PERCEPTUAL SOUNDFIELD RECONSTRUCTION IN THREE DIMENSIONS VIA SOUND FIELD EXTRAPOLATION

Ege Erdem[1,*], *Enzo De Sena*[2], *Hüseyin Hacıhabiboğlu*[1,*], *Zoran Cvetković*[3]

[1] Graduate School of Informatics, METU, Ankara, TR-06800, Turkey
[2] Institute of Sound Recording (IoSR), University of Surrey, Guildford, GU2 7XH, UK
[3] Division of Informatics, King's College London, WC2R 2LS, London, UK

ABSTRACT

Perceptual sound field reconstruction (PSR) is a spatial audio recording and reproduction method based on the application of stereophonic panning laws in microphone array design. PSR allows rendering a perceptually veridical and stable auditory perspective in the horizontal plane of the listener, and involves recording using near-coincident microphone arrays. This paper extends the PSR concept to three dimensions using sound field extrapolation carried out in the spherical-harmonic domain. Sound field rendering is performed using a two-level loudspeaker rig. An active-intensity-based analysis of the rendered sound field shows that the proposed approach can render direction of monochromatic plane waves accurately.

Index Terms— Perceptual soundfield reconstruction, spatial audio, sound field extrapolation

1. INTRODUCTION

Recreating the auditory experience of an acoustic performance has been of interest for broadcasters, artists and academics for over a century. Blumlein was the first to attempt rendering of sound sources accurately in space using a pair of figure-eight microphones. Since then, many solutions have been proposed [1], exploring different microphone configurations, directivity patterns, and channel-mixing strategies. Apart from approaches based on Ambisonics recording, these methods are predominantly heuristic, and thus fundamentally reliant on the skills and taste of sound engineers.

Higher-order Ambisonics (HOA) is an elegant approach based on the spherical harmonics decomposition of the sound field [2]. Spherical microphone arrays can be used to record spherical harmonic components up to a certain order. The commercially-available Soundfield microphone allows recording up to the first order, whereas the Eigenmike up to the fourth-order [3] harmonics. Once the sound field has been recorded, it can be reproduced using various decoding methods. The mode-matching decoding, for instance, aims to reconstruct a certain number of spherical harmonic components in the center of the loudspeaker array. The main issue with HOA is that it requires expensive microphone arrays as well as a large number ($(N + 1)^2$ for N-th order) of carefully positioned and calibrated loudspeakers. Methods of lower complexity, such as SIRR [4, 5], DirAC [6], SDM [7], and Perceptual soundfield reconstruction (PSR) [8, 9] leverage psychoacoustic phenomena [10].

PSR uses a set of microphones of specifically designed directivity patterns, each connected to a loudspeaker in a corresponding direction without additional mixing [9]. The microphones are positioned on a circular array the radius of which is left as a free parameter, that allows a control over time-intensity difference trade-offs. A radius of 15.5 cm was used in [9], and it was shown that PSR performs on a par with VBAP [11] and second-order Ambisonics in the centre of the sweet-spot, but has a more graceful performance degradation away from the sweet-spot. More specifically, PSR provides better locatedness of phantom sources than techniques based on intensity alone. Using results of a computational model [12], the increased locatedness was attributed to the higher naturalness of the presented binaural cues [13].

The current formulation of PSR has two main limitations: i) it is confined to the horizontal plane, ii) each channel requires a dedicated (2nd or higher order) microphone. The latter becomes particularly problematic for extending PSR to the full 3D case, as the number of needed channels/loudspeakers increases substantially. Towards overcoming these limitations, this paper investigates how the PSR array can be implemented using a single, spherical-harmonics-based coincident microphone array, such as the Eigenmike. This is carried out in two steps. First, the pressure is extrapolated at positions of the corresponding PSR microphones. This is done in the spherical harmonic domain in a manner similar to other prior art methods [14]. The direction of the active intensity field is then also extrapolated at the same positions. Second, the directions of the active intensity vectors at locations of PSR microphones are used to weigh the pressure signal according to the PSR directivity patterns. Numerical simulations using the proposed method are carried out in a 3D case, and results suggest that even a straightforward extension of PSR directivity patterns designed for the 2D case achieves very good results.

Section 2 presents the relevant background on spherical harmonic theory and on PSR. Section 3 introduces the proposed approach, which is then numerically analyzed in Section 4. Section 5 concludes the paper and identifies future work directions.

2. BACKGROUND

2.1. Perceptual Soundfield Reconstruction (PSR)

PSR is based on the summing localisation effect [15] and uses time-intensity stereophonic panning curves as established by Franssen [16] and Williams [17]. These curves provide the pairs of time-level differences between loudspeaker pairs that result in the phantom image being perceived in the direction of one loudspeakers or the other. In [9], directivity patterns are designed that interpolate between such psychoacoustic laws to record and render acoustic sources at locations between loudspeaker pairs in the context of multichannel systems. To that end, first, the maximal time delay between channels due to a source in the direction of one of the loudspeakers is

ICASSP 2019

calculated; this delay is specified by the radius of the array and the particular angular placement of channels. Then, the level difference in that direction is set to be the level difference that is needed in combination with the time delay to create a phantom source in that direction. Finally, an equal-loudness constraint is imposed, giving the following directivity pattern which effectively interpolates time-level difference pairs between the end points:

$$\Gamma_d(\Theta) = \begin{cases} \left[1 + \frac{\sin^2(|\Theta|+\beta)}{\sin^2(|\Theta|-(\phi_0+\beta))}\right]^{-1/2} & \Theta \in [-\phi_0, \phi_0] \\ 0 & \text{elsewhere} \end{cases} \quad (1)$$

where ϕ_0 is the angle between loudspeakers in the horizontal plane and $\beta = \arctan \frac{\eta \sin(\phi_0)}{1-\eta \cos(\phi_0)}$, and where η is a value calculated from psychoacoustic curves to achieve the desired level differences at loudspeaker directions. The parameters used in the original PSR formulation are $\phi_0 = 2\pi/5$ and $\eta = 0.302$. The reader is referred to [9] for details of how these parameters are calculated. Whereas different interpolating functions can be used, the formulation in (1) has a form of a generalized tangent panning law, *i.e.* the tangent panning law is its special case for $\beta = 0$, which physically means infinite level differences (expressed in dB) in loudspeaker directions.

2.2. Spherical Harmonic Decomposition

A sound pressure $p(kr, \theta, \phi)$ of a frequency $\omega = kc$, where c is the speed of sound, on a spherical surface of radius r, can be represented as a linear combination of spherical harmonics [18]:

$$p(kr, \theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} p_{nm}(kr) Y_n^m(\theta, \phi) \quad (2)$$

where $\theta \in [0, \pi]$ the inclination angle with respect to the $+z$ axis, $\phi \in [0, 2\pi)$ is the azimuth angle defined from the $+x$ axis,

$$Y_n^m(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{jm\phi} \quad (3)$$

are the spherical harmonics, and $P_n^m(\cdot)$ are the associated Legendre polynomials of order $n \in \mathbb{Z}^+$, and degree $-n \leq m \leq n$. Coefficients p_{nm} can be calculated via the spherical harmonic decomposition (SHD) as:

$$p_{nm}(kr) = \int_0^{2\pi} \int_0^{\pi} p(kr, \theta, \phi) Y_n^m(\theta, \phi) \sin\theta d\theta d\phi. \quad (4)$$

2.3. Rigid Spherical Microphone Arrays

SHD coefficients can be obtained from a finite set of Q pressure signals captured on a rigid sphere of radius r_a using rigid spherical microphone arrays (RSMAs). This involves the calculation of an approximation to the surface integral using a numerical quadrature:

$$\widetilde{p}_{nm}(k) = \sum_{q=1}^{Q} w_q p(\theta, \phi) \left[Y_n^m(\theta, \phi)\right]^* \quad (5)$$

where w_q are the quadrature weights. If discrete orthonormality conditions [19] are satisfied, SHD coefficients up to an order N can be obtained using and RSMA with $Q \geq (N+1)^2$ elements.

The SHD coefficients calculated this way include the effect of the rigid sphere which depends both on frequency and on the radius of the sphere, r_a. This effect can be eliminated according to:

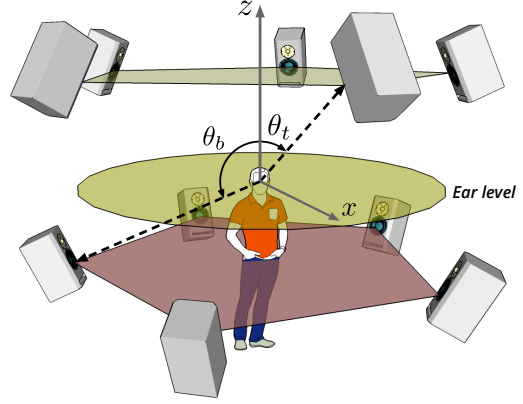$$p_{nm}(k) = \widetilde{p}_{nm}(k) / \left[4\pi i^n b_n(kr_a)\right], \quad (6)$$

Fig. 1. The proposed reproduction setup with 10 loudspeakers.

where

$$b_n(kr) = j_n(kr) - \frac{j_n^{'}(kr_a)}{h_n^{(2)'}(kr_a)} h_n^{(2)}(kr), \quad (7)$$

and $j_n(\cdot)$ and $h_n^{(2)}(\cdot)$ are the spherical Bessel function of the first kind and spherical Hankel function of the second kind, respectively. The derivatives of these functions with respect to their arguments are given as $j_n^{'}(\cdot)$ and $h_n^{(2)'}(\cdot)$, respectively.

3. 3D-PSR RECORDING VIA EXTRAPOLATION

3.1. Motivation

The 3D extension of PSR studied here uses a reproduction setup of 10 loudspeakers, arranged in two horizontal layers above and below the ear level, as shown in Fig. 1. Each layer consists of 5 pentagonally placed loudspeakers, rotated by $\pi/5$ with respect to each other.

The corresponding 3D-PSR microphone array requires microphones positioned on the surface of an open sphere of radius 15.5 cm, with their acoustic axes pointing towards the loudspeakers. Design of a bespoke 10 channel near-coincident microphone array is challenging. If the design is carried out using differential microphone arrays (DMAs) [20] of order M, at least $10M + 1$ microphones would be needed (assuming that the center microphone is shared among all channels), that are equalised for their frequency responses. The positioning of these DMAs would also present practical problems. The same directivity patterns can also be designed via steered beamforming using the SHD coefficients obtained from an RSMA. However, the lack of inter-channel time delays [9] eliminates one of the fundamental premises of the original PSR design. In order to keep the robustness and flexibility offered by the RSMAs whilst capturing interchannel time differences, we propose an approach based on sound field extrapolation using SHD coefficients.

3.2. Sound Field Extrapolation

A pressure field composed of plane waves can be approximated using a linear combination of spherical harmonic functions as:

$$p(k, \mathbf{r}) = \sum_{n=0}^{N} \sum_{m=-n}^{n} p_{nm} 4\pi i^n j_n(kr) Y_n^m(\theta, \phi). \quad (8)$$

where $\mathbf{r} = (r, \theta, \phi)$ represents a point in spherical coordinates. Note that $\lim_{N\to\infty} p(k, r, \theta, \phi)$ characterises the sound field exactly

at all points while the truncated series given in (8) will provide a good approximation only up to a finite radius around the origin [21].

Particle velocity and pressure are related by conservation of momentum which can be expressed in time and frequency domains as:

$$-\nabla p(t, \mathbf{r}) = \rho_0 \frac{\partial \mathbf{u}(t, \mathbf{r})}{\partial t} \overset{\mathcal{F}}{\leftrightarrow} -\nabla p(k, \mathbf{r}) = j\rho_0 kc\mathbf{u}(k, \mathbf{r}), \quad (9)$$

where ρ_0 is the ambient density. Particle velocity can be calculated at an arbitrary point around the origin using this relation as:

$$\mathbf{u}(k, \mathbf{r}) = -\frac{1}{j\rho_0 kc} \left[\frac{\partial p}{\partial r} \hat{\mathbf{u}}_r + \frac{1}{r} \frac{\partial p}{\partial \theta} \hat{\mathbf{u}}_\theta + \frac{1}{r \sin \theta} \frac{\partial p}{\partial \phi} \hat{\mathbf{u}}_\phi \right] \quad (10)$$

where $\hat{\mathbf{u}}_r$, $\hat{\mathbf{u}}_\theta$, and $\hat{\mathbf{u}}_\phi$ are the unit vectors in the radial, inclination and azimuth directions, respectively. The partial derivatives are:

$$\frac{\partial p}{\partial r} = \sum_{n=0}^{N} \sum_{m=-n}^{n} p_{nm} 4\pi i^n k j_n'(kr) Y_n^m(\theta, \phi) \quad (11)$$

$$\frac{\partial p}{\partial \theta} = \sum_{n=0}^{N} \sum_{m=-n}^{n} p_{nm} 4\pi i^n j_n(kr) \frac{\partial Y_n^m(\theta, \phi)}{\partial \theta} \quad (12)$$

$$\frac{\partial p}{\partial \phi} = \sum_{n=0}^{N} \sum_{m=-n}^{n} p_{nm} 4\pi i^{n+1} m Y_n^m(\theta, \phi) \quad (13)$$

where

$$\frac{\partial Y_n^m(\theta, \phi)}{\partial \theta} = \quad (14)$$

$$\left[m \cot \theta Y_n^m(\theta, \phi) + \sqrt{(n-m)(n+m+1)} e^{-i\phi} Y_n^{m+1}(\theta, \phi) \right]$$

and $Y_n^m = 0$ for $|m| > n$.

Once particle velocity is extrapolated, active intensity, that represents the direction and strength of energy, can be obtained at any point at which the approximation is sufficiently accurate as:

$$\mathbf{I}_a(k, \mathbf{r}) = \frac{1}{2} \operatorname{Re} \{p(k, \mathbf{r})\mathbf{u}^*(k, \mathbf{r})\} \ . \quad (15)$$

This will be used to obtain directional responses of emulated microphones located at desired positions.

3.3. Microphone Directivity Patterns

Once the direction of the sound field is calculated via the extrapolation, it is possible to obtain virtual microphone recordings at points around the sphere, and thus emulate near-coincident recording setups. A microphone directivity pattern can be represented as a linear combination of spherical harmonic functions as:

$$\Gamma(\theta', \phi') = \sum_{n=0}^{N} \sum_{m=-n}^{n} \alpha_{nm} Y_n^m(\theta', \phi') \quad (16)$$

where (θ', ϕ') is the local spherical coordinates defined with respect to the microphone axis and the coefficients $\alpha_{nm} \in \mathbb{C}$ have to satisfy $\alpha_{nm} = \pm\alpha_{nm}^*$ to obtain an axisymmetric pattern. In order for $\Gamma(\theta, \phi)$ to be real the coefficients should satisfy $\alpha_{nm} = \alpha_{nm}^*$. The design space can be constrained to use the spherical harmonics of degree $m = 0$, resulting in [22]:

$$\Gamma(\Theta) = \sum_{n=0}^{N} \alpha_n \left(\frac{2n+1}{4\pi} \right) P_n(\cos \Theta) = \sum_{n=0}^{N} \beta_n \cos^n \Theta \quad (17)$$

where Θ is the angle between the acoustic axis of the microphone and the wave of a plane wave. The directivity pattern can be constrained to have unit response in the direction of its acoustic axis by imposing $\sum_{n=0}^{N} \beta_n = 1$. Whereas the method proposed here is not limited to axisymmetric directivity patterns, we employ the time-intensity directivity pattern described in Section 2.1 [9].

The PSR pattern, $\Gamma_d(\Theta)$, is approximated here as a pattern in the form (17) by jointly minimizing the L^2-distance in the pick-up region and the L^2-norm in the rejection region:

$$\underset{\beta_1, \beta_2, \ldots, \beta_N}{\operatorname{argmin}} \quad \lambda \int_0^{\phi_0} |\Gamma(\Theta) - \Gamma_d(\Theta)|^2 + (1 - \lambda) \int_{\phi_0+\epsilon}^{\pi} |\Gamma(\Theta)|^2 \, d\Theta \quad (18)$$

The resulting pattern for $\lambda = \frac{1}{2}$, $\epsilon = \frac{\pi}{10}$, and $N = 4$ has coefficients $\beta_0 = 0.001$, $\beta_1 = 0.458$, $\beta_2 = 0.536$, $\beta_3 = 0.040$ and $\beta_4 = -0.126$ (obtained using Mathematica 11).

3.4. Emulated off-centre microphone recordings for 3D-PSR

The local pressure, $p(k, \mathbf{r})$ and the active intensity vector, $\mathbf{I}_a(k, \mathbf{r})$ can be calculated at any point $\mathbf{r}$ within a region where the sound field extrapolation is accurate. The directional response of the emulated microphone depends on the angle between the local intensity vector and the acoustic axis of the microphone:

$$\Theta_{\mathbf{r}} = \arccos \frac{\langle \mathbf{r}, \mathbf{I}_a(k, \mathbf{r}) \rangle}{|\mathbf{I}_a(k, \mathbf{r})||\mathbf{r}|} \quad (19)$$

Then, the emulated microphone signal is obtained as

$$p_{\text{rec}}(k, \mathbf{r}) = \Gamma(\Theta_{\mathbf{r}})p(k, \mathbf{r}).$$

3D PSR involves the calculation of emulated microphone signals at a distance of 15.5 cm with the directivity pattern given above. The acoustic axes of the emulated microphones are radially outwards. The microphones at the top and bottom layers have common inclination angles of $\theta_t \approx 0.352\pi$ and $\theta_b \approx 0.648\pi$, respectively. The azimuth angles for the top and bottom layers are $\phi_t = 2m\pi/5$ and $\phi_b = (2m+1)\pi/5$ for $m = 0, \ldots, 4$. We assume that the listener's head would be positioned at the centre of the loudspeaker rig.

4. NUMERICAL EVALUATIONS

We evaluated the 3D-PSR proposed here using numerical simulations in terms of their directional reproduction accuracy for monochromatic plane wave fields. Absolute angular error used in the discussion below is defined as:

$$\epsilon(\mathbf{r}) = \arccos \langle \mathbf{n}_{PSR}(\mathbf{r}), \mathbf{n}_{pw}(\mathbf{r}) \rangle \quad (20)$$

where $\mathbf{n}_{PSR}$ is the unit vector in the direction of the reproduced field and $\mathbf{n}_{pw}$ is the unit vector in the direction of the plane wave.

We evaluate the directional accuracy of reproduced sound fields due to monochromatic plane waves of frequencies 250 Hz, 500 Hz, and 1 kHz, incident from $\mathbf{n}_k = (\pi/2, \pi/4)$ and reconstructed using 3D-PSR. The simulations use emulated microphone recordings using the sound field extrapolated 15.5 cm away from the origin. The maximum order of extrapolation used were $N = 1$, $N = 2$, and $N = 3$, for 250 Hz, 500 Hz, and 1 kHz waves, respectively. Fig. 2 shows the angular reproduction error as a contour plot and intensity vectors as a vector plot for the tested cases. The average directional errors calculated within a spherical volume of radius 0.2 m around the center of the simulated volume are $0.66°$, $3.68°$, $27.75°$, for the
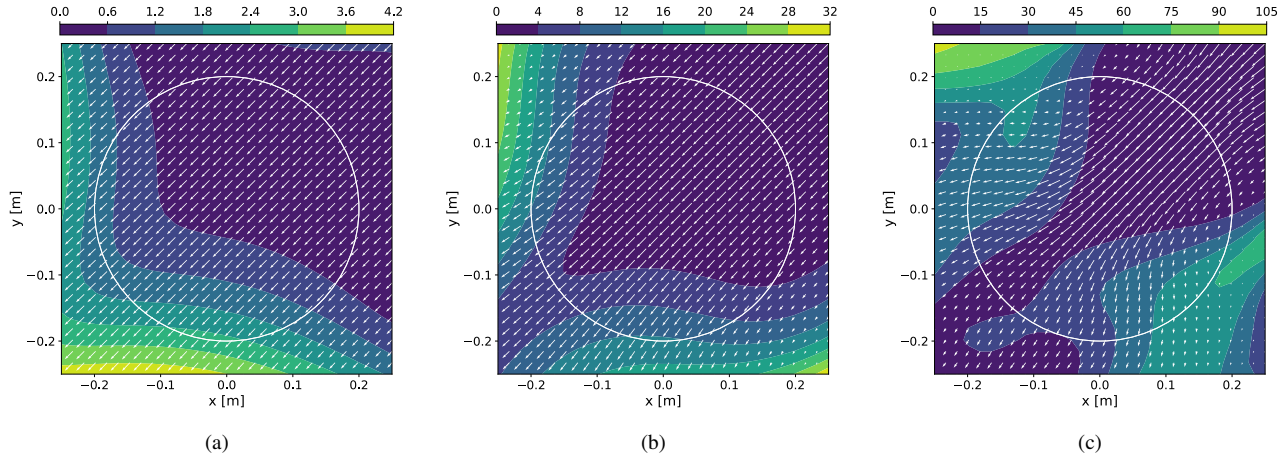
Fig. 2. Reconstructed sound field around the optimal listening area for a monochromatic plane wave with (a) $f = 250$ with $N = 1$, (b) $f = 500$ with $N = 2$, and (c) $f = 1$ kHz with $N = 3$. The contour plot shows the direction error (in degrees). The vector plot shows the local active intensity. The circle shows a circular region with a radius of 0.2 m. Note that the color bars are different across figures.
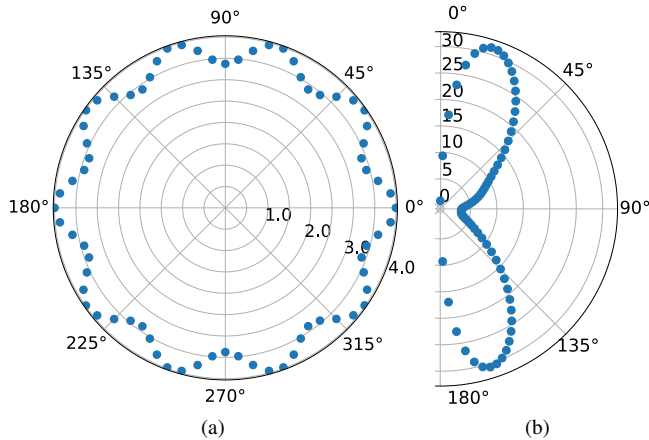

Fig. 3. Average absolute angular error (in degrees) for different directions of incidence for a monochromatic plane wave with $f = 500$ Hz and $N = 2$ in (a) horizontal plane ($\theta = \pi/2$ and $\phi \in [0, 2\pi)$), and (b) median plane ($\theta \in [0, \pi]$ and $\phi = 0$)

three tested frequencies respectively. Figures also indicate that the directional accuracy decreases as the frequency increases.

In order to assess the dependence of system accuracy on the direction of incidence we simulated monochromatic plane waves ($f = 500$ Hz) incident from different directions in the horizontal and the median planes. Fig. 3 shows average absolute angular errors (in degrees) within a sphere of radius 0.2 m as polar plots for incidences in the horizontal and median planes. The maximum order used in the extrapolation was $N = 2$. The figures indicate a good reproduction accuracy in terms of active intensity directions, specifically in the horizontal plane. Potentially improved performance could be achieved by including two loudspeakers positioned at the apex and nadir of the sphere, respectively. While average absolute angular error can be high for elevated sources, the perceptual impact is not likely to be high since localisation blur for directions above the horizontal plane are generally higher [15].

5. CONCLUSIONS AND FUTURE WORK

An extension of perceptual soundfield reconstruction to three dimensions was presented in this paper. Such an extension is not trivial with differential microphone arrays due to the required number of microphones and the necessity to equalize them. We proposed a conceptual framework that enables 3D-PSR recordings over rigid spherical microphone arrays via sound field extrapolation. Numerical simulations with monochromatic sound fields using the directivity pattern designed for 2D-PSR show that 3D-PSR via sound field extrapolation is feasible. Notice also that the proposed extrapolation method is not limited to PSR microphone arrays, but can be used with any near-coincident array, including those typically used in the audio engineering community [1].

The following directions of future work are envisioned. The directivity pattern used here was designed using interchannel time and level differences needed for accurate horizontal localization. Obtaining an appropriate directivity pattern taking into account psychoacoustics in the vertical-dimension is the subject of future work. Future work will also include different extrapolation methods as well as a formal analysis of the error measure of the extrapolation step, including white noise gain (WNG) analysis. We expect a larger error in extrapolated positions further away from the centre. This means there may be a trade-off between noise gain/extrapolation error and taking advantage of the perceptual improvement of time-intensity reproduction. Finally, a comparison with other recording/reproduction systems is necessary to demonstrate the utility of the proposed technique. Different loudspeaker arrangements will be explored too.

6. REFERENCES

[1] J. Eargle, *The microphone book*, Focal Press, 2004.

[2] J. Daniel, S. Moreau, and R. Nicol, "Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging," in *Audio Engineering Society Convention 114*. Audio Engineering Society, 2003.

[3] J. Meyer and G. W. Elko, "A spherical microphone array for spatial sound recording," *J. Acoust. Soc. Amer.*, vol. 111, no. 5, pp. 2346–2346, 2002.

[4] J. Merimaa and V. Pulkki, "Spatial impulse response rendering i: Analysis and synthesis," *J. Audio Eng. Soc.*, vol. 53, no. 12, pp. 1115–1127, 2005.

[5] V. Pulkki and J. Merimaa, "Spatial impulse response rendering ii: Reproduction of diffuse sound and listening tests," *J. Audio Eng. Soc.*, vol. 54, no. 1/2, pp. 3–20, 2006.

[6] V. Pulkki, "Spatial sound reproduction with directional audio coding," *J. Audio Eng. Soc.*, vol. 55, no. 6, pp. 503–516, 2007.

[7] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, "Spatial decomposition method for room impulse responses," *J. Audio Eng. Soc.*, vol. 61, no. 1/2, pp. 17–28, 2013.

[8] J. D. Johnston and Y. H. Lam, "Perceptual soundfield reconstruction," presented at the 109th Audio Eng. Soc. Conv., Preprint #2399, Los Angeles, CA, USA, September 2000.

[9] E. De Sena, H. Hacıhabiboğlu, and Z. Cvetković, "Analysis and design of multichannel systems for perceptual sound field reconstruction," *IEEE Trans. Audio, Speech, Language Proc.*, vol. 21, no. 8, pp. 1653–1665, August 2013.

[10] H. Hacıhabiboglu, E. De Sena, Z. Cvetkovic, J. Johnston, and J.O. Smith, "Perceptual spatial audio recording, simulation, and rendering: An overview of spatial-audio techniques based on psychoacoustics," *IEEE Signal Processing Magazine*, vol. 34, no. 3, pp. 36–54, 2017.

[11] V. Pulkki, "Virtual sound source positioning using vector-base amplitude panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, June 1997.

[12] E. De Sena and Z. Cvetković, "A computational model for the estimation of localisation uncertainty," in *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Process. (ICASSP-13)*, Vancouver, Canada, May 2013, pp. 388–392.

[13] E. De Sena, *Analysis, Design and Implementation of Multichannel Audio Systems*, Ph.D. thesis, King's College London, 2013.

[14] P. Samarasinghe, T. Abhayapala, and M. Poletti, "Wavefield analysis over large areas using distributed higher order microphones," *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 22, no. 3, pp. 647–658, 2014.

[15] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT Press, Cambridge, Massachusetts, 1997.

[16] N. V. Franssen, *Stereophony*, Philips Research Laboratories, 1964.

[17] M. Williams and G. Le Du, "Microphone array analysis for multichannel sound recording," presented at the 107th Audio Eng. Soc. Conv., Preprint #4997, New York, USA, September 1999.

[18] E. G. Williams, *Fourier acoustics: sound radiation and nearfield acoustical holography*, Academic Press, 1999.

[19] J. Meyer and G. W. Elko, "Spherical microphone arrays for 3d sound recording," in *Audio signal processing for next-generation multimedia communication systems*, Benesty J Huang Y, Ed., pp. 67–89. Springer US, New York, NY, USA, 2004.

[20] J. Benesty and C. Jingdong, *Study and design of differential microphone arrays*, vol. 6, Springer Science & Business Media, 2012.

[21] B. Rafaely, *Fundamentals of Spherical Array Processing*, vol. 8 of *Springer Topics in Signal Processing*, Springer-Verlag, Berlin, Heidelberg, Oct. 2015.

[22] E. De Sena, H. Hacıhabiboğlu, and Z. Cvetković, "A generalized design method for directivity patterns of spherical microphone arrays," in *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Process. (ICASSP-11)*, Prague, Czech Republic, May 2011.