The Performance of A Personal Sound Zone System with Generic and Individualized Binaural Room Transfer Functions

Yue Qiao, Edgar Choueiri
The Performance of A Personal Sound Zone System with Generic and Individualized Binaural Room Transfer Functions

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

Presented at the 152nd AES Convention
May 17, 2022

*E-mail: yqiao@princeton.edu
Concept of PSZ\textsuperscript{[1]}
The Pressure Matching (PM) Method

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

PSZ Setup

PSZ Playback

Mismatched TFs → Degraded PSZs

Sources of TF Mismatches

- Head-Related Transfer Functions (HRTFs)
- Binaural Room Transfer Functions (BRTFs)
- Sound speed, Ambient temperature, background noise…
- Loudspeaker/Microphone Positions & Responses
- Number of Passengers, Seat Positions…

Please refer to the paper for a full list of references.
BRTF Mismatch in PSZ

Head-Related Transfer Functions (HRTFs)

Why not use actual listeners’ BRTFs in both phases?

Binaural Room Transfer Functions (BRTFs)

Setup Phase

Cho and Chang, ICA, 2019

Vindrola et al., JAES, 2020

Playback Phase

Ebri et al., AES Conv., 2020

Molés-Cases et al., JASA, 2022
Individualizing BRTFs

Benefits

- For binaural reproduction
  - Improves localization accuracy with headphone-based systems[3]
  - Improves performance for crosstalk cancellation systems[4]
- For PSZ systems?

Potential Barriers

- Time, cost, and equipment
- Additional measurements & signal processing
- Availability
- ...

Does individualization for PSZ make a difference?

If so, is the improvement worth the efforts?

Robust, but not too robust

Regularized PM Solution

\[ \tilde{g}^* = (H^H H + \beta I)^{-1} H^H p_T \]

Optimize regularization s.t. PSZ filters are robust to

- small head misalignments of a single listener
- BRTF mismatch between listeners
- larger head movements, torso movements, etc.

Robustness

Best Performance

N/A
Robust, but not too robust

A Probabilistic approach\cite{Moller2019}

\( H_{ml} = A_{ml} e^{i\phi_{ml}} \)

\( A_{ml} \sim N(\hat{A}_{ml}, \sigma_{A,ml}^2) \)

\( \phi_{ml} \sim N(\hat{\phi}_{ml}, \sigma_{\phi,ml}^2) \)

Modified Cost Function

\( J_{\text{prob}} = \mathbb{E}\{\|Hg - p_T\|^2\} \)

Optimal Solution

\[
\mathbf{g}^*_{\text{prob}} = (\hat{\mathbf{H}}^H \hat{\mathbf{H}} + \sum_{m=1}^{M} \Sigma_m)^{-1} \hat{\mathbf{H}}^H \mathbf{p}_T
\]

Variance matrix \( \Sigma_m \): empirically derived

Only account for head misalignments

\cite{Moller2019}
PSZ system for evaluation

Loudspeaker Array (200~7000Hz)
B&K HATS dummy head
Head Tracker (Infrared depth sensor)
Head Position Display

Allowable head movement range: 1 cm/10 deg
2 Listener Setups

Generic Setup

Individualized Setup

$P_{T,L}, P_{T,R}$

HATS

Human Listener

HATS
2 Sound Zone Configurations (SZCs)

Each SZC is evaluated with both generic and individualized filters.
2 Types of Measurements

- **BRTF Measurements**
  - Prior to filter evaluation
  - Used for deriving TF variance & generating PSZ filters

- **System TF Measurements (BRTFs*PSZ Filters)**
  - After filter generated & loaded
  - *In situ*: minimal head movements
  - *Ex situ*: additional head misalignments

- Exponential sine sweeps used for TF measurements
Performance Evaluation Metrics

• Acoustic Contrast (AC)\[^6\]
  - Quantifies the isolation level between two PSZs
  \[
  AC = \frac{g^HH_B^H H_B g}{g^H H_D^H H_D g}
  \]

• Robustness against head misalignments
  - Variation in AC level from multiple \textit{ex situ} measurements

• Other metrics (e.g., Array Effort, Reproduction Error) not adopted

Results - BRTF Differences

- Difference between the HATS and human listener’s BRTFs
- Difference increases with frequency

<table>
<thead>
<tr>
<th>Individualized BRTF</th>
<th>Generic BRTF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{11}$</td>
<td>$H_{18}$</td>
</tr>
</tbody>
</table>

- Mean
- Standard Deviation
Results - BRTF Differences

<table>
<thead>
<tr>
<th>Individualized BRTF</th>
<th>Generic BRTF</th>
</tr>
</thead>
</table>

- $H_{41}$
- $H_{48}$

• Difference still exists even though the HATS is fixed
• Potentially due to change in head scattering
PSZ Performance in Best-Case Scenario

AC Spectrum from the *in situ* case

- Max: 9.4 dB
- Avg: 5.3 dB

- AC increases with individualized filters
- More noticeable above 2kHz
PSZ Performance in Best-Case Scenario

AC Spectrum from the *in situ* case

- *DZ* for HATS degraded by replacing the other listener
- Change in one => degradation in both!
Robustness Against Head Misalignments

AC from 20 \textit{ex situ} takes

Robustness $\sim$ width of the curve

- Robustness decrease for individualized filters above 3kHz
- Individualized is still better
- AC decrease for individualized filters compared to the \textit{in situ} case
Robustness Against Head Misalignments

AC from 20 *ex situ* takes

Robustness $\sim$ width of the curve

- Similar robustness for both filters
- Same AC decrease for individualized filters as in SZC 1
Conclusions

• Does individualization for PSZ make a difference?
  - Yes, it does!
  - AC improvement at all frequencies, both in situ and ex situ
  - Slight decrease in robustness at high frequencies

• What else?
  - Replacing one listener can affect performance for both listeners
  - Best performance may be hard to retain… (without dynamic reproduction)

• Is the improvement worth the efforts?
  - Not known yet…
  - Subjective evaluation required (to be presented in a future work)
The Performance of A Personal Sound Zone System with Generic and Individualized Binaural Room Transfer Functions

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

Presented at the 152nd AES Convention
May 17, 2022

*E-mail: yqiao@princeton.edu