

3D SOUND FIELD ANALYSIS USING CIRCULAR HIGHER-ORDER MICROPHONE ARRAY

Hanchi Chen, Thushara D. Abhayapala, and Wen Zhang

Research School of Engineering, CECS, The Australian National University, Canberra, Australia
Email: {hanchi.chen, thushara.abhayapala, wen.zhang}@anu.edu.au

ABSTRACT

This paper proposes the theory and design of circular higher-order microphone arrays for 3D sound field analysis using spherical harmonics. Through employing the spherical harmonic translation theorem, the local spatial sound fields recorded by each higher-order microphone placed in the circular arrays are combined to form the sound field information of a large global spherical region. The proposed design reduces the number of the required sampling points and the geometrical complexity of microphone arrays. We develop a two-step method to calculate sound field coefficients using the proposed array structure, i) analytically combine local sound field coefficients on each circular array and ii) solve for global sound field coefficients using data from the first step. Simulation and experimental results show that the proposed array is capable of acquiring the full 3D sound field information over a relatively large spherical region with decent accuracy and computational simplicity.

Index Terms— Microphone arrays, higher-order microphones, spherical harmonics

1. INTRODUCTION

Analyzing and controlling sounds over large regions of space has become an important problem in practice with a broad range of applications, such as security and surveillance systems, environment monitoring, active noise cancellation, and 3D audio rendering. Existing sound field recording techniques using zeroth order (omni-directional) microphones typically require a large set of spatial sampling points as well as complex 3D sampling schemes in order to cover a large spatial region [1]. Reducing the necessary spatial sampling points as well as simplifying the microphone array geometry can significantly reduce implementation difficulty of large microphone arrays. For this purpose, irregular and distributed microphone arrays have been investigated for sound acquisition and sound field measurement within a region [2–4].

Spherical harmonic decomposition of sound fields, also known as mode domain representation of sound field, can be used to interpolate the value of the sound field over a continuous spatial region from the array output [5]. However, 3D sound field analysis using spherical harmonics requires the sensors to be placed uniformly in the three dimensional space; especially, spherical microphone arrays are preferred

given analytical methods exist to calculate the sound field coefficients [1, 6–9]. For high operating frequencies and large observation regions, a massive number of omnidirectional microphones are required to capture the sound field.

Recently, a method of spatial sound field recording using higher-order microphones has been proposed [10, 11]. It has been shown that with the use of multiple higher-order microphones arranged in a spherical array, less sampling points are required for sound field recording compare to omni-directional microphone arrays. However, this method involves the inversion of a very large matrix, which results in high computational effort and risk of ill-conditioning.

The purpose of this paper is to propose a more compact microphone array structure using higher-order microphones. Through following a systematic approach, a microphone array aperture consisting of one or more circular higher-order microphone arrays is proposed. The proposed array aperture only requires a small number of spatial sampling points on a circular geometry. We propose a two-step method for calculating spherical harmonic coefficients using the proposed array structure, with the first step combining the local sound field coefficients on each circular array, and the second step calculating the global coefficients through a series of matrix operations. We show that the proposed method only requires small-sized matrix inversions, thus improving the robustness and computational effectiveness of the system.

2. BACKGROUND THEORY

2.1. Sound field model

This paper uses spherical harmonic decomposition to describe and analyze the sound field within a spherical region of interest. It is assumed that this region is a free space with no sound sources inside. The sound waves propagating inside the region are due to sources outside the region. Defining a spherical coordinate system with its origin \mathcal{O} located at the center of the sphere, the sound pressure $P(R, \vartheta, \varphi, k)$ at any point and frequency within the sphere can be represented as a weighted sum of spherical harmonics [12],

$$P(R, \vartheta, \varphi, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n C_{nm}(k) j_n(kR) Y_{nm}(\vartheta, \varphi), \quad (1)$$

This work was supported under the Australian Research Councils Discovery Projects funding scheme (project no. DP140103412).

where $k = 2\pi f/c$ is the wave number, f and c are the frequency and the wave propagation speed, respectively. C_{nm} are the spherical harmonic coefficients, $j_n(kR)$ is the spherical Bessel function of order n , and $Y_{nm}(\vartheta, \varphi)$ denotes the spherical harmonic of order n and degree m , written as [13]

$$Y_{nm}(\vartheta, \varphi) = \mathcal{P}_{nm}(\cos \vartheta) E_m(\varphi), \quad (2)$$

where $E_m(\varphi) = e^{im\varphi}/\sqrt{2\pi}$ is the normalized circular harmonic,

$$\mathcal{P}_{nm}(\cos \vartheta) = (-1)^m \sqrt{\frac{(2n+1)(n-m)!}{2(n+m)!}} P_{nm}(\cos \vartheta), \quad (3)$$

\mathcal{P}_{nm} denotes the normalized associated Legendre Polynomial of order n and degree m .

In this paper, the sound field with respect to \mathcal{O} is considered the global sound field, and the corresponding coefficients C_{nm} are the global sound field coefficients.

Local sound field

Defining a local origin \mathcal{O}_q whose position with respect to \mathcal{O} is $\mathbf{R}_q = (R_q, \vartheta_q, \varphi_q)$, then the sound pressure at a point $\mathbf{r} = (r, \theta, \phi)$ with respect to \mathcal{O}_q can be expressed by

$$P(r, \theta, \phi) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} B_{\nu\mu}(k) j_{\nu}(kr) Y_{\nu\mu}(\theta, \phi), \quad (4)$$

where $B_{\nu\mu}(k)$ represent the sound field coefficients with respect to the local origin \mathcal{O}_q . The sound field with respect to \mathcal{O}_q is called the local sound field. For simplicity, the wave number k is omitted for the rest of the paper.

2.2. Spherical harmonic translation theorem

The relationship between the local sound field coefficients with respect to \mathcal{O}_q and the global coefficients with respect to \mathcal{O} can be described by the Spherical harmonic translation theorem [13]. The relationship can be written as [10]

$$B_{\nu\mu} = \sum_{n=0}^{\infty} \sum_{m=-n}^n C_{nm} \hat{S}_{n\nu}^{m\mu}(\mathbf{R}_q), \quad (5)$$

where

$$\hat{S}_{n\nu}^{m\mu}(\mathbf{R}_q) = 4\pi i^{\nu-n} \sum_{\ell=|\mu-m|}^{n+\nu+1} i^{\ell} (-1)^{2m-\mu} j_{\ell}(kR_q) Y_{\ell(\mu-m)}^*(\vartheta_q, \varphi_q) W, \quad (6)$$

$$W = \sqrt{\frac{(2n+1)(2\nu+1)(2\ell+1)}{4\pi}} W_1 W_2. \quad (7)$$

Here, W_1 and W_2 denote Wigner 3-j symbols, with

$$W_1 = \begin{pmatrix} n & \nu & \ell \\ 0 & 0 & 0 \end{pmatrix}, W_2 = \begin{pmatrix} n & \nu & \ell \\ m & -\mu & \mu - m \end{pmatrix}. \quad (8)$$

In (5), $B_{\nu\mu}$ are the local sound field coefficients in (4) and C_{nm} are the global sound field coefficients in (1). It can be seen that by substituting (5) into (4), one can derive the sound pressure decomposition of a given point with respect to \mathcal{O}_q using the spherical harmonic coefficients with respect to \mathcal{O} .

3. HIGHER-ORDER MICROPHONE ARRAY

3.1. Higher-order microphone

A higher-order microphone is capable of measuring the local sound field within its proximity, and extracting the sound field coefficients up to a certain spherical harmonics order. Thus if a higher-order microphone of order V is placed at a local origin \mathcal{O}_q , the sound pressure at a point close to \mathcal{O}_q can be expressed by a limited summation of spherical harmonics

$$P_q(r, \theta, \phi) = \sum_{\nu=0}^V \sum_{\mu=-\nu}^{\nu} B_{\nu\mu} j_{\nu}(kr) Y_{\nu\mu}(\theta, \phi). \quad (9)$$

A total of $(V+1)^2$ spherical harmonics and their respective weighing coefficients are present in the summation.

3.2. Continuous circular higher-order microphone array

A concept of continuous circular microphone array has been proposed in [14]. In this paper, this concept is extended for the higher-order microphone case.

Consider a continuous distribution of V th order microphones are placed along a circle (R_s, ϑ_s) , then each higher-order microphone, at a particular azimuth angle φ , is able to detect its local sound field coefficients, denoted as $B_{\nu\mu}(\varphi)$. The relationship between $B_{\nu\mu}(\varphi)$ and C_{nm} is given by the following theorem:

Theorem 1. *Given a set of local sound field coefficients $B_{\nu\mu}(\varphi)$ which are measured along a circle, and an integer m' , their relationship with the global sound field coefficients can be given by*

$$\int_0^{2\pi} B_{\nu\mu}(\varphi) E_{m'}(\varphi) d\varphi = \sum_{n=|\mu-m'|}^{\infty} C_{n(\mu-m')} H_{n\nu}^{(\mu-m')\mu}(R_s, \vartheta_s). \quad (10)$$

where

$$H_{n\nu}^{m\mu}(R_s, \vartheta_s) = 4\pi i^{\nu-n} \sum_{\ell=|\mu-m|}^{n+\nu+1} i^{\ell} (-1)^{2m-\mu} j_{\ell}(kR_s) \mathcal{P}_{\ell(\mu-m)}(\vartheta_s) W. \quad (11)$$

Proof. Using (2), (6) can be rewritten with $\mathbf{R}_s = (R_s, \vartheta_s, \varphi)$,

$$\hat{S}_{n\nu}^{m\mu}(R_s, \vartheta_s, \varphi) = H_{n\nu}^{m\mu}(R_s, \vartheta_s) E_{(m-\mu)}(\varphi), \quad (12)$$

where $H_{n\nu}^{m\mu}(R_s, \vartheta_s)$ is given by (11). Substituting (12) into (5) yields

$$B_{\nu\mu}(\varphi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n C_{nm} H_{n\nu}^{m\mu}(R_s, \vartheta_s) E_{(m-\mu)}(\varphi). \quad (13)$$

Multiplying both sides of (13) with $E_{m'}(\varphi)$ and integrating with respect to φ over $[0, 2\pi)$, due to the orthogonality property of complex exponential functions

$$\int_0^{2\pi} E_{(m-\mu)}(\varphi) E_{m'}(\varphi)^* d\varphi = \delta_{m-\mu, m'}, \quad (14)$$

the integration $\int_0^{2\pi} C_{nm} H_{n\nu}^{m\mu}(R_s, \vartheta_s) E_{(m-\mu)}(\varphi) E_{m'}(\varphi) d\varphi$ is non-zero only when $m = \mu - m'$, thus (13) reduces to (10), which completes the proof. \square

By replacing $B_{\nu\mu}(\varphi)$ with $B_{\nu\mu}(\varphi_q)$, the discrete form of (10) can be written as

$$\frac{1}{Q} \sum_{q=1}^Q B_{\nu\mu}(\varphi_q) E_{\mu-m}(\varphi_q) \approx \sum_{n=|m|}^{\infty} C_{nm} H_{n\nu}^{m\mu}(R_s, \vartheta_s), \quad (15)$$

where Q is the number of sampling points evenly distributed on the circle. In (15), the variable m' has been replaced by $(\mu - m)$ to illustrate the direct relationship between $B_{\nu\mu}$ and C_{nm} . Due to the spatial sampling, an upper bound for the range of $(\mu - m)$ that can be evaluated is given by

$$|\mu - m| \leq \lfloor \frac{(Q-1)}{2} \rfloor. \quad (16)$$

3.3. Solving for global coefficients

A method for calculating the global sound field coefficients C_{nm} up to order N using the local coefficients $B_{\nu\mu}(\varphi_q)$ can be formulated based on (15).

Step 1 of the method is to evaluate the summation on the left hand side of (15). For each existing global sound field mode m , evaluate the summation for all combinations of $B_{\nu\mu}(\varphi_q)$ and m that satisfy (16). Denote the summation as $\alpha_{\nu\mu}^m$, then

$$\alpha_{\nu\mu}^m = \frac{1}{Q} \sum_{q=1}^Q B_{\nu\mu}(\varphi_q) E_{(m-\mu)}(\varphi_q). \quad (17)$$

The second step is to solve a matrix inversion problem to find C_{nm} . Using (15) and (17), the relationship between C_{nm} and $\alpha_{\nu\mu}^m$ can be represented in matrix form as

$$\boldsymbol{\alpha}_m = \mathbf{H}_m \mathbf{C}_m, \quad (18)$$

where $\boldsymbol{\alpha}_m = [\alpha_{00}^m \ \alpha_{1-1}^m \ \alpha_{10}^m \ \dots \ \alpha_{\nu\mu}^m]^T$, and $\mathbf{C}_m = [C_{|m|m} \ C_{(|m|+1)m} \ \dots \ C_{Nm}]^T$ is the set of global coefficients of mode m .

$$\mathbf{H}_m = \begin{bmatrix} H_{|m|0}^{m0} & H_{(|m|+1)0}^{m0} & \dots & H_{N0}^{m0} \\ H_{|m|1}^{m(-1)} & H_{(|m|+1)1}^{m(-1)} & \dots & H_{N1}^{m(-1)} \\ \vdots & \vdots & \ddots & \vdots \\ H_{|m|\nu}^{m\mu} & H_{(|m|+1)\nu}^{m\mu} & \dots & H_{N\nu}^{m\mu} \end{bmatrix}$$

is the matrix that contains the weights for spherical harmonics translation. A solution for \mathbf{C}_m can be found by calculating the Moore-Penrose Pseudo Inverse of \mathbf{H}_m . The size of \mathbf{H}_m is $(V+1)^2$ by $(N-|m|+1)$, which is significantly smaller than the $(N+1)^2$ -by- $(N+1)^2$ matrix inversion proposed in [10], thus both the computational simplicity and the condition of the matrix inversion are significantly better compared to the method in [10].

The complete set of global sound field coefficients is found by solving (18) for $m = [-N : N]$, where N is the maximum order of the global sound field.

Implementing multiple circular higher-order microphone arrays in the global region can improve the robustness and precision of the microphone system. Assuming a total number of K circular arrays are implemented, then in order to calculate the global coefficients, one needs to formulate (18) for each circular array, denoted as $\boldsymbol{\alpha}_{m;K} = \mathbf{H}_{m;K} \mathbf{C}_m$, then the solution for \mathbf{C}_m can be expressed as

$$\mathbf{C}_m = (\hat{\mathbf{H}}_m^* \hat{\mathbf{H}}_m + \lambda \mathbf{I})^{-1} \hat{\mathbf{H}}_m^* \hat{\boldsymbol{\alpha}}_m \quad (19)$$

where $\hat{\mathbf{H}}_m = [\mathbf{H}_{m;1}^T \ \mathbf{H}_{m;2}^T \ \dots \ \mathbf{H}_{m;K}^T]^T$, λ is the regularization parameter, and $\hat{\boldsymbol{\alpha}}_m = [\boldsymbol{\alpha}_{m;1}^T \ \boldsymbol{\alpha}_{m;2}^T \ \dots \ \boldsymbol{\alpha}_{m;K}^T]^T$. Evaluating (19) for $m = [-N : N]$ yields the complete set of global sound field coefficients.

3.4. Dimensionality analysis

Due to the nature of Spherical Bessel functions, only a number of $j_\ell(kR_s)$ are active within a certain radius. A commonly used rule for deciding the active orders of $j_\ell(kR_s)$ is given by [15], i.e.

$$\ell \leq \lceil \frac{ekR_s}{2} \rceil. \quad (20)$$

From (20) and the range of ℓ in (11), we can derive the maximum global spherical harmonic order detectable by a circular V_{th} order microphone array

$$N = V + \lceil \frac{ekR_s}{2} \rceil, \quad (21)$$

where N is the maximum global sound field order detectable and R_s is the radius of the circular array.

The minimum number of sampling points Q on a circle can be derived from eqs (16) and (20), using $\ell \geq |\mu - m|$,

$$Q \geq 2 \lceil \frac{ekR_s}{2} \rceil + 1. \quad (22)$$

4. SIMULATION RESULTS

A series of simulations have been conducted to validate the performance of the proposed array structure. Two instances of the proposed array structure are used in the following simulations. Both array configurations are designed to capture sound fields up to 700 Hz within a sphere with 0.5m radius, with their dimensions determined based on eqs (21) and (22). Multiple circular arrays are employed in both cases to guarantee the quality of the matrix inversion in (19). One design consists of first order microphones arranged into four circular arrays, positioned at $(R_s, \vartheta_s) = (0.4, 90^\circ)$, $(0.34, 72^\circ)$, $(0.28, 108^\circ)$ and $(0.22, 72^\circ)$, the number of first order microphones on each array is 17, 15, 13 and 11, respectively. The second design utilizes only second order microphone arrays, with two circular arrays located at $(R_s, \vartheta_s) = (0.4, 90^\circ)$ and $(0.2, 72^\circ)$, with 17 and 9 second order microphones placed on each array. AWGN is added to the microphone input of all simulations with a SNR of 40 dB. A point source is placed at $(R, \vartheta, \varphi) = (1.6, 60^\circ, -60^\circ)$ for all the simulation setups.

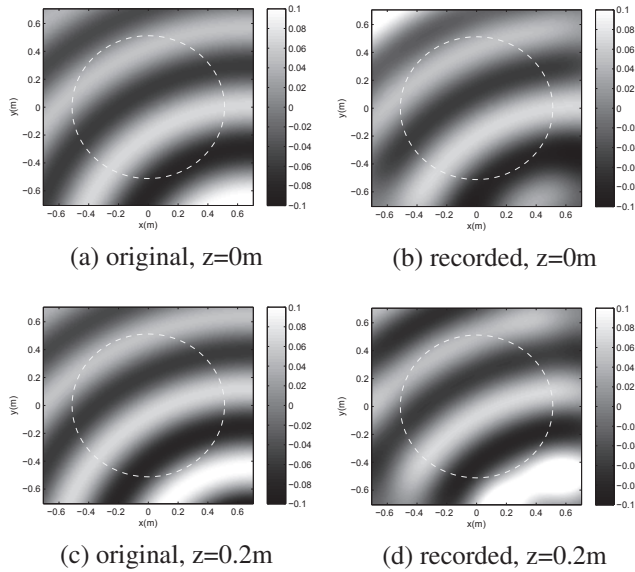


Fig. 1. Comparison of original and recorded sound field due to 700 Hz point source, reconstructed at $z=0$ and $z=0.2$ m plane

Figure 1 shows the simulation result for the first order array configuration. In this simulation, the sound field generated by the point source is recorded by the array, and the resulting sound field coefficients are used to reconstruct the sound field. The sound field is plotted for two layers: the $z = 0$ plane and $z = 0.2$ m plane. Plots (a) and (c) show the original sound field at these two planes, and plots (b) and (d) show the reconstruction of the sound field coefficients obtained from the microphone array. The result shows that the microphone array is capable of accurately capture the sound field within its coverage (yellow circle).

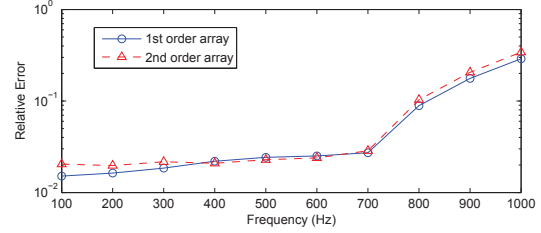


Fig. 2. Reproduction error at different frequencies for first and second order array configurations

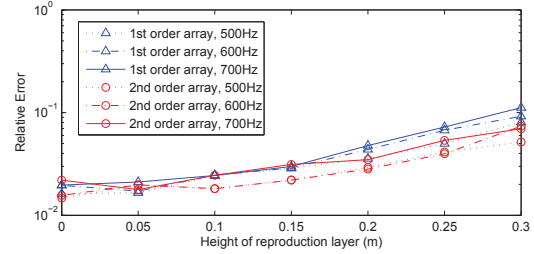


Fig. 3. Reproduction error at different elevations and frequencies for first and second order array configurations

Figure 2 depicts the error performance for two different array configurations at a frequency range of 100 – 1000 Hz. For this figure, the error is calculated by averaging the amplitude error over the entire region of interest, and normalizing by the average sound pressure in the same region. Since both array configurations are designed to operate at up to 700 Hz, it can be seen from Fig. 2 that the reproduction error for both configurations are low for frequencies below 700 Hz, and the error increases rapidly once the frequency becomes higher than the design frequency.

The reproduction error is also evaluated at different planes using the same method, but with the region limited to horizontal planes within the spherical area. The results are shown in Fig. 3. The recorded sound field is reconstructed on planes of different heights, ranging from $z = 0$ to $z = 0.3$ m. The simulation shows that the reproduction error is smaller around the equator compared to that near the poles of the sphere, which is due to the fact that the microphones are clustered around the equator plane.

5. EXPERIMENTAL RESULT

In order to further validate the proposed method, we conducted an experiment of recording a three-dimensional sound field using higher order microphone. We use a single Eigenmike as a 4th order microphone, which consists of 32 condenser microphone capsules placed on a rigid sphere of 4.2cm radius. The goal of the experiment is to test the robustness of the algorithm with the presence of noise and interference in a real-life system.

The region of interest is set to be a sphere of 25cm diameter, a loudspeaker is placed at $(R, \theta, \phi) = (1.5\text{m}, 90^\circ, 30^\circ)$ with respect to the center of the region of interest. The Eigenmike is placed on the equatorial plane of the spherical region and is moved around a circle of 10cm radius. A total of 25 sampling points are evenly distributed along the circle. At each sampling point, the Eigenmike records a sweeping signal played by the loudspeaker, which is then converted to a set of 4th order spherical harmonic coefficients. The 25 sets of local coefficients are then combined using the proposed method to compute the 13th order global coefficients. A visualization of the reconstructed sound field at 3500 Hz within the region of interest is shown in Fig. 4.

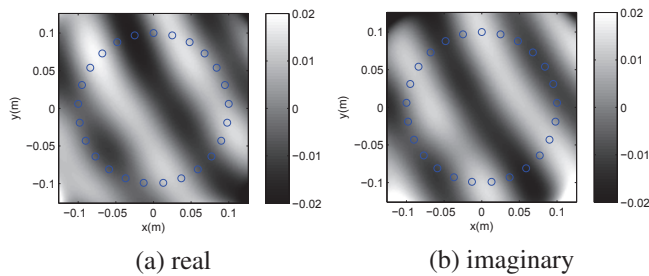


Fig. 4. Reconstructed sound field at 3500Hz due to a loudspeaker placed at $(1.5\text{m}, 90^\circ, 30^\circ)$, sampling points are indicated by blue circles.

We have identified the primary causes of error to be the sensitivity variation of the microphone capsules, and the reverberation inside the laboratory. Despite said interferences, the Eigenmike was able to record the sound field with acceptable accuracy.

We believe that the proposed spatial sampling method allows for easier implementation of sound field recording systems compared to spherical sampling methods, especially when combined with the recording technique used in this experiment, and for applications such as room response modelling over a large space.

6. CONCLUSION

This paper presents a circular higher-order microphone array structure and an associated analytical algorithm for sound field analysis based on spherical harmonics decomposition. Simulations and experiments show that the proposed array architecture offers decent accuracy and robustness, and has the potential of simplifying sound field recording systems in certain applications.

REFERENCES

- [1] 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)*, May 2002, vol. 2, pp. II-1949–II-1952.
- [2] P.-A. Gauthier, C. Camier, Y. Pasco, A. Berry, E. Chambatte, R. Lapointe, and M.-A. Delalay, "Beamforming regularization matrix and inverse problems applied to sound field measurement and extrapolation using microphone array," *Journal of Sound and Vibration*, vol. 330, no. 24, pp. 5852 – 5877, 2011.
- [3] S.M. Golan, S. Gannot, and I Cohen, "Performance analysis of a randomly spaced wireless microphone array," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 121–124.
- [4] T. Takagi, H. Noguchi, K. Kugata, M. Yoshimoto, and H. Kawaguchi, "Microphone array network for ubiquitous sound acquisition," in *Proc. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, March 2010, pp. 1474–1477.
- [5] Y. Haneda, K. Furuya, S. Koyama, and K. Niwa, "Close-talking spherical microphone array using sound pressure interpolation based on spherical harmonic expansion," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 604–608.
- [6] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2002, vol. 2, pp. II-1781–II-1784.
- [7] B. Rafaely, "The spherical-shell microphone array," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 4, pp. 740–747, May 2008.
- [8] C. Jin, A. Parthy, and A. Van Schaik, "Optimisation of centred rigid and open spherical microphone arrays," in *Proc. 120th Audio Engineering Society Convention*, Paris, France, May, 2006 2006, p. 6 pages, Audio Engineering Society.
- [9] T. D. Abhayapala and M.C.T. Chan, "Limitation and error analysis of spherical microphone arrays," in *Proc. 14th International Congress on Sound and Vibration (ICSV)*, Cairns, Australia, July 2007, p. 8 pages.
- [10] P. N. Samarasinghe, T. D. Abhayapala, and M. A Poletti, "3D spatial soundfield recording over large regions," in *Proc. International Workshop on Acoustic Signal Enhancement (IWAENC)*, Sep. 2012, pp. 1–4.
- [11] P. N. Samarasinghe, T. D. Abhayapala, and M. A Poletti, "Wavefield analysis over large areas using distributed higher order microphones," *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, vol. 22, no. 3, pp. 647–658, Mar. 2014.
- [12] E. G. Williams, *Fourier Acoustics: Sound Radiation and Near field Acoustical Holography*, USA: Academic, 1999.
- [13] P. A. Martin, *Multiple scattering: interaction of time harmonic waves with N obstacles*, Cambridge Univ., 2006.
- [14] T. D. Abhayapala and A Gupta, "Spherical harmonic analysis of wavefields using multiple circular sensor arrays," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 6, pp. 1655–1666, Aug 2010.
- [15] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2542–2556, 2007.