MODELLING SOURCE DIRECTIVITY IN ROOM IMPULSE RESPONSE SIMULATION FOR SPHERICAL MICROPHONE ARRAYS

Sina Hafezi, Alastair H. Moore and Patrick A. Naylor

Department of Electrical and Electronic Engineering, Imperial College London, UK
{s.hafezi14, alastair.h.moore, p.naylor}@imperial.ac.uk

ABSTRACT

In this work we present a new room impulse response simulation for spherical microphone arrays taking into account source directivity. We calculate the emission angle of the sound ray leaving the source based on the location of the image and the receiver using Allen and Berkley's image method. We provide an implementation of a room impulse response simulator for a spherical microphone array including a directional source with arbitrary directivity. We validate our implementation considering the zeroth and the first-order reflections. Our results show a worst-case directional gain error of 7% in comparison with theoretical predictions.

Index Terms— directivity, directional, simulation, spherical microphone arrays, room impulse response

1. INTRODUCTION

Allen and Berkley's image method [1] is a widely used approach for Room Impulse Response (RIR) simulation for an omnidirectional sound source and a receiver in a reverberant environment. Several extensions to the image method have been proposed [2, 3]. Recently the use of Spherical Microphone Arrays (SMA) has had growing interest due to their ability to capture sound fields in three dimensions [4–8]. RIR simulation for SMAs has been presented for omnidirectional sources [9, 10]. However, in real-world situations, the sound source almost always has a specific non-omnidirectional directivity pattern. It is therefore highly valuable to be able to include source directivity in RIR simulators.

Early works on source directivity are described in [11–13]. In contrast, in this paper we give general RIR simulation techniques, full analytical details and the corresponding implementations in downloadable software. We begin by analyzing the effect of source directivity on Allen and Berkley's image method and then derive the directivity-based equations to be used in the simulation. We consider the cases of microphones in the free field and also on a rigid baffle as used in SMAs.

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 609465

In the rest of this paper, we first briefly summarise Allen and Berkley's image method and the RIR simulation for SMA in Section 2. We then in Section 3 analyze the effect of source directivity on the image method, and provide extensions to the existing open-source RIR simulator for SMA available at [14]. Finally, in Section 4, we describe our implementation and validate it for some standard source directivity patterns and draw conclusion in Section 5.

2. REVIEW

2.1. Allen and Berkley's image method

Consider a rectangular room illustrated in Figure 1 with length L_x, width L_y, and height L_z. The acoustic reflection coefficients [15] of the four walls, floor and ceiling are β_{x1}, β_{x2}, β_{y1}, β_{y2}, β_{z1} and β_{z2} respectively, where ν_1 coefficients ($\nu \in \{x, y, z\}$) correspond to the boundaries at $\nu = 0$ and the ν_2 coefficients correspond to the boundaries at $\nu = L_\nu$. For a sound source located at $\mathbf{r_s} = (x_s, y_s, z_s)$ and receiver located at $\mathbf{r} = (x, y, z)$, the location vector (with respect to the receiver) for all the images obtained due to the walls can be expressed as a vector $\mathbf{r_{p,m}} = \mathbf{r_p} + \mathbf{r_m}$,

$$\mathbf{r_p} = [x_s - x - 2p_x x_s, y_s - y - 2p_y y_s, z_s - z - 2p_z z_s] \tag{1}$$

$$\mathbf{r_m} = [2m_x L_x, 2m_y L_y, 2m_z L_z] \tag{2}$$

where each of the elements in $p = (p_x, p_y, p_z)$ can take a value of 0 or 1 and each element in $m = (m_x, m_y, m_z)$ is integer-valued. $\mathbf{P}$ and $\mathbf{M}$ are all the possible sets of p and m. The Room Transfer Function (RTF) is given by [1]

$$H(\mathbf{r}|\mathbf{r_s}, k) = \sum_{p \in \mathbf{P}} \sum_{m \in \mathbf{M}} \beta_{p,m} \frac{e^{+ik\|\mathbf{r_{p,m}}\|}}{4\pi\|\mathbf{r_{p,m}}\|} \tag{3}$$

$$\beta_{p,m} = \beta_{x1}^{|m_x - p_x|} \beta_{x2}^{|m_x|} \beta_{y1}^{|m_y - p_y|} \beta_{y2}^{|m_y|} \beta_{z1}^{|m_z - p_z|} \beta_{z2}^{|m_z|} \tag{4}$$

where $\|.\|$ denotes the ℓ^2-norm and the wavenumber k is related to frequency f (Hz) and the speed of sound c (m/s) via the relationship $k = 2\pi f/c$.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

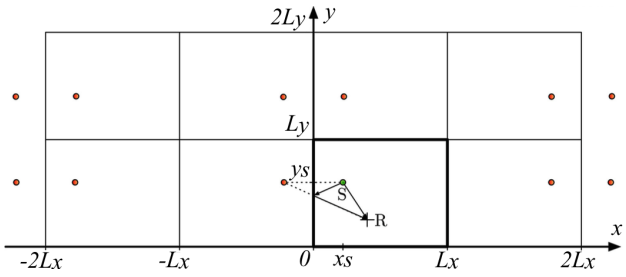

Fig. 1. A 2D illustration [10] of the image space showing source S, receiver R and associated images (red). An example first-order reflection is shown by the dashed line.

2.2. Extension to a Rigid Spherical Microphone Array

The rigid body of the SMA blocks the direct paths from images to the microphones whose journey passes through the body of the SMA. This effect has been studied in [9, 10]. Given a rigid SMA centred at $\mathbf{r_a}$, the RTF from a sound source at $\mathbf{r_s}$ to a point on the surface of the SMA, $\mathbf{r_a} + \mathbf{r}$, is

$$H(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_\mathbf{s}, k) = \sum_{p\in\mathbf{P}} \sum_{m\in\mathbf{M}} \beta_{p,m} G_N(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_{\mathbf{p,m}}, k) \tag{5}$$

$$G_N(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_{\mathbf{p,m}}, k) = \frac{ik}{4\pi} \sum_{l=0}^{\infty} \frac{ih_l^{(1)}(kr_{pm})}{h_l^{(1)\prime}(kr_{pm})(kr_{pm})^2} \times (2l+1) P_l(\cos\Theta_{\tilde{\mathbf{r}},\tilde{\mathbf{r}}_{\mathbf{p,m}}}) \tag{6}$$

$$\mathbf{r_p} = [x_s - x_a - 2p_x x_s, y_s - y_a - 2p_y y_s, z_s - z_a - 2p_z z_s] \tag{7}$$

where $G_N(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_{\mathbf{p,m}}, k)$ is the Neumann Green's Function [16], which describes the sound propagation between a source at $\mathbf{r_{p,m}}$ and a receiver at $\mathbf{r}$ on the rigid sphere, both with respect to the centre of SMA as illustrated in Figure 2. The angle between $\mathbf{r}$ and $\mathbf{r_{p,m}}$ is $\Theta_{\tilde{\mathbf{r}},\tilde{\mathbf{r}}_{\mathbf{p,m}}}$, $h_l^{(1)}$ and P_l are respectively the spherical Hankel function of the first kind and the Legendre polynomial of order l and $\tilde{\cdot}$ denotes a location vector in spherical coordinates. Note that $\mathbf{r_{p,m}} = \mathbf{r_p} + \mathbf{r_m}$, where $\mathbf{r_p}$ and $\mathbf{r_m}$ can respectively be found using (7) and (2).

3. MODELLING SOURCE DIRECTIVITY

In the following we use "emission angle" to refer to the angle by which the sound leaves the source and impinges on the observer. The RTF in (5) is the weighted sum of individual responses for each image. In the case of a directional source, which omits sound with different gain depending on the direction of emission, we will have an extra image-dependent gain for each expression inside the summation.

Therefore, for a directional sound source we can rewrite

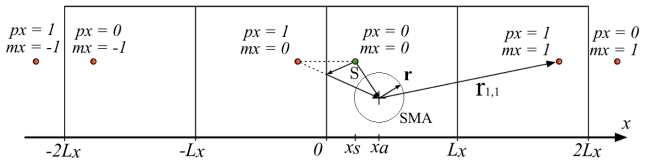

Fig. 2. A 2D illustration [10] of the image space showing source S, SMA and associated images (red). An example first-order reflection is shown (dashed) for the image with $p_x = 1$ and $m_x = 0$.

the RTF as follow:

$$H(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_\mathbf{s}, k) = \sum_{p\in\mathbf{P}} \sum_{m\in\mathbf{M}} \beta_{p,m} g(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_{\mathbf{p,m}},k) G_N(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_{\mathbf{p,m}}, k) \tag{8}$$

where $g(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_{\mathbf{p,m}},k)$ is the frequency-dependent source directivity gain associated with the image at $\tilde{\mathbf{r}}_{\mathbf{p,m}}$ and the receiver at $\tilde{\mathbf{r}}$, both with respect to the centre of the SMA.

A source directivity pattern is a function of emission angle with respect to the source orientation. We therefore need to transform $g(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_{\mathbf{p,m}},k)$ into $g(\theta_{\mathbf{e,pm,o}}, \varphi_{\mathbf{e,pm,o}},k)$ where $\theta_{e,pm,o}$ and $\varphi_{e,pm,o}$ are respectively the azimuth and the elevation emission angles with respect to the source orientation.

For an arbitrary image, we define N_ν ($\nu \in \{x, y, z\}$) as the number of reflections the image had by walls perpendicular to axis ν,

$$N_\nu = |2m_\nu - p_\nu|. \tag{9}$$

In Section 3.1 and 3.2, we determine the emission angle as a function of the image and the receiver location.

3.1. Analysis of azimuth angle

For an arbitrary azimuth angle θ in the xy plane in the range of $[0, 2\pi)$, using the reflection matrices for reflections in the x and y axes, θ' is the reflected angle after N_x and N_y reflections respectively by walls perpendicular to the x and y axes. Hence we can write

$$\begin{bmatrix} \cos(\theta') \\ \sin(\theta') \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^{N_y} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}^{N_x} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}. \tag{10}$$

In (10), for an arbitrary image, since the reflection matrix only affects the sign of x and y coordinate, we only need the parity of N_x and N_y which are respectively determined by the parity of p_x and p_y using (9). Therefore we replace N_x and N_y respectively with p_x and p_y. This simplification results in four possible outcomes depending on the value of the set (p_x, p_y),

$$\theta' = \Theta(\theta|p) = \begin{cases} \theta & (p_x, p_y) = (0, 0) \\ 2\pi - \theta & (p_x, p_y) = (0, 1) \\ wrap(\pi - \theta) & (p_x, p_y) = (1, 0) \\ wrap(\pi + \theta) & (p_x, p_y) = (1, 1) \end{cases} \tag{11}$$

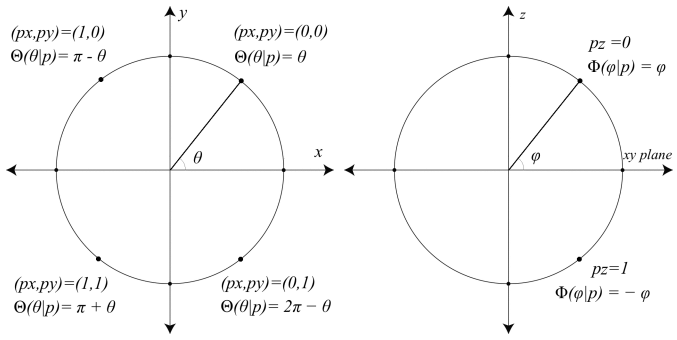

Fig. 3. Illustration of possible outcomes of reflection function for an arbitrary azimuth and elevation angles (θ, φ).

where $wrap(.)$ denotes wrapping to $[0, 2\pi)$ and $\Theta(.|p)$ represents the reflection function on an azimuth angle after all reflections by walls perpendicular to x and y which result in the creation of an image with the set p.

3.2. Analysis of elevation angle

For an arbitrary elevation angle φ made with xy plane in the range of $[-\frac{\pi}{2}, \frac{\pi}{2}]$, φ' is the reflected angle after N_z reflections by walls perpendicular to z axis. Hence we can write

$$\begin{bmatrix} \cos(\varphi') \\ \sin(\varphi') \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^{N_z} \begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix}. \quad (12)$$

In (12), for an arbitrary image, having done the same simplification on N_z as we did in Section 3.1 for N_x and N_y, we have

$$\varphi' = \Phi(\varphi|p) = \begin{cases} \varphi & p_z = 0 \\ -\varphi & p_z = 1 \end{cases} \quad (13)$$

where $\Phi(.|p)$ represents the reflection function on an elevation angle after all reflections by walls perpendicular to z axis which result in the creation of an image with the set p.

Figure 3 illustrates the possible outcomes of the reflection functions $\Theta(\theta|p)$ and $\Phi(\varphi|p)$ for an arbitrary azimuth θ and elevation φ after N_x, N_y and N_z number of reflections by walls perpendicular to x, y and z axes respectively.

We define the impinging azimuth angle, θ_{pm}, and the impinging elevation angle, φ_{pm} as

$$(\theta_{pm}, \varphi_{pm}) = \angle(\mathbf{r} - \mathbf{r_{p,m}}) \quad (14)$$

where $\angle(.)$ denotes the spherical angle of a vector as a set of azimuth and elevation angles. These are the angles by which the sound ray impinges on the receiver.

Since the impinging angle is the reflected emission angle after all reflections, using (11) and (13), $\Theta(\theta_{pm}|p)$ and $\Phi(\varphi_{pm}|p)$ are respectively the emission azimuth and elevation angles, $(\theta_{e,pm}, \varphi_{e,pm})$.

$$(\theta_{e,pm}, \varphi_{e,pm}) = (\Theta(\theta_{pm}|p), \Phi(\varphi_{pm}|p)). \quad (15)$$

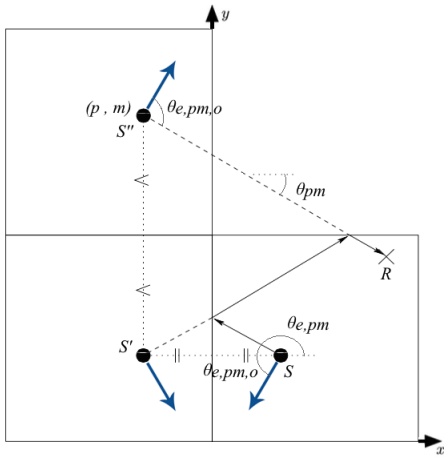

Fig. 4. A 2D illustration of an example of a second-order reflected path from source S to receiver R. First-order image S' and the second-order image S" is also shown. The blue arrows represent the associated orientation of the source directivity pattern. The solid line is the actual path the sound travels from the source to the receiver.

Using (14) and (15), for an arbitrary image, having the orientation of the source as a set of (θ_o, φ_o), the emission angles with respect to the orientation of source, $(\theta_{e,pm,o}, \varphi_{e,pm,o})$ are

$$(\theta_{e,pm,o}, \varphi_{e,pm,o}) = (\Theta(\theta_{pm}|p) - \theta_o, \Phi(\varphi_{pm}|p) - \varphi_o). \quad (16)$$

Alternatively, expanding functions Θ and Φ, (16) can be rewritten as

$$\begin{aligned}(\theta_{e,pm,o}, \varphi_{e,pm,o}) &= (\theta_{pm} - \Theta(\theta_o|p), \varphi_{pm} - \Phi(\varphi_o|p)) \\ &= ((\theta_{pm} - \theta_{o,pm}), (\varphi_{pm} - \varphi_{o,pm}))\end{aligned} \quad (17)$$

where $\theta_{o,pm}$ and $\varphi_{o,pm}$ are respectively the orientation azimuth and elevation of the directional image. Since $\Theta(\theta_o|p)$ and $\Phi(\varphi_o|p)$ represent the orientation of a directional image, then the emission angle with respect to source orientation can also be seen as the difference between the impinging angles and the image orientation as illustrated in Figure 4.

Having the orientation of the sound source and the source directivity pattern with respect to that orientation, we can rewrite (8) as

$$\begin{aligned}H(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_\mathbf{s}, k) = &\sum_{p\in\mathbf{P}} \sum_{m\in\mathbf{M}} \beta_{p,m} g(\theta_{e,pm,o}, \varphi_{e,pm,o}, k) \\ &\times G_N(\tilde{\mathbf{r}}|\tilde{\mathbf{r}}_{\mathbf{p,m}}, k).\end{aligned} \quad (18)$$

4. IMPLEMENTATION AND VALIDATION

The proposed method is implemented as an extension to the RIR simulators for a free-field microphone and SMA respectively available at [17] and [14]. It follows the same

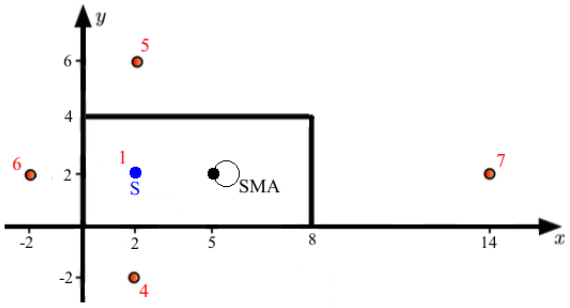

Fig. 5. A 2D partial illustration of the image space showing the position of the images (red) for the source S (blue) and SMA (black) considering the zeroth and the first-order reflections only. Note that image 2 at $(2, 2, -4)$ and image 3 at $(2, 2, 12)$ are not displayed since they lie on the source S in xy plane representation.

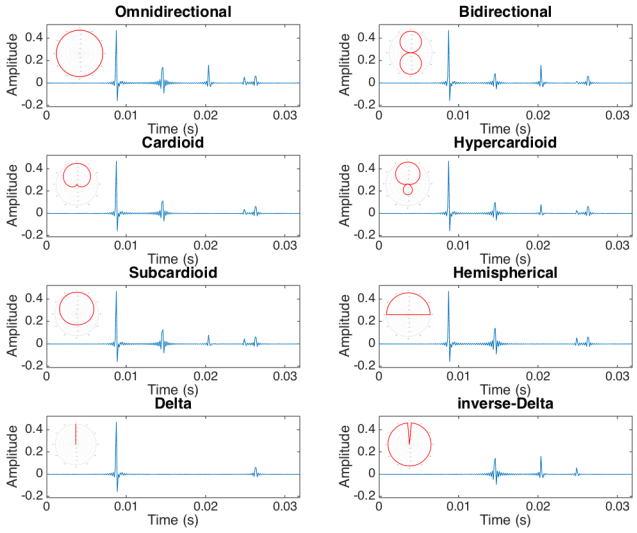

Fig. 6. RIRs (blue) for each source directivity pattern (red)

structure explained in [10] while additionally considering a frequency-dependent directivity function and orientation for a directional source. Since the image orientation is also calculated in the implementation, we used (17) for the calculation of emission angles although (16) may alternatively be used.

In operation of these tools, the user can either choose a standard directivity pattern or specify a gain table describing a frequency-dependent directivity pattern across discrete azimuth and elevation angles. The standard directivity patterns provided include Omnidirectional (Omni), Bidirectional (Bi), Cardioid (Card), Hypercardioid (Hyper), Subcardioid (Sub), Hemispherical (Hemi), Delta with emission only in the direction of orientation, and inverse-Delta (iDelta) with emission everywhere equally except in the direction of orientation. When the user enters a sampled directivity pattern, the tool also provides optional interpolation methods from a list of linear, nearest, spline or cubic. Further details are available in our MATLAB implementation for free-field microphone at [18] and for SMA at [19].

Due to deeper interest in validation for the SMA, we produced the simulated impulse response for the SMA case for all directivity patterns provided in Table 2. As illustrated in Figure 5, we set a source at $(2, 2, 4)$ m with orientation azimuth and elevation set of $(0, 0)$ radian, a rigid SMA at $(5.042, 2, 4)$ m with radius of 4.2 cm and a microphone position azimuth and elevation set of $(\pi, 0)$ radian, in a room with dimensions of $(8, 4, 8)$ m and reflection coefficients for the walls perpendicular to the x, y and z axes respectively equal to 0.8, 0.4 and 0.2. The sampling rate was 8 kHz and $c = 343$ m/s.

In order to validate the accuracy of source directivity modelling, it is sufficient to consider only the zeroth and first-order reflections, which result in six images plus the source itself as shown in Figure 5. Since the propagation distances for images 2 and 3 are equal to the propagation distances for images 3 and 5 respectively, the receiving impulses in the RIR for these four images are seen as two coincident impulses which results in the total of five peaks in the RIR for the omnidirectional case as shown in Figure 6.

Table 2 shows the emission angle for each image and its associated directivity gain calculated by $g = |\ \alpha + (1 - \alpha)\cos(\theta)\cos(\varphi)\ |$ where α is provided in Table 1.

Pattern	Omni	Bi	Card	Hyper	Sub
α	1	0	0.5	0.25	0.75

Table 1. Alpha used in standard directivity formula

At the RIR samples where we have peaks in the omnidirectional case, we define the scaling factor that is applied over each peak due to the source directivity as the amplitude in the directional case divided by the amplitude in the omnidirectional case (see Table 3). Comparing the results in Table 3 and Table 2, it can be seen that the maximum directional gain error is 0.07. The errors arise in part at least since the values in Table 3 are calculated using the amplitude at the nearest sample to the peak.

5. CONCLUSION

We have presented the modelling of source directivity in RIR simulation for SMAs using an extension to the image method. Part of the method involves the derivation of the relationship between the emission angle and the image location. The overall simulation of the point-to-point RIRs have been integrated into a freely downloadable software tool. Finally a validation has been performed that confirms the accuracy of our implementation of such modelling. A number of standard directivity patterns have been tested to compare the theoretical di-

Image no.	1	4&5	6	2&3	7
$\theta_{pm}\,(deg)$	0	± 53.13	180	0	0
$\varphi_{pm}\,(deg)$	0	0	0	± 69.44	0
Omni	1.00	1.00	1.00	1.00	1.00
Bi	1.00	0.60	1.00	0.35	1.00
Card	1.00	0.80	0.00	0.67	1.00
Hyper	1.00	0.70	0.50	0.51	1.00
Sub	1.00	0.90	0.50	0.83	1.00
Hemi	1.00	1.00	0.00	1.00	1.00
Delta	1.00	0.00	0.00	0.00	1.00
iDelta	0.00	1.00	1.00	1.00	0.00

Table 2. Theoretical gain associated with each image and pattern

Image no.	1	4&5	6	2&3	7
Bi	1.00	0.60	0.99	0.30	0.99
Card	1.00	0.79	0.00	0.67	0.98
Hyper	1.00	0.69	0.50	0.48	0.99
Sub	1.00	0.90	0.50	0.84	0.99
Hemi	1.00	0.99	0.00	1.02	0.99
Delta	1.00	-0.01	-0.01	-0.07	0.98
iDelta	0.00	1.01	1.01	1.07	0.02

Table 3. Scaling factor of each image and pattern.

rectional gain with the results of our simulation. Worst-case error in our experiments has been shown to be less than 7%.

REFERENCES

[1] 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.

[2] P. M. Peterson, "Simulating the response of multiple microphones to a single acoustic source in a reverberant room," *J. Acoust. Soc. Am.*, vol. 80, no. 5, pp. 1527–1529, Nov. 1986.

[3] E.A. Lehmann and A.M. Johansson, "Diffuse reverberation model for efficient image-source simulation of room impulse responses," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 6, pp. 1429–1439, Aug. 2010.

[4] E. Fisher and B. Rafaely, "The nearfield spherical microphone array," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2008, pp. 5272–5275.

[5] Zhibin Lin and Wei Qingyu, "Localization of multiple acoustic sources using optimal spherical microphone arrays," in *Proc. 9th International Conference on Signal Processing ICSP 2008*, Oct. 2008, pp. 349–352.

[6] Y. Peled and B. Rafaely, "Method for dereverberation and noise reduction using spherical microphone arrays," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2010, pp. 113–116.

[7] H. Sun, S. Yan, and U. P. Svensson, "Robust minimum sidelobe beamforming for spherical microphone arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 4, pp. 1045–1051, May 2011.

[8] J. Meyer and T. Agnello, "Spherical microphone array for spatial sound recording," in *Proc. Audio Eng. Soc. Convention*, New York, NY, USA, Oct. 2003, pp. 1–9.

[9] D. P. Jarrett, E. A. P. Habets, M. R. P. Thomas, and P. A. Naylor, "Simulating room impulse responses for spherical microphone arrays," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 129–132.

[10] 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.

[11] M. Kompis and N. Dillier, "Simulating transfer functions in a reverberant room including source directivity and head-shadow effects," *J. Acoust. Soc. Am.*, vol. 93, pp. 2779–2787, 1993.

[12] C. Jin A. Wabnitz, N. Epain and A. Schaik, "Room acoustics simulation for multichannel microphone arrays," in *Proc. Intl. Symposium on Room Acoustics (ISRA)*, Melbourne, Australia, 2010.

[13] T. Betlehem and M. Poletti, "Sound field of a directional source in a reverberant room," *Journal of the Acoustical Society of New Zealand*, vol. 25, no. 4, pp. 12–22, 2012.

[14] D. P. Jarrett, "Spherical Microphone array Impulse Response (SMIR) generator," http://www.ee.ic.ac.uk/sap/smirgen/.

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

[16] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, first edition, 1999.

[17] E. A. P. Habets, "Room impulse response (RIR) generator," http://home.tiscali.nl/ehabets/rirgenerator.html, May 2008.

[18] S. Hafezi, "Room Impulse Response for Directional source (RIRD) generator," http://www.ee.ic.ac.uk/sap/rirdgen/.

[19] S. Hafezi, "Spherical Microphone array Impulse Response for Directional source (SMIRD) generator," http://www.ee.ic.ac.uk/sap/smirdgen/.