DIRECTION OF ARRIVAL ESTIMATION USING PSEUDO-INTENSITY VECTORS WITH DIRECT-PATH DOMINANCE TEST

Alastair H. Moore, Christine Evers*, Patrick A. Naylor*, David L. Alon†and Boaz Rafaely†*

*Imperial College London
Dept. of Electrical and Electronic Engineering

†Ben-Gurion University of the Negev
Dept. of Electrical and Computer Engineering

ABSTRACT

The accuracy of direction of arrival estimation tends to degrade under reverberant conditions due to the presence of reflected signal components which are correlated with the direct path. The recently proposed direct-path dominance test provides a means of identifying time-frequency regions in which a single signal path is dominant. By analysing only these regions it was shown that the accuracy of the FS-MUSIC algorithm could be significantly improved. However, for real-time implementation a less computationally demanding localisation algorithm would be preferable. In the present contribution we investigate the direct-path dominance test as a pre-processing step to pseudo-intensity vector-based localisation. A novel formulation of the pseudo-intensity vector is proposed which further exploits the direct path dominance test and leads to improved localisation performance.

Index Terms— direction of arrival estimation, spherical harmonic domain, pseudo-intensity vectors

1. BACKGROUND

Direction-of-Arrival (DOA) estimation (also known as bearing-only source localisation) is a fundamental problem in acoustic signal processing, particularly as a preprocessing step for beamforming and speech dereverberation [1]. In the context of robot audition it is important for estimation to be computationally efficient such that source localisation can be performed in real time and with low latency on relatively low cost hardware. In this paper we assume a spherical microphone array such that processing can be performed in the Spherical Harmonic Domain (SHD). This is advantageous because beampatterns can be created which are independent of the look direction making the system equally adept at localising sources to the rear and above as those to the front.

Localisation using Steered Response Power (SRP) measures the output power from a beamformer as it is steered in turn at a grid of possible source directions and selects one or more maxima as DOA(s). Under anechoic conditions the Minimum Variance Distortionless Response (MVDR) beamformer achieves the optimal measure of signal power in the look direction. However, when reverberation is present correlated signal components arriving from other directions degrade the beampattern and with it the spatial resolution of the SRP map. Subspace localisation methods similarly suffer in the presence of coherent reflections.

Frequency smoothing [2] is a well established technique for decorrelating coherent reflections by combining information across multiple frequency bands. It has an elegant formulation in the Spherical Harmonic Domain (SHD) due to frequency independence of the array manifold and has been applied to localisation of individual reflections using Multiple Signal Classification (MUSIC) [3], Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [4] and SRP [5].

For MUSIC [6] the noise space of the correlation matrix is used to evaluate the spatial spectrum for a grid of possible source directions. The effective rank of the correlation matrix yields the number of independent sources. In principle, exploiting the sparsity of speech, one can perform accurate DOA estimation for multiple sources by computing the MUSIC spectrum independently for each TF-region whose correlation matrix has effective rank one. However, coherent reflections are not independent and so the effective rank is less than the total number of sources and reflections. To improve the identification of TF-regions whose effective rank is one, the coherence test proposed in [7] averages the correlation matrix over time. The Direct-Path Dominance (DPD) test [8] leads to further improvement using frequency smoothing by also averaging the correlation matrix over frequency thus decorrelating coherent reflections.

In contrast to SRP and MUSIC, the Pseudo-Intensity Vector (PIV) approach to DOA estimation [9] does not evaluate a cost function over a grid of possible source directions. Instead it calculates a direct estimate of an intensity vector which represents the direction and magnitude of the propagating wavefront. It therefore offers DOA estimation with no trade-off between spatial resolution and computational cost. In general it has been found that averaging PIVs over time and frequency is sufficient to produce DOA estimates which are robust to noise and moderate levels of diffuse reverberation. Further-

more, multiple sources can be localised, for example, using spherical k-means clustering [10]. However, like other approaches, its accuracy degrades in the presence of strong coherent reflections. One might hypothesise that using the DPD test as a preprocessing step would improve the accuracy of PIV-based DOA estimation.

The contributions of this paper are 1) to evaluate the accuracy of PIVs using only TF-regions which pass the DPD test and 2) to propose a new formulation of the PIVs calculated directly from the time and frequency smoothed correlation matrices.

2. PROBLEM FORMULATION

2.1. Spherical harmonic domain signal model

Consider a soundfield composed of L plane waves sampled by a spherical microphone array of radius r with Q sensors. The sound pressure at the q-th microphone, located at angle $\Omega_q = (\theta_q, \phi_q)$ is given in the frequency domain by

$$p(k, r, \Omega_q) = \sum_{l=1}^{L} v(k, \Omega_q, \Psi_l)s_l(k) + n_q(k) \qquad (1)$$

where k is the wavenumber, $s_l(k)$ denotes the complex amplitude of the l-th plane wave at the origin and the array manifold, $v(k, \Omega, \Psi)$, describes the response at Ω, relative to that at the origin, to a plane wave with incident angle Ψ. Spatially white sensor noise is represented by $n_q(k)$.

The Spherical Fourier Transform (SFT) over the surface of the sphere can be approximated up to spherical harmonic order N as the weighted sum of the microphone signals in the frequency domain

$$p_{nm}(k, r) \approx \sum_{q=1}^{Q} w_q p(k, r, \Omega_q) \left[Y_n^m(\Omega_q)\right]^*, n \le N, |m| \le n \qquad (2)$$

where Y_n^m is the spherical harmonic of order n and degree m and $\{w_q\}_1^Q$ are the weights of the sampling scheme [11]. The approximation is valid provided $kr < N$, the $Q \ge (N+1)^2$ sensors are approximately equally distributed over the sphere and the sampling weights are chosen appropriately [12].

The SFT of a single plane wave $s_l(k)$ is [13]

$$p_{nm}^{(l)}(k, r) = b_n(kr) \left[Y_n^m(\Psi_l)\right]^* s_l(k) \qquad (3)$$

where the mode strength $b_n(kr)$ is determined by the array configuration (e.g. whether it is rigid or open). Dividing out the mode strength removes the frequency dependence of the array manifold yielding mode strength compensated eigenbeams whose direction dependence is independent of frequency

$$\tilde{p}_{nm}^{(l)}(k, r) = \left[Y_n^m(\Psi_l)\right]^* s_l(k) \qquad (4)$$

though care must be taken that $b_n(kr)$ is invertible at the frequencies of interest. We combine all $(N+1)^2$ eigenbeams into a vector

$$\tilde{\mathbf{p}}_{\mathbf{nm}}^{(l)}(k) = \left[\tilde{p}_{00}^{(l)} \, \tilde{p}_{1(-1)}^{(l)} \, \tilde{p}_{10}^{(l)} \, \tilde{p}_{11}^{(l)} \, \cdots \tilde{p}_{NN}^{(l)}\right]^T \qquad (5)$$

where superscript T denotes the matrix transpose operator.

Considering now the spatially white sensor noise $n_q(k)$, performing the SFT followed by mode strength compensation gives spherical harmonic coefficients

$$\tilde{n}_{nm}(k) = \frac{1}{b_n(kr)} \sum_{q=1}^{Q} w_q \left[Y_n^m(\Omega_q)\right]^* n_q(k)$$

which can similarly be expressed as a vector

$$\tilde{\mathbf{n}}_{\mathbf{nm}}(k) = \left[\tilde{n}_{00} \, \tilde{n}_{1(-1)} \, \tilde{n}_{10} \, \tilde{n}_{11} \, \ldots \tilde{n}_{NN}\right]^T. \qquad (6)$$

The original signal model in (1) can thus be expressed in the SHD as

$$\mathbf{a}_{\mathbf{nm}}(k) = \sum_{l=1}^{L} \tilde{\mathbf{p}}_{\mathbf{nm}}^{(l)}(k) + \tilde{\mathbf{n}}_{\mathbf{nm}}(k). \qquad (7)$$

For the analysis of time varying signals it is convenient to work in the Short Time Fourier Transform (STFT) domain. We therefore reformulate (7) as

$$\mathbf{a}_{\mathbf{nm}}(\nu, \tau) = \sum_{l=1}^{L} \tilde{\mathbf{p}}_{\mathbf{nm}}^{(l)}(\nu, \tau) + \tilde{\mathbf{n}}_{\mathbf{nm}}(\nu, \tau) \qquad (8)$$

where the discrete frequencies and frame times, denoted by ν and τ, respectively, follow directly from the choice of frame length and hop size.

2.2. Direct path dominance

An estimate of the spatial correlation matrix is given by [8]

$$\tilde{\mathbf{R}}_a(\nu, \tau) = \frac{1}{J} \sum_{j_\nu = -J_\nu}^{J_\nu} \sum_{j_\tau = -J_\tau}^{J_\tau} \mathbf{a}_{\mathbf{nm}}(\nu + j_\nu, \tau + j_\tau)$$
$$\times \mathbf{a}_{\mathbf{nm}}^H(\nu + j_\nu, \tau + j_\tau) \qquad (9)$$

where $J = (2J_\nu + 1)(2J_\tau + 1)$. Averaging across time alone approximates the expected value but the maximum rank of the matrix would be limited to the number of non-coherent sources (since a direct path and its reflection(s) are linearly dependent over τ.) Local frequency smoothing increases the rank of the matrix up to the total number of plane waves present (whether coherent or not) since the energy in each frequency band at a given time is independent of the other frequency bands. The fact that this frequency smoothing can be achieved simply as an average over frequency is due to the frequency-independence of the array manifold SHD [3]. If the effective rank of $\tilde{\mathbf{R}}_a(\nu, \tau)$ is unity then we assume that

the spatial correlation matrix is dominated by energy from a single direction. Reflections always arrive after the direct path and tend to be at a reduced amplitude so are unlikely to be observed in isolation. It is therefore reasonable to assume that the direction corresponds to the direct path hence the test is named the Direct-Path Dominance (DPD) test. The set of time-frequency indices which pass the DPD test is defined as

$$\mathcal{A}_{\text{DPDtest}} = \left\{ (\nu, \tau) : \text{erank}\left(\tilde{\mathbf{R}}_a(\nu, \tau) \right) = 1 \right\} \quad (10)$$

where erank denotes the effective rank.

For those TF-regions which pass the DPD test, $\tilde{\mathbf{R}}_a(\nu, \tau)$ can be separated by Singular Value Decomposition (SVD) into a one-dimensional signal space, $\mathbf{U}_s$, and $(N+1)^2 - 1$ dimensional noise space, $\mathbf{U}_n$, according to

$$\tilde{\mathbf{R}}_a = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^H = [\mathbf{U}_s \mathbf{U}_n] \begin{bmatrix} \boldsymbol{\Sigma}_s & 0 \\ 0 & \boldsymbol{\Sigma}_n \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \quad (11)$$

where the dependence on (ν, τ) is omitted for brevity. In [8] the spatial spectrum is calculated from $\mathbf{U}_n$ using MUSIC for all $\tilde{\mathbf{R}}_a(\nu, \tau)$ whose effective rank is unity. These individual spectra are fused to form a single spectrum whose peaks correspond to the DOAs of the multiple sound sources. The resolution of the DOAs estimation is therefore dependent on the number of directions in which the spatial spectrum is evaluated.

2.3. Pseudo-Intensity Vectors

Acoustic intensity describes the flow of energy per unit area in terms of its magnitude and direction. In [9] an approximation based on 0- and 1-order eigenbeams called the Pseudo-Intensity Vector (PIV) was defined as

$$\mathbf{I_a}(\nu, \tau) = \frac{1}{2} \mathcal{R} \left\{ \tilde{p}_{00}^*(\nu, \tau) \begin{bmatrix} D_x(\nu, \tau) \\ D_y(\nu, \tau) \\ D_z(\nu, \tau) \end{bmatrix} \right\} \quad (12)$$

where $\tilde{p}_{00}(\nu, \tau)$ is the omnidirectional pressure and

$$D_\alpha(\nu, \tau) = \sum_{m=-1}^{1} Y_1^m(\varphi_\alpha)\tilde{p}_{1(m)}(\nu, \tau), \qquad \alpha \in x, y, z \quad (13)$$

are dipoles steered in the negative x, y and z directions, given by $\varphi_x = (\pi/2, \pi)$, $\varphi_y = (\pi/2, -\pi/2)$ and $\varphi_z = (\pi, 0)$. Assuming a single plane wave is present, $\mathbf{I_a}(\nu, \tau)$ points in the direction of propagation so a unit vector indicating the DOA is simply $-\mathbf{I_a}(\nu, \tau)/\|\mathbf{I_a}(\nu, \tau)\|$, where $\|\cdot\|$ indicates the ℓ_2 norm. Averaging $\mathbf{I_a}(\nu, \tau)$ across frequency provides some robustness to noise and diffuse reverberation [9], but strong coherent reflections lead to systematic errors.

We consider two approaches to exploiting the DPD test to improve PIV-based source localisation. The first, DPD-PIV, calculates the mean PIV for each TF-region selected by DPD. The second, DPD-SS-PIV, formulates a new PIV using the signal space of the selected correlation matrices.

2.3.1. DPD-PIV

Let the i-th entry in $\mathcal{A}_{\text{DPDtest}}$ correspond to the TF-bin given by (ν_i, τ_i). From (9) the corresponding TF-region is given by $\nu_i - J_\nu \leq \nu_i + j_\nu \leq \nu_i + J_\nu$, $\tau_i - J_\tau \leq \tau_i + j_\tau \leq \tau_i + J_\tau$. We define the set of Direct-Path Dominant Pseudo-Intensity Vectors (DPD-PIVs) as

$$\mathbf{I}_{\text{DPD}}(\nu_i, \tau_i) = \frac{1}{J} \sum_{j_\nu=-J_\nu}^{J_\nu} \sum_{j_\tau=-J_\tau}^{J_\tau} \mathbf{I_a}(\nu_i + j_\nu, \tau_i + j_\tau),$$
$$\forall i \in \mathcal{A}_{\text{DPDtest}}. \quad (14)$$

2.3.2. DPD-SS-PIV

The signal space $\mathbf{U}_s(\nu, \tau)$ obtained from (11) describes the (single) dominant signal in the SHD

$$\mathbf{U}_s(\nu, \tau) = \begin{bmatrix} \tilde{u}_{00} \ \tilde{u}_{1(-1)} \ \tilde{u}_{10} \ \tilde{u}_{11} \ \ldots \ \tilde{u}_{NN} \end{bmatrix}^T. \quad (15)$$

We formulate the Direct-Path Dominant Signal Space Pseudo-Intensity Vectors (DPD-SS-PIVs) directly as

$$\mathbf{I}_{\text{SS-DPD}}(\nu_i, \tau_i) = \frac{1}{2} \mathcal{R} \left\{ \tilde{u}_{00}^*(\nu_i, \tau_i) \begin{bmatrix} \dot{D}_x(\nu_i, \tau_i) \\ \dot{D}_y(\nu_i, \tau_i) \\ \dot{D}_z(\nu_i, \tau_i) \end{bmatrix} \right\},$$
$$\forall i \in \mathcal{A}_{\text{DPDtest}} \quad (16)$$

where

$$\dot{D}_\alpha(\nu, \tau) = \sum_{m=-1}^{1} Y_1^m(\varphi_\alpha)\tilde{u}_{1(m)}(\nu, \tau), \qquad \alpha \in x, y, z. \quad (17)$$

In contrast to MUSIC, which evaluates a potentially dense grid of candidate directions to find that which is "most orthogonal" to the noise space, DPD-SS-PIV directly estimates the source direction from the signal space.

3. EXPERIMENTAL VALIDATION

The performance of the DPD-PIV and DPD-SS-PIV methods are evaluated using two illustrative simulation setups. In the first, localisation for a single source with one coherent reflection is considered. In the second we demonstrate localisation of three simultaneously active sources in a reverberant room. Results are compared to localisation using PIVs without DPD preselection of TF-bins and to DPD-MUSIC as presented in [8]. To allow a graphical comparison with the DPD-MUSIC spatial spectrum, PIV DOA estimates are visualised as 2D histograms.

Acoustic Impulse Responses (AIRs) for a 32-element rigid spherical microphone array with radius 4.2 cm were simulated with sampling frequency 8 kHz using a modification of the image-source method [14, 15] and convolved with
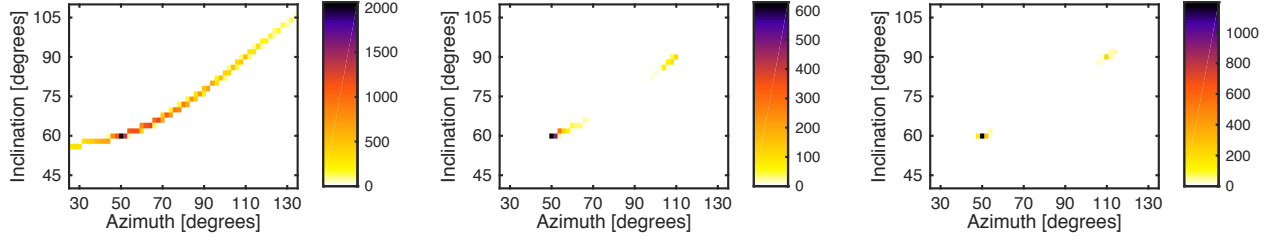
Fig. 1. 2D histograms showing number of PIV DOA estimates in each DOA bin for experiment 1 with raw PIVs (left), DPD-PIVs (middle) and DPD-SS-PIVs (right). The direct path is from $(60°, 50°)$ and single coherent reflection from $(90°, 110°)$.

speech signals from the APLAWD database [16] with level normalised according to ITU-T P.56 [17]. To focus on the effect of reverberation the noise-free case was considered. The microphone signals were transformed to the STFT domain using Hamming windowed frames of length 128 samples (16 ms) and frame increment of 32 samples (4 ms). Retaining frequency bins between 1500 and 3875 Hz, the SFT (up to order $N = 3$) and mode strength compensation yield $\mathbf{a_{nm}}(\nu, \tau)$ from which $\mathbf{I_a}(\nu, \tau)$ follows directly. To calculate $\tilde{\mathbf{R}}_a(\nu, \tau)$ we use $J_\nu = 2$ and $J_\tau = 12$. The DPD test criteria was [8]

$$\mathrm{erank}\left(\tilde{\mathbf{R}}_a(\nu, \tau)\right) = 1 \quad \mathrm{if} \quad \frac{\sigma_1(\nu, \tau)}{\sigma_2(\nu, \tau)} > \varepsilon$$

where $\sigma_1(\nu, \tau)$ and $\sigma_2(\nu, \tau)$ are the largest and second largest singular values, respectively, of $\tilde{\mathbf{R}}_a(\nu, \tau)$. The optimal choice of ε depends on the combination of STFT window parameters, J_ν and J_τ and is a trade-off between the number of selected TF-regions and the possibility of erroneously accepting TF-regions whose signal space is greater than unity. For the parameter combinations used in this study $\varepsilon = 10$ was found to offer a good compromise.

Experiment 1: To simulate a coherent reflection the same 4-second speech signal was convolved with the anechoic AIR of a plane wave from $(60°, 50°)$ and $(90°, 110°)$ with the latter 'reflection' delayed by 20 ms and attenuated with gain of 0.6.

Experiment 2: To simulate multiple speakers in a reverberant room, the microphone array was centred at cartesian co-ordinates (2.5 m, 3.0 m, 1.5 m) in a 4 m×5 m×3 m rectangular room with reverberation time (T_{60}) of 0.5 s. Eight seconds of continuous speech was generated for three different female speakers by concatenating several utterances. These were presented from a distance of 1.5 m with DOAs $(80°, 140°)$, $(100°, 180°)$ and $(80°, 220°)$.

4. RESULTS AND DISCUSSION

Figure 1 shows the localisation results for experiment 1. Without any preselection of TF-bins the PIV produces estimates which are the vector sum of the direct and reflection

DOAs. Both the DPD-PIV and DPD-SS-PIV methods produce DOA estimates which are localised to mainly the direct path DOA and to a lesser extent the reflected path DOA. Note that the reflected path is only localised because in this artificial case the reflection is the only signal present at speech offsets. Under natural reverberation no single reflection path would dominate in this way.

Figure 2 shows the localisation results for experiment 2 using the three different PIV methods. Without DPD preselection of TF-bins the PIVs form a single broad cluster around the three direct path DOAs. For the DPD-PIV and DPD-SS-PIV methods the DOA estimates form three clusters around the true source directions, with the DPD-SS-PIV producing tighter clusters. As a comparison, Fig. 3 shows the spatial spectrum produced by DPD-MUSIC [8] using a $2°$ grid in azimuth and inclination (16,200 look directions). The three clear peaks correspond precisely with the source positions. The accuracy of the DOA estimates can be quantified by the number of selected TF-regions which produce DOA estimates within $5°$ of a true source. Thus defined, the four methods achieve 1.5% (PIV), 50% (DPD-PIV), 72% (DPD-SS-PIV) and 86% (DPD-MUSIC) accuracy. DPD-MUSIC is the most accurate because it uses information from the 2- and 3-order spherical harmonics, whereas are limited to 0- and 1-order. The benefit of DPD-SS-PIV is that, following SVD, it requires only 4 complex multiplications per TF-region making it a strong candidate for real time implementation.

5. CONCLUSIONS

Two new PIV-based methods for DOA estimation in reverberant acoustics have been proposed which exploit the DPD test proposed in [8]. Experimental results presented for the challenging case of multiple sources demonstrate significantly improved localisation accuracy for DPD-PIV (50%) and DPD-SS-PIV (72%) over a baseline (1.5%) in which DOAs are obtained from individual TF-bins without DPD preselection. It is proposed that computational savings of our new method significantly outweigh the slight reduction in accuracy compared to DPD-MUSIC (86%) for many resource-constrained real-world applications, such as robot audition.
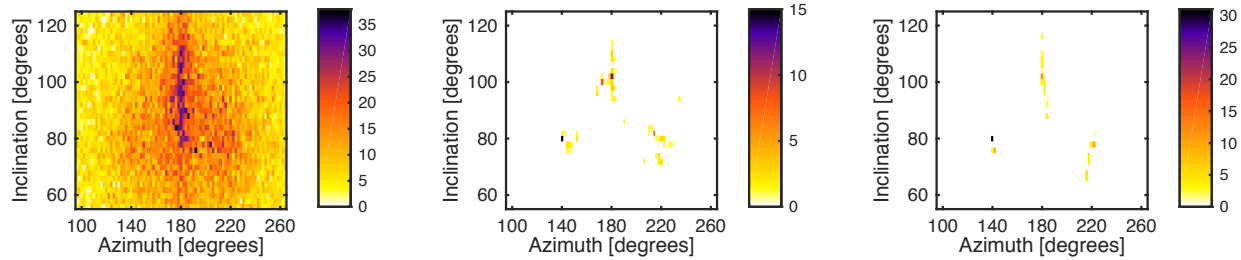
Fig. 2. 2D histograms showing number of PIV DOA estimates in each DOA bin for experiment 2 with raw PIVs (left), DPD-PIVs (middle) and DPD-SS-PIVs (right). Ground truth DOAs are (80°, 140°), (100°, 180°) and (80°, 220°).
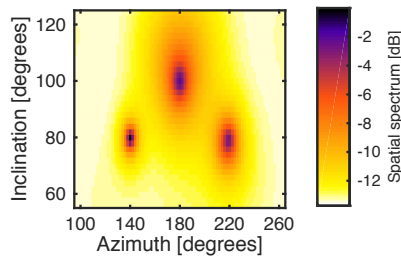
Fig. 3. DPD-MUSIC spectrum for experiment 2 with ground truth DOAs: (80°, 140°), (100°, 180°) and (80°, 220°).

REFERENCES

[1] P. A. Naylor and N. D. Gaubitch, Eds., *Speech Dereverberation*, Springer, 2010.

[2] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 823–831, Aug. 1985.

[3] D. Khaykin and B. Rafaely, "Coherent signals direction-of-arrival estimation using a spherical microphone array: Frequency smoothing approach," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Oct. 2009, pp. 221–224.

[4] H. Sun, H. Teutsch, E. Mabande, and W. Kellermann, "Robust localization of multiple sources in reverberant environments using EB-ESPRIT with spherical microphone arrays," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

[5] H. Sun, E. Mabande, K. Kowalczyk, and W. Kellermann, "Localization of distinct reflections in rooms using spherical microphone array eigenbeam processing," *J. Acoust. Soc. Am.*, vol. 131, pp. 2828–2840, 2012.

[6] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[7] S. Mohan, M. E. Lockwood, M. L. Kramer, and D. L. Jones, "Localization of multiple acoustic sources with small arrays using a coherence test," vol. 123, pp. 2136–2147, 2008.

[8] O. Nadiri and B. Rafaely, "Localization of multiple speakers under high reverberation using a spherical microphone array and the direct-path dominance test," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 10, pp. 1494–1505, Oct. 2014.

[9] D. P. Jarrett, E. A. P. Habets, and P. A. Naylor, "3D source localization in the spherical harmonic domain using a pseudointensity vector," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Aalborg, Denmark, Aug. 2010, pp. 442–446.

[10] C. Evers, A. H. Moore, and P. A. Naylor, "Multiple source localisation in the spherical harmonic domain," in *Proc. Intl. Workshop Acoust. Signal Enhancement (IWAENC)*, Nice, France, July 2014.

[11] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 1, pp. 135–143, Jan. 2005.

[12] B. Rafaely, B. Weiss, and E. Bachmat, "Spatial aliasing in spherical microphone arrays," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 1003–1010, Mar. 2007.

[13] B. Rafaely, "Plane-wave decomposition of the pressure on a sphere by spherical convolution," *J. Acoust. Soc. Am.*, vol. 116, no. 4, pp. 2149–2157, Oct. 2004.

[14] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.

[15] D. P. Jarrett, E. A. P. Habets, M. R. P. Thomas, and P. A. Naylor, "Rigid sphere room impulse response simulation: algorithm and applications," *J. Acoust. Soc. Am.*, vol. 132, no. 3, pp. 1462–1472, Sept. 2012.

[16] G. Lindsey, A. Breen, and S. Nevard, "SPAR's archivable actual-word databases," Technical report, University College London, June 1987.

[17] ITU-T, "Objective measurement of active speech level," Mar. 1993.