Ear Centering for Near-Distance Head-Related Transfer Functions

Ayrton Urviola
Perception Research
Lima, Peru
aurviola@perception3d.com

Shuichi Sakamoto
Tohoku University
Sendai, Japan
saka@ais.riec.tohoku.ac.jp

César D. Salvador
Perception Research
Lima, Peru
salvador@perception3d.com

PERCEPTION
IMMERSIVE 3D EXPERIENCES

Acoustic Information Systems
Sakamoto Laboratory

RIEC

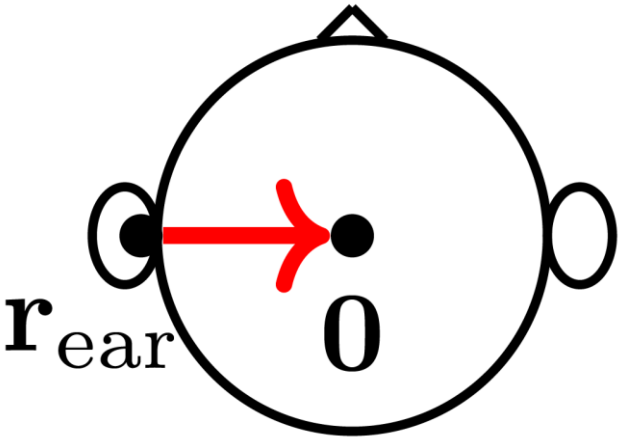
TOHOKU UNIVERSITY

Outline

1. Introduction
2. Ear centering for near-distance head-related transfer functions (HRTFs)
3. Evaluation of plane-wave (PW) and spherical-wave (SW) ear centering
4. Considerations for practical implementations
5. Conclusion

1. Introduction

- **Necessity**: HRTFs for near-field auditory displays
- **Method**: The spherical Fourier transform (SFT) is widely used in near-distance HRTF synthesis
- **Problem**: Mismatch between the SFT center (head center) and the measurement positions (ears)



Ear centering using a translation operator

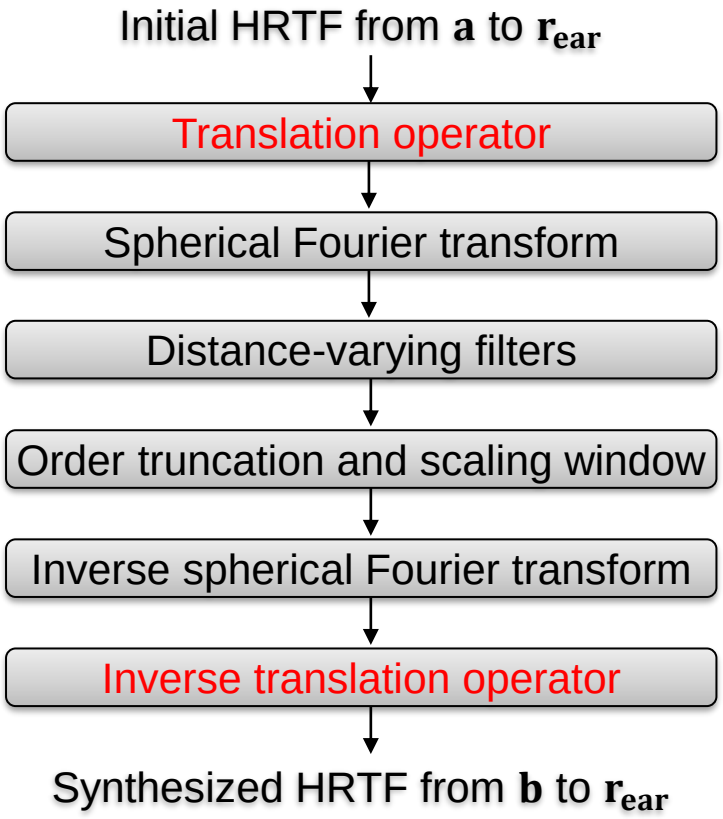
1. Introduction: Review of ear centering methods

Reference	Distance	Translation point	Domain	Translation operator
Richter, 2014	Yes	Optimized point around the ear	SFT	Ratio of hankel functions
Zaunschirm, 2018	No	Y axis with 8.5 cm radius	Unit sphere	Plane-wave
Ben-hur, 2019	No	Y axis with 8.75 cm radius	Unit sphere	Plane-wave
Porschmann, 2020	No	Y axis with 9.19 cm radius	Unit sphere	Plane-wave with rigid sphere
Arend, 2021	No	Y axis with 9.19 cm radius	Unit sphere	Plane-wave with rigid sphere

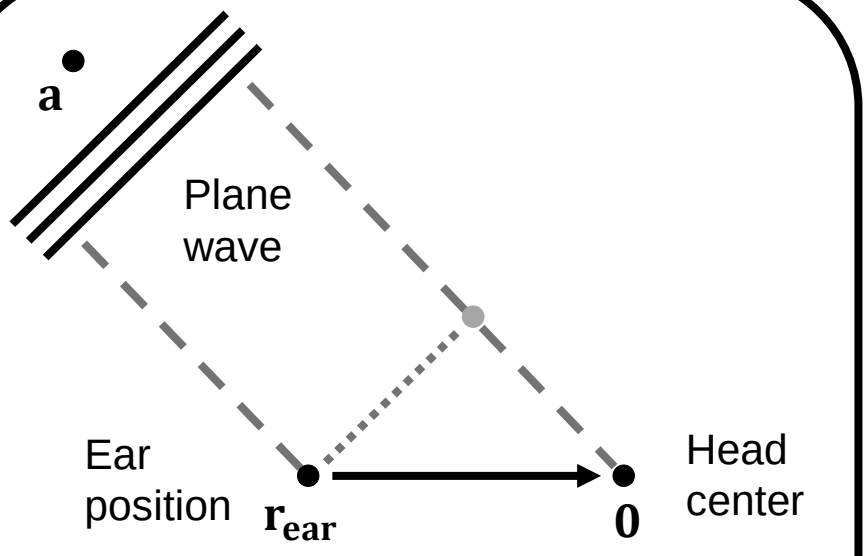
Proposal: Include distance information using a **free-field spherical-wave translation operator**

1. Introduction: Near-distance HRTF synthesis

- Input: distribution of free-field HRTFs from $\mathbf{a}$ to $\mathbf{r_{ear}}$
- Output: synthesized free-field HRTF from $\mathbf{b}$ to $\mathbf{r_{ear}}$
- Our proposal focuses on translation operators
- Intermediate steps are similar to conventional methods

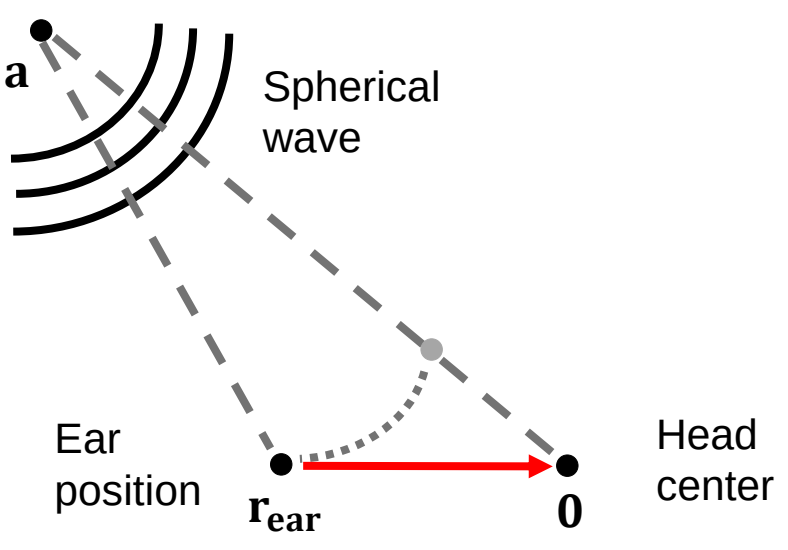


1. Introduction: PW and SW ear centering



PW translation operator

$$\mathcal{T}_{\text{PW}}(\mathbf{a}, \mathbf{r}_{\text{ear}} \mapsto \mathbf{0}, k) = e^{-jkr_{\text{ear}} \cos \Theta_{\mathbf{a},\mathbf{r}_{\text{ear}}}}$$



SW translation operator

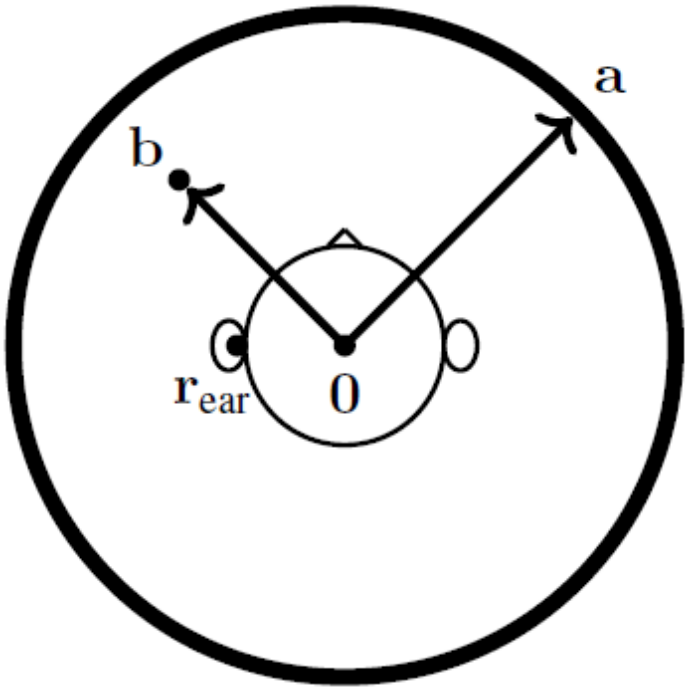
$$\mathcal{T}_{\text{SW}}(\mathbf{a}, \mathbf{r}_{\text{ear}} \mapsto \mathbf{0}, k) = \frac{\|\mathbf{a} - \mathbf{r}_{\text{ear}}\|}{a} e^{-jk(a - \|\mathbf{a} - \mathbf{r}_{\text{ear}}\|)}$$

[1] Z. Ben-Hur *et al.*, *IEEE Trans. Audio, Speech, Language Process.*, 2019.

Proposal

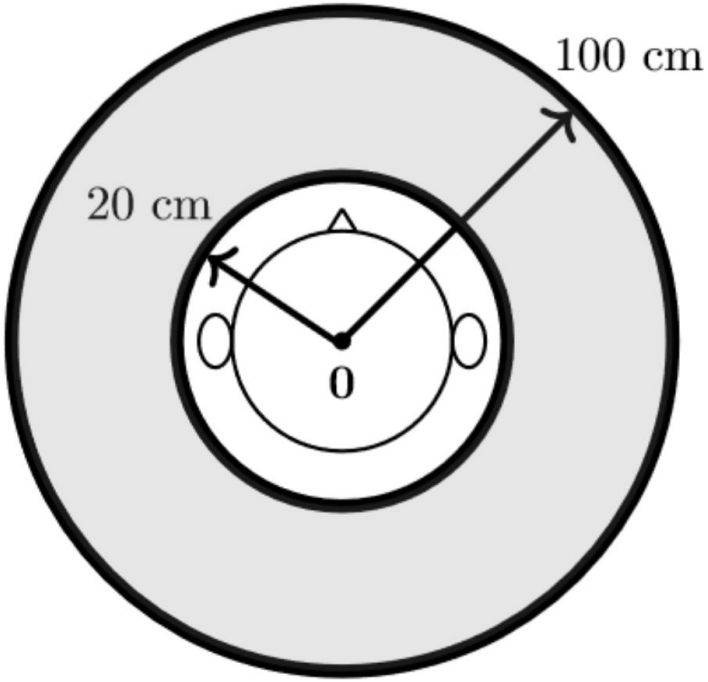
2. Ear centering for near-distance HRTFs

- $\mathbf{0}$: Origin, head center
- $\mathbf{r_{ear}}$: Ear position
- $\mathbf{a} = (a, \theta_a, \phi_a)$: Point in a continuous, spherical distribution at a far-distance a.
- $\mathbf{b} = (b, \theta_b, \phi_b)$: Arbitrary point at a near-distance b.

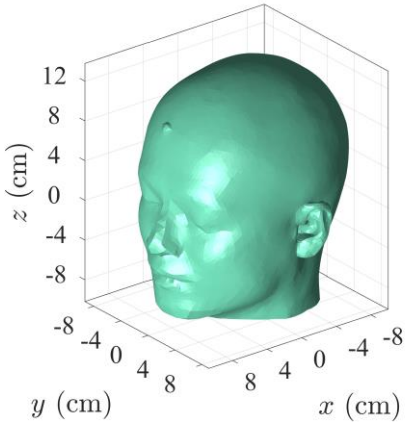


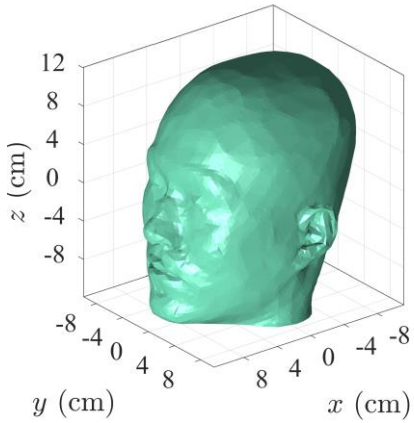
3. Evaluation of PW and SW ear centering: Conditions

Region of interest



Head Models





3. Evaluation of PW and SW ear centering: Conditions

- Initial positions
  - Icosahedral grids
  - Distance: 100 cm
  - 12 points ($N_{grid} = 2$)
  - 252 points ($N_{grid} = 14$)
- Final positions
  - Icosahedral grids
  - Distance: 20 – 100 cm
  - 642 points ($N_{grid} = 24$)

- Maximum frequency

$$f_{\max} = \frac{cN_{\text{grid}}}{2\pi r_{\text{h}}}$$

  - $r_h = 16\ cm$
- Maximum SFT order

$$N = \min(\lceil kr_{\text{h}} \rceil, N_{\text{grid}})$$

3. Evaluation of PW and SW ear centering: Conditions

- Synthesis error

$$E(b_i, f_\kappa) = \underset{s_\ell}{\mathrm{RMS}}\left\{ \frac{\underset{\Omega_j}{\mathrm{RMS}}\{\boldsymbol{H} - \hat{\boldsymbol{H}}\}}{\underset{\Omega_j}{\mathrm{RMS}}\{\boldsymbol{H}\}} \right\}$$

- Target HRTF

$$\boldsymbol{H}(b_i, \Omega_j, f_\kappa, s_\ell)$$

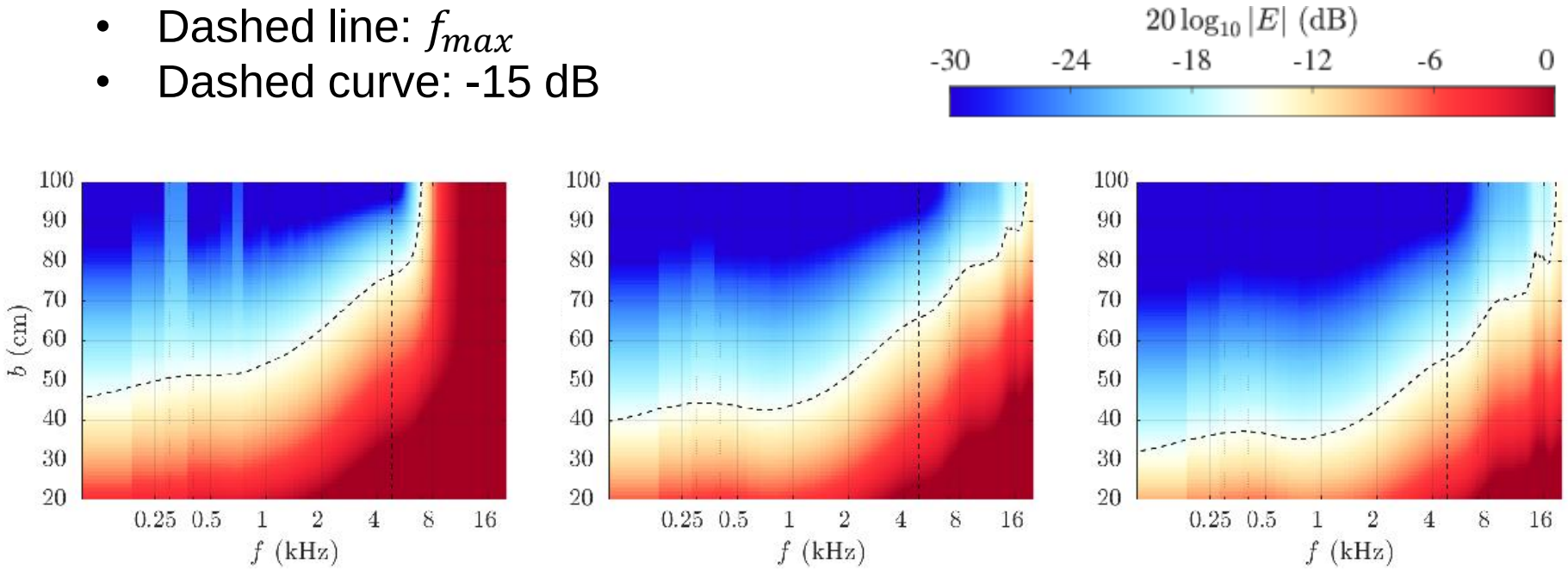
- Synthesized HRTF

$$\hat{\boldsymbol{H}}(b_i, \Omega_j, f_\kappa, s_\ell)$$

- Parameters
  - b_i 81 radial distances
  - Ω_j 642 directions
  - f_κ 257 frequency bins
  - s_l 2 head models

3. Evaluation of PW and SW ear centering: Results

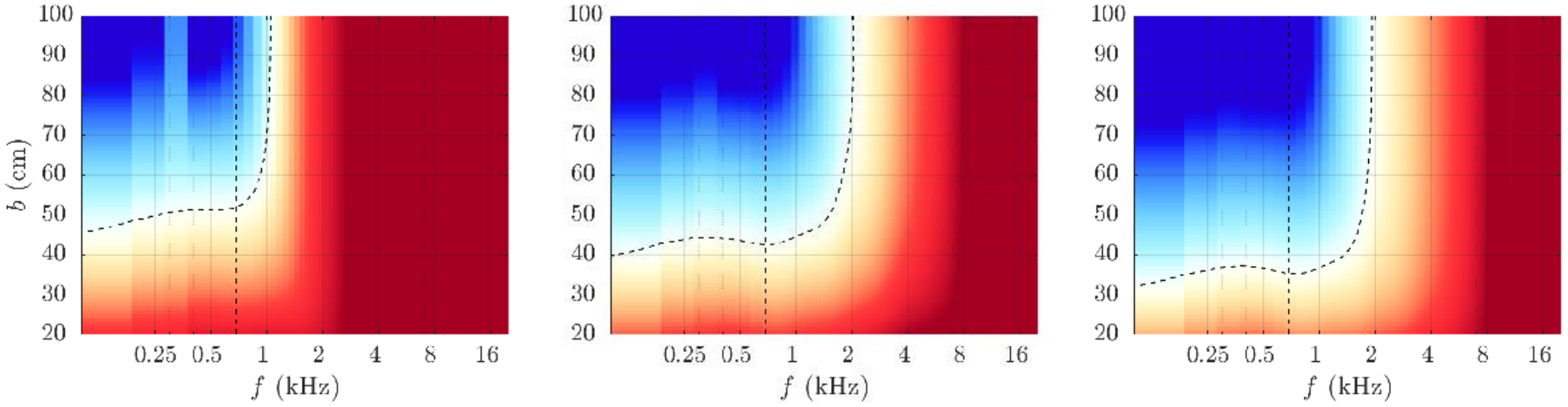
- Number of initial HRTF points: $P = 252$
- SFT order of initial HRTF grid: $N_{grid} = 14$
- Dashed line: f_{max}
- Dashed curve: -15 dB



No ear centering | PW ear centering | SW ear centering

3. Evaluation of PW and SW ear centering: Results

- Number of initial HRTF points: $P = 12$
- SFT order of initial HRTF grid: $N_{grid} = 2$
- Dashed line: f_{max}
- Dashed curve: -15 dB



No ear centering | PW ear centering | SW ear centering

3. Evaluation of PW and SW ear centering: Summary

- Below f_{max}, SW ear centering outperforms No ear centering across all distances, overall improvement of 6 dB
- Below f_{max}, SW ear centering outperforms PW ear centering across all distances, overall improvement of 3 dB
- Below 30 cm, 3 dB enhancement holds beyond f_{max}

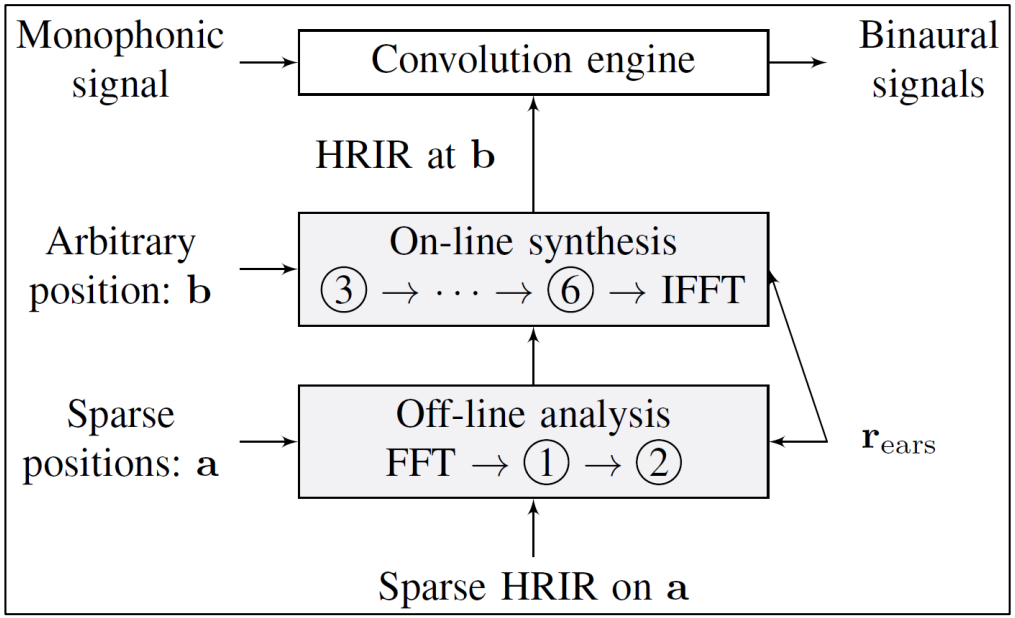
4. Considerations for practical implementations

Off-line analysis

- FFT, direct ear centering, and SFT
- $O(N_{grid}^4 N^2)$

On-line synthesis

- DVFs, order truncation and scaling window, ISFT, and IFFT
- $O(N^2)$



5. Conclusion

- Synthesis accuracy increased consistently when comparing SW to PW translation operators
- Enhancements at near distances within the frequency range of operation and even beyond f_{max}
- Extensions to this work might include:
  - Regularization techniques to optimize the bandwidth of the SFT
  - Perceptual evaluations by means of detectability of differences, and localization tests

Thanks for your attention

PERCEPTION

IMMERSIVE 3D EXPERIENCES

http://perception3d.com

Acoustic Information Systems

Sakamoto Laboratory

RIEC

TOHOKU UNIVERSITY