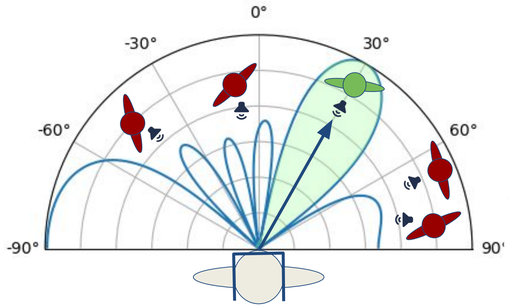
Viability of Eye-tracking Glasses for Beamforming Hearing-Aids

Emily Steiner
Department of Electrical Engineering, Stanford University

Motivation

- Current assistive hearing technology is limited in aiding selective attention in complex auditory 'cocktail party' environments
- Beamforming leverages time delay information between sensors in a phased array to spatially direct sound amplification
- We want to evaluate the **relative benefit of using eye-tracking versus head-movement** as a steering technique for the hardware limitations of a beamforming pair of glasses



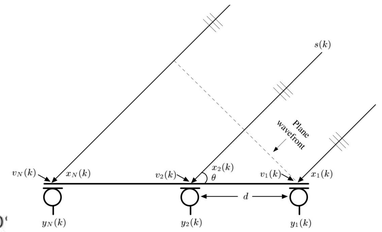


image from [4]

Related Work

Nuance Hearing [1]
- Collar microphone array

Practical utility of a head-mounted gaze-directed beamforming system [2]
- User study evaluating non-adaptive beamforming

A visually guided beamformer to aid listening in complex acoustic environments [3]
- 4x4 microphone grid, three beam selection to maintain binaural information

References

[1] Y. H. et al., "Directional hearing aid," U.S. Patent 0 567 888B2, Feb. 18, 2020.
[2] F. Culling, et al. "Practical utility of a head-mounted gaze-directed beamforming system," The Journal of the Acoustical Society of America, 2023
[3] T. R. Jennings and G. Kidd, "A visually guided beamformer to aid listening in complex acoustic environments," in Proceedings of Meetings on Acoustics, 2018.
[4] J. Benesty, J. Chen, and Y. Huang, Microphone array signal processing.Springer Science & Business Media, 2008, vol. 1.
[5] H. Van Trees, Optimum Array Processing. John Wiley Sons, Ltd, 2002.

Simulation & Implementation

Evaluation of Minimum Variance Distortionless Response (MVDR) beamforming (adaptive) in simulations of predicted use cases — evaluation of ability to resolve sources at small angles

Hardware Model

8 mic linear array
14cm total length

Anechoic Acoustic Model [4]

database of music and voice recordings

White Noise

angular phase delay

Broadband Beamforming [4][5]

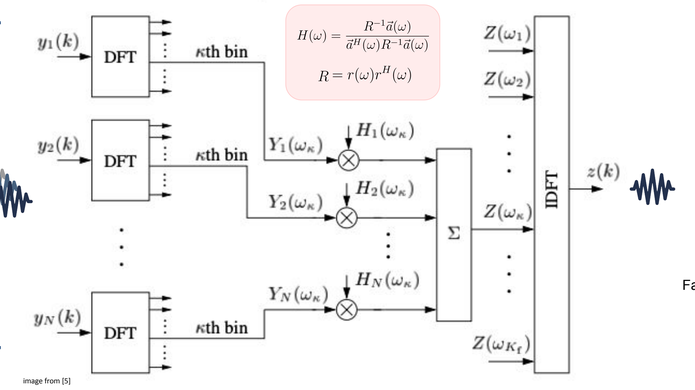
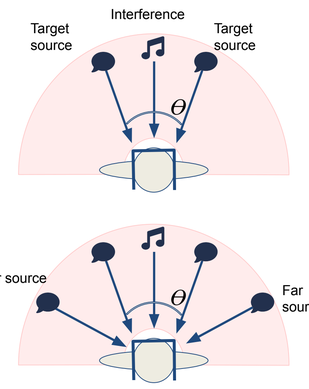


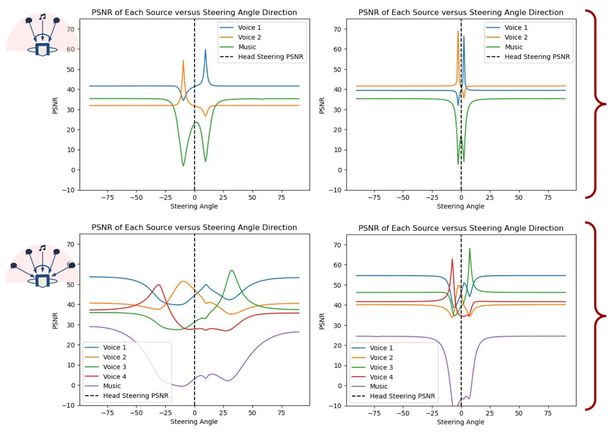
image from [5]

Scenarios



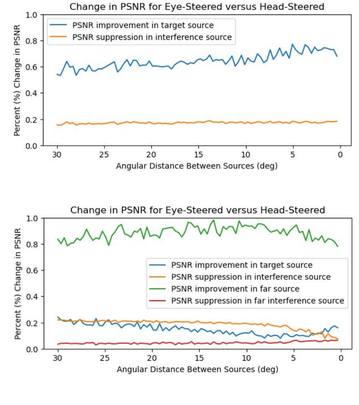
Experimental Results

20° between sources



5° between sources

Varying distance between sources



Discussion:
- High variation in measurements (dependant on audio clips used)
- Number of sources far outweighs distance between sources
- Qualitative output does not always match PSNR improvement

Limitations: simulation model, real-time time windowing, assumption that eye-tracking represents ground truth, far-field point sources, number of audio clips

Future Work
- Simulation (image-source model, time-windowing) & metrics (directivity index, word error rate)
- Prototype and user study for evaluation of ease and convenience
- Investigation into NN/probabilistic attention models, SSS