Optimal Spatial Sampling of Plant Transfer Functions for Head-Tracking Personal Sound Zones

Yue Qiao* (presenter) & Edgar Choueiri
3D Audio and Applied Acoustics (3D3A) Lab
Princeton University

Presented at the 154th AES Convention
May 15, 2023

*E-mail: yqiao@princeton.edu
Personal Sound Zones\textsuperscript{[1]}

Audio Programs \quad Loudspeakers \quad Listeners

DZ

BZ

\textsuperscript{[1]} Druyvesteyn and Garas, JAES, 1997
The Pressure Matching (PM)\textsuperscript{[2]} Method

\[ g^* = \arg \min_g \| p_T - H \cdot g \|^2 \]

PSZ Setup

- PSZ Filters
- Target Pressure
- Plant Acoustic Transfer Functions (ATFs)

PSZ Playback

- Actual Pressure
- Actual ATFs

Mismatch

Difference

The Importance of Head Tracking in PSZ

Head movements $\rightarrow$ Mismatched plant ATFs $\rightarrow$ Degraded PSZ isolation

Especially true for **ear-targeting PSZ systems**

[3] Vindrola et al., JAES, 2020

[4] Qiao and Choueiri, AES Conv. 152, 2022
Solutions for Head-Tracker Reproduction

**Approaches**
- Dynamic loudspeaker beamforming\(^5\)
- Adaptive filtering (Filtered-x Least-Mean-Square\(^6\), Recursive Least Square)
- Filter cross-fading with plant spatial sampling\(^7\)

---

\(^5\) Qiao and Choueiri, AES Conv. 151, 2021

\(^6\) Vindrola et al., JASA, 2021

\(^7\) Lindfors et al., JAES, 2022
Challenges with implementing head-tracked PSZ systems...

<table>
<thead>
<tr>
<th></th>
<th>Single-listener Crosstalk Cancellation systems</th>
<th>Two-listener PSZ systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degrees-of-freedom in head movements</td>
<td>$6$</td>
<td>$6^2$</td>
</tr>
<tr>
<td>Number of plant ATF channels</td>
<td>$2 \times 2$</td>
<td>$4 \times N$</td>
</tr>
<tr>
<td>Performance Requirement$^8$</td>
<td>$\sim 20.7$ dB for envelopment</td>
<td>$\sim 25.6$ dB for non-distraction</td>
</tr>
</tbody>
</table>

The implementation of head-tracked PSZ seems practically impossible!

What is the minimum required spatial sampling resolution?
What are the rules for optimizing the spatial sampling process?

$^8$ Canter and Coleman, AES Conv. 150, 2021
Experimental system setup

- Loudspeaker array (200~7000Hz)
- B&K HATS dummy head
- Mechanical translation stage
- Binaural Microphones
Plant sampling resolution: $\Delta x = \Delta y = 1 cm$
Evaluation Procedure

Step I.
Plant spatial sampling

Step II.
PSZ filter generation

Step III.
Performance evaluation

System response

Performance Metrics

PM for better phase control & audio quality

\[ g^* = (\hat{H}^H\hat{H} + \sum_{m=1}^{M} \Sigma_m)^{-1}\hat{H}^Hp_T \]
Evaluation Metrics

Two aspects of isolation\cite{9}

Between BZ and DZ:

**Inter-Zone Isolation (IZI)**

\[ IZI_2 = \frac{||H_2g_2^*||^2}{||H_1g_1^*||^2} \]

Between target and interfering programs:

**Inter-Program Isolation (IPI)**

\[ IPI_2 = \frac{||H_2g_2^*||^2}{||H_1g_1^*||^2} \]

Only consider right listener being in BZ

For the left moving listener: IZI ~ moving DZ, IPI ~ moving BZ

\[9\] Qiao et al., JASA Express Lett., 2022
Results — Full Resolution

• X movements

Scattering from the right listener
• Y movements
Takeaways

• In the best case scenario, high isolation can be preserved over a large area

• Isolation is inherently lower at low frequencies due to room modes, setup limitations, etc.
Results — Spatial Downsampling

• X movements

IPI (moving $BZ$) is more robust than IZI (moving $DZ$)

3cm for both IZI and IPI
• Y movements

3cm for IZI and 5cm for IPI
Results — Shifting Positions

- X movements, shifting in Y

Filter robustness
Near — Front — Mid — Back — Far from the array
• Y movements, shifting in X

Far from A

Filter robustness

Near Listener B

Left

Mid

Right
What are the rules for optimizing the spatial sampling process?

- 2 Distances
  - the distance between two listeners $\downarrow$, sampling resolution $\uparrow$
  - the distance between the listener and the array $\downarrow$, sampling resolution $\uparrow$
- Temporal frequency: frequency $\uparrow$, sampling resolution $\uparrow$
- $BZ/DZ$: for moving DZ, sampling resolution $\uparrow$; for moving BZ, resolution $\downarrow$
The qualitative rules are generalizable to other PSZ systems of similar dimensions

The findings suggest a sampling of $BZ$ and $DZ$ at different resolutions

The findings are also insightful for implementing interpolation/adaptive filtering

The observed crossover frequency (1500 Hz) can be used for splitting approaches

Future work: investigating spatial sampling of head rotations
Optimal Spatial Sampling of Plant Transfer Functions for Head-Trackerd Personal Sound Zones

Yue Qiao* (presenter) & Edgar Choueiri
3D Audio and Applied Acoustics (3D3A) Lab
Princeton University

Presented at the 154th AES Convention
May 15, 2023

*E-mail: yqiao@princeton.edu