AES 139 Convention Paper: 9421

Comparison of Techniques for Binaural Navigation of Higher-Order Ambisonic Soundfields

Joseph G. Tylka and Edgar Y. Choueiri

PRINCETON UNIVERSITY

3D Audio and Applied Acoustics Laboratory, Princeton University

Contact: josephgt@princeton.edu
Website: www.princeton.edu/3D3A

This work is sponsored by the Sony Corporation of America

1 Introduction

Binaural navigation of higher-order ambisonic (HOA) soundfields enables a listener to virtually explore a recorded acoustic space in 3D, but necessarily generates errors as the listener navigates away from the recording point. Applications of binaural navigation may be found in virtual-reality 3D-audio reproductions of real-world spaces and acoustic performances. It is the aim of this work to compare, through numerical simulations, existing binaural navigation techniques in terms of the errors introduced by each technique and their effects on localization.

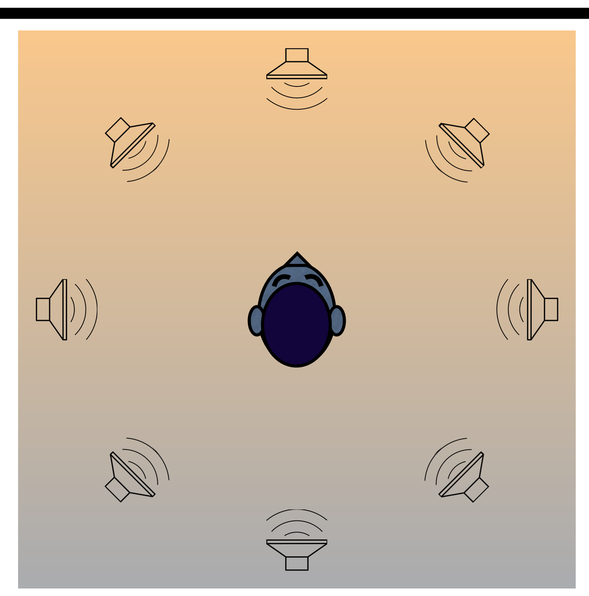
2 Navigation Techniques

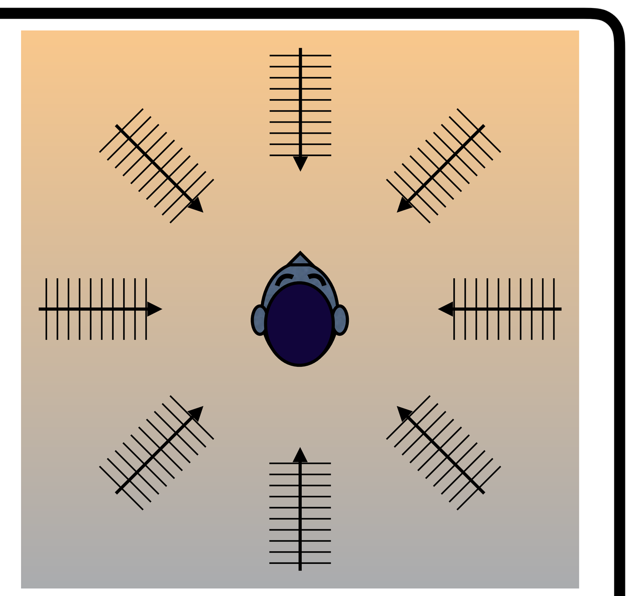
1. **Virtual HOA (VHOA)**:
   - decode to fixed set of virtual (point-source) loudspeakers
   - attenuate, delay, and filter by HRTFs based on listener position
2. **Plane-wave Translation (PWT)**: [1]
   - transform to plane-wave expansion
   - delay and filter by HRTFs based on listener position

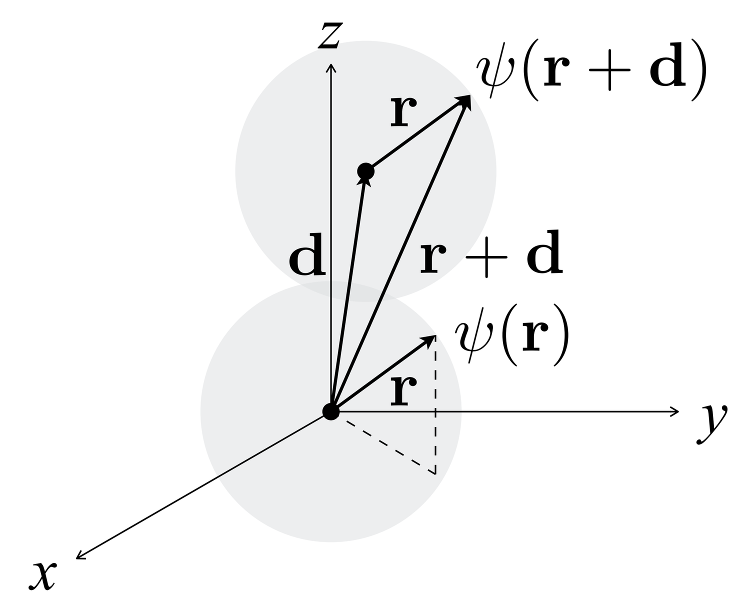
$$\underbrace{\psi_{\mathrm{PW}}^{L,R}(k,\mathbf{d})}_{\text{binaural potential}} = \sum_{p=0}^{N_P-1} \underbrace{e^{ik\hat{\mathbf{s}}_p\cdot\mathbf{d}}}_{\text{plane-wave}} \underbrace{w_p\mu(k,\hat{\mathbf{s}}_p)}_{\text{signal}} \underbrace{H^{L,R}(k,-\hat{\mathbf{s}}_p)}_{\text{HRTF}}$$

3. **Soundfield Re-expansion (SRE)**:
   - re-expand (as spherical harmonics) soundfield about the listener's position [2]
   - render to binaural using plane-wave expansion and HRTFs

$$\underbrace{R_n^m}_{j_n\cdot Y_n^m}(k,\mathbf{r}+\mathbf{d}) = \sum_{n'=0}^{\infty}\sum_{m'=-n'}^{n'} \underbrace{\Gamma_{n',n}^{m',m}(k,\mathbf{d})}_{\text{translation filter}} R_{n'}^{m'}(k,\mathbf{r}) \Longrightarrow \underbrace{C_{n'}^{m'}(k;\mathbf{d})}_{\text{translated}} = \sum_{n=0}^{N_S}\sum_{m=-n}^{n} \underbrace{\Gamma_{n',n}^{m',m}(k,\mathbf{d})}_{\text{translation filter}} \underbrace{A_n^m(k)}_{\text{original}}$$


Virtual HOA


Plane-wave Translation



Soundfield Re-expansion

3 Metrics

- *Normalized reconstruction error*: the difference between the reconstructed field and the original (band-limited) expansion.

$$\epsilon_R(k,\mathbf{r};\mathbf{d}) = \frac{|\psi_{\mathrm{BL}}(k,\mathbf{r}+\mathbf{d}) - \psi'(k,\mathbf{r};\mathbf{d})|^2}{|\psi_{\mathrm{BL}}(k,\mathbf{r}+\mathbf{d})|^2}$$

- *Velocity and energy vectors:* predicted localization due to low-frequency ITDs and high-frequency ILDs, respectively [3].

$$\mathbf{r}_{\mathrm{V}}(k) = \mathrm{Re}\left[\frac{\sum_n G_n(k)\hat{\mathbf{r}}_n}{\sum_n G_n(k)}\right], \quad \mathbf{r}_{\mathrm{E}}(k) = \frac{\sum_n |G_n(k)|^2\hat{\mathbf{r}}_n}{\sum_n |G_n(k)|^2}$$

4 Simulation Parameters

- Point-source 2.5 m in front of the listener
- Original expansion up to order $N_S = 4$
- Plane-wave expansion: $N_P = 100$ terms arranged on Fliege nodes [4]
- Re-expansion up to order $N_S' = 4$

5.1 Localization Vectors

- All techniques accurately reproduce low-frequency velocity vectors.
- PWT is unable to accurately reproduce energy vectors with lateral translation.
- SRE accurately reproduces energy vectors, but over a limited translation range.

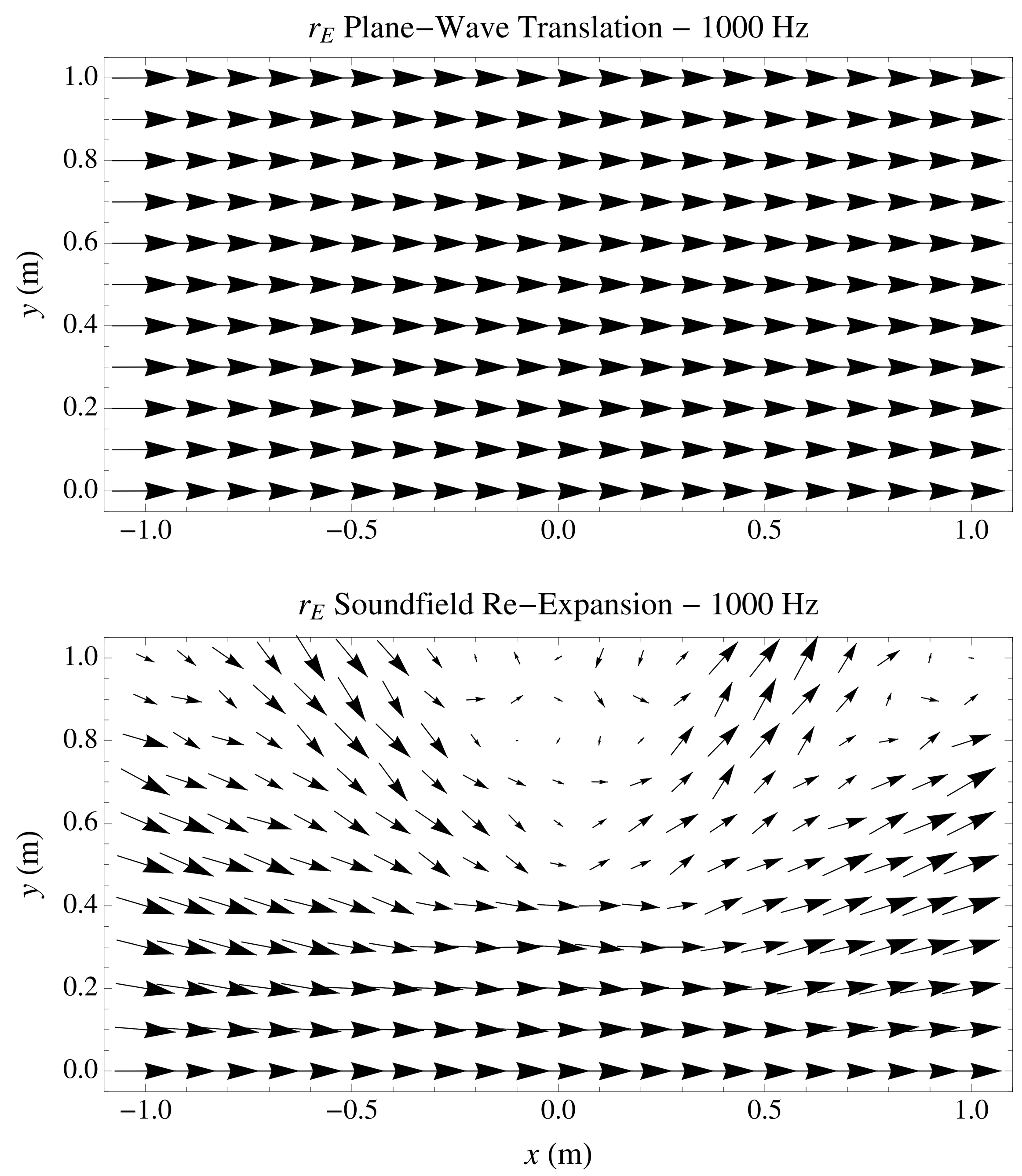


Figure 1: Reproduced energy vectors for PWT (top) and SRE (bottom) plotted on a rectangular 10 cm × 10 cm grid in the *x*-*y* plane.

5.2 Reconstruction Errors

- VHOA and PWT create static "sweet-spots" centered on the origin.
- SRE creates a "moving" sweet-spot centered on the listener.

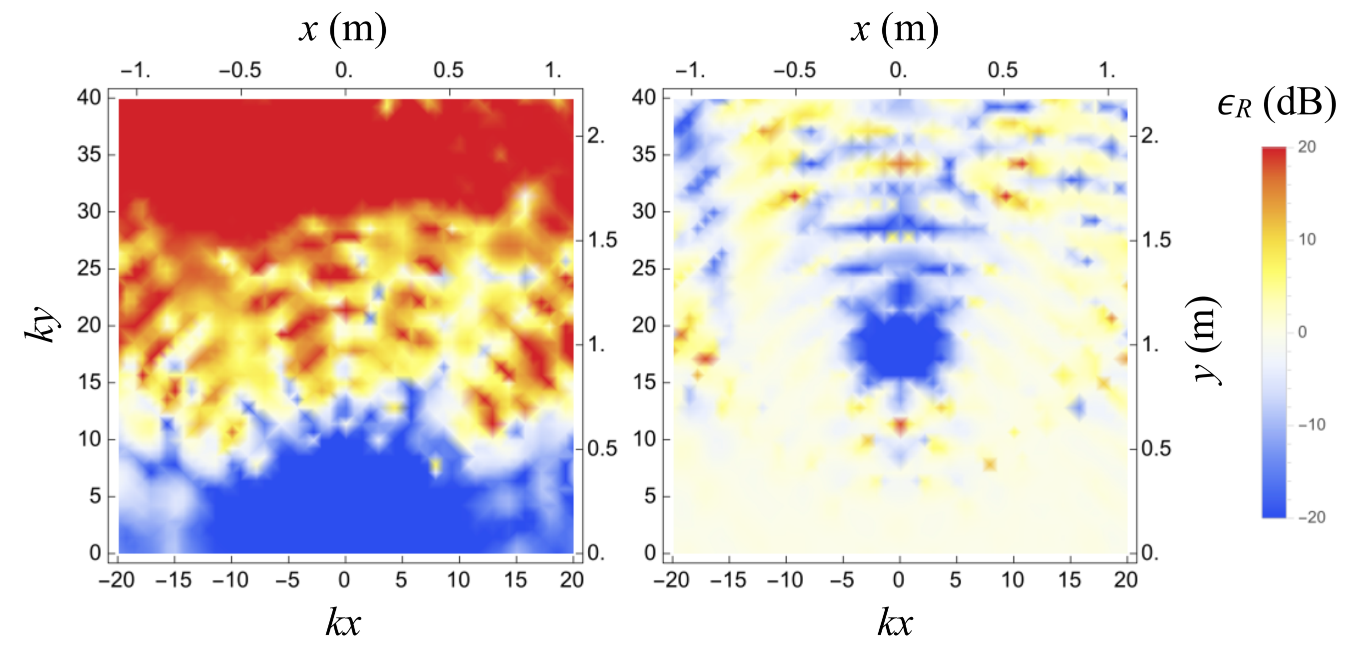


Figure 2: Normalized reconstruction error in the *x*-*y* plane at 1 kHz for translation to **d** = (0,1,0) m. Left and bottom axes show nondimensionalized distance at 1 kHz; right and top axes show distance in meters.

6 Conclusions

- SRE achieves arbitrarily low reconstruction errors in the vicinity of the listener; VHOA and PWT achieve low errors only near the recording position.
- SRE appears to be the only technique capable of accurately generating high-frequency localization cues with lateral translation, although the range over which this is possible is strictly limited by the original expansion order.

References

[1] F. Schultz and S. Spors. "Data-Based Binaural Synthesis Including Rotational and Translatory Head-Movements". Presented at the AES 52nd International Conference, September 2013.
[2] N. A. Gumerov and R. Duraiswami. *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Elsevier Science, 2005.
[3] M. A. Gerzon. "General Metatheory of Auditory Localisation". Presented at the AES 92nd Convention, March 1992.
[4] J. Fliege and U. Maier. "The distribution of points on the sphere and corresponding cubature formulae". *IMA J. Numer. Anal.*, 19(2):317–334, 1999.