Evaluation of Sensor Self-Noise in Binaural Rendering of Spherical Microphone Array Signals

Hannes Helmholz, Jens Ahrens, David L. Alon, Sebastià V. A. Garí, Ravish Mehra

May 4-8 2020, International Conference on Acoustics, Speech and Signal Processing (ICASSP)
Introduction

Sound field capture
Spherical Microphone Array

Sound field reproduction
Headphones (head-tacked binaural)

(Pictures: Audio Signal Processing Group, TH Köln)
Rendering Method

**ReTiSAR (Real-Time Spherical Array Renderer)**


\[
E_{L,R}(\omega) = \sum_{n=0}^{N} \sum_{m=-n}^{n} (-1)^m d_n(\omega) \hat{H}_n^m(\omega) \hat{S}_{n-m}^r(\omega) e^{-jm\alpha}
\]  

(1)
Instrumental Evaluation (Uniform Contribution)

Sensor signal ↓ Ear signal (left and right)

Sound field (wanted / target) = Self-noise (unwanted)

• Based on simulated plane wave array IRs
• Referenced sensor in direction of source
• Rendered head orientation 0° (facing source)
Instrumental Evaluation (Uniform Contribution)

Sensor signal

↓

Ear signal (left and right)

Sound field (wanted / target) ≠ Self-noise (unwanted)

- Referenced sensor unchanged
- Rendered head orientation 90° (lateral source)
Instrumental Evaluation (Uniform Contribution)

- Influence of radial filter soft-limiting

110ch sh8 18dB  (thick)

110ch sh8 0dB   (thin)
Instrumental Evaluation (Uniform Contribution)

- Influence of matched rendering order (identical array radius)

32ch sh4 8.75cm (thick)

110ch sh8 8.75cm (thin)
Instrumental Evaluation (Non-Uniform Contribution)

- Influence of boost level from single channel

32ch  sh4  +0dB

32ch  sh4  +12dB

Rendered self-noise in ear signals: **full 360° head rotation over ≈3.5s**
Perceptual Evaluation (Non-Uniform Contribution)

- 3AFC paradigm
- 2-down-1-up staircase (adaptive step size)
- Loudness differences equalized (based uniform contribution levels)
- Initial conditions tuned according to pre-test

110ch sh4 8.75cm

32ch sh4 8.75cm
Perceptual Evaluation (Non-Uniform Contribution)

- Grouping for 110 channel (I-IV) and 32 channel (V-VIII) configurations
- Extensive breaks between 8 conditions
- 11 expert listeners with thorough instructions
Conclusions

• Uniform self-noise contributions are perceived diffuse and externalized identical for all head orientations.
• Rendered self-noise level and coloration is strongly influenced by
  • Array configuration – radius, number of sensors, spherical sampling grid
  • Rendering configuration – radial filter limitation, (un)mached SH rendering order

• Non-uniform self-noise causes changes in the noise timbre with head orientation.
• For large arrays, a single sensor may have a 6 dB higher noise floor before a timbre change with head orientation is perceived. With small arrays, it is 2 dB.
(No target sound field present; channel positioned on equator with virtual head movement restricted to rotation around vertical axis)

Adjacent materials: https://doi.org/10.5281/zenodo.3661422

• Future work
  • Instrumental evaluation of emerging SNR in ear signals (Forum Acusticum 2020).
  • Instrumental evaluation for large variety of SMA sampling grids and non-uniform noise contributions (IEEE TASLP).

Thank you for your attention!

hannes.helmholz@chalmers.se