SHARPENING OF ANGULAR SPECTRA BASED ON A DIRECTIONAL RE-ASSIGNMENT APPROACH FOR AMBISONIC SOUND-FIELD VISUALISATION

Leo McCormack, Archontis Politis, Ville Pulkki

Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland

ABSTRACT

A method for computing and sharpening angular spectra, derived from low-order ambisonic signals, is presented in this paper, which is intended for high-resolution directional sound-field visualisation. The method relies on a re-assignment principle, whereby the directional energy for each grid point is assigned to a new direction, which corresponds to a direction-of-arrival (DoA) estimate within a spatially-localised region, centred around the respective grid point. This leads to the concentration of energy around the true sources, and hence, to sharper angular spectra than that of steered response power (SRP) beamformers of maximum directivity, with the same order of ambisonic input. It is demonstrated that the proposed method, when using low-order input, can achieve similar results to the SRP approach of much higher order.

Index Terms— sound-field visualisation, spatial sharpening, Ambisonics, spherical harmonic domain

1. INTRODUCTION

Many popular sound-field visualisation methods rely on computing the angular spectrum, via a suitable localisation function evaluated over a dense grid of directions surrounding the reference point. For convenience, these localisation functions are often formulated in the spherical harmonic domain (SHD), whereby the input signals form an orthonormal representation of the sound-field, and are either synthesised or derived from microphone array signals [1, 2]. Examples of localisation functions adapted from microphone array processing to the SHD include: the steered-response power (SRP) [3], the MUltiple SIgnal Classification (MUSIC) pseudo-spectrum [4], and the side-lobe suppressed Cross-Pattern Coherence (CroPaC) algorithm [5]. The output of these localisation functions may be projected onto a two-dimensional (2D) representation of the sphere, and the relative energies, or statistical likelihood measures, may be depicted using a colour gradient. Bright spots in the angular spectrum infer directions with potential sources, or prominent early reflections, and peak finding algorithms may be employed to extract them numerically. Alternatively, the angular spectrum may be combined with a video stream from the same perspective; such systems are commonly referred to as *acoustic cameras* [5, 6].

This paper is primarily concerned with localisation functions that can model contributions from multiple narrow-band sources simultaneously. Subspace-methods, such as MUSIC, are often well-suited to this task; however, they require preliminary source number detection and are sensitive to off-grid DoAs and to coherent sources or reflections [7]. The CroPaC algorithm is generally more robust to coherent interferes [8], but requires higher orders to sufficiently suppress the side-lobes [5]. Regardless, selection of a dense grid resolution may render MUSIC and CroPaC computationally prohibiting for real-time applications, especially for large arrays. In such cases, alternative approaches should be employed, such as iterative grid refinement around the expected DoAs [9].

The simplest angular spectrum may be generated utilising SRP beamforming, which reduces to a weight-and-sum operation on the input spherical harmonic (SH) signals [3]; also referred to as ambisonic signals. This fundamental approach is computationally efficient and is unaffected by correlations between sources, as it is purely an estimation of directional power at certain directions. However, it is inherently limited by the spatial resolution of the available SH components, as determined by the ambisonic order, which often results in blurred spatial images at lower orders.

In this work, a method is proposed that attempts to retain the simplicity and efficiency of the elementary SRP approach, however, with an additional operation that sharpens the angular spectrum. The result is a higher resolution spatial image, which would normally have required a much higher order of input and a finer beamforming grid to produce. The approach is formulated in the SHD and is inspired by the time-frequency re-assignment principle for high-resolution imaging of time-frequency spectrograms [10]. Similar to time-frequency re-assignment, the proposed method first estimates the signal power for a certain point on the 2D manifold of directions, as given by the beamforming operation, while also estimating a new DoA for each focusing point. The energy of each point is then re-assigned to its new DoA estimate, resulting in a sharper angular spectrum. The re-assigned DoA is based on a spatially weighted acoustic intensity vector, which expresses propagating energy flow

Thanks to the Aalto ELEC Doctoral School for funding this research.

due to waves incident around the beamforming direction. Such a spatially-localised intensity vector was first used for directional analysis and sound-field reproduction in [11], and further formulated in [12, 13]. Recently, it has been used for DoA estimation using a histogram-clustering approach [14]. Herein, the sharpening operation is both non-parametric and preserves the directional energy of the recording, which is a desirable characteristic for visualisation and acoustic analysis purposes.

2. BACKGROUND

2.1. Sound-field model and angular spectrum

It is assumed that all sources are in the far-field, with respect to the array, and that the Ambisonic representation of the sound-field is modelled as a continuous distribution of plane waves with amplitudes $a(\boldsymbol{\gamma})$ incident to the array, where $\boldsymbol{\gamma}$ denotes the direction vector of incidence $\boldsymbol{\gamma} \equiv (\phi, \theta)$ at azimuth ϕ and elevation θ, respectively. The Nth-order spherical harmonic representation of this distribution is given by the spherical harmonic transform (SHT) as

$$\mathbf{a}_{\mathrm{N}}(t) = \mathscr{SHT}\{a(t,\boldsymbol{\gamma})\} = \int_{\boldsymbol{\gamma}} a(t,\boldsymbol{\gamma})\mathbf{y}_{\mathrm{N}}(\boldsymbol{\gamma})\,\mathrm{d}\boldsymbol{\gamma}, \quad (1)$$

where t is the time index, $\mathbf{a}_{\mathrm{N}} = [a_{00}, ..., a_{nm}, ..., a_{\mathrm{NN}}]^{\mathrm{T}}$ are the sound-field coefficients, and $\int_{\boldsymbol{\gamma}} \cdot \mathrm{d}\boldsymbol{\gamma}$ denotes an integration over all directions. The vector $\mathbf{y}_{\mathrm{N}}$ denotes the $(\mathrm{N}+1)^2$ real SHs Y_{nm} of mode-number (n,m), up to a maximum order N, as

$$\mathbf{y}_{\mathrm{N}}(\boldsymbol{\gamma}) = [Y_{00}(\boldsymbol{\gamma}), \cdots, Y_{nm}(\boldsymbol{\gamma}), \cdots, Y_{\mathrm{NN}}(\boldsymbol{\gamma})]^{\mathrm{T}}, \quad (2)$$

where $n = 0, 1, \cdots, \mathrm{N}$, $m = -n, \cdots, n$, and $(\cdot)^{\mathrm{T}}$ denotes the matrix or vector transpose. Spherical array processing in the SHD operates on the plane-wave coefficients $\mathbf{a}_{\mathrm{N}}$.

Beamforming in the SHD reduces to a weight-and-sum operation on the ambisonic signals $\mathbf{a}_{\mathrm{N}}$. More specifically, a frequency-invariant beamformer with a beampattern $c(\boldsymbol{\gamma})$ applied to the amplitude distribution, resulting in

$$s(t) = \int_{\boldsymbol{\gamma}} a(t,\boldsymbol{\gamma})c(\boldsymbol{\gamma})\,\mathrm{d}\boldsymbol{\gamma} = \mathbf{c}_{\mathrm{N}}^{\mathrm{T}}\mathbf{a}_{\mathrm{N}}(k), \quad (3)$$

where $\mathbf{c}_{\mathrm{N}} = \mathscr{SHT}\{c(\boldsymbol{\gamma})\}$ is the SHT of the beampattern. Axisymmetric patterns with their main lobe oriented at $\boldsymbol{\gamma}_0$, reduce to a simpler representation

$$c_{nm}(\boldsymbol{\gamma}_0) = c_n Y_{nm}(\boldsymbol{\gamma}_0), \quad (4)$$

or equivalently

$$\mathbf{c}_{\mathrm{N}}(\boldsymbol{\gamma}_0) = \mathbf{C}\mathbf{y}_{\mathrm{N}}(\boldsymbol{\gamma}_0), \quad (5)$$

where the diagonal matrix $\mathbf{C} = \mathrm{diag}\{c_0, c_1, c_1, c_1, ..., c_N\}$. An expression for the beampattern is

$$c(\boldsymbol{\gamma}_0, \boldsymbol{\gamma}) = \mathbf{y}_{\mathrm{N}}^{\mathrm{T}}(\boldsymbol{\gamma}_0)\mathbf{C}\mathbf{y}_{\mathrm{N}}(\boldsymbol{\gamma}) = \sum_{n=0}^{\mathrm{N}} \frac{(2n+1)}{4\pi} c_n P_n(\boldsymbol{\gamma}_0 \cdot \boldsymbol{\gamma}), \quad (6)$$

where P_n denotes the Legendre polynomials of degree n. When the beamforming coefficients are simply

$$c_n = 4\pi/(\mathrm{N}+1)^2 \quad \text{and} \quad \mathbf{c}_{\mathrm{N}}(\boldsymbol{\gamma}_0) = \frac{4\pi}{(\mathrm{N}+1)^2}\mathbf{y}_{\mathrm{N}}(\boldsymbol{\gamma}_0), \quad (7)$$

the beampatterns express a directional delta function band-limited to order N, and have the maximum directivity factor attainable for the given order $Q_{\mathrm{N}} = (\mathrm{N}+1)^2$. A plane wave decomposition (PWD) of the amplitude distribution can then be performed using such beamformers at a grid of $K \geq (\mathrm{N}+1)^2$ directions $\mathbf{G}_{\mathrm{PWD}} = [\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, ..., \boldsymbol{\gamma}_K]$ as

$$\mathbf{s}_{\mathrm{PWD}}(t) = \frac{4\pi}{(\mathrm{N}+1)^2}\mathbf{Y}_{\mathrm{G}}^{\mathrm{T}}\mathbf{a}_{\mathrm{N}}(t), \quad (8)$$

where $\mathbf{s}_{\mathrm{PWD}}(t) = [a(t,\boldsymbol{\gamma}_1), ..., a(t,\boldsymbol{\gamma}_K)]^{\mathrm{T}}$ are the estimated plane wave signals at the grid directions, and the matrix $\mathbf{Y}_{\mathrm{G}} = [\mathbf{y}_{\mathrm{N}}(\boldsymbol{\gamma}_1), ..., \mathbf{y}_{\mathrm{N}}(\boldsymbol{\gamma}_K)]$ contains SH values evaluated at the same grid directions.

For a dense grid of measurements $K >> (\mathrm{N}+1)^2$, an SRP-based estimate of the angular spectrum $\mathbf{p}_{\mathrm{SRP}}(t) = [p_1(t), ..., p_{\mathrm{K}}(t)]^{\mathrm{T}}$ can be computed based on the power of these plane wave signals

$$p_k(t) = \frac{1}{2T+1}\sum_{i=t-T}^{t+T} a^2(i,\boldsymbol{\gamma}_k), \quad (9)$$

evaluated at $\mathbf{G}_{\mathrm{PWD}}$, with $2T$ denoting the window length for the local power averaging. Subsequently, the angular spectrum $\mathbf{p}_{\mathrm{SRP}}(t)$ can be used for acoustic analysis, source localisation, and sound-field visualisation. It is evident that the spatial resolution of the spatial image, obtained through PWD, depends on the available order of the sound-field coefficients. Due to the width of lower-order beampatterns, sound sources can often appear spatially larger than in reality. The side-lobes also result in erroneous depictions of sound energy for directions that do not coincide with a source; these aberrations may be observed in the leftward examples of Fig. 1.

2.2. Spatially-localised acoustic intensity

The active-intensity vector expresses the mean flow of energy due to the effect of all sources and multi-path contributions in the sound scene, and has been extensively used for spatial sound reproduction [15, 16], and for localisation of dominant sound sources, through the analysis of its directional statistics [17, 18]. The notion of an intensity vector due to contributions in a spatially-localised region of the sound-field was introduced in [11]. It is expressed by the product of the acoustic pressure and acoustic particle velocity, due to a sound-field weighted by a desired beampattern $c(\boldsymbol{\gamma}_0, \boldsymbol{\gamma})$

$$\mathbf{i}_c(t,\boldsymbol{\gamma}_0) = -\frac{1}{2T+1}\sum_{i=t-T}^{t+T} s(i,\boldsymbol{\gamma}_0)\mathbf{s}_{\mathrm{xyz}}(i,\boldsymbol{\gamma}_0), \quad (10)$$

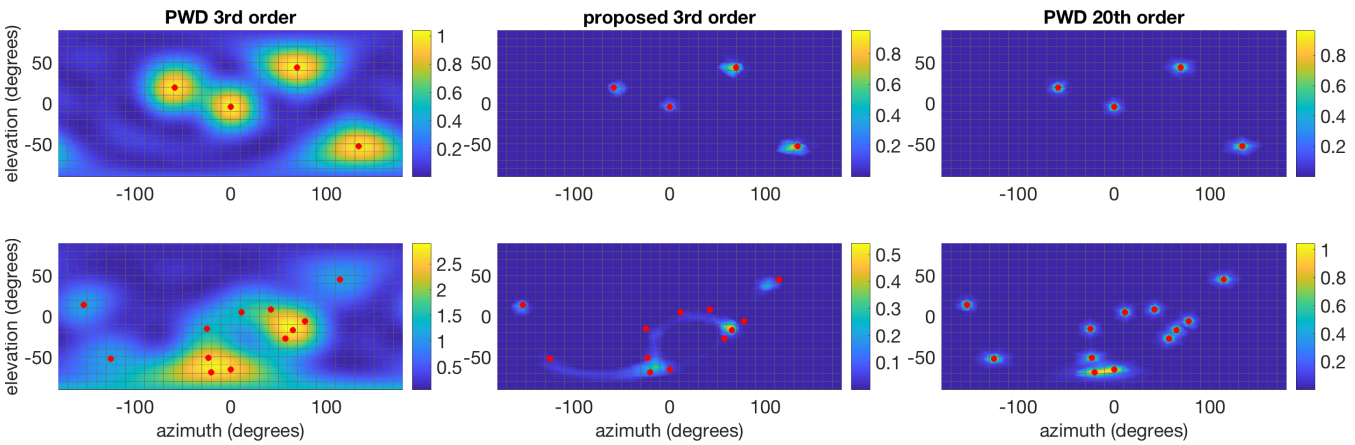

Fig. 1: Third-order Ambisonic sound-field visualisations with PWD (left) and the proposed method up-scaled to 20th order (middle), using a scanning grid comprised of 900 uniformly distributed points. A 20th order PWD map (right) is provided as a visual reference. Two infinite SNR sound scenes are depicted with 4 (top) and 12 (bottom) sources, respectively.

where $s(t,\boldsymbol{\gamma}_0)$ denotes the beamformer output for the focusing direction $\boldsymbol{\gamma}_0$. The signal vector $\mathbf{s}_{\text{xyz}} = [s_{\text{X}}, s_{\text{y}}, s_{\text{z}}]^{\text{T}}$ expresses the acoustic velocity signals for the weighted sound-field and is given by

$$\mathbf{s}_{\text{xyz}}(t,\boldsymbol{\gamma}_0) = \int_{\boldsymbol{\gamma}} c(\boldsymbol{\gamma}_0,\boldsymbol{\gamma}) a(t,\boldsymbol{\gamma}) \boldsymbol{\gamma}\, \text{d}\boldsymbol{\gamma}. \tag{11}$$

It is evident from (11) that to acquire these velocity signals, the required beampatterns are given by

$$\mathbf{c}_{\text{xyz}}(\boldsymbol{\gamma}_0,\boldsymbol{\gamma}) = c(\boldsymbol{\gamma}_0,\boldsymbol{\gamma})\boldsymbol{\gamma}, \tag{12}$$

which are essentially the three orthogonally-oriented dipoles $\boldsymbol{\gamma}$ modulated by the beampattern of the beamformer $c(\boldsymbol{\gamma}_0,\boldsymbol{\gamma})$. This beamforming operation can be expressed directly in the SHD by a $(\text{N}+1)^2 \times 3$ beamforming matrix $\mathbf{C}_{\text{xyz}}(\boldsymbol{\gamma}_0)$, such that

$$\mathbf{s}_{\text{xyz}}(t,\boldsymbol{\gamma}_0) = \mathbf{C}^{\text{T}}_{\text{xyz}}(\boldsymbol{\gamma}_0)\mathbf{a}_{\text{N}}(t). \tag{13}$$

Conveniently, due to the fixed nature of the three orthogonal dipoles, the matrix itself can be expressed as a product between a direction-independent fixed part $\mathbf{D}_{\text{x,y,z}}$ and the coefficients of the beamformer $\mathbf{c}_{\text{N}}$

$$\begin{aligned}\mathbf{C}_{\text{xyz}}(\boldsymbol{\gamma}_0) &= [\mathbf{c}_{\text{x}}(\boldsymbol{\gamma}_0)\, \mathbf{c}_{\text{y}}(\boldsymbol{\gamma}_0)\, \mathbf{c}_{\text{z}}(\boldsymbol{\gamma}_0)] \\ &= [\mathbf{D}_{\text{x}}\mathbf{c}(\boldsymbol{\gamma}_0)\, \mathbf{D}_{\text{y}}\mathbf{c}(\boldsymbol{\gamma}_0)\, \mathbf{D}_{\text{z}}\mathbf{c}(\boldsymbol{\gamma}_0)].\end{aligned} \tag{14}$$

The matrices $\mathbf{D}_{\text{x,y,z}}$ depend only on the SH order of the input signals, and may be pre-computed up to very high orders of interest for a given application. More information on the structure and construction of $\mathbf{D}_{\text{x,y,z}}$ can be found in [12], while MatLab code, which can generate them for arbitrary orders, can be found in [19].

3. DIRECTIONAL RE-ASSIGNMENT-BASED SRP

The proposed method is based on the SRP approach in the SHD, using the powers (9) of the extracted plane waves. However, instead of using the resulting angular spectrum for localisation or acoustic visualisation and analysis, each plane wave signal is re-assigned to a new DoA, computed through the spatially-localised intensity vector of (10). More specifically, for each PWD direction $\boldsymbol{\gamma}_k$, the following steps are performed:

(a) extract the amplitude $a(t,\boldsymbol{\gamma}_k)$ (8) and power P_k (9)

(b) compute the spatially-localised velocity signals $\mathbf{s}_{\text{xyz}}(t,\boldsymbol{\gamma}_k)$ (13)

(c) compute the spatially-localised active-intensity vector: $\mathbf{i}_c(t,\boldsymbol{\gamma}_k)$ from (10) and extract the re-assignment DoA for that grid direction by $\mathbf{n}_k = -\mathbf{i}_c(t,\boldsymbol{\gamma}_k)/||\mathbf{i}_c(t,\boldsymbol{\gamma}_k)||$

(d) re-assign a_k or P_k to the new DoA $\mathbf{n}_k$.

The manner to which this re-assignment is conducted depends on the target application. The most computationally efficient option is to assign the powers P_k to the target visualisation grid by quantising the re-assignment DoA $\mathbf{n}_k$ to the closest target grid point. The visualisation grid, in this case, can be of much higher resolution than the original SRP grid resolution. In cases where two re-assignment DoAs coincide, the respective powers are simply summed to the target grid point. In the trivial case of a single plane wave source in the scene, all re-assignment DoAs will point to the source direction, independently of the beamforming direction, and hence all SRP powers are combined at the same point [11]. For a proper power preservation in this case, and assuming a uniformly arranged set of PWD/SRP directions, the beamformers should

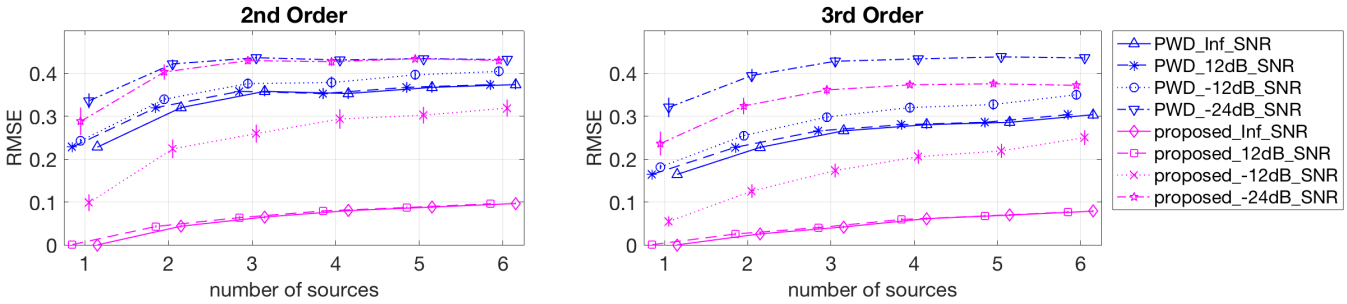

Fig. 2: Means and 95% confidence intervals for the RMSE of the residual, derived by subtracting the angular spectra of the PWD and proposed method from a 20th order PWD reference.

be normalised with the energy-preserving condition

$$\beta_{\mathrm{EP}} \sum_{k=1}^{K} |c(\boldsymbol{\gamma}_k, \boldsymbol{\gamma})|^2 = 1 \quad \text{with} \quad \beta_{\mathrm{EP}} = \frac{(\mathrm{N}+1)^2}{K}. \tag{15}$$

A more computationally costly alternative, which results in a smoother visualisation, is to re-encode the re-assigned plane wave signals into a higher-order $\mathrm{L} > \mathrm{N}$ of spatial resolution, and subsequently compute the angular spectrum via a normal SRP on the resulting sharpened SH signals $\mathbf{a}_{\mathrm{L}}$

$$\mathbf{a}_{\mathrm{L}}(t) = \sum_{k=1}^{K} a_k(t, \boldsymbol{\gamma}_k) \mathbf{y}_{\mathrm{L}}(\mathbf{n}_k). \tag{16}$$

In this case, the original SRP beamformers should meet an amplitude-preserving condition, such that

$$\beta_{\mathrm{AP}} \sum_{k=1}^{K} c(\boldsymbol{\gamma}_k, \boldsymbol{\gamma}) = 1 \quad \text{with} \quad \beta_{\mathrm{AP}} = \frac{4\pi}{K}. \tag{17}$$

Some examples, when utilising third-order input and a 900 point uniform scanning grid, are shown in Fig. 1, for both PWD and the proposed method up-scaled to 20th order.

4. EVALUATION AND DISCUSSION

The evaluation consisted of a comparison between the angular spectra derived using a low-order PWD and the proposed method, against a high-order PWD reference case. Angular spectra were generated from synthetic ambisonic sound scenes, with between 1 and 6 sources and differing signal-to-noise ratios (SNR), using both second and third-order representations. The diffuse component of the sound-field was synthesised using 1442 approximately uniformly distributed uncorrelated noise sources, and introduced into the simulation in such a manner as to attain the target SNRs.

The proposed method case was up-scaled to 20th order as in (16). The angular spectra computed for each scenario were normalised and subtracted from the 20th order PWD spectra, and the root-mean-square error (RMSE) of the residual was computed. In total, 100 randomly generated source combinations were used for the simulation; the directions for which were derived via random indices into a 5100 point t-design [20]. The means and 95% confidence intervals for the RMSE for each source number and SNR are shown in Fig. 2.

It can be observed that the proposed approach yields consistently lower error than PWD for almost all test cases, however, the performance benefit of the proposed method is inversely proportional to the SNR of the scene. The proliferation of error for the proposed approach in low SNR sound scenes, may be attributed to the reduction in DoA estimation accuracy in such conditions; however, the proposed approach appears to never be worse than PWD. Furthermore, for cases in which the sources are in close proximity to each other, the directional energy may be migrated to an area between them, which can be observed for the 12 sources case in Fig. 1. The DoA estimates under low SNR conditions are also generally biased towards the centre of the spatially-localised beampattern; an analytical explanation of this phenomenon is a topic of future research.

5. CONCLUSION

This work has presented an ambisonic sound-field visualisation approach, which combines the simplicity and efficiency of the SRP approach with an additional sharpening operation, based on the re-assignment of directional energy to local estimates of the DoA. It is demonstrated that, when compared with a high-order reference SRP case, the proposed method performs consistently better than the SRP approach for the majority of cases.

Unlike many high-resolution alternatives, the proposed method does not require an estimation of the number of sources and makes no assumptions regarding the sound-field conditions. Since the method is purely based on energy, the method also works with coherent sources. Furthermore, much of the required computations may be performed during an initialisation stage, therefore, the run-time computational complexity remains relatively low.

6. REFERENCES

[1] Boaz Rafaely, *Fundamentals of spherical array processing*, Springer, 2015.

[2] Daniel P Jarrett, Emanuël AP Habets, and Patrick A Naylor, *Theory and applications of spherical microphone array processing*, Springer, 2017.

[3] Daniel P Jarrett, Emanuël AP Habets, and Patrick A Naylor, "3D source localization in the spherical harmonic domain using a pseudointensity vector.," in *EUSIPCO*, 2010.

[4] Heinz Teutsch, *Modal array signal processing: principles and applications of acoustic wavefield decomposition*, Springer, 2007.

[5] Leo McCormack, Symeon Delikaris-Manias, and Ville Pulkki, "Parametric acoustic camera for real-time sound capture, analysis and tracking," in *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17)*, 2017, pp. 412–419.

[6] Adam O'Donovan, Ramani Duraiswami, and Dmitry Zotkin, "Imaging concert hall acoustics using visual and audio cameras," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE, 2008, pp. 5284–5287.

[7] Björn Ottersten, Mats Viberg, Petre Stoica, and Arye Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing," in *Radar array processing*, pp. 99–151. Springer, 1993.

[8] Symeon Delikaris-Manias and Ville Pulkki, "Cross pattern coherence algorithm for spatial filtering applications utilizing microphone arrays," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 11, pp. 2356–2367, 2013.

[9] Kaluri V Rangarao and Shridhar Venkatanarasimhan, "gold-MUSIC: A variation on music to accurately determine peaks of the spectrum," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 2263–2268, 2013.

[10] Patrick Flandrin, Francois Auger, and Eric Chassande-Mottin, "Time-frequency reassignment: from principles to algorithms," in *Applications in Time-Frequency signal processing*, p. 102. CRC Press, 2003.

[11] Archontis Politis, Juha Vilkamo, and Ville Pulkki, "Sector-based parametric sound field reproduction in the spherical harmonic domain," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 5, pp. 852–866, 2015.

[12] Archontis Politis and Ville Pulkki, "Acoustic intensity, energy-density and diffuseness estimation in a directionally-constrained region," *arXiv:1609.03409*, 2016.

[13] Leo McCormack, Symeon Delikaris-Manias, Angelo Farina, Daniel Pinardi, and Ville Pulkki, "Real-time conversion of sensor array signals into spherical harmonic signals with applications to spatially localised sub-band sound-field analysis," in *Audio Engineering Society Convention 144*. Audio Engineering Society, 2018.

[14] Symeon Delikaris-Manias, Despoina Pavlidi, Athanasios Mouchtaris, and Ville Pulkki, "DOA estimation with histogram analysis of spatially constrained active intensity vectors," in *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, 2017, pp. 526–530.

[15] Ville Pulkki, Archontis Politis, Mikko-Ville Laitinen, Juha Vilkamo, and Jukka Ahonen, "First-order directional audio coding (dirac)," in *Parametric Time-Frequency Domain Spatial Audio*, pp. 89–138. John Wiley & Sons, 2017.

[16] Archontis Politis, Leo McCormack, and Ville Pulkki, "Enhancement of ambisonic binaural reproduction using directional audio coding with optimal adaptive mixing," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics,*, 2017.

[17] Sakari Tervo, "Direction estimation based on sound intensity vectors," in *17th European Signal Processing Conference (EUSIPCO)*, 2009, pp. 700–704.

[18] Dovid Levin, Emanuël AP Habets, and Sharon Gannot, "On the angular error of intensity vector based direction of arrival estimation in reverberant sound fields," *The Journal of the Acoustical Society of America*, vol. 128, no. 4, pp. 1800–1811, 2010.

[19] Archontis Politis, *Microphone array processing for parametric spatial audio techniques*, Ph.D. thesis, Aalto University, 2016, https://github.com/polarch/Spherical-Array-Processing.

[20] Manuel Gräf and Daniel Potts, "On the computation of spherical designs by a new optimization approach based on fast spherical fourier transforms," *Numerische Mathematik*, vol. 119, no. 4, pp. 699–724, 2011.