BLIND ESTIMATION OF DIRECTIONAL PROPERTIES OF ROOM REVERBERATION USING A SPHERICAL MICROPHONE ARRAY

Prasanga N. Samarasinghe, Thushara D. Abhayapala *

Acoustics and Audio Group
Research School of Engineering, CECS,
The Australian National University,
ACT 2600, Canberra, Australia.

ABSTRACT

This paper presents an experimental study on a novel technique to blindly estimate the directional properties of room reflections using a spherical microphone array. The algorithm is developed based on a spatial correlation model formulated in the spherical harmonics domain. This model expresses the cross correlation matrix of the recorded soundfield coefficients in terms of direct sound and reflections. The directional gain of the reflected path is estimated from the above model, which provides information on the DOAs of dominant wall reflections. The practical feasibility of the proposed algorithm is evaluated using a subset of the speech corpus from the ACE (Acoustic Characterization of Environments) Challenge.

1. INTRODUCTION

Soundfield analysis of reverberant enclosures is a topic of interest because it provides knowledge on how the soundfield varies in terms of direction and time. Knowledge of the directions of arrival (DOAs) of dominant wall reflections is particularly desirable in areas such as auditorium/room assessment [1–3], inference of room geometry [4–7], dereverberation [8], determination of psychoacoustic indicators [9] and validation of diffuse field assumption [10]. This paper presents the results from the use of a spherical microphone array to blindly estimate the DOAs of dominant wall reflections in actual rooms.

Initially, soundfield analysis inside rooms was purely based on room impulse response (RIR) measurements from a single microphone [11]. Such measurements only yields omnidirectional pressure, thus the spatial information of the incident reflections are often lost. However, by studying the times of arrival and the room geometry, it is occasionally possible to identify the reflecting surfaces and DOAs. To overcome the shortcomings of a single microphone recording, multiple microphone arrays were later utilized in a variety of ways to derive the directional properties of room reverberation [5, 12, 13]. Some of these microphone array solutions obtained directional information using directional microphones, beamforming [14, 15] and geometric approaches [7].

Recently, the spherical microphone array was introduced for three dimensional (3D) scene analysis inside rooms [16, 17]. In particular, the measurements from a spherical microphone array can be processed to derive the spherical harmonic components or eigenbeams, which characterizes the spatial properties of the soundfield. Gover *et al.* [18] used a 32−channel spherical microphone array to derive the Directional properties of room reverberation via beamforming, and Park and Rafaely [2] applied planewave decomposition using a virtual 98−element spherical array to perform soundfield analysis in an auditorium. Rafaely *et al.* [3] also proposed a dual cocentered open sphere microphone array design for the above method, which provided high resolution due to the large aperture size. The main limitation of these approaches is the high number of microphones required. More recently, with the introduction of portable commercial spherical microphone arrays such as the Eigenmike™, there has been an increased interest in utilizing them for reverberant room analysis. In [19] Sun *et al.* studied the accuracy of optimum array processing methods (steered beamforming and subspace methods) for localization of an unknown source and several dominant room reflections from a single Eigenmike measurement.

In this paper, we propose a novel method to estimate the directional properties of room reflections using a multi-channel coherence estimator [20] based on a spatial correlation model formulated in the spherical harmonics domain. The spatial correlation model is formulated through a spatial correlation matrix containing the spherical harmonic coefficients/eigenbeams derived from a spherical microphone array[1]. The main advantage of this method is its capability to estimate additional information such as the Direct to Reverberant Ration (DRR) as given in [23].

2. PROBLEM FORMULATION

We first consider a spherical array of Q omnidirectional microphones recording the incident soundfield caused by a single source inside the room enclosure of interest. The observed soundfield at the q^{th} $(q = 1, 2, \cdots, Q)$ microphone can be expressed in the time-frequency domain as

$$P(\boldsymbol{x}_q, k, t) = S(k, t) H(\boldsymbol{x}_q, \boldsymbol{y}_o, k) \tag{1}$$

where $k = 2\pi f/c$ is the wavenumber with f and c representing the frequency in Hz and speed of sound in ms^{-1} respectively, $S(k, t)$ is the Short Time Fourier transform of the source signal, t is the temporal frame index, and $H(\boldsymbol{x}_q, \boldsymbol{y}_o, k)$ is the room transfer function (RTF) between the source location $\boldsymbol{y}_o = (r_0, \theta_o, \phi_0)$ and the receiver location $\boldsymbol{x}_q = (r, \theta_q, \phi_q)$. Note that from now on, we omit the time dependency (t) for notational convenience. In a reverberant enclosure, $P(\boldsymbol{x}_q, k)$ would contain the direct path from the source as well as the room response. This decomposition can be reflected in the RTF as

$$H(\boldsymbol{x}_q, \boldsymbol{y}_o, k) = H_{\text{dir}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k) + H_{\text{rvb}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k) \tag{2}$$

*This work is supported by Australian Research Council (ARC) Discovery Projects funding scheme (project no. DP140103412).

[1] Or other alternative structures as given in [21, 22]

where $H_{\text{dir}}(\cdot)$ and $H_{\text{rvb}}(\cdot)$ represent the direct and reflected components of the room impulse response, respectively. Note that, for convenience, we include both early and late reflections in $H_{\text{rvb}}(\cdot)$.

Assuming the aperture size of the microphone array is sufficiently small compared to the distance to the source[2], the direct path of (2) can be considered to be a single plane wave of the form

$$H_{\text{dir}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k) = G_D(k) e^{ik\hat{\boldsymbol{y}}_o \cdot \boldsymbol{x}_q} \tag{3}$$

where $G_D(k)$ denotes the direct path gain and $\hat{\boldsymbol{y}}_o$ denotes the unit vector along the incoming direction. Similarly, the reflected path from all directions can be given by

$$H_{\text{rvb}}(\boldsymbol{x}_q, \boldsymbol{y}_o, k) = \int_{\hat{\boldsymbol{y}}} G_R(k, \hat{\boldsymbol{y}}) e^{ik\hat{\boldsymbol{y}} \cdot \boldsymbol{x}_q} d\hat{\boldsymbol{y}} \tag{4}$$

where $G_R(k, \hat{\boldsymbol{y}})$ is the gain of the reflected plane wave arriving from the direction $\hat{\boldsymbol{y}} = (1, \theta, \phi)$ for $\theta \in [0 : \pi]$ and $\phi \in [0 : 2\pi)$ and $\int_{\hat{\boldsymbol{y}}} d\hat{\boldsymbol{y}} = \int_0^{2\pi} \int_0^{\pi} \sin\theta d\theta d\phi$. From (1), (2), (3) and (4), we re-write $P(\boldsymbol{x}_q, k)$ as

$$P(\boldsymbol{x}_q, k) = S(k)\Big(G_D(k) e^{ik\hat{\boldsymbol{y}}_o \cdot \boldsymbol{x}_q} + \int_{\hat{\boldsymbol{y}}} G_R(k, \hat{\boldsymbol{y}}) e^{ik\hat{\boldsymbol{y}} \cdot \boldsymbol{x}_q} d\hat{\boldsymbol{y}}\Big). \tag{5}$$

By observing the above derivation, the total reflected field power P_R at the observation point is

$$P_R = E\{|S(k)|^2\} \int_{\hat{\boldsymbol{y}}} E\{|G_R(k, \hat{\boldsymbol{y}})|^2\} d\hat{\boldsymbol{y}} \tag{6}$$

where $E\{\cdot\}$ denotes the expectation operator. Our objective is to estimate the directional properties of P_R, which requires the estimation of the spherical function $E\{|G_R(k, \hat{\boldsymbol{y}})|^2\}$ in all look directions $\hat{\boldsymbol{y}}$. To simplify this, we decompose the above function in terms of a set of spatial basis functions called the spherical harmonic functions[3], such that a limited number of coefficients are capable of sufficiently representing the continuous function of interest. This decomposition takes the form of

$$E\Big\{|G_R(k, \hat{\boldsymbol{y}})|^2\Big\} = \sum_{v=0}^{\infty} \sum_{u=-v}^{v} \gamma_{vu}(k) Y_{vu}(\hat{\boldsymbol{y}}), \tag{7}$$

where $Y_{vu}(\cdot)$ represent the v^{th} order u^{th} mode spherical harmonic function and $\gamma_{vu}(k)$ is the corresponding spherical harmonic coefficient. The main task at hand is to estimate these coefficients, which describe the reflected power in any arbitrary look direction in the 3D space.

3. ESTIMATION OF THE DIRECTIONAL PROPERTIES OF ROOM REVERBERATION

In this section, we utilize the spherical harmonic decomposition of soundfields to formulate a spatial correlation matrix, which leads to the estimation of the desired parameters $\gamma_{vu}(k)$.

[2]When the aperture size is much smaller than the radiating wavelength, it is said to be a *point source*, which radiates power equally in all directions with a spherical radiation pattern. At great distances with respect to wavelength from the source, the spherically spreading waves can be regarded as plane waves forming a far-field. A common rule of thumb is far-field sources are located at a distance of $r > 2L^2/\lambda$ where L is the aperture radius and λ is the operating wavelength.

[3]Spherical harmonics are a set of orthonormal spatial basis functions, which can be used to represent functions defined over a sphere. Thus, any spherical function $f(\theta, \phi)$ may be expanded as a linear combination of these basis functions.

3.1. Spatial correlation in the spherical harmonic domain

Formulation of a spatial correlation model requires the soundfield of interest to be represented in the spatial domain. For this purpose, we derive a similar relationship to (5) in terms of the spherical harmonic decomposition of functions on the sphere. The left-hand-side of (5) is the incident soundfield over a spherical surface outlined by the microphone array, and therefore can be expressed in a similar form to (7) as [24]

$$P(\boldsymbol{x}_q, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{\alpha_{nm}(k) b_n(kr)}_{a_{nm}(k)} Y_{nm}(\theta_q, \phi_q). \tag{8}$$

where $a_{nm}(k)$ are the corresponding spherical harmonic coefficients, which are further simplified by $a_{nm}(k) \triangleq \alpha_{nm}(k) b_n(kr)$ for the observed incident soundfield based on the assumption that it is a homogeneous incident soundfield [16, 25] and

$$b_n(kr) = \begin{cases} j_n(kr) & \text{for an open array} \\ j_n(kr) - \frac{j_n'(kr)}{h_n'(kr)} h_n(kr) & \text{for a rigid array} \end{cases} \tag{9}$$

with $j_n(\cdot)$ and $h_n(\cdot)$ denoting the spherical Bessel and Hankel functions of order n respectively and r denoting the radius of the spherical microphone array (e.g., Eigenmike). Note that $\alpha_{nm}(k)$, the incident soundfield coefficients, can be derived up to order $N = \lceil kr \rceil$ using the microphone array recordings $P(\boldsymbol{x}_q, k)$ for $q = 1, 2 \cdots Q$ [16, 25].

Similarly, the spherical functions in the right-hand-side of (5) can be decomposed in terms of spherical harmonics. These include the reverberant gain function $G_R(k, \hat{\boldsymbol{y}})$ distributed over all possible look directions, and the plane wave soundfields $e^{ik\hat{\boldsymbol{y}}_o \cdot \boldsymbol{x}_q}$ and $e^{ik\hat{\boldsymbol{y}} \cdot \boldsymbol{x}_q}$, as observed by the spherical microphone array. We write $G_R(k, \hat{\boldsymbol{y}})$ in terms of

$$G_R(k, \hat{\boldsymbol{y}}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \beta_{nm}(k) Y_{nm}(\theta, \phi). \tag{10}$$

where $\beta_{nm}(k)$ are the respective spherical harmonic coefficients, and $e^{ik\hat{\boldsymbol{y}} \cdot \boldsymbol{x}_q}$ by

$$e^{ik\hat{\boldsymbol{y}} \cdot \boldsymbol{x}_q} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \underbrace{i^n Y_{nm}^*(\theta, \phi) j_n(kr)}_{d_{nm}(k)} Y_{nm}(\theta_q, \phi_q). \tag{11}$$

where the spherical harmonic coefficients are given by by $d_{nm}(k) = i^n Y_{nm}^*(\theta, \phi) j_n(kr)$ for a given planewave incident direction $\hat{\boldsymbol{y}}$ [25].

By substituting (8), (10) and (11) in (5), we derive a modal domain relationship analogous to (5) as

$$\alpha_{nm}(k) = S(k) i^n \big(G_D(k) Y_{nm}^*(\theta_0, \phi_0) + \beta_{nm}(k)\big). \tag{12}$$

where the soundfield coefficients recorded by the microphone array are now represented in terms of their respective direct and reflected components. This relationship serves as the basis for the spatial correlation matrix formulated below. Based on (12), the cross correlation between α_{nm} and $\alpha_{n'm'}$ is

$$\begin{aligned} E\Big\{\alpha_{nm}(k)\alpha_{n'm'}^*(k)\Big\} &= i^n (-i)^{n'} E\Big\{|S(k)|^2\Big\} \\ &\quad \Big(E\Big\{|G_D(k)|^2\Big\} Y_{nm}^*(\theta_0, \phi_0) Y_{n'm'}(\theta_0, \phi_0) \\ &\quad + E\Big\{\beta_{nm}(k)\beta_{n'm'}^*(k)\Big\}\Big). \end{aligned}$$

(13)

When deriving the above result, the cross correlation between the direct path gain and reverberant path gain coefficients was considered to be negligible (the reflection gain from reflective surfaces are independent from the direct path gain).

Assuming that the reflection gains from different incoming directions are uncorrelated, the term $E\Big\{\beta_{nm}(k)\beta^*_{n'm'}(k)\Big\}$ of the above equation can be simplified using (10) and (7) as

$$E\Big\{\beta_{nm}(k)\beta^*_{n'm'}(k)\Big\} = \sum_{v=0}^{\infty}\sum_{u=-v}^{v}\gamma_{vu}(k) \Big(\frac{(2v+1)(2n+1)(2n'+1)}{4\pi}\Big)^{1/2} W_1 W_2 \tag{14}$$

where W_1 and W_2 are Wigner coefficients [26], representing

$$W_1 = \begin{pmatrix} v & n & n' \\ 0 & 0 & 0 \end{pmatrix} \text{ and} \tag{15}$$

$$W_2 = \begin{pmatrix} v & n & n' \\ u & m & -m' \end{pmatrix}. \tag{16}$$

By substituting (14) in (13) we arrive at

$$E\Big\{\alpha_{nm}(k)\alpha^*_{n'm'}(k)\Big\} = i^n(-i)^{n'}E\{|S(k)|^2\}\big(E\{|G_D(k)|^2\} Y^*_{nm}(\theta_0,\phi_0)Y_{n'm'}(\theta_0,\phi_0) + \sum_{v,u}\gamma_{vu}(k) \Big(\frac{(2v+1)(2n+1)(2n'+1)}{4\pi}\Big)^{1/2} W_1 W_2\big) \tag{17}$$

The above result provides a comprehensive expression for the spatial correlation between two spherical harmonic coefficients of an enclosed soundfield, in terms of its direct path component $G_D(k)$ and reverberant path components $\gamma_{vu}(k)$. It can be utilized in any room acoustic application that seeks the separation of direct and reverberant soundfields. In the following section, we define a spatial correlation matrix based on the above result, which leads us to the estimation of the desired coefficients $\gamma_{vu}(k)$.

3.2. Spatial Correlation matrix

We define the *modal domain spatial correlation matrix* $\boldsymbol{R}(k)$ by

$$\boldsymbol{R}(k) \triangleq E\big\{\boldsymbol{\alpha}(k)\boldsymbol{\alpha}^H(k)\big\} \tag{18}$$

where $\boldsymbol{\alpha}(k) = \big[\alpha_{00}(k) \quad \alpha_{1-1}(k) \quad .. \quad \alpha_{NN}(k)\big]^T_{1\times(N+1)^2}$.

By substituting (17) into (18), we obtain

$$\boldsymbol{R}(k) = P_D \begin{bmatrix} b_{0000} & b_{001-1} & \cdots & b_{00NN} \\ b_{1-100} & b_{1-11-1} & \cdots & b_{1-1NN} \\ \vdots & \vdots & \vdots & \vdots \\ b_{NN00} & b_{NN1-1} & \cdots & b_{NNNN} \end{bmatrix} + E\Big\{|S(k)|^2\Big\} \begin{bmatrix} \boldsymbol{d}_{0000} & \boldsymbol{d}_{001-1} & .. & \boldsymbol{d}_{00NN} \\ \boldsymbol{d}_{1-100} & \boldsymbol{d}_{1-11-1} & .. & \boldsymbol{d}_{1-1NN} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{d}_{NN00} & \boldsymbol{d}_{NN1-1} & .. & \boldsymbol{d}_{NNNN} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{1-1} \\ \vdots \\ \gamma_{VV} \end{bmatrix} \tag{19}$$

where $P_D = E\big\{|S(k)|^2|G_D(k,\hat{\boldsymbol{y}})|^2\big\}$ is the direct path power, V is the truncation limit of (7),

$$b_{nmn'm'} = Y^*_{nm}(\theta_0,\phi_0)Y_{n'm'}(\theta_0,\phi_0),$$

$$\boldsymbol{d}_{nmnm'} = [d_{nmnm'00}\; d_{nmnm'1-1}\; d_{nmnm'10}\; d_{nmnm'11} \cdots d_{nmnm'VV}]_{(N+1)^2(V+1)^2\times 1}, \text{ and}$$

$$d_{nmnm'vu} = \Big(\frac{(2v+1)(2n+1)(2n'+1)}{4\pi}\Big)^{1/2} W_1 W_2.$$

Since the spherical microphone array characteristics are initially known, $\boldsymbol{\alpha}(k)$ in (19) can be calculated following the method given in [16, 17]. If the direction of arrival of the source is known or estimated, then $b_{nmn'm'}$ can also be calculated. The terms $\boldsymbol{d}_{nmnm'}$ are composed of known functions. Thus, we can estimate the desired parameters $\gamma_{vu}(k)$ by solving the following set of equations, which were derived by reformulating (19) as

$$\underbrace{\begin{bmatrix} R_{0000} \\ R_{001-1} \\ \vdots \\ R_{00NN} \\ R_{1-100} \\ \vdots \\ R_{NNNN} \end{bmatrix}}_{\widetilde{\boldsymbol{r}}(k)} \approx \underbrace{\begin{bmatrix} b_{0000} & d_{000000} & \cdots & d_{0000VV} \\ b_{000-1} & d_{000-100} & \cdots & d_{000-1VV} \\ \vdots & \vdots & \vdots & \vdots \\ b_{00NN} & d_{00NN00} & \cdots & d_{00NNVV} \\ b_{1-100} & d_{1-10000} & \cdots & d_{1-100-1VV} \\ \vdots & \vdots & \vdots & \vdots \\ b_{NNNN} & d_{NNNN00} & \cdots & d_{NNNNVV} \end{bmatrix}}_{\boldsymbol{B}(k)} \times \underbrace{\begin{bmatrix} P_D \\ E\Big\{|S(k)|^2\Big\}\gamma_{00} \\ E\Big\{|S(k)|^2\Big\}\gamma_{1-1} \\ \vdots \\ E\Big\{|S(k)|^2\Big\}\gamma_{VV} \end{bmatrix}}_{\hat{\boldsymbol{p}}(k)}. \tag{20}$$

Here, $R_{nmn'm'}$ in $\widetilde{\boldsymbol{r}}(k)$ denotes the $(n^2+n+m+1)^{th}$ row and $(n'^2+n'+m'+1)^{th}$ column components of $\boldsymbol{R}(k)$, which can be calculated from the spherical microphone measurements. The desired parameters $\gamma_{vu}(k)$ can be derived by solving (20) using the least-squares method

$$\hat{\boldsymbol{p}}(k) = \boldsymbol{B}^{\dagger}(k)\widetilde{\boldsymbol{r}}(k) \tag{21}$$

where $[\cdot]^{\dagger}$ and $\hat{[\cdot]}$ represents the Pseudo-inverse and estimated value, respectively. To avoid an underdetermined system, the condition $(N+1)^4 > (V+1)^2+1$ has to be satisfied, thus the maximum solvable order of (7) is $V = \Big\lfloor((N+1)^4-1)^{1/2}\Big\rfloor$. Once $\gamma_{vu}(k)$ is derived, the directional properties of the reflected field can be studied by analyzing (6) for each look direction $\hat{\boldsymbol{y}}$. Note that, the term P_D (direct path power) is also derived as a by product of solving (20). In [23], the authors provide a comprehensive study on the use of this result to estimate the DRR of a given room.

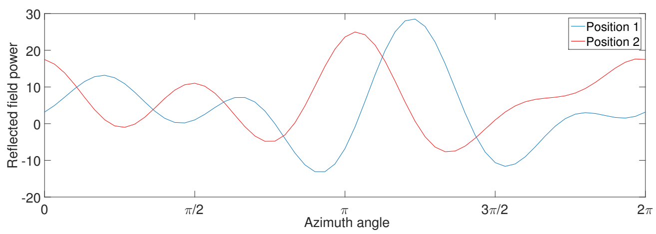

Fig. 1. Estimated reflected field directivity in a Lecture room for two receiver (eigenmike) positions.

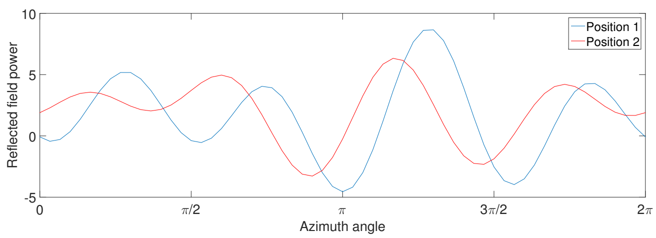

Fig. 2. Estimated reflected field directivity in a Meeting room for two receiver (eigenmike) positions.

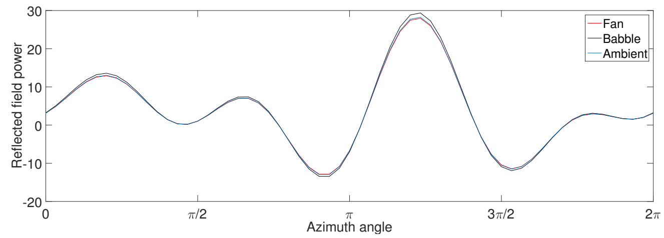

Fig. 3. Estimated reflected field directivity in a lecture room for different noise types at 18 dB SNR.

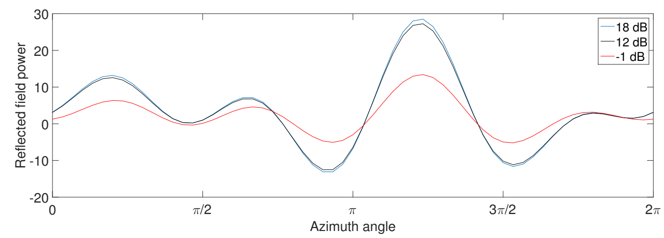

Fig. 4. Estimated reflected field directivity in a lecture room for different noise levels.

4. EXPERIMENTAL RESULTS

We evaluate the proposed method with real acoustic data. We used the ACE Challenge database [27, 28] to retrieve a large corpus of multi-channel recordings spanned over a variety of rooms, speakers and environmental conditions. The ACE Challenge database[4] is a recently developed database to stimulate research in non-intrusive acoustic parameter estimation in realistic environments including noise and reverberation. The results and analysis provided in this paper are based on a subset of the above corpus, which were recorded using an Eigenmike. Included in this subset are 32 channel noisy reverberant speech recordings obtained in 2 rectangular rooms at 2 different Eigenmike positions per room (near and far from the source position) for 5 male talkers and 5 female talkers, each uttering 5 different English phrases, in the presence of 3 types of noise recordings (babble, fan, ambient) at 3 different SNR levels (-1 dB, 12 dB and 18 dB). The reflected field power was estimated up to order $V = 6$ (7).

In the first example, we consider a lecture room of size $(6.93 \times 9.73 \times 3)$ m. The source and eigenmike are both located in the same height, therefore we only plot the reflected field directivity in the horizontal plane across the eigenmike. Figure 1 shows the directional properties of the reflected field power for two different receiver locations. The results are averaged over the frequency band $2000 - 3000$ Hz for 5 male talkers and 5 female talkers, each uttering 5 different English phrases, each in the presence of 3 types of noise recordings (babble, fan, ambient) with a SNR of 18 dB. Note that the reflected field power appears negative at times, which is due to the truncation of the infinite summation in (7). The estimated reflected power for a given direction is thus, only a fraction of the real power. However, for a sufficiently high order V, the truncation effect on the resulting directional pattern is expected to be minimal. It is observed in Fig. 1, that the reflected field power present at the spherical microphone array is directional with distinct peaks occurring at $\pi/2$ intervals. These peaks clearly represent the wall reflections, and the height of each peak suggests the reflective properties of each wall. For different microphone positions, the directional pattern appears to remain similar, with only a fixed rotation according to the microphone movement.

In the second example, we consider a meeting room of size $(6.61 \times 5.11 \times 2.95)$ m, where the source and eigenmike are located in the same horizontal plane. Figure 2 shows the resulting directional properties averaged over the frequency band $2000 - 3000$ Hz for 5 male talkers and 5 female talkers, each uttering 5 different English phrases, each in the presence of 3 types of noise recordings (babble, fan, ambient) with a SNR of 18 dB. Similar to the previous results, the directional pattern shows clear peaks at $\pi/2$ intervals representing the room walls. When the microphone position is changed from Position 1 to Position 2, there appears a slight increase in the height of the second peak, which is due to the microphone moving towards that particular wall.

In the next two examples, we observe the robustness of the proposed estimation method in the presence of noise. In Fig. 3, we show the estimated reflected field directivity in a lecture room for different noise types at 18 dB SNR. In Fig. 4, we show the estimated reflected field directivity averaged over all 3 noise types at different SNR levels. From Fig. 3 it is apparent that the the estimation method is robust for different noise types, however, from Fig. 4, it appears that the the peak strength of directivity degrades with decreasing SNR. This is due to the noise power dominating the reflected field power.

5. CONCLUSION

In this paper, we presented a blind estimation algorithm that utilizes recordings from a spherical microphone to estimate the directional properties of the reflected soundfield. Based on experimental results obtained through real acoustic data, it was shown that the proposed method is capable of clearly recognizing the dominant reflective directions, which is a promising tool for a plethora of applications that require prior knowledge of the reflection characteristics for a given room.

[4]Available at www.ace-challenge.org

6. REFERENCES

[1] L. L. Beranek, “Music, acoustics, and architecture,” *Bulletin of the American Academy of Arts and Sciences*, vol. 45, no. 8, pp. 25–46, 1992.

[2] M. Park and B. Rafaely, “Sound-field analysis by plane-wave decomposition using spherical microphone array,” *The Journal of the Acoustical Society of America*, vol. 118, no. 5, pp. 3094–3103, 2005.

[3] B. Rafaely, I. Balmages, and L. Eger, “High-resolution plane-wave decomposition in an auditorium using a dual-radius scanning spherical microphone array,” *The Journal of the Acoustical Society of America*, vol. 122, no. 5, pp. 2661–2668, 2007.

[4] H. Sun, E. Mabande, K. Kowalczyk, and W. Kellermann, “Joint doa and tdoa estimation for 3d localization of reflective surfaces using eigenbeam mvdr and spherical microphone arrays,” in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2011, pp. 113–116.

[5] E. Mabande, K. Kowalczyk, H. Sun, and W. Kellermann, “Room geometry inference based on spherical microphone array eigenbeam processing,” *The Journal of the Acoustical Society of America*, vol. 134, no. 4, pp. 2773–2789, 2013.

[6] D. Markovic, F. Antonacci, A. Sarti, and S. Tubaro, “Estimation of room dimensions from a single impulse response,” in *2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. IEEE, 2013, pp. 1–4.

[7] F. Antonacci, J. Filos, M. Thomas, E. Habets, A. Sarti, P. A. Naylor, and S. Tubaro, “Inference of room geometry from acoustic impulse responses,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 10, pp. 2683–2695, 2012.

[8] Y. Peled and B. Rafaely, “Method for dereverberation and noise reduction using spherical microphone arrays,” in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2010, pp. 113–116.

[9] L. Cremer and H. A. Müller, *Principles and applications of room acoustics*, vol. 1, Chapman & Hall, 1982.

[10] A. D. Pierce et al., *Acoustics: an introduction to its physical principles and applications*, vol. 20, McGraw-Hill New York, 1981.

[11] H. Kuttruff, *Room acoustics*, CRC Press, 2009.

[12] Y. Yamasaki and T. Itow, “Measurement of spatial information in sound fields by closely located four point microphone method.,” *Journal of the Acoustical Society of Japan (E)*, vol. 10, no. 2, pp. 101–110, 1989.

[13] J. Becker, M. Sapp, and O. Schmitz, “Four-microphone-array measurements combined with geometrical room acoustic simulation technique,” *The Journal of the Acoustical Society of America*, vol. 105, no. 2, pp. 1368–1368, 1999.

[14] T. Nishi, “Relation between objective criteria and subjective factors in a sound field, determined by multivariate analyses,” *Acta Acustica united with Acustica*, vol. 76, no. 4, pp. 153–162, 1992.

[15] H. Okubo, M. Otani, R. Ikezawa, S. Komiyama, and K. Nakabayashi, “A system for measuring the directional room acoustical parameters,” *Applied Acoustics*, vol. 62, no. 2, pp. 203–215, 2001.

[16] T.D. Abhayapala and D.B. Ward, “Theory and design of high order sound field microphones using spherical microphone array,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2002, vol. II, pp. 1949–1952.

[17] J. Meyer and G. Elko, “A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2002, vol. 2, pp. 1781–1784.

[18] B. N. Gover, J. G. Ryan, and M. R. Stinson, “Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array,” *The Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 2138–2148, 2004.

[19] H. Sun, E. Mabande, K. Kowalczyk, and W. Kellermann, “Localization of distinct reflections in rooms using spherical microphone array eigenbeam processing,” *The Journal of the Acoustical Society of America*, vol. 131, no. 4, pp. 2828–2840, 2012.

[20] Y. Hioka, K. Niwa, S. Sakauchi, K. Furuya, and Y. Haneda, “Estimating direct-to-reverberant energy ratio using d/r spatial correlation matrix model,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 8, pp. 2374–2384, 2011.

[21] H. Chen, T. D. Abhayapala, and W. Zhang, “Theory and design of compact hybrid microphone arrays on two-dimensional planes for three-dimensional soundfield analysis,” *The Journal of the Acoustical Society of America*, vol. 138, no. 5, pp. 3081–3092, 2015.

[22] A. Gupta and T. D. Abhayapala, “Double sided cone array for spherical harmonic analysis of wavefields,” in *2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*. IEEE, 2010, pp. 77–80.

[23] P. N. Samarasinghe, T. D. Abhayapala, and H. Chen, “Estimating the direct-to-reverberant energy ratio using a spherical harmonics based spatial correlation model,” *Transactions on Audio, Speech and Language Processing*, under review 2016.

[24] P. N. Samarasinghe, T. D. Abhayapala, and M. A. Poletti, “3d spatial soundfield recording over large regions,” in *Acoustic Signal Enhancement; Proceedings of IWAENC 2012; International Workshop on*. IEEE, 2012, pp. 1–4.

[25] E.G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*, pp. 115–125, Academic Press, London, UK, 1999.

[26] P.A. Martin, *Multiple scattering: interaction of time-harmonic waves with N obstacles*, Number 107. Cambridge University Press, 2006.

[27] J. Eaton, N. D. Gaubitch, A. H. Moore, and P. A. Naylor, “The ace challengecorpus description and performance evaluation,” in *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015 IEEE Workshop on*. IEEE, 2015, pp. 1–5.

[28] J. Eaton, N. D. Gaubitch, A. H. Moore, and P. A. Naylor, “Estimation of room acoustic parameters: The ace challenge,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 10, pp. 1681–1693, Oct 2016.