

# Direction of Arrival Estimation in front of a Reflective Plane Using a Circular Microphone Array

Nikolaos Stefanakis\* and Athanasios Mouchtaris\*,†

\*FORTH-ICS, Heraklion, Crete, Greece, GR-70013

†University of Crete, Department of Computer Science, Heraklion, Crete, Greece, GR-70013

**Abstract**—The presence of reflecting surfaces inside an enclosure is generally known to have an adverse effect in acoustic source localization and Direction of Arrival (DOA) estimation performance. In this paper, we focus on the problem of indoor multi-source DOA estimation along the horizontal plane, considering a circular sensor array which is placed just in front of one of the vertical walls of the room. We present a modification in the propagation model, which traditionally accounts for the direct path only, by incorporating also the contribution of the earliest reflection introduced by the adjacent vertical wall. Based on the traditional and the modified model, a Matched Filter and a Minimum Variance Distortionless Response beamformer are designed and tested for DOA estimation. Results with simulated and real data demonstrate the validity of the proposed model and its superiority in comparison to the traditional one.

## I. INTRODUCTION

In real acoustic environments, the transmitted signal is often received via multiple paths due to reflection, diffraction and scattering by objects in the transmission medium. The multipath effect can be understood as mirror image sources which produce multiple wavefronts interfering with each other, a fact that unfavourably affects direct path localization techniques [1]. The image sources tend to widen the estimated DOA distributions around the true DOA, an effect that grows in proportion to the Reverberation Time (RT) of the acoustic environment [2].

To some degree, DOA estimation and localization in reverberant rooms is still possible if it can be assumed that the energy of the direct wavefronts predominate over the contributions of early reflections, reverberations and noise. The performance can be improved to some extent by pre-selecting the signal portions that are less severely distorted with multipath signals and noise [3], [4], as well as signal portions where one source is significantly more dominant than others [5]. On the other hand, several works propose to employ a propagation model which is aware of some of the early reflections introduced by the acoustic environment. Bergamo et al. [6] were among the first to test this idea in a lightly reverberant environment by exploiting the image source principle [7]. In several works that followed, it was shown that single reflections convey additional information which can be

exploited in order not only to make sound source localization possible in reverberant rooms [8]–[10], but also in order to extract additional important spatial information regarding the sound sources. For example, in [11], [12] the authors demonstrate an approach for estimating the orientations of the sound sources, while in [13], the additional information is exploited in order to make range and elevations estimates, something that would not be possible with the given sensor array under free field conditions.

In this paper, we consider the very common problem, that due to practical constraints, the user is forced to place the array very close to one of the walls of the room. Due to the image source introduced by the wall, it is then expected that there will always be a secondary acoustic path which is coherent and comparable in strength to the direct path [7]. The motivation for this research is similar to the work in [14], which also considered the case of a planar reflector, and proposed to use a hemi-spherical array for sound capturing and beamforming in the half-space. In this paper, we consider the use of a circular sensor array but we impose no modifications to the array design. We modify the propagation model instead, by incorporating the earliest reflection introduced by the adjacent wall together with the direct path. Using the modified propagation model, we then perform DOA estimation by steering a Matched Filter beamformer and a Minimum Variance Distortionless Response beamformer [15] over a grid of possible angular locations and by processing the peaks in the histogram resulting by accumulating all local DOA estimates. Results with real and synthetic data demonstrate that circular arrays – well known for their ability to provide full-space 2D coverage [5] – are also capable to operate successfully in the half-space when required due to space constraints.

## II. PROPAGATION MODEL

Signals are represented in the Time-Frequency (TF) domain with  $\omega \in \mathbb{R}$  and  $\tau \in \mathbb{Z}$  denoting the angular frequency and the time-frame index respectively. Let us denote by  $\mathbf{x}(\tau, \omega) = [X_1(\tau, \omega), \dots, X_M(\tau, \omega)]^T$  and  $S_j(\tau, \omega)$ ,  $j = 1, \dots, J$  the STFTs of the observed signals and the  $j$ th source signal. With these notations, the observation signal can be modeled as

$$\mathbf{x}(\tau, \omega) = \sum_{j=1}^J \mathbf{a}(\omega, \theta_j) S_j(\tau, \omega) + \mathbf{z}(\tau, \omega) \quad (1)$$

This research has been funded in part by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 644283, Project LISTEN, and in part by EU and Greek national funds through the National Strategic Reference Framework(NSRF), grant agreement No 11SYN\_6\_1381, Project SeNSE.

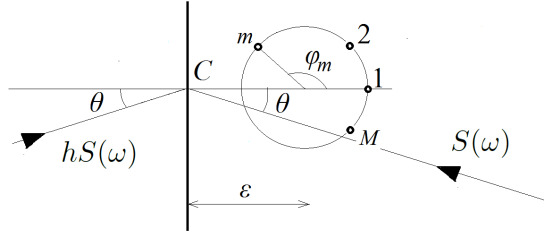


Fig. 1. A circular array of  $M$  sensors receiving a plane wave and its specular reflection. The reference point for defining the incident angle, denoted with  $C$ , coincides with the projection of the array center on the wall.

where  $\mathbf{a}(\omega, \theta_j) = [a_1(\omega, \theta_j), \dots, a_M(\omega, \theta_j)]^T$  is the so-called steering vector associated with the  $j$ th source at angle  $\theta_j$  and  $\mathbf{z}(\tau, \omega) = [z_1(\tau, \omega), \dots, z_M(\tau, \omega)]^T$  models additive noise and the reverberant part of the signal which is not included in  $\mathbf{a}$ .

The  $m$ th component of the classical steering vector for a circular sensor array of radius  $R$  given a single plane wave impinging at angle  $\theta$  and at frequency  $\omega$  can be written as [16]

$$a_m(\omega, \theta) = e^{jkR \cos(\phi_m - \theta)} \quad (2)$$

where  $k = \omega/c$  is the wavenumber and  $c$  is the speed of sound. Here,  $\phi_m$  is the angle of the  $m$ th sensor which similar to  $\theta$  is defined with respect to the center of the circle. It is assumed here that both the array plane and the plane wave direction are in the azimuth plane.

The model of (2) accounts for the direct path of the sound only and ignores any distinct reflections that may occur by the listening environment. Of course, estimating all the secondary paths is difficult in practice, as it would require detailed knowledge of the room geometry. Assuming however that the distance of the sensor array from a particular wall is much smaller in comparison to that from the other walls, it can be expected that the earliest reflection carries a relatively large portion of the energy of the reverberant part of the signal. Moreover, assuming far field conditions and a perfect specular reflection, we may determine this component deterministically, by considering an array of known orientation and distance from the closest wall.

In Fig. 1 we present the proposed half-space geometric model. A plane wave of strength  $S(\omega)$  and at angle  $\theta$  impinges on a circular array of radius  $R$  which has its center at a distance of  $\epsilon$  from the closest wall. The reflected component can be estimated in the frequency domain by accounting for a mirror plane wave, of strength  $hS(\omega)$ , arriving from angle  $\theta' = \pi - \theta$ . The quantity  $h \in \mathbb{R}^+$  is called the Image Source Relative Gain (ISRG) and expresses the relative gain with which the image source contributes to the sound field. It should be observed that the incident angle  $\theta$  is defined here with respect to the projection of the array center on the reflective plane and not with respect to the array center itself, which is the case for the propagation model in (2).

Based on this design, we introduce the vector  $\hat{\mathbf{a}}(\omega, \theta) = [\hat{a}_1(\omega, \theta), \dots, \hat{a}_M(\omega, \theta)]^T$  with its  $m$ th component defined as

$$a_m(\omega, \theta) = e^{jkR \cos(\phi_m - \theta)} e^{jk\epsilon \cos \theta} + h e^{jkR \cos(\phi_m - \pi + \theta)} e^{-jk\epsilon \cos \theta}. \quad (3)$$

We then normalize  $\hat{\mathbf{a}}(\omega, \theta)$  in order to construct the *modified steering vector*  $\mathbf{a}$  as

$$\mathbf{a}(\omega, \theta) = \hat{\mathbf{a}}(\omega, \theta) / \|\hat{\mathbf{a}}(\omega, \theta)\|_2, \quad (4)$$

where  $\|\cdot\|_2$  denotes the Euclidean norm. To be noticed that  $h$  does not account for the difference in the time of arrival of the primary plane wave and its reflection; this is explicitly taken into account by the two phasors  $e^{jk\epsilon \cos \theta}$  and  $e^{-jk\epsilon \cos \theta}$  in (3). This is convenient as it allows ISRG to be independent from the incident angle  $\theta$ . Also,  $h$  is assumed real and constant with frequency, which seems to be a non-trivial simplification considering that the reflectivity of the surface might in practice vary significantly with frequency. However, it will be shown that the presented model is robust to deviations between the assumed and the actual ISRG value and that a single reasonable “guess” about the value of  $h$  might work sufficiently well for a wide range of acoustic conditions. Intuitively, a value of ISRG close to 1 would correspond to a rigid surface, implying that the energy of the incident wave is equal to the energy of the reflected wave.

### III. MULTIPLE-SOURCE DOA ESTIMATION

The modified and the classical steering vector provide with two alternative propagation models which we can use in order to compare the performance of DOA estimation. In addition, we may consider different algorithmic approaches to DOA estimation. In this paper, DOA estimation is performed by steering a Matched Filter (MF) and a Minimum Variance Distortionless Response (MVDR) beamformer across a grid of potential source locations in 2D. The so-called *angular-spectrum* of the MF beamformer is defined at each time-frequency (TF) point as [15]

$$P_{MF}(\tau, \omega, \theta) = \frac{\mathbf{a}^H(\omega, \theta) \hat{\mathbf{R}}(\tau, \omega) \mathbf{a}(\omega, \theta)}{\mathbf{a}^H(\omega, \theta) \mathbf{a}(\omega, \theta)} \quad (5)$$

and that of the MVDR beamformer as

$$P_{MVDR}(\tau, \omega, \theta) = \frac{1}{\mathbf{a}^H(\omega, \theta) \hat{\mathbf{R}}^{-1}(\tau, \omega) \mathbf{a}(\omega, \theta)}. \quad (6)$$

where  $\hat{\mathbf{R}}(\tau, \omega)$  is the time-averaged empirical covariance matrix. This matrix is obtained at each TF point using a simple recursive formula

$$\hat{\mathbf{R}}(\tau, \omega) = (1 - q) \hat{\mathbf{R}}(\tau - 1, \omega) + q \mathbf{x}(\tau, \omega) \mathbf{x}(\tau, \omega)^H \quad (7)$$

where  $0 \leq q \leq 1$  is the forgetting factor. For both equations, vector  $\mathbf{a}$  can be constructed in accordance to the modified steering vector, as in (4), or in accordance to the classical steering vector, as in (2). In what follows, we will use the notations  $\text{MF}^*$  and  $\text{MVDR}^*$  when referring to beamformers based on the modified steering vector and MF and MVDR for those based on the classical steering vector.

Now, for DOA estimation we follow the assumption of one predominant source per time-frequency point, which is valid for signals with a sparse time-frequency representation such as speech [17], [18], at least until low reverberant conditions. The process consists of using a grid search to find the most energetic DOA at each time-frequency point, processing the collection of DOAs across time in order to form a histogram and then localizing the most prominent peaks in the histogram. For each combination of steering vector and beamformer, one local DOA at each time-frequency point can be estimated by searching over a grid of  $L$  possible angles as

$$\tilde{\theta}(\tau, \omega) = \operatorname{argmax}_{\theta} P(\tau, \omega, \theta), \quad (8)$$

where the azimuth angle  $\theta$ , in degrees, varies uniformly in the range  $[-180^\circ, 180^\circ)$  and  $\omega$  is considered along all STFT bins in the range  $\omega_{LB} \leq \omega \leq \omega_{UB}$ , where  $\omega_{LB}$  and  $\omega_{UB}$  correspond to a lower- and upper- frequency limit respectively. Considering the constraints imposed by the physical boundary, it is impossible to have a source at  $(90, 180)$  or at  $(-180, -90)$  and although angular locations in this range are scanned in (8), a particular time-frequency point is assigned a DOA or not according to the rule

$$\hat{\theta}(\tau, \omega) = \begin{cases} \tilde{\theta}(\tau, \omega), & \text{if } -90^\circ + \delta\theta < \tilde{\theta}(\tau, \omega) < 90^\circ - \delta\theta \\ \emptyset, & \text{otherwise} \end{cases} \quad (9)$$

where  $\delta\theta$  is a user defined threshold in degrees. The angle  $\hat{\theta}(\tau, \omega)$  is then stored together with all other estimations in the collection

$$\Theta(\tau) = \cup_{\omega=\omega_{LB}}^{\omega_{UB}} \hat{\theta}(\tau, \omega). \quad (10)$$

The sources' direction can then be found by localizing the peaks in the histogram which is formed with the estimated DOAs in  $\Theta(\tau)$ . This may extend not only across many frequency bins, as (10) implies, but also across multiple time frames. In this case, the DOA estimation is derived from a set of estimates in a block of  $B$  consecutive time-frames

$$C(\tau) = \cup_{t=\tau-B}^{\tau} \Theta(t), \quad (11)$$

with  $B$  being an integer denoting the *History Length* (HL). The collection  $C(\tau)$  is updated at each time-frame and the resulting histogram is smoothed as described in [19]. Assuming that the number of sources  $J$  is known, the  $J$  highest peaks are selected from the histogram under the constraint that they are “distant-enough”, i.e. separated by a user defined threshold  $dA$ . We note here that unlike the work in [19], we make no pre-selection in order to identify the signal portions where one source is isolated. Depending on how much sources overlap, this may lead to spurious peaks in the histogram, impairing the DOA estimation performance.

#### IV. EXPERIMENTAL VALIDATION

Experimental results with both simulated data and real recordings are presented for a circular array of four omnidirectional sensors and radius of  $R = 0.049$  m. The sensor angles  $\phi_m$  for this array were at  $-135^\circ$ ,  $-45^\circ$ ,  $45^\circ$  and  $135^\circ$ . The values of several parameters were kept the same for both

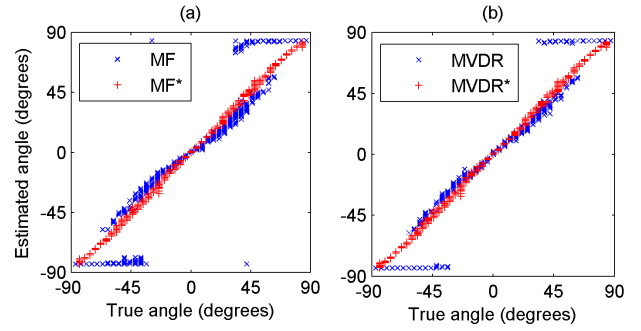


Fig. 2. Scatter plot with the estimated angle versus true angle for a single source, with MF and MF\* beamformer in (a) and for MVDR and MVDR\* beamformer in (b).

experiments based on simulated and real data and were as follows; we used an FFT size of 1024 samples at a sampling rate of 22.05 kHz and a squared Hanning window of 50% overlap. The history length was dynamically varying at each time-frame as  $B = \min\{\tau, 45\}$  where  $\tau$  is the running time frame index,  $dA$  was set equal to  $16^\circ$  and  $\delta\theta$  equal to  $3^\circ$ . The modified steering vector was constructed according to (3) with  $\epsilon = 0.063$  m and for a value of ISRG constant with frequency and equal to  $h = 0.9$ . The empirical covariance matrix was calculated using  $q = 0.7$  and local DOAs were searched upon a uniform grid with spacing of  $1^\circ$ . As a metric of the DOA estimation performance, we used the Mean Absolute Estimation Error (MAEE) which measures the absolute difference between the true DOA and the estimated DOA, in degrees, averaged over all sources, orientations and time-frames of the source signals [5].

A first series of results is based on a simulated rectangular room with dimensions of  $L_x \times L_y \times L_z = 5 \times 6 \times 4$  m and with the center of the circular array placed at  $[0.063 \ 3.2 \ 1.6]$  m, just in front of the vertical wall at  $x = 0$ . For these simulations, the image source method of Allen and Berkley [7] was implemented in Matlab using the toolbox provided by Habets [20]. For the source signals we used recordings of continuous speech from different subjects of 6 sec duration each. The sound sources were placed at the height of  $z = 1.75$  m and at a distance of 1.2 m from reference point  $C$ . To focus on the effect of reverberation, the SNR was set at a fixed value of 26 dB, in all cases, by adding white Gaussian noise to the observation signals. Finally, for DOA estimation we used all the frequency bins from 150 to 4000 Hz.

At first we studied the case of one single speaker at  $RT_{60} = 0.5$  sec by spanning  $-84^\circ$  to  $84^\circ$  with a step of  $4^\circ$ . In Fig. 2, we provide a scatter plot with the values of the estimated angles as a function of a true angle at each location, across all time frames excluding the first 44, for the MF and MF\* beamformer in (a) and for the MVDR and MVDR\* beamformer in (b). The graphs demonstrate clearly that the two approaches based on the classical steering vector fail systematically in localizing the source, which becomes more and more prominent as the sources are located further

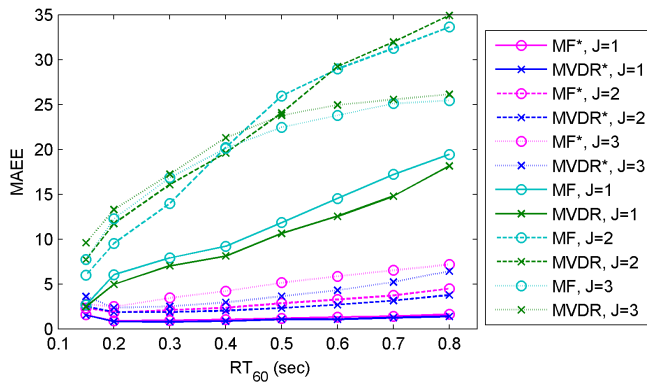


Fig. 3. MAEE as a function of the reverberation time inside a  $5 \times 6 \times 4$  m rectangular room.

away from  $0^\circ$ . On the other hand, there is a good agreement between the true and the estimated angle for both the MF\* and MVDR\* beamformer.

Additional results are presented in terms of MAEE in Fig. 3. The DOA estimation performance of each beamformer in the previous experiment ( $J = 1$ ) is shown here as a function of the reverberation time varying from 0.15 to 0.8 sec. For the same range of  $RT_{60}$ , the experiments corresponding to  $J = 2$  and 3 simultaneous sources were designed by considering 40 different random orientations for each value of  $J$ , with the sources' angles allowed to vary from  $-84$  to  $84$  degrees and with the restriction that no source is closer than  $20^\circ$  to another source.

In general, the results are not surprising for the anechoic propagation model; the performance drops significantly with the increment of the  $RT_{60}$ . Only at low  $RT_{60}$  and for the case of a single source are the MF and the MVDR beamformers capable of achieving an acceptable performance. Also, the MVDR beamformer performs slightly better than the MF beamformer, but this advantage is lost for higher number of sources. To note here that both MF and MVDR approach achieve a satisfactory performance when the same experiment is repeated in anechoic conditions; the MAEE for  $J = 3$  is 2.78 for MF and 2.44 for MVDR, which provides further evidence that the failure of the classical approach is due to the acoustic environment.

On the other hand, the two beamformers based on the modified steering vector achieve a good performance across the entire range of  $RT_{60}$ , with the MVDR\* showing a clear advantage in comparison to the MF\*, especially for  $J = 2$  and 3. It is important to observe that while a fixed ISRG value of  $h = 0.9$  has been used for constructing the propagation model in all cases, the method appears robust to deviations between the actual and the assumed wall reflectivity. In fact, although the reflectivity in the simulation model varies significantly with respect to the reverberation time, the exact association between  $RT_{60}$  and ISRG is unknown to us. This argument is also reinforced with the real data experiments conducted below.



Fig. 4. Experimental setup. The sensor array is tangent to the plasterboard wall at  $x = 0$ .

Experiments with real recordings were performed using a circular array with characteristics identical to those in the simulations. To note that while this circular array is originally designed to operate with 8 sensors, only the 4 of them are used in the experiment. The experiment took place in a small rectangular office with dimensions of  $L_x \times L_y \times L_z = 2.9 \times 4.4 \times 2.8$  m and with an estimated  $RT_{60}$  of 0.3 sec. The center of the circular array was placed at  $[0.063 \ 2 \ 0.6]$  m, tangent to the vertical wall made of plasterboard at  $x = 0$ . As it can be seen from Fig. 4, the array was lying on a small rectangular wooden surface which was also tangent to the wall. The conditions of this experiment are representative of actual difficulties that are to be found in a real life application; for example, the wall is not a perfect planar surface, as there is a rectangular plastic pipe extended horizontally just a little above the array (see Fig. 4). Also, the wooden table is expected to introduce additional reflections which are not considered in the propagation model. However, we believe that this was not a source for serious degradation in the performance, as the image sources introduced by the table are expected to have the same azimuth angle as the actual sources.

Speech signals from three male and three female speakers were recorded at 17 different angles and in particular at  $-75$  degrees, from  $-70$  to  $0$  with a step of  $10$  degrees and from  $5$  to  $75$  with a step of  $10$  degrees. The speakers were seated during the experiment at a distance of approximately  $1.25$  m from reference point  $C$ . We note that due to space limitations, the angular locations for the speakers were limited within  $\pm 75^\circ$ . The parameters for DOA estimation were similar as with the simulated data, except from the fact that the low frequency bound used for DOA estimation was set equal to  $500$  Hz and  $\delta\theta$  was set equal to  $10^\circ$ . The MAEE values for each approach and number of speakers  $J$  are shown in Table I. These values were averaged over all 17 angular locations for  $J = 1$  while for 2 and 3 simultaneous speakers, they were averaged over all possible angle combinations excluding those where adjacent sources were less than  $20^\circ$  apart.

The results once more illustrate the validity of the mod-

TABLE I  
MAEE VALUES IN THE REAL ACOUSTIC ENVIRONMENT.

J	MF*	MVDR*	MF	MVDR
1	1.50	1.24	9.34	9.12
2	4.05	2.30	10.85	13.82
3	6.17	4.54	13.03	14.63

ified propagation model and the superiority of the MVDR\* approach against MF\*. It is worth noting that, implemented in Matlab with a 3.4 GHz processor, MF\* and MVDR\* operate at 65% and 90% real-time respectively. A disappointing fact is that MAEE increases quite abruptly with the number of active sources. Given some technique to determine the time-frequency bins of single active sources would seem promising for avoiding this deterioration, but it should be expected that well known time-frequency point pre-selection techniques might prove inappropriate for the considered conditions. For example, the Single Source Zone (SSZ) criterion proposed for the case of a circular array in [5] has demonstrated important improvements to DOA estimation in reverberant conditions but with the array placed close to the center of the room. However, applying the same bin selection technique to our case did not improve performance, neither for the case of the anechoic nor for the case of the modified propagation model. This provides further evidence on the necessity of a re-formulated propagation model - as the one proposed in this paper - as well as for a re-formulated time-frequency point selection technique in the case of the half-space problem.

## V. CONCLUSION

We have presented a modification in the propagation model and used it for 2D DOA estimation in the half-space. In contrast to other environment aware approaches that require detailed knowledge of the room geometry, our approach involves only knowledge of the distance of the sensor array from the closest wall. The proposed modified beamformers, MF\* and MVDR\*, were shown to provide a viable solution for DOA estimation, with the latter showing a clear advantage in terms of performance.

In both simulated and real experiments in this paper, the ISRG was set to a fixed value. Intuitively, the DOA estimation accuracy can be improved by deriving an estimation of the ISRG as a function of frequency, using for example in-situ estimation of the reflection coefficient as described in [21]. Also, in the presented formulation, the number of active sources was assumed to be known. More sophisticated processing of the resulting histogram is expected to improve not only the DOA estimation accuracy, but also to allow for counting of the number of sources, following for example the work presented in [5]. Finally, the presented propagation model has a straightforward connection to other planar sensor array configurations as well as to other sensor array processing applications, such as signal enhancement, de-noising and source separation.

## REFERENCES

- [1] C. Blandin, A. Ozerov, and E. Vincent, "Multi-source tdoa estimation in reverberant audio using angular spectra and clustering," *Signal Processing*, vol. 92, pp. 1950–1960, 2012.
- [2] M. Cobos, J. Lopez, and S. Spors, "Analysis of room reverberations effects in source localization using small microphone arrays," in *Proc. of 4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, 2010.
- [3] M. Akatas, T. Akgun, and H. Ozkan, "Acoustic direction finding in highly reverberant environment with single acoustic vector sensor," in *Proc. of 23rd European Signal Processing Conference (EUSIPCO)*, 2014, pp. 2346–2350.
- [4] A. Moore, C. Evers, P. Naylor, D. Alon, and B. Rafaely, "Direction of arrival estimation using pseudo-intensity vectors with direct path dominance test," in *Proc. of 23rd International Signal Processing Conference (EUSIPCO)*, 2014, pp. 2341–2345.
- [5] D. Pavlidis, A. Griffin, M. Puigt, and A. Mouchtaris, "Real-time multiple sound source localization and counting using a circular microphone array," *IEEE Trans. on Audio, Speech, and Lang. Process.*, vol. 21, no. 10, pp. 2193–2206, 2013.
- [6] P. Bergamo, S. Asgari, H. Wang, and D. Maniezzo, "Collaborative sensor networking towards real-time acoustical beamforming in free-space and limited reverberance," *IEEE Trans. on Mobile Computing*, vol. 3, no. 3, pp. 211–224, 2004.
- [7] J. Allen and D. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, 1979.
- [8] T. Korhonen, "Acoustic localization using reverberation with virtual microphones," in *Proc. of International Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2008, pp. 211–223.
- [9] O. Öçal, I. Dokmanić, and M. Vetterli, "Source localization and tracking in non-convex rooms," in *Proc. of ICASSP*, 2014, pp. 1443–1447.
- [10] P. Svaizer, A. Brutti, and M. Omologo, "Use of reflected wavefronts for acoustic source localization with a line array," in *Joint Workshop on Hands-free Speech Communications and Microphone Arrays*, 2011, pp. 165–169.
- [11] —, "Environment aware estimation of the orientation of acoustic sources using a line array," in *Proc. of 20th European Signal Processing Conference (EUSIPCO)*, 2012, pp. 1024–1028.
- [12] K. Niwa, Y. Hioka, S. Sakauchi, K. Furuya, and Y. Haneda, "Estimation of sound source orientation using eigenspace of spatial correlation matrix," in *Proc. of ICASSP*, 2010, pp. 129–132.
- [13] F. Ribeiro, D. Ba, C. Zhang, and D. Florêncio, "Turning enemies into friends: using reflections to improve sound source localization," in *IEEE Int. Conf. Multimedia and Expo (ICME)*, 2010, pp. 731–736.
- [14] Z. Li and R. Duraiswami, "Hemispherical microphone arrays for sound capture and beamforming," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2005, pp. 106–109.
- [15] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.
- [16] A. Alexandridis, A. Griffin, and A. Mouchtaris, "Capturing and reproducing spatial audio based on a circular microphone array," *Journal of Electrical and Computer Engineering*, vol. 2013, article ID 718574.
- [17] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. on Signal Process.*, vol. 52, pp. 1830–1847, 2004.
- [18] S. Rickard and O. Yilmaz, "On the approximate w-disjoint orthogonality of speech," in *Proc. of ICASSP*, 2002, pp. 529–532.
- [19] D. Pavlidis, M. Puigt, A. Griffin, and A. Mouchtaris, "Real-time multiple sound source localization using a circular microphone array based on single-source confidence measures," in *Proc. of ICASSP*, 2012, pp. 2625–2628.
- [20] D. Jarrett, E. Habets, M. Thomas, and P. Naylor, "Rigid sphere room impulse response simulation: Algorithm and applications," *J. Acoust. Soc. Am.*, vol. 132, no. 3, pp. 1462–1472, audiobabs-erlangen.de/fau/professor/habets/software/smirgenerator 2012.
- [21] Y. Zhang, W. Lin, and C. Bi, "A technique based on the equivalent source method for measuring the surface impedance and reflection coefficient of a locally reacting material," in *Proc. of International Congress on Noise Control Engineering*, 2014.