CHALMERS
UNIVERSITY OF TECHNOLOGY

Spatial audio signal processing for augmented telepresence applications

Science Talks 13 (2025) 100421

Contents lists available at ScienceDirect

Science Talks

journal homepage: www.elsevier.es/sctalk

Spatial audio signal processing for augmented telepresence applications

Thomas Deppisch

Division of Applied Acoustics, Chalmers University of Technology, Gothenburg, Sweden

ARTICLE INFO

Keywords:
Augmented reality
Room impulse response
Spatial audio
Telepresence
Virtual acoustics

ABSTRACT

During the COVID-19 pandemic, the shift to remote communication, particularly through video calls, led to both opportunities and challenges. While initially a welcome alternative to in-person meetings, virtual gatherings became increasingly overwhelming, culminating in the term “zoom fatigue.” However, reduced travel highlighted the potential environmental benefits of online meetings. My PhD research focuses on improving the naturalness of remote communication to enhance the appeal of virtual meetings. Specifically, I develop signal processing techniques that preserve spatial and acoustic cues important for natural speech perception, such as the cocktail party effect. By modeling microphone array signals, particularly those integrated into smart glasses or augmented reality headsets, I estimate and apply spatial room transfer functions to create natural binaural audio experiences. My work also addresses challenges posed by head movement, using continuous-space domain estimation to update room transfer functions during head rotations. First results show the effectiveness of the method under controlled conditions. Future work will investigate the approach in more realistic scenarios.

Video to this article can be found online at https://doi.org/10.1016/j.sctalk.2025.100421.

E-mail address: thomas.deppisch@chalmers.se.

http://dx.doi.org/10.1016/j.sctalk.2025.100421
Received 9 September 2024; Accepted 12 January 2025

Figures and tables

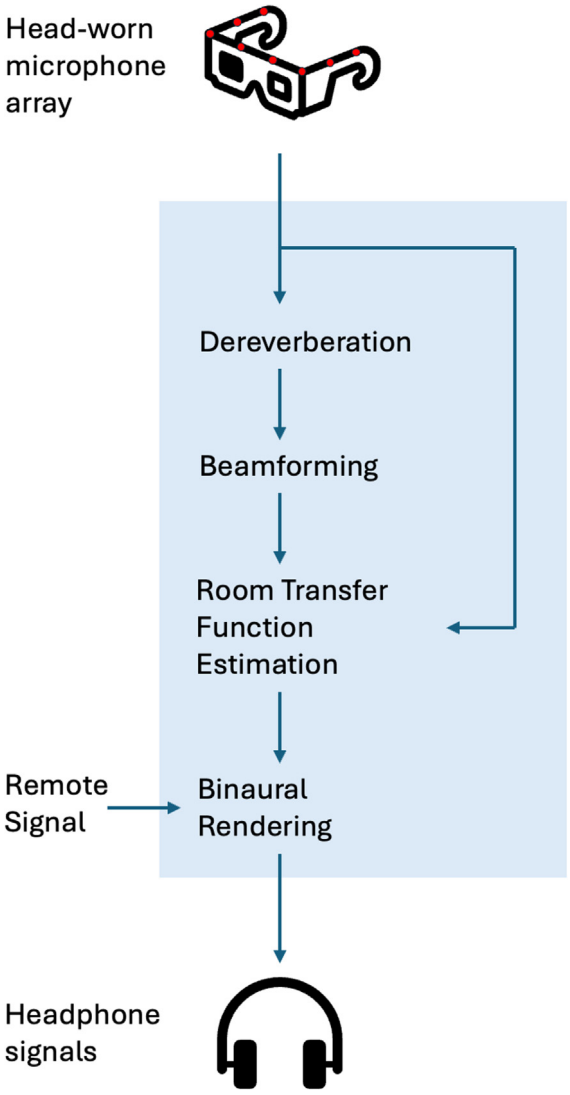

Fig. 1. Proposed approach for the blind estimation of spatial room transfer functions from head-worn microphone arrays [1,2]. Dereverberation and beamforming is used to obtain a pseudo reference signal which is related to the array signals to estimate the room transfer function. The spatial room transfer function is then applied to the remote signal to add spatio-spectral cues of the near-end environment. Binaural rendering creates ear signals that are played back via headphones.

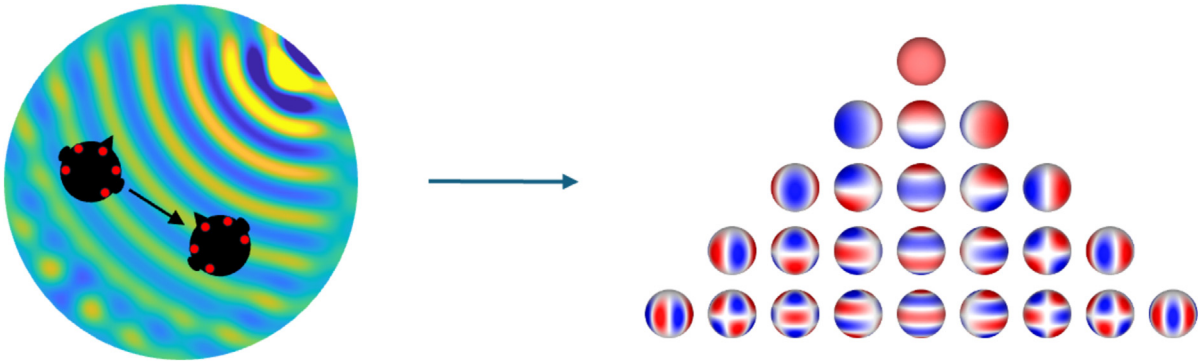

Fig. 2. Estimation during head rotation and movement. Head-movement poses an additional challenge to the estimation. Estimation in a space-continuous domain, expressed via circular or spherical harmonics basis functions, facilitates updating the transfer function estimate during arbitrary head rotations [3].

CRediT authorship contribution statement

Thomas Deppisch: Conceptualization, Methodology, Visualization, Writing – original draft.

Acknowledgments

We thank Reality Labs Research at Meta for funding this research.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] T. Deppisch, J. Ahrens, S.V. Amengual Garí, P. Calamia, Blind estimation of spatial room impulse responses using a pseudo reference signal, 2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), Seoul, Republic of Korea 2024, pp. 470–474, https://doi.org/10.1109/ICASSPW62465.2024.10626717.

[2] T. Deppisch, N. Meyer-Kahlen, S.V.A. Garí, Blind identification of binaural room impulse responses from smart glasses, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2024, https://doi.org/10.1109/TASLP.2024.3454964.

[3] T. Deppisch, J. Ahrens, S.V. Amengual Garí, P. Calamia, Spatial room impulse response identification from rotating equatorial microphone arrays, 32nd European Signal Processing Conference (EUSIPCO), Lyon, France 2024, pp. 116–120.