NON-SPHERICAL MICROPHONE ARRAY STRUCTURES FOR 3D BEAMFORMING AND SPHERICAL HARMONIC ANALYSIS

Thushara D. Abhayapala and Aastha Gupta

Applied Signal Processing Group
Department of Information Engineering, RSISE, CECS
The Australian National University
Email: {Thushara.Abhayapala, Aastha.Gupta}.anu.edu.au

ABSTRACT

Decomposition of a soundfield into spherical harmonics is a fundamental problem in acoustic signal processing. This paper shows how to design non-spherical microphone array structures in 3D space to decompose a soundfield into spherical harmonic components. We use the mode limiting property of the Bessel functions and zeros of the associated Legendre functions together with orthoganality of the exponential functions over circles to construct an array of parallel circles of microphones. The result provides flexible design guidelines to construct 3D arrays than traditional spherical arrays. A simple beamforming example is given to verify the result.

***Index Terms*—** Spherical array of microphones, spherical harmonics, circular array, beamforming, soundfield

1. INTRODUCTION

Array of microphone structures, which are capable of processing acoustic signals to extract useful spatial information of three dimensional surroundings, are important in a plethora of applications such as beamforming, direction of arrival estimation, and spatial soundfield recording. The spherical harmonic analysis of the 3D acoustic field, which is commonly known as modal analysis [1] technique, has been shown to be a useful tool to design signal processing algorithms for these applications [2–5]. The spherical array of microphones are suited in decomposing a 3D acoustic field into spherical harmonic components [2, 6–8], which could then be processed as necessary to form beams in desired directions, to estimate source locations, or to record spatial sounds. In this paper, we show an alternative 3D structure consisting of circular arrays that is different to a traditional spherical array structure, to decompose acoustic field into spherical harmonic coefficients.

Meyer and Elko [9] proposed a method to use circular arrays of microphones on the x-y plane together with a centre microphone at the origin to extract spherical harmonic coefficients. This is a novel use of circular arrays. Typically the use of circular arrays is to decompose a sound field in to cylindrical harmonics, which is more suited for height invariant 2D soundfields. Although, Meyer's work gives some flexibility in controlling the vertical spatial response, fundamentally a 2D array on a x-y plane is not able to determine all of the spherical harmonic coefficients. We show this fact in Section 3. In this paper, we investigate the spherical harmonic decomposition and propose a systematic way to build a 3D flexible array structure consisting of circular arrays in a set of parallel planes, which are parallel to x-y plane.

2. SOUNDFIELD ANALYSIS

2.1. Spherical harmonic expansion

A soundfield at a point (r, θ, ϕ) in a source free region can be expressed in terms of the spherical harmonic expansion as

$$S(r, \theta, \phi; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \alpha_{nm}(k) j_n(kr) Y_{nm}(\theta, \phi) \qquad (1)$$

where m and n (≥ 0) are integers, $\alpha_{nm}(k)$ are the spherical harmonic coefficients of the soundfield, $k = 2\pi f/c$ is the wavenumber, f is the frequency, c is the speed of sound, $j_n(\cdot)$ are the spherical Bessel functions of order n, and the spherical harmonics

$$Y_{nm}(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_{n|m|}(\cos\theta) e^{jm\phi} \qquad (2)$$

which are defined in terms of the associated Legendre functions $P_{n|m|}(\cdot)$ and the exponential functions. Knowing the soundfield over angles, harmonic coefficients can be calculated using

$$\alpha_{nm}(k) = \frac{1}{j_n(kr)} \int_0^{2\pi} \int_0^{\pi} S(r, \theta, \phi; k) Y_{nm}^*(\theta, \phi) \qquad (3)$$

provided $j_n(kr) \neq 0$. For convenience, we express (1) in terms of the normalized associated Legendre and exponential functions as

$$S(r, \theta, \phi; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \alpha_{nm}(k) j_n(kr) \mathcal{P}_{n|m|}(\cos\theta) E_m(\phi) \qquad (4)$$

where $E_m(\phi) \triangleq (1/\sqrt{2\pi}) e^{jm\phi}$ and

$$\mathcal{P}_{n|m|}(\cos\theta) \triangleq \sqrt{\frac{2n+1}{2}} \sqrt{\frac{(n-|m|)!}{(n+|m|)!}} P_{n|m|}(\cos\theta), \qquad (5)$$

which form orthonormal basis sets in azimuth $\phi \in [0, 2\pi)$ and elevation $\theta \in [0, \pi]$.

2.2. Truncation

The representation series (4) can be safely truncated [10] to a finite number using the properties of the Bessel function (see Fig. 1) provided that all sources are located outside the region of interest (e.g., aperture of the array structure). Let R be the radius of a spherical region of interest, then the soundfield inside this sphere can be represented by (4), with summation over n truncated to $N = \lceil ekR/2 \rceil$ terms [10]. We approximate this bound to $N = \lceil kR \rceil$ to use in this paper.

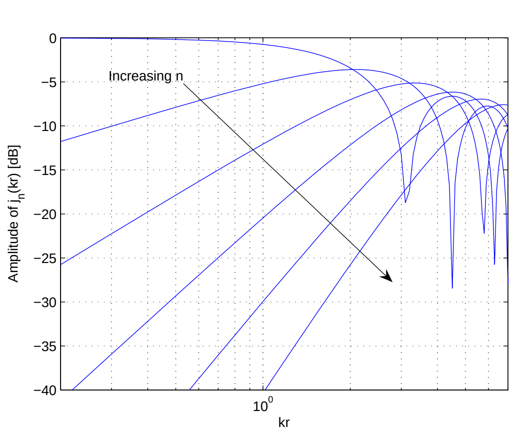

Fig. 1: Magnitude of the spherical Bessel functions of order $n = 0, 1, 2, 3, 4, 5$ in dB showing the characteristics as a function of the argument.

2.3. Soundfield Coefficients

Suppose the soundfield is restricted to have only first $N+1$ coefficients due to finite restriction of the region of interest. Then, we can see from (4) that there are a total of $(N+1)^2$ coefficients to be determined. Table 1 depicts the growth of number of coefficients to be determined as order n grows with modes ranging from $-n$ to n. The soundfield coefficients can be estimated by sampling the space

$m \diagdown n$	0	1	2	$\dots$	N
N					α_{NN}
2			α_{22}		$\vdots$
1		α_{11}	α_{21}		
$m=0$	α_{00}	α_{10}	α_{20}	$\dots$	α_{N0}
-1		$\alpha_{1(-1)}$	$\alpha_{2(-1)}$		
-2			$\alpha_{2(-2)}$		$\vdots$
$-N$					$\alpha_{N(-N)}$

Table 1: Soundfield coefficients arranged with order n and mode m.

using an array of sensors. Spherical microphone arrays approximate the analysis equation (3) by a sum of signal samples taken over the spherical surface to perform this task. However, there are known limitation of the spherical arrays [7] such as strict orthogonality condition and inflexibility with the sensor geometry. In this paper, we develop an alternative array structure to estimate soundfield coefficients.

3. CIRCULAR APERTURE

In this section, we investigate the soundfield on a circle parallel to the x-y plane. Let $S(r_q, \theta_q, \phi; k)$ be the the soundfield on a circle given by $\theta = \theta_q$ and $r = r_q$, where θ_q and r_q are suitably chosen constants. We use (4) to write

$$S(r_q, \theta_q, \phi; k) = \sum_{n=0}^{N} \sum_{m=-n}^{n} \alpha_{nm}(k) j_n(kr_q) \mathcal{P}_{n|m|}(\cos\theta_q) E_m(\phi) \tag{6}$$

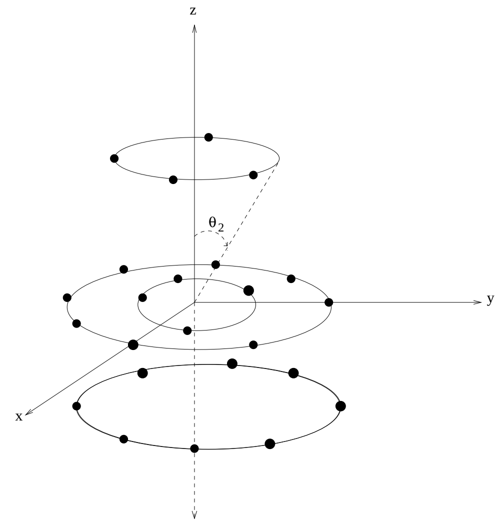

Fig. 2: Circular array structure for 3rd order spherical harmonic decomposition

where $N = \lceil kr_q \rceil$ is due to the natural truncation property (see Section 2.2). We multiply (6) by $E_{-m}(\phi)$ and integrate with respect to ϕ over $[0, 2\pi)$ to get

$$\sum_{n=|m|}^{N} \alpha_{nm}(k) j_n(kr_q) P_{n|m|}(\cos\theta_q) = \int_0^{2\pi} S(r_q, \theta_q, \phi; k) E_{-m}(\phi) d\phi. \tag{7}$$

We have the following comments regarding (7):

- Left hand side (LHS) of (7) is a weighted sum of soundfield coefficients along a row for a given m in Table 1.
- Equation (7) can be evaluated for $m = -N, \dots, N$, where the truncation number N is dependent on the radius r_q of the circle.
- For a given circle (r_q, θ_q) some of the spherical harmonics could be zero if either $j_n(kr_q) = 0$ or $\mathcal{P}_{n|m|}(\cos\theta_q) = 0$. Thus, care needs to be taken while using (7) for coefficient calculations. Figures 3 and 4 depict the magnitude of normalized associated Legendre functions $\mathcal{P}_{n|m|}(\cos\theta)$ in dB.
- The main contribution of this paper is to extract spherical harmonic coefficients $\alpha_{nm}(k)$ by exploiting (7), using a number of carefully placed circles on different (r_q, θ_q) together with properties of spherical Bessel functions and associated Legendre functions.

Suppose we are interested in designing a Nth order microphone array to estimate $(N+1)^2$ spherical harmonic coefficients. Consider there are $Q \geq (N+1)$ circles of microphones located on planes given by (r_q, θ_q), $q = 1, \dots, Q$. Also choose r_q such that $n = \lceil kr_q \rceil$, where $n = 0, \dots, N$, i.e., each r_q corresponds to a specific n. This condition is strictly not necessary but significantly simplifies the estimation of α_{nm} as smaller circles do not contain higher order harmonic coefficients; e.g., a single sensor at the origin only contains α_{00}.

By writing (7) for a specific m for all applicable circles, we have

$$\boldsymbol{J}_m \boldsymbol{\alpha}_m = \boldsymbol{a}_m, \text{ for } m = -N, \dots, N \tag{8}$$

where $\boldsymbol{J}_m =$

$$\begin{bmatrix} j_{|m|}(kr_1)\mathcal{P}_{|m||m|}(\cos\theta_1) & \cdots & j_N(kr_1)\mathcal{P}_{N|m|}(\cos\theta_1) \\ \vdots & \ddots & \vdots \\ j_{|m|}(kr_Q)\mathcal{P}_{|m||m|}(\cos\theta_Q) & \cdots & j_N(kr_Q)\mathcal{P}_{N|m|}(\cos\theta_Q) \end{bmatrix}, \tag{9}$$

$\boldsymbol{\alpha}_m = [\alpha_{|m|m}, \alpha_{|m+1|m}, \ldots, \alpha_{Nm}]^T$, and $\boldsymbol{a}_m = [a_{1m}, \ldots, a_{Qm}]^T$, with $a_{qm} = \int_0^{2\pi} S(r_q, \theta_q; k) E_{-m}(\phi) d\phi$.

If (r_q, θ_q), $q = 1, \ldots, Q$ are chosen such that $\boldsymbol{J}_m$ has a valid Moore-Penrose inverse $\boldsymbol{J}_m^+$, then $\boldsymbol{\alpha}_m$ can be calculated for each m by solving (8) in the least squares sense as

$$\boldsymbol{\alpha}_m = \boldsymbol{J}_m^+ \boldsymbol{a}_m. \tag{10}$$

However, we can be a bit more creative in placing circles to exploit the underlying structure of the wave propagation rather than relying on the ability of the least squares to do the job. To do this, we inspect (8) for specific values of (r_q, ϕ_q) and m:

- For a single sensor at the origin, $(r_q, \phi_q) = (0, 0)$, the only available mode is $m = 0$, thus $\alpha_{00} = a_{00}(0, 0)/\mathcal{P}_{0|0|}(1)$.
- For $\theta = \pi/2$, $\mathcal{P}_{n|m|}(\cos(\pi/2)) = 0$ if $n + |m|$ is odd. This can be seen in Fig. 4 which depicts the normalized associated Legendre functions up to order 3 for odd $n+|m|$ values. Also note that from Fig. 3, for $n + |m|$ even values, normalized associated Legendre functions do not greatly attenuate the corresponding harmonic coefficients at $\theta = \pi/2$. In this paper, we use this property in our design example to illustrate the power of this technique.
- When $m = N$, (8) reduces to a single equation with one unknown α_{NN}. However, if we use this equation alone to calculate α_{NN} there may be significant errors involved as there could be contributions from $\alpha_{(N+1)N}$ in this equation. Note that from Fig. 1, dB reduction from the first mode to the second mode at r (with $1 = \lceil kr \rceil$) is 7dB, from 2 to 3 at the circle with $2 = \lceil kr \rceil$ is about 5dB, and so on. Thus, to improve the accuracy of this calculation, we may need measurements from an additional circle. However, for a set of circles on the $\theta = \pi/2$ plane, this is not a a problem since there is no contribution from the next mode because its sum of order and mode is odd.
- Also note that any sensor on the z-axis only contains $m = 0$ coefficients.

Above are some of the properties we can observe from (7) together with the characteristics of spherical Bessel functions (Figure 1 and normalized associated Legendre functions (Figures 3 and 4). We assert that a design engineer can use these properties to construct a 3D array structure to estimate spherical harmonic coefficients of spatial soundfields, without relying on the spherical array structure as has been done in literature [2, 6–8].

Until now we showed how to calculate sound field coefficients given the soundfield on a number of circular apertures. In practice, we can not obtain soundfield at every point on these circles. In the following section, we show how to use only samples of the soundfield on these circles. This will enable us to give design guidelines to build practical microphone arrays.

4. IMPLEMENTATION

Let the design objective be to construct an array capable of decomposing soundfield coefficients up to N modes. We use a set of circular arrays to calculate the spherical harmonic coefficients.

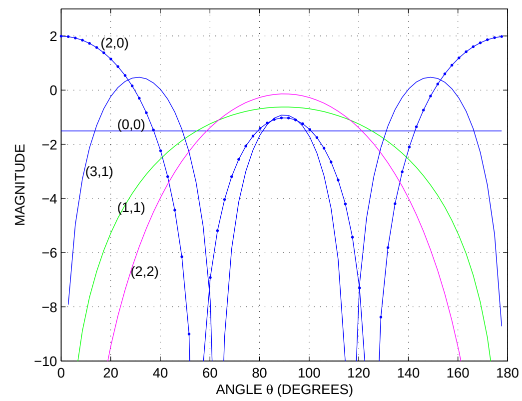

Fig. 3: Magnitude of the normalized associate Legendre functions $\mathcal{P}_{n|m|}(\cos\theta)$ in dB, where the addition of order n and mode m are even: $(n, |m|) = (0, 0); (2, 0); (1, 1); (2, 2); (3, 1); (3, 3)$

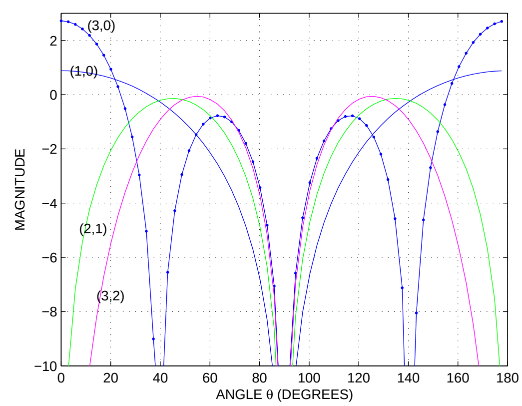

Fig. 4: Magnitude of the normalized associate Legendre functions $\mathcal{P}_{n|m|}(\cos\theta)$ in dB, where the addition of order n and mode m are even: $(n, |m|) = (1, 0); (2, 1); (3, 0); (3, 2)$

4.1. Sampling of circles

For the circular aperture at r_q, we need to evaluate the integral in (7) with a summation since in practice we can only have a finite number of samples of $S(r_q, \theta_q, \phi; k)$. For this radius, if the field is limited to n orders (i.e. $n = \lceil kr_q \rceil$), the maximum mode m involved is n. Thus, $S(r_q, \theta_q, \phi; k)$ is mode limited to n, i.e., it contains terms with $e^{jm\phi}$ with $m = 0, \ldots, n$. According to Shannon's sampling theorem for periodic functions, $S(r_q, \theta_q, \phi; k)$ can be reconstructed by its samples over $[0, 2\pi]$ with at least $2n + 1$ samples. Hence we can approximately replace the right hand side (RHS) of (7) by

$$\text{RHS of (7)} \approx \frac{2\pi}{V_n} \sum_{v=1}^{V_n} S(r_q, \theta_q, \phi_v; k) E_{-m}(\phi_v), \tag{11}$$

where $V_n \geq 2n + 1$. Thus, to calculate coefficients of all orders and modes up to Nth order, we need at least $(N+1)^2$ microphones over $N + 1$ circles. That is, a single sensor at the origin, 3 sensors on a the next circle, followed by 5 sensors, and ends with $2N + 1$ sensors on the $N + 1$th circle.

4.2. Placing circles

By following the description given in Section 2.3, we place alternative circles on the x-y plane (i.e. $\theta_q = pi/2$) corresponding to $n = \lceil kr_q \rceil$ with $n = 1, 3, 5, \ldots$. Then, we can write (8) as $\boldsymbol{J}^{\mathrm{e}}_m \boldsymbol{\alpha}^{\mathrm{e}}_m = \boldsymbol{a}_m$, where $\boldsymbol{J}^{\mathrm{e}}_m$ and $\boldsymbol{\alpha}^{\mathrm{e}}_m$ contain elements corresponding to even $n + |m|$. This equation could be solved to find coefficients α_{nm} with $n + |m|$ even. Observe from Fig. 4 that we can place the second set of circles corresponding to $n = 2, 4, \ldots$ at $\theta_q \in \{[50^\circ :, 70^\circ], [110^\circ, 130^\circ]\}$ to calculate the rest of the coefficients. In this case, we can rewrite (8) as $\boldsymbol{J}^{\mathrm{o}}_m \boldsymbol{\alpha}^{0}_m = \boldsymbol{a}_m - \boldsymbol{J}^{\mathrm{e}}_m \boldsymbol{\alpha}^{\mathrm{e}}_m$ where $\boldsymbol{J}^{\mathrm{o}}_m$, $\boldsymbol{\alpha}^{0}_m$ and $\boldsymbol{J}^{\mathrm{e}}_m$, $\boldsymbol{\alpha}^{\mathrm{e}}_m$ can be defined accordingly to represent odd and even $n + |m|$ respectively.

4.3. Broadband

To make the array work for broadband without adding infinitely many circles for each frequency, we need to reuse existing set of circles for higher frequencies. Suppose the desired frequency band is $[k_{\mathrm{l}}, k_{\mathrm{u}}]$. Let $r_n(k_{\mathrm{l}})$ be the radial distance to the nth circle from the origin corresponding to frequency k_{l}, which we use to calculate coefficients of order n. Since we use the constraint $n = \lceil kr_n \rceil$ to choose $r_n(k_{\mathrm{l}})$, it is easy to see that $r_n(k_{\mathrm{l}}) = r_{sn}(s \times k)$, where s is a positive integer. That is, we can use the the same circle for realizing higher order coefficients of the integer multiplier of the initial frequency. However, we need more sensors on the circle to calculate higher orders. This concept could be considered as similar to nested arrays in the line array literature for broadband. We stop short of giving detail broadband design in this paper due to space constraints.

5. SIMULATIONS

To illustrate our theoretical array design, we show an example of a 3rd order ($N = 3$) array of circular arrays of microphones in 3D beamforming. Let the design frequency be 1.5KHz and the speed of sound propagation $c = 340\mathrm{ms}^{-1}$. First, we place a microphone at the origin to estimate α_{00}. Next we have two circular arrays on the x-y plane ($\theta_q = 0$) with radii $1/k = 3.6$cm and $3/k = 10.8$cm respectively. Two circles have 4 and 8 uniformly separated microphones, which are one more than the minimum required. The output from these microphones on the x-y plane can be used to calculate, all α_{nm} coefficients where $n + |m|$ is an even number (up to $n = 3$). Finally, we place two more circles with $(r_3, \theta_3) = (1/k, 0.9)$ and $(r_4, \theta_4) = (3/k, \pi - 0.9)$ with 4 and 8 microphones respectively. We use the output of the second set of microphones together with already estimated coefficients to calculate the remaing coefficient set as outlined in Section 4.2. Figure 2 shows an example of array geometry.

By weighting each estimated coefficient by

$$\big(i^{-n}/(2n+1)\big)\mathcal{P}_{n|m|}(\cos\theta_\ell)E_m(\phi_\ell)$$

and summing together will give a beamformer in the look direction (θ_ℓ, ϕ_ℓ). Figure 5 depicts the response of the above array structure steered to the look dirction of $(\pi/3, \pi/3)$.

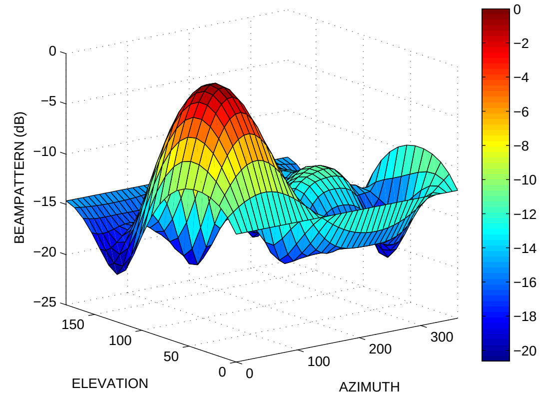

Fig. 5: Response of the Beamformer example in Section 5 where the look direction $(\pi/3, \pi/3) = (60^\circ, 60^\circ)$.

6. REFERENCES

[1] T.D. Abhayapala, *Modal Analysis and Synthesis of Broadband Nearfield Beamforming Arrays*, Ph.D. thesis, Australian National University, Canberra, Australia, 1999.

[2] 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*, May 2002, vol. 2, pp. 1781–1784.

[3] H. Teutsch and W. Kellerman, "Estimation of the number of wideband sources in an acoustic wave field using eigen-beam processing for circular apertures," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2005, pp. 110–113.

[4] T.D. Abhayapala, R.A. Kennedy, and R. Williamson, "Nearfield broadband array design using a radially invariant modal expansion," *Journal of the Acoustical Society of America*, vol. 107 (1), pp. 392–403, 2000.

[5] T.D. Abhayapala and H. Bhatta, "Coherent broadband source localization by modal space processing," in *10th International Conference on Telecommunications, TCT 2003*, 2003, vol. 2, pp. 1617–1623.

[6] T.D. Abhayapala and D.B. Ward, "Theory and design of higher order sound field microphones using spherical microphone array," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2002, vol. 2, pp. 1949–1952.

[7] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Acoust. Speech Sig. Proc.*, vol. 13, pp. 135–143, Jan. 2005.

[8] Z. Li, R. Duraiswami, and L.S. Davis, "Flexible layout and optimal cancellation of the orthonormality error for spherical microphone arrays," in *IEEE International Conference Acoustics, Speech and Signal Processing,*, May 2004, vol. 4, pp. 41–44.

[9] J. Meyer and G. Elko, "Spherical harmonic modal beamforming for an augmented circular microphone array," in *IEEE International Conference Acoustics, Speech and Signal Processing,*, March 2008, pp. 5280–5283.

[10] T.D. Abhayapala T.S. Pollock R.A. Kennedy, "Characterization of 3d spatial wireless channels," in *IEEE 58th Vehicular Technology Conference, VTC 2003-Fall*, 2003, vol. 1, pp. 123–127.