reveals that MEMS type microphones have more regular and symmetric characteristics than electret type microphones.

The main part of the comparison between the two microphones is to find if there is an audible difference between recordings made with microphones and which are eventually rated better. Test stimuli for listening experiments, e.g., sound-field recordings made with a microphone with electret capsules, were recorded using external preamplifiers and ADCs in the Focusrite 18i20 audio interface. Recordings made with a microphone with MEMS capsules were recorded using miniDSP MCHstreamer entirely in the digital domain. Audio samples were recorded simultaneously using two microphones at every location and normalized to -20 LUFS. As listeners were not experts, the MUSHRA test was conducted first to exclude subjects that rated anchors too high or reference too low. Post-screening of subjects was performed in accordance with Recommendation ITU-R BS.1534-3 [12]. Then, a preference test was performed to find if there was an audible difference between MEMS and electret type of microphone. The preference test results indicate a significant difference in overall sound quality between recordings made with MEMS and electret microphones (72 out of 120 answers rated electret microphones higher), but listeners' comments revealed that MEMS microphone recordings have better perceived spatial quality. It was probably because MEMS Infineon IM69D130 are omnidirectional capsules and electret JLI-2590A are unidirectional capsules.

Ambisonic microphones are available from FOA (first-order ambisonics, e.g., Sennheiser AMBEO VR Mic) to HOA (up to fourth-order ambisonics, e.g., MH Acoustics Eigenmike) and are designed with different capsules, from MEMS transducers (e.g., Zylia ZM-1) and electret capsules to high-grade condenser microphones (e.g., NEVATON VR). The aim of this study was to indicate whether the microphone capsule's type affects the recorded sound field's overall quality (timbral and spatial quality). The preference tests indicated that there are significant differences in overall sound and spatial quality between digital and analog type of microphones which confirmed observations from audio engineers (especially regarding an overall sound quality between recordings made with Zylia ZM-1 and Sennheiser Ambeo microphones). Although similar experiments have been conducted for example in [13–15], there is no comparison of custom made ambisonics microphones with the same housing but different microphone capsules. Future work will focus on modifying the microphone's design to attenuate interferences between capsules and include localization tests combined with visual VR component.

Additional information

The authors declare: no competing financial interests and that all material taken from other sources (including their own published works) is clearly cited and that appropriate permits are obtained.

References

1. M.A. Gerzon; Periphony: With-height sound reproduction; J. Audio Eng. Soc., 1973, 21(1), 2–10
2. JLI Electronics; https://www.jlielectronics.com/content/JLI-2590A.pdf (accessed on: 2023-08-08)
3. Infineon; https://www.infineon.com/dgdl/Infineon-IM69D130-DataSheet-v01_00-EN.pdf?fileId=5546d462602a9dc801607a0e46511a2e (accessed on: 2023-08-08)
4. Fusion 360; https://www.autodesk.com/products/fusion-360/overview (accessed on: 2023-08-31)
5. miniDSP; https://www.minidsp.com/images/documents/Product%20Brief-MCHStreamer.pdf (accessed on: 2023-08-08)
6. AFMG EASERA; https://www.afmg.eu/en/afmg-easera (accessed on: 2023-08-31)
7. REAPER; https://www.reaper.fm/ (accessed on: 2023-08-31)
8. L. McCormack, S. Delikaris-Manias, A. Farina, D. Pinardi, V. Pulkki; Real-Time Conversion of Sensor Array Signals into Spherical Harmonic Signals with Applications to Spatially Localised Sub-Band Sound-Field Analysis; In: Proceedings of the 144th Audio Engineering Society International Convention 2018, Milan, Italy, 23–26 May 2018
9. M. Schoeffler et al.; webMUSHRA – A Comprehensive Framework for Web-based Listening Tests; J. Open Res. Softw., 2018, 6(1), 8; DOI: 10.5334/jors.187
10. A.E. Milne et al.; An online headphone screening test based on dichotic pitch; Behav. Res. Methods, 2021, 53(4), 1551–1562; DOI: 10.3758/s13428-020-01514-0.
11. L. McCormack, N. Meyer-Kahlen, A. Politis; Spatial Reconstruction-Based Rendering of Microphone Array Room Impulse Responses; J. Audio Eng. Soc., 2023, 71(5), 267–280
12. Recommendation ITU-R BS.1534-3; Method for the subjective assessment of intermediate quality level of audio systems; International Telecomunication Union, 2015
13. E. Bates, et al.; Comparing Ambisonic Microphones–Part 1; In: Proceedings of the Audio Engineering Society Conference on Sound Field Control, Guildford, UK, 18–20 July, 2016
14. E. Bates, et al.; Comparing Ambisonic Microphones–Part 2; In: Proceedings of the 142th Audio Engineering Society International Convention 2017, Berlin, Germany, 20–23 May 2017
15. G. Zalles, Y. Kamel, I. Anderson, M.Y. Lee, C. Neil, M. Henry, S. Cappiello, C. Mydlarz, M. Baglione, A. Roginska; A Low-Cost, High-Quality MEMS Ambisonic Microphone; In: Proceedings of the 143rd Audio Engineering Society International Convention 2017, New York, USA, 18–20 October 2017

© 2024 by the Authors. Licensee Poznan University of Technology (Poznan, Poland). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).