DPA4560 vs Meta Rayban: a binaural comparison

Adriano Farina
CIRI-EC
University of Bologna
Adriano.farina3@unibo.it

Lamberto Tronchin
Department of Architecture
University of Bologna
lamberto.tronchin@unibo.it

Angelo Farina
Department of Engineering for Industrial Systems and Technologies
University of Parma
angelo.farina@unipr.it

***Abstract*— Rayban | Meta is the second generation of smart glasses developed by Meta and Luxottica. They are one of the first mass-market all-in-one consumer devices allowing users to record and reproduce sounds binaurally. Traditionally, binaural recording systems use two microphones, one in each hearing canal, belonging either to a person or to a dummy head. In both cases, the incoming sound reflects on the body, shoulders, and ear pinnae, thus physically encoding several binaural cues.**

Rayban | Meta, instead, relies on a 5-microphone array, none of which enter the ear canal, and therefore is devoid of the information encoded by the pinnae. The binaural signal is obtained through a beamforming algorithm, about which nothing has been published in the literature.

For this reason, we evaluated the quality of the binaural signals through impulse response measurements. Wearing a pair of Rayban | Meta and a set of DPA4560 binaural microphones, we used the exponential sine sweep method, sampling every 10°. Using the Aurora plugins, we obtained values for IACC (Inter-Aural Cross-Correlation), ITD (Interaural Time Difference), and ILD (Interaural Level Difference).

As frequency response tests, especially regarding sound reproduction, are widely available, we focused on the binaural parameters only.

Keywords—acoustic measurements, signal processing, binaural response.

I. Introduction

The Rayban | Meta is one of the most popular wearable recording devices on the market. They produce binaural audio using a scarcely documented microphone array.

II. Measurement procedure

A test subject was fitted with a DPA4560 binaural kit and with a pair of Meta Rayban glasses, making sure that there was no mechanical interference between the two devices. The Meta Rayban recorded a series of videos on its internal memory, making sure not to split measurements between different takes. Multiple takes were necessary, as videos have a maximum length of 3 minutes. Surprisingly, despite being recorded as videos, the audio files had a 44.1kHz sample rate. Even more surprisingly, according to their metadata, they were stored in 32-bit. For ease of processing, they were converted to 48kHz using the "convert sample type" function in Adobe Audition.

The audio from the DPA4560 kit was recorded using an Android smartphone running USB Audio Tool Pro as a 48kHz 24-bit WAV file.

The subject sat on an office chair fitted with an angular measurement device, obtained by fixing a caliper to a circular sheet of paper, on which 10° increments had been marked. While the sheet of paper did flex vertically, it remained unperturbed horizontally.

The vertical alignment of the subject's head was confirmed visually, using an observer (the thesis supervisor) in a fixed position and a vertical segment of the wall as reference.

Fig. 1. Test subject wearing both the Meta Rayban and the DPA4560

The test signal was an Exponential Sine Sweep signal (20-20000Hz, 25s, with 5s of silence between repetitions) played at 64dB(A) from a Bedrock BTB115 Advanced Talkbox, positioned at 1.5m of height, and 1m of horizontal distance from the centre of the segment connecting the ears of the test subject.

Fig. 2. Base of the chair with angle markings.

We obtained 36 binaural impulse responses for each recording system. These were convolved with the appropriate inverse sweep processed with the Acoustical Parameters

979-8-3315-5828-4/25/$31.00 ©2025 IEEE

plugin from the Aurora suite. Of particular interest were the binaural parameters: IACC, ITD, and ILD.

III. Analysis of the Results

The main difficulty of this study was the attempt to evaluate the differences between the two systems without resorting to user testing, which is the most common method in the literature [1]. The chosen parameters are closely related to externalisation and localisation; they rely on the same mechanisms employed by the brain, but they are still not direct measurements of human sensations. We compared proxies for the qualia, not the qualia themselves.

A. IACC

In the following figure, we can observe the polar plot for IACC values at different octaves with the two recording systems. We can observe that at lower frequencies, both systems are *approximately* omnidirectional, while at higher frequencies, they start beamforming more and more. The DPA4560 seems to be somewhat more regular in angular terms, with the Rayban presenting slightly backwards-facing lobes.

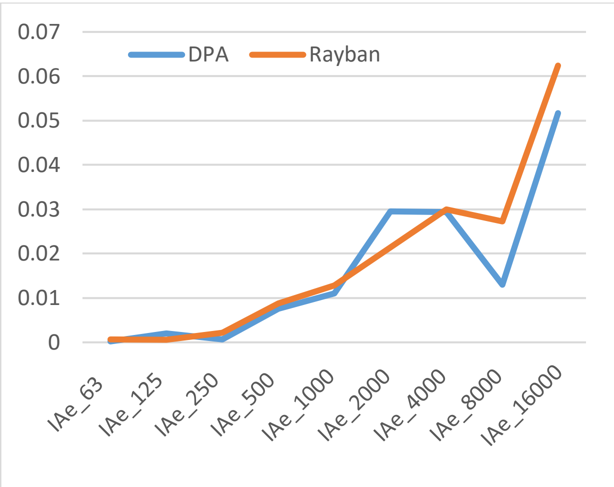

Fig. 3. IACC Frequency Variance.

In Figures 4 and 5, we see plots with IACC values on the y-axis and octaves on the x-axis. We can observe that in both cases the values tend to 1 under 250Hz, and they rapidly diverge at increasing frequencies. We also plotted the variance of the two systems.

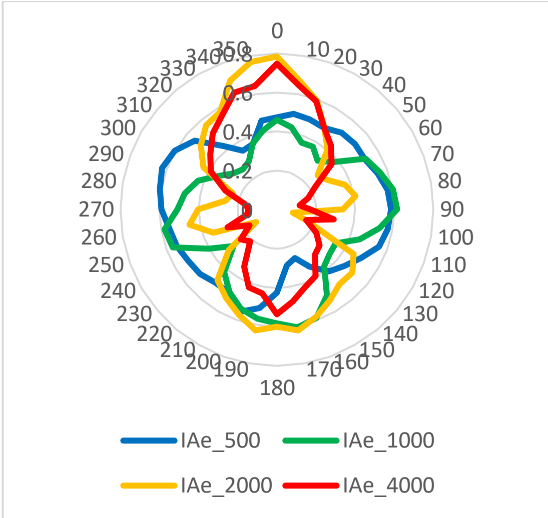

Fig. 4. IACCe DPA4560 polar pattern.

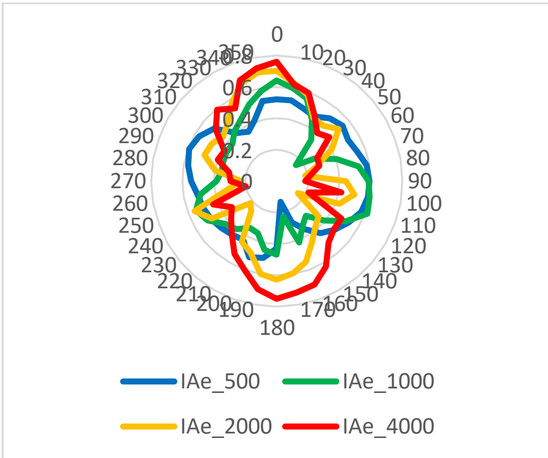

Fig. 5. IACCe Meta Rayban polar patterns.

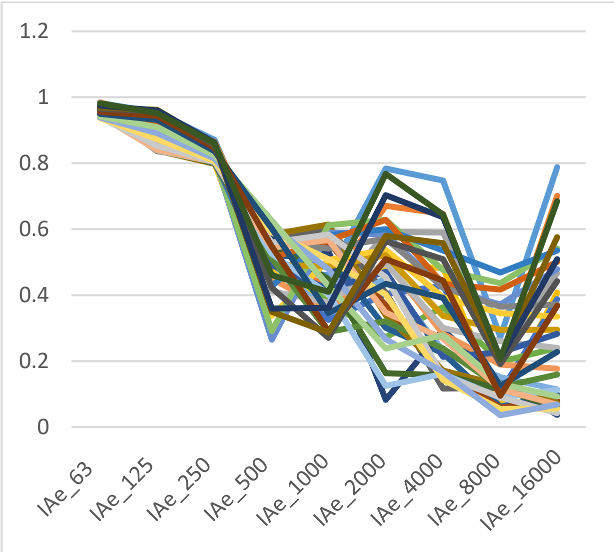

Fig. 6. IACCe DPA4560 in the frequency domain.

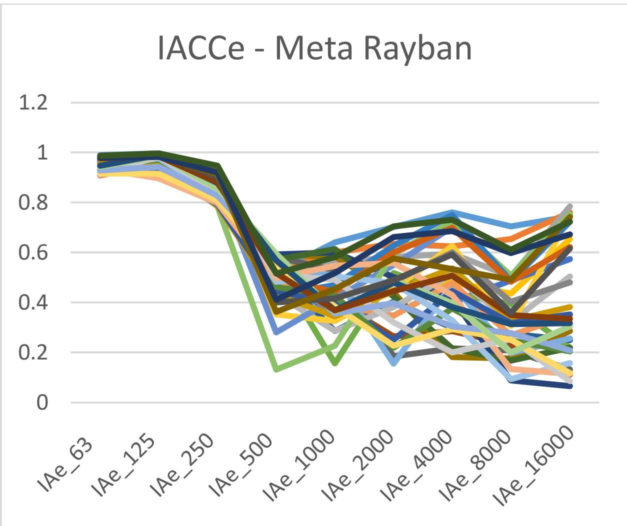

Fig. 7. IACCe Meta Rayban in the frequency domain.

B. ITD

In the figures below, we see a comparison between the τ_{IACC} Values of the two systems. This is fundamentally the same thing as ITD, but for a clear illustration of the definition of τ_{IACC}, see Figure 5 from [2]. Both systems follow the expected sinusoidal profile.

Just like for IACC, ITD presents significant variance on a frequency basis, but very little coherent information can be gleaned by analysing variance on an angular basis.

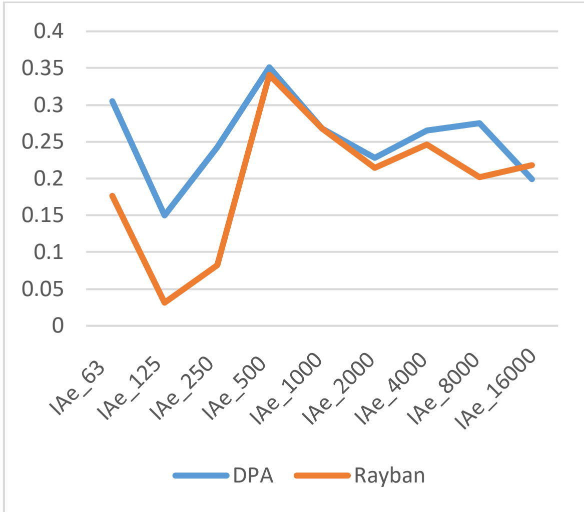

Fig. 8. ITD Variance in the frequency domain.

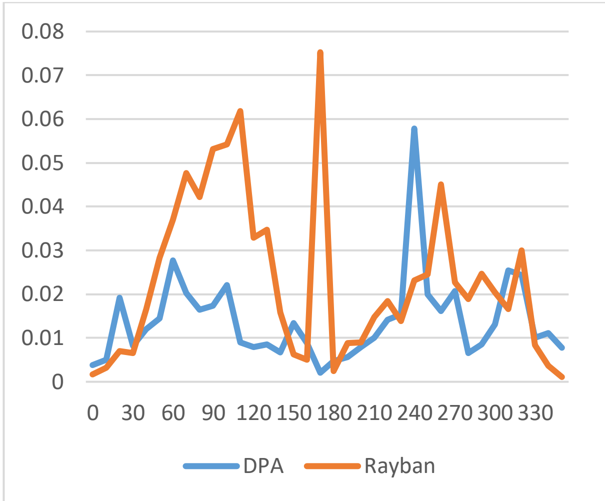

Fig. 9. ITD comparison by angle.

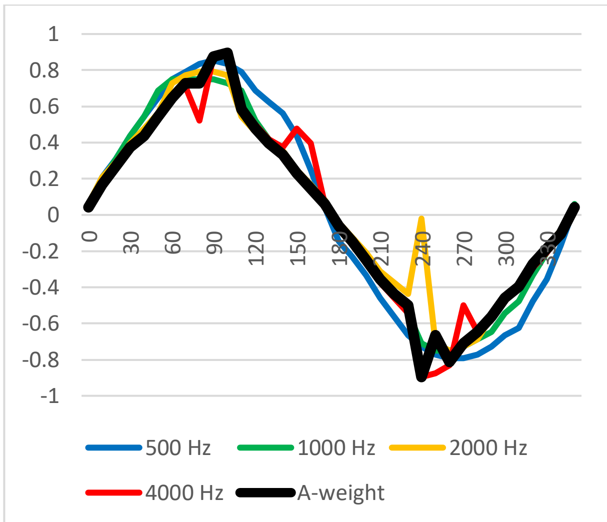

Fig. 10. ITD values: Tau (ms) DPA4560.

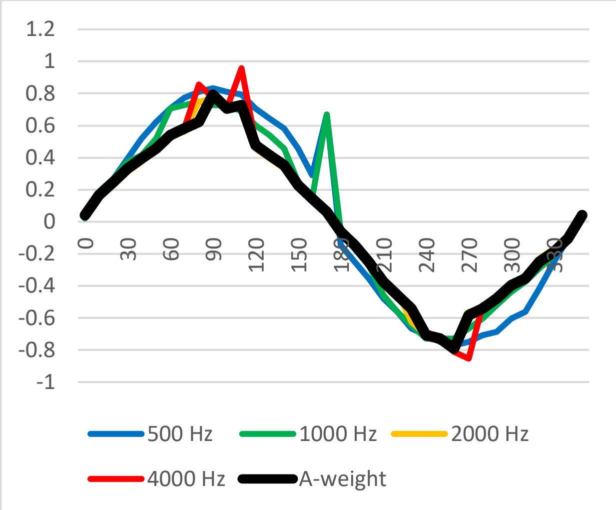

Fig. 11. ITD values: Tau (ms) Meta Rayban.

C. ILD

This is the parameter where the difference between the two systems is most evident. In particular, the Rayban system presents a noticeable flattening of the curve between 230° and 310°, compared to the DPA, and the geometrically expected sinusoid-like shape. This asymmetry in the results is a serious flaw, and it should be verified more thoroughly, as it could be the result of a measurement error, however unlikely.

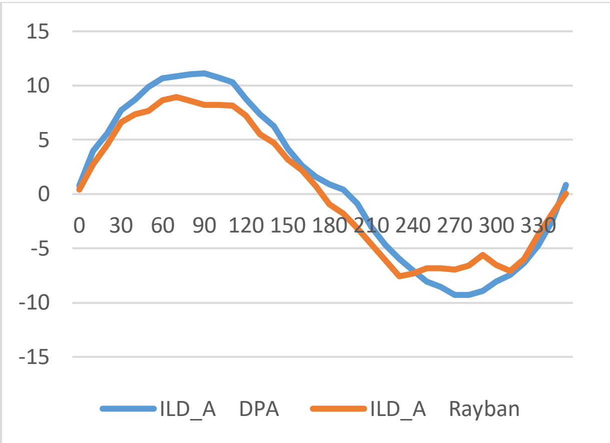

Fig. 12. Comparison of A-weighted ILD values, by angle of incoming sound

The Meta Rayban has consistently lower values of ILD, at every angle, but especially when the sound source is perpendicular to the subject.

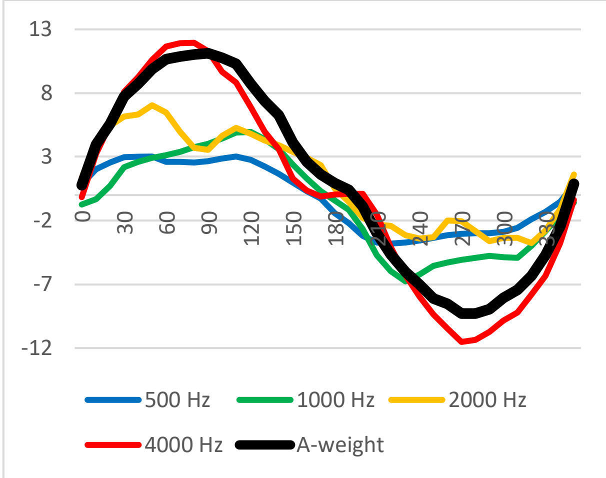

Fig. 13. ILD Angular values of DPA4560.

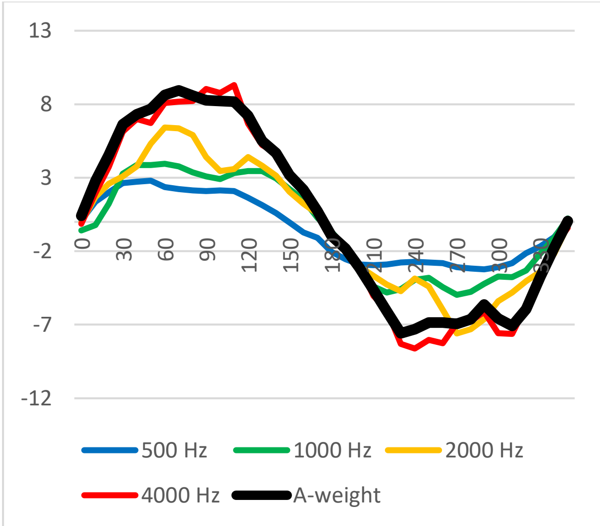

Fig. 14. ILD Angular values of Meta Rayban.

IV. Conclusions

Both systems can capture convincing binaural sound scenes, as evaluated through purely quantitative analysis.

It is plain to see that the processed parameters have some amount of variance, which is quite consistent with the existing literature, but which could be statistically mitigated if each measurement were repeated several times.

Several studies on the relationship between these quantitative measurements and psychoacoustic effects of source localization are available in the scientific literature. The comparatively lower ILD in the Rayban recording points to a slightly worse externalisation and localisation performance compared to the DPA.

However, the Meta Rayban is a playback as well as a recording system, and it would be interesting to perform further experiments, comparing their performance as binaural reproduction devices with a pair of reference headphones. Plausibly, some of the localisation cues that the Rayban does not record could be provided by the listener's very head, as they are physically present both during the recording and during playback. To test this, a dummy head should be used. The same test signal should be recorded on the dummy head, on the Meta Rayban positioned on the dummy head, and on the Rayban positioned on the subject's head. It should also be recorded on the DPA microphones, positioned both on the human subject and on the dummy head. In this way, a full separation will be achieved between the effects of the recording system and the effects of the HRTF.

Most importantly, further work is needed to validate the results by testing on actual human subjects. In particular, it would be important to evaluate the localization and externalization effects achieved by the two systems, using the techniques described in [3].

ACKNOWLEDGMENT

This research has been granted by the Emilia-Romagna European Regional Development Fund (ERDF) Programme 2021-2027 under AGORA project (un Atelier sonoro per il GiocO, la Ricerca, l'Acustica – progetto innovativo per l'inclusione delle persone vulnerabili tramite il suono immersivo 3D delle piazze), CUP E97G22000580007.

REFERENCES

[1] K. DiPasquale, "Comparing Externalization Between the Neumann KU100 Versus Low Cost DIY Binaural Dummy Head," *Journal of the Audio Engineering Society,* 2019.

[2] D. D'Orazio, P. Guidorzi and M. Garai, "A Matlab Toolbox for the analysis of Ando's factors," in *AES 126th Convention*, Munich, 2009.

[3] J. Cubick, C. Sánchez Rodríguez, W. Song, and E. N. MacDonald, "Comparison of binaural microphones for externalization of sounds," in *Proceedings of International Conference on Spatial Audio*, Graz, Austria, 2015.

[4] Bedrock, "Bedrock BTB115 Advanced TalkBox," 2024. [Online]. Available: https://bedrock-audio.com/index.php/bedrock-btb115.

[5] eXtream Software Development, "USB Audio Recorder PRO," 2025. [Online]. Available: https://play.google.com/store/apps/details?id=com.extreamsd.usbaudiorecorderpro&hl=en-US.

[6] Dennis Pro Audio Microphones, "4560 Binaural Headset Microphone," 2016. [Online]. Available: https://www.dpamicrophones.com/microphones/headset-and-earset/4560?variant=24.

[7] Meta Corporation, "Ray-Ban | Meta Smart Glasses Collection: Now Available," 17 October 2023. [Online]. Available: https://www.meta.com/en-gb/blog/ray-ban-meta-smart-glasses-collection-pre-orders/.

[8] *ISO 3382-1:2009, Measurement of room acoustic parameters,* 2009.

[9] "MMA-A Digital Audio Interface - Flawless mobile audio," 07 09 2019. [Online]. Available: https://web.archive.org/web/20190907095155/https:/www.dpamicrophones.com/mma-a-digital-audio-interfaces.

[10] T. Nishino, N. Inoue, K. Takeda and F. Itakura, "Estimation of HRTFs on the horizontal plane using physical features," *Applied Acoustics,* vol. 68, no. 8, pp. 897-908, doi: 10.1016/j.apacoust.2006.12.010, 2007.

[11] V. Jayaram, I. Kemelmacher-Shlizerman, and S. M. Seitz, "HRTF Estimation in the Wild," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, New York, 2023.

[12] S. I. E. E. Limited, "Methods and systems for synthesising an HRTF". United States of America Patent 2025/0099854 A1, 27 03 2025.

[13] T. Koumura, M. Nakatani, H. I. Liao and H. M. Kondo, "Dark, loud, and compact sounds induce frisson," *Quarterly Journal of Experimental Psychology,* vol. 74, no. 6, pp. 1140-1152, doi: 10.1177/1747021820977174, 2020.

[14] R. Wan, N. I. Durlach and H. S. Colburn, "Application of an extended equalization-cancellation model to speech intelligibility with spatially distributed maskers," *The Journal of the Acoustical Society of America,* vol. 128, no. 6, pp. 3678–3690, doi: 10.1121/1.3502458, 2010.

[15] M. A. Akeroyd, "An overview of the major phenomena of the localization of sound sources by normal-hearing, hearing-impaired, and aided listeners," *Trends in Hearing,* no. 18, p. doi: 10.1177/2331216514560442, 2014.

[16] D. Wang and G. J. Brown, "Binaural Sound Localization," in *Computational Auditory Scene Analysis*, John Wiley & Sons, 2005, p. chapter 5.

[17] M. Zaunschirm, M. Frank and F. Zotter, "Binaural Rendering with Measured Room Responses: First-Order Ambisonic Microphone vs. Dummy Head," *Applied Sciences,* p. doi: 10.3390/app10051631, 2020.

[18] I. B. Witew, A. Lindau, J. V. Schuitmann, M. Vorländer, S. Weinzierl, and D. D. Vries, "Uncertainties of IACC related to dummy head orientation," *DAGA,* 2010.

[19] L. Tronchin, A. Farina, T. Hotehama, R. Shimokura and V. Tarabusi, "The calculation of Binaural Parameters by means of HRTFs and Ando's method," *Proceedings of ICA2004 - Acoustical Science and Technology for Quality of Life,* p. doi: 11585/18596, 2004.

[20] T. P. Franken, P. X. Joris and P. H. Smith, "Principal cells of the brainstem's interaural sound level detector are temporal differentiators rather than integrators," *eLife,* p. doi: 10.7554/eLife.33854, 2018.

[21] T. Leclère, M. Lavandier and J. F. Culling, "Speech intelligibility prediction in reverberation: Towards an integrated model of speech transmission, spatial unmasking, and binaural de-reverberation," *Journal of the Acoustical Society of America,* pp. 3335-3345, doi: 10.1121/1.4921028, 2015.

[22] W. L. Martens and D. Stepanavicius, "Interaural Cross-Correlation Affects Perceived Elevation of Auditory Images Presented via Height Channels in Multichannel Audio Reproduction Systems," in *The 2016 Joint Conference of The Australian Acoustical Society and The Acoustical Society of New Zealand*, Brisbane, Australia, 2016.

[23] N. Gößling, D. Marquardt and S. Doclo, "Perceptual Evaluation of Binaural MVDR-Based Algorithm to Preserve the Interaural Coherence of Diffuse Noise Fields," *Trends Hearing,* 2020.

[24] A. Farina, P. Martignon, A. Capra, and S. Fontana, "Measuring impulse responses containing complete spatial information," in *4th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan*, Honolulu, 2006.

[25] A. Farina and R. Ayalon, "Recording concert hall acoustics for posterity," *24th AES Conference on Multichannel Audio,* pp. http://pcfarina.eng.unipr.it/Public/Papers/185-AES24.PDF, 26-28 June 2003.

[26] D. Thery and B. F. G. Katz, "Auditory perception stability evaluation comparing binaural and loudspeaker Ambisonic presentations of dynamic virtual concert auralizations," *Journal of the Acoustical Society of America,* pp. 246-258, doi:10.1121/10.0002942, 2021.

[27] P. Antoniuk, S. K. Zielinski and L. H, "Ensemble width estimation in HRTF-convolved binaural music recordings using an auditory model and a gradient-boosted decision trees regressor," *Journal of Audio Speech Music Processing,* pp. 53, doi:10.1186/s13636-024-00374-2, 2024.

[28] Y. Ando, Concert Hall Acoustics, Berlin: Springer, 1985, pp. doi: 10.1007/978-3-642-69810-1.

[29] L. Rayleigh, "On our perception of sound direction," *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science,* pp. 214-232, doi: 10.1080/14786440709463595, 1907.

[30] N. I. Durlach, "Binaural signal detection: Equalization and cancellation Theory", in *Foundations of Modern Auditory Theory*, New York, Academic Press, 1972, pp. 369-462.