128th AES Convention, London, UK, May 22-25, 2010. Paper # 8064

Design of a Circular Microphone Array for Panoramic Audio Recording and Reproduction: Array Radius

Enzo De Sena, Hüseyin Hacıhabiboğlu and Zoran Cvetković

Centre for Digital Signal Processing Research (CDSPR), King's College London

A method for quantification and prediction of the perceived naturalness is introduced. This method is used to select the microphone array radius of Johnston and Lam's perceptual sound field reconstruction system. Simulation results support array diameters close to that proposed by Johnston and Lam in the sense that they achieve optimal naturalness in the centre of the listening area, but also suggest that larger arrays might provide a more graceful degradation of the naturalness for listening positions away from the centre.

Perceptual sound field reconstruction system

- The system aims at the perceptual reconstruction of the sound field [1].
- Circular array of five microphones situated at vertices of a regular pentagon.
- Reproduction using five loudspeakers in the same regular configuration. Each microphone drives the corresponding loudspeaker.

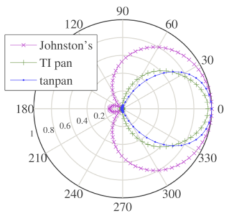

Figure 1 - The three different microphone directivities under test.

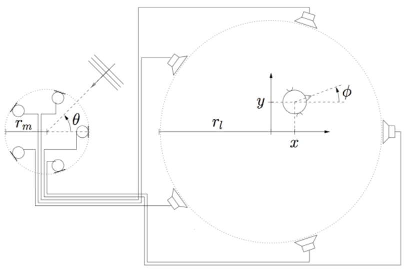

Figure 2 - The proposed system

Employed auditory model

The auditory model employed in the simulations is composed of:
1) Spherical head model to simulate scattering by and diffraction around the head;
2) Gammatone filterbank of 24 critical bands covering a range between 20 Hz - 15.5 kHz;
3i) ILD in each critical band computed as logarithmic ratio of energy in left and right ears;
3ii)ITD in each critical band computed as the maximum of the cross-correlation function.

Natural correlation between ITD and ILD

- ILD is mainly due to the diffraction of the sound wave around the head, whereas ITD is the result of the physical distance between the ears.
- ITD and ILD are naturally correlated [3]: high ITD usually corresponds to high ILD and vice-versa.
- The function that associates a given ITD with its "natural ILD" is obtained in each critical band. Such functions can be called *naturalness functions*.
- In figure 3 the naturalness functions obtained with our auditory model (black lines) are compared with the ones presented in [3] whose main difference lie in that real HRTF are employed. Their similarity validates the use of the spherical head model.

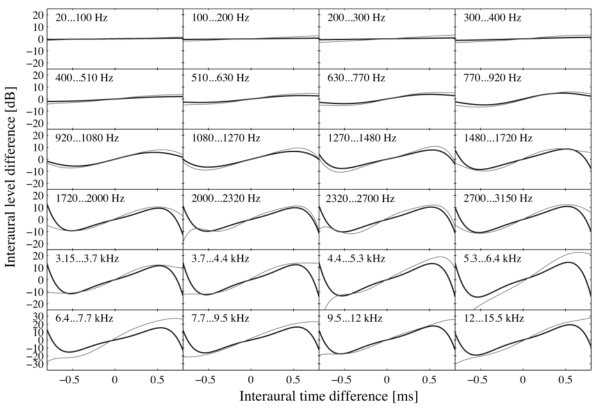

Figure 3 - Comparison between the naturalness functions obtained with our auditory model (black lines), with the ones given by [3] (grey lines) in 24 critical bands.

Perceived naturalness

- The "naturalness error" is taken as the average over the 24 critical bands of the absolute difference between the measured ILD - ΔL_i - and the natural ILD associated to the measured ITD - $F_i(\tau_i)$:

$$\epsilon \triangleq \frac{1}{24} \sum_{i=1}^{24} |\Delta L_i - F_i(\tau_i)| \qquad (1)$$

Selection of the array radius

- Natural sound sources trigger ILD and ITD cues that are highly correlated - see Figure 3. Phantom images reproduced by any multichannel system should ideally preserve this property.
- During informal listening tests we observed that the radius of the microphone array clearly influences the perceived naturalness of the phantom images.
- In order to analyse such an impact, the system is simulated in both the recording and reproduction phases for given source direction and array radius, and the "naturalness error" is computed by means of equation (1).

Results: user in the centre

- Results are averaged over 60 source directions. Listener heading is $\phi = 0$.
- ✓ *Tanpan* is more natural than *TI pan*, which is more natural than *Johnston's*: narrower directivities - see figure 1 - deliver more natural but less accurate sound fields (refer to [2]).
- ✓ The naturalness error is minimum around 10 cm for all the directivity patterns.
- ✓ The same simulation scenario is run with different head radii. Figure 5 shows the optimal array radius is approximately a linear function of the head radius.

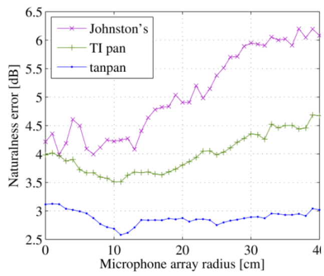

Figure 4 - Naturalness error as a function of the microphone array radius.

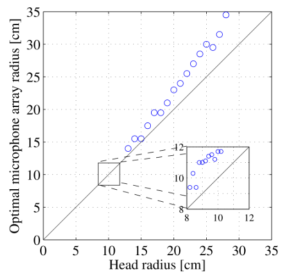

Figure 5 - Optimal microphone array radius as a function of listener's head radius.

Results: User off-centre

- ✓ As the radius increases, the naturalness error becomes more uniform.
- ✓ This suggests that an array radius that is optimal in the central position cause the performance to degrade more rapidly as the listener moves, whereas higher radii deliver a non-optimal but more stable sweet spot.

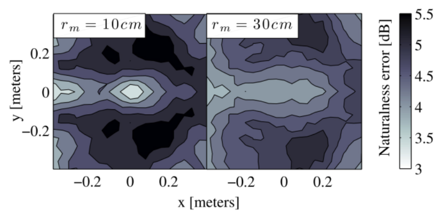

Figure 6 - Average naturalness error (TI pan) as a function of user position.

References

[1] J.Johnston and Y.Lam, "Perceptual soundfield reconstruction," AES 109 Convention - Preprint # 5202, Los Angeles, USA, September 2000.

[2] H. Hacıhabiboglu, E. De Sena, and Z. Cvetkovic, "Design of a circular microphone array for panoramic audio recording and reproduction: Microphone directivity," AES 128 Convention - Preprint # 8063, London, UK, May 2010.

[3] W. Gaik, "Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling," J. Acoust. Soc. Am., 94:98, 1993.

Acknowledgments

This work is supported by the Engineering and Physical Sciences Research Council (EPSRC) Research Grant EP/F001142/1 "Perceptual Sound Field Reconstruction and Coherent Emulation".