Multipath Enabled Private Audio with Noise

Anadi Chaman*, Yu-Jeh Liu*, Jonah Casebeer†, Ivan Dokmanić*

*Department of Electrical and Computer Engineering, †Department of Computer Science
University of Illinois at Urbana-Champaign, USA

Private Audio Communication

- Given L loudspeakers, the goal is to communicate distinct audio messages to K users in a room.
  - Users only receive their intended audio messages.
  - An eavesdropper at any other location is jammed.

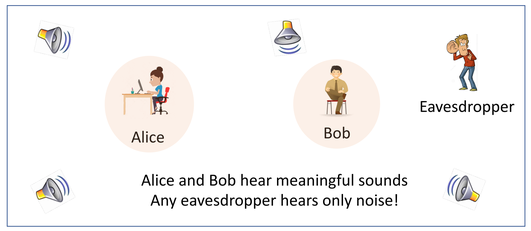

Fig. 1: Problem setup.

Conventional approach

- Traditional multi-zone sound field reproduction systems send linearly filtered signals to boost SNR in target zones.

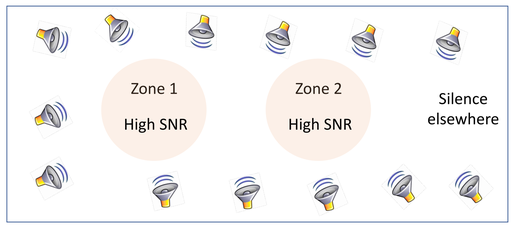

Fig. 2: Conventional approaches to personal audio zones

- Silencing sound everywhere else is hard: needs a large number of speakers.
- SNR could be more easily reduced by sending noise!

Key idea

- Loudspeakers emit appropriately structured noise streams that echo across the room and sum up to yield intelligible audio only at the target locations.
- Unlike traditional sound field reproduction systems, the performance of our approach is enhanced by the presence of echoes.

Our Approach

1. Multichannel convolutional synthesis by noise

- The i^{th} speaker emits x_i, a convolution of filter g_i with white Gaussian noise n_i.

$$x_i = n_i * g_i \qquad y_k = \sum_{i=1}^{L} (h_{ki} * x_i)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_K \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & \cdot & \cdot & H_{1L} \\ H_{21} & H_{22} & \cdot & \cdot & H_{2L} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ H_{K1} & H_{K2} & \cdot & \cdot & H_{KL} \end{bmatrix} \begin{bmatrix} N_1 & 0 & 0 & \cdot & 0 \\ 0 & N_2 & 0 & \cdot & \cdot \\ 0 & 0 & N_3 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & N_L \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ \cdot \\ \cdot \\ g_L \end{bmatrix}$$

y H N g

$$H_{ki} = \begin{bmatrix} h_{ki} & 0 & 0 & \cdot \\ 0 & h_{ki} & 0 & \cdot \\ 0 & 0 & h_{ki} & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot \end{bmatrix} \qquad N_i = \begin{bmatrix} n_i & 0 & 0 & \cdot \\ 0 & n_i & 0 & \cdot \\ 0 & 0 & n_i & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot \end{bmatrix}$$

$$y = Hx = HNg \qquad g = \underset{f}{\operatorname{argmin}} \, || \, y - HNf \, ||_2^2$$

- At non-user locations, RIRs are different ⇒ incorrect *descrambling* of x_i.
  - Inherent randomness in n_i then ensures unintelligibility.

2. Noise in the nullspace approach

- Noise signals projected on the nullspace of H are sent in addition to message carrying signals.

$$\begin{aligned} y &= Hx \\ &= H(s + w) = Hs \, , \end{aligned}$$

where $s = \underset{z}{\operatorname{argmin}} \, || \, y - Hz \, ||_2^2 \, , \;\; w = Proj_{Null(H)}(v), \;\; v \sim N(0, \sigma^2 I_{NK \times Nk})$.

- Projected noise sums to zero at the intended locations while continuing to mask the audio signals elsewhere.

Why are echoes useful?

- Echoes enhance the spatial diversity of acoustic channels.
  - This improves the conditioning of H ⇒ better signal reconstruction at the intended users.
  - High diversity in RIRs also ensures that noise is cancelled only at the desired locations.

References

[1] M. Poletti, “An investigation of 2-D multizone surround sound systems,” in Proc. 125th Audio Eng. Soc. (AES) Conv., San Francisco, CA, Oct. 2008.

[2] T. Betlehem, W. Zhang, M. Poletti, and T. D. Abhayapala, “Personal Sound Zones: Delivering interface-free audio to multiple listeners,” in IEEE Signal Process. Mag., vol. 32, pp. 81–91, 2015.

[3] Y. J. Wu and T. D Abhayapala, “Spatial multizone soundfield reproduction: Theory and design,” in IEEE Trans. Audio, Speech, Lang. Process., vol. 19, pp. 1711–1720, 2011.

Results

- Experiments done with 6 loudspeakers.
- Toeplitz structure of matrices exploited to efficiently solve all least squares design problems using FFT.

Numerical Simulations

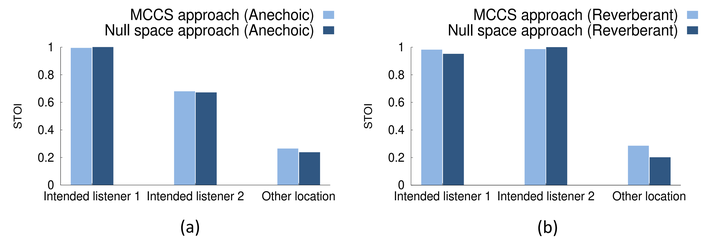

Fig. 3: STOI scores at 2 intended listeners and one additional location in (a) anechoic and (b) reverberant setting.

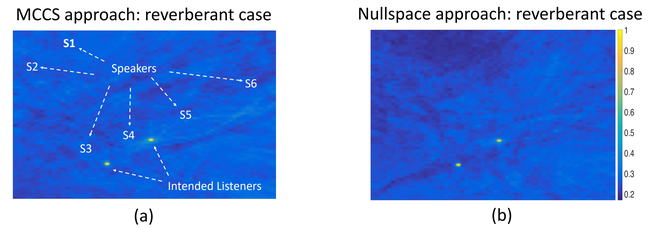

Fig. 4: Heat maps reflecting STOI scores at 4200 locations in a simulated room for (a) MCCS and (b) Null space approach.

Real world experiments

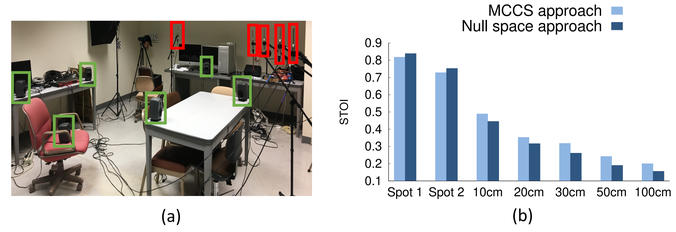

Fig. 5: (a) Experimental setup. (b) STOI scores vs distance from the focus spots.

Conclusion

- Our approaches involve emitting noise signals from loudspeakers and then utilizing echoes to ensure intelligibility only at selected locations.
- With merely 6 loudspeakers and a few impulse response measurements, we show that private audio communication is feasible.