

FLEXIBLE LAYOUT AND OPTIMAL CANCELLATION OF THE ORTHONORMALITY ERROR FOR SPHERICAL MICROPHONE ARRAYS

Zhiyun Li, Ramani Duraiswami, Elena Grassi, Larry S. Davis

Perceptual Interfaces and Reality Laboratory, UMIACS, University of Maryland, College Park

ABSTRACT

This paper describes an approach to achieving a flexible layout of microphones on the surface of a spherical microphone array for beamforming. Our approach achieves orthonormality of spherical harmonics to higher order for relatively distributed layouts. This gives great flexibility in microphone layout on the spherical surface. One direct advantage is it makes much easier to build a real world system, such as those with cable outlets and a mounting base, with minimal effects on the performance. Simulation results are presented.

1. INTRODUCTION

Spherical arrays of microphones are recently becoming the subject of some study as they allow omnidirectional sampling of the sound-field, and may have applications in sound-field capture. The paper [2] presented a preliminary analysis of such arrays, and showed how sound can be analyzed using them. This paper performed an elegant separation of the analysis and beamforming parts by using a modal beamformer structure. Here, we propose an extension of this approach that allows flexible microphone placements.

This paper is organized into four parts. First, we will present the basic principle of beamformer of spherical microphone array. Second, we will give a theoretical analysis of the discrete system. In that part, we analyze how the orthonormality error is introduced into the system, how it gets amplified and how it affects on performance. To cancel the error noise, we design a flexible and optimal system in the third part. In the fourth part, we will provide the design example and simulation results.

2. BACKGROUND

The basic principle of Ref. [2] is to make use of the orthonormality of spherical harmonics to decompose the sound-field arriving at a spherical array. Then the orthogonal components of the soundfield are linearly combined to approximate a desired beampattern.

This work was partially supported by NSF Award 0205271. We thank Dr. Nail A. Gumerov for some discussions.

2.1. Scattering Theory

For a unit magnitude plane wave \mathbf{k} , incident from direction (θ_k, φ_k) , the incident field at a point $(\theta_s, \varphi_s, r_s)$ is

$$p_i = e^{i\mathbf{k}\cdot\mathbf{r}_s} = 4\pi \sum_{n=0}^{\infty} i^n j_n(kr_s) \sum_{m=-n}^n Y_n^m(\theta_k, \varphi_k) Y_n^{m*}(\theta_s, \varphi_s); \quad (1)$$

where j_n is the spherical Bessel function of order n , Y_n^m is the spherical harmonics of order n and degree m . At the same point, the field scattered by the rigid sphere of radius a is [1]:

$$p_s = -4\pi \sum_{n=0}^{\infty} i^n \frac{j'_n(ka)}{h'_n(ka)} h_n(kr_s) \sum_{m=-n}^n Y_n^m(\theta_k, \varphi_k) Y_n^{m*}(\theta_s, \varphi_s). \quad (2)$$

The total field on the surface ($r_s = a$) of the rigid sphere is:

$$p_t = (p_s + p_i)|_{r_s=a} \quad (3)$$

$$= 4\pi \sum_{n=0}^{\infty} i^n b_n(ka) \sum_{m=-n}^n Y_n^m(\theta_k, \varphi_k) Y_n^{m*}(\theta_s, \varphi_s),$$

$$b_n(ka) = j_n(ka) - \frac{j'_n(ka)}{h'_n(ka)} h_n(ka), \quad (4)$$

where h_n is the spherical Hankel function of the first kind.

2.2. Soundfield Decomposition and Beamforming

If we assume that the pressure recorded at each point (θ_s, φ_s) on the surface of the sphere Ω_s , is weighted by

$$W_{n'}^{m'}(\theta_s, \varphi_s, ka) = \frac{Y_{n'}^{m'}(\theta_s, \varphi_s)}{4\pi i^{n'} b_{n'}(ka)}. \quad (5)$$

Then making use of orthonormality of spherical harmonics:

$$\int_{\Omega_s} Y_n^{m*}(\theta_s, \varphi_s) Y_{n'}^{m'}(\theta_s, \varphi_s) d\Omega_s = \delta_{nn'} \delta_{mm'} \quad (6)$$

the total output from a pressure-sensitive spherical surface is:

$$P = \int_{\Omega_s} p_t W_{n'}^{m'}(\theta_s, \varphi_s, ka) d\Omega_s = Y_{n'}^{m'}(\theta_k, \varphi_k) \quad (7)$$

This shows the gain of the plane wave coming from (θ_k, φ_k) , for a continuous pressure-sensitive spherical microphone is $|Y_n^{m'}(\theta_k, \varphi_k)|$. Since an arbitrary real function $F(\theta, \varphi)$ can be expanded in terms of complex spherical harmonics, we can implement arbitrary beampatterns. For example, an ideal beampattern directed at a direction (θ_0, φ_0) :

$$F(\theta, \varphi) = \begin{cases} 1, & (\theta, \varphi) = (\theta_0, \varphi_0) \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

can be expanded into:

$$F(\theta, \varphi) = 2\pi \sum_{n=0}^{\infty} \sum_{m=-n}^n Y_n^{m*}(\theta_0, \varphi_0) Y_n^m(\theta, \varphi) \quad (9)$$

The advantage of this system is that it can be steered into any 3D directions *digitally* with the same beampattern. This is for an ideal continuous microphone array on spherical surface.

3. ORTHONORMALITY ERROR ANALYSIS

For a real-world system, however, we have a discretely sampled array with S microphones mounted at (θ_s, φ_s) , $s = 1, 2, \dots, S$. To keep the same performance, Ref. [2] uses special discrete points at which the following discrete analog of orthonormality holds:

$$\frac{4\pi}{S} \sum_{s=1}^S Y_n^{m*}(\theta_s, \varphi_s) Y_{n'}^{m'}(\theta_s, \varphi_s) = \delta_{nn'} \delta_{mm'} \quad (10)$$

(Note that the formula (4) in [2], for discrete orthonormality has a typographical error.)

Unfortunately, it can be easily proved that for any possible layout with limited number of points, (10) cannot be satisfied exactly even for low order spherical harmonics (See Appendix A). Instead, we have:

$$\frac{4\pi}{S} \sum_{s=1}^S Y_n^{m*}(\theta_s, \varphi_s) Y_{n'}^{m'}(\theta_s, \varphi_s) = \delta_{nn'} \delta_{mm'} + \epsilon_{nn'}^{mm'}, \quad (11)$$

where $\epsilon_{nn'}^{mm'}$ is the error caused by discreteness. Ideally, we want the error to go to zero in the limit of an infinite number of microphones on each continuous points on the spherical surface:

$$\lim_{S \rightarrow \infty} \epsilon_{nn'}^{mm'} = 0, \quad \forall n, m, n', m'. \quad (12)$$

Now, we will see how this error could degrade the performance of soundfield decomposition. To extract the component of order n' and degree m' from the soundfield (3), we consider individual integrals of (7):

$$P_n^m = \int_{\Omega_s} p_{t_n}^m \frac{Y_n^{m'}(\theta_s, \varphi_s)}{4\pi i^{n'} b_{n'}(ka)} d\Omega_s \quad (13)$$

where:

$$p_{t_n}^m = 4\pi i^n b_n(ka) Y_n^m(\theta_k, \varphi_k) Y_n^{m*}(\theta_s, \varphi_s). \quad (14)$$

Using S discrete points, we have:

$$P_n^m = Y_n^m(\theta_k, \varphi_k) \left\{ \left[\frac{i^n b_n(ka)}{i^{n'} b_{n'}(ka)} \right] (\delta_{nn'} \delta_{mm'} + \epsilon_{nn'}^{mm'}) \right\}. \quad (15)$$

We notice that:

$$\left[\frac{i^n b_n(ka)}{i^{n'} b_{n'}(ka)} \right] \delta_{nn'} \delta_{mm'} = \delta_{nn'} \delta_{mm'} \quad (16)$$

So, (15) can be rewritten as:

$$P_n^m = Y_n^m(\theta_k, \varphi_k) \{ \delta_{nn'} \delta_{mm'} + \left[\frac{i^n b_n(ka)}{i^{n'} b_{n'}(ka)} \right] \epsilon_{nn'}^{mm'} \}. \quad (17)$$

The second term is the noise caused by orthonormality error. We call it the *orthonormality error noise* (OEN) which is inherent with the microphone array layout. To prevent it from damaging the orthonormality, we must have:

$$\left| \frac{i^n b_n(ka)}{i^{n'} b_{n'}(ka)} \epsilon_{nn'}^{mm'} \right| \ll 1 \quad (18)$$

So, we get:

$$|\epsilon_{nn'}^{mm'}| \ll \left| \frac{b_{n'}(ka)}{b_n(ka)} \right|, \quad \forall n, n', m, m' \quad (19)$$

Since b_n decays very quickly with respect to n as shown in Fig.1, for a given microphone number and layout, we cannot extract the high order component of soundfield if (19) fails. In addition, we can see (19) is independent of magnitude of the incoming sound wave. That means: even if the microphones have recorded the high order components, the system may be unable to extract them.

There exist several schemes which might be used to minimize this error. In [2], the 32 microphones are positioned at the center of the faces of a truncated icosahedron. In [3], several options are mentioned including equiangular grid layout [5] and an intuitive equidistance layout [4]. The common limitation of those schemes is that they are inflexible. If a patch of the spherical surface is inappropriate for mounting microphones, the orthonormality error may be large. This will destroy the beampattern.

4. OPTIMAL LAYOUT AND WEIGHTS

In this section, we will formally define the optimal layout. We proposed an iterative approach to find it. Then we solve for the optimal weights. The OEN can be cancelled completely to order N using $S = (N + 1)^2$ microphones positioned on the spherical surface. We denote the positions as (θ_s, φ_s) , $s = 1, 2, \dots, (N + 1)^2$.

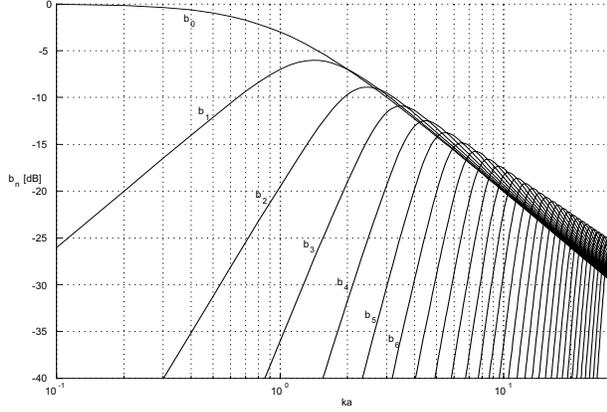


Fig. 1. $b_n(ka)$ for orders from 0 to 30.

For each $Y_{n'}^{m'}(\theta_s, \varphi_s)$, we introduce a correcting coefficient. In total we have $(N + 1) \times (N + 1)^2$ coefficients, $C_{n'}^{m'}(\theta_s, \varphi_s)$, for $(s = 1, 2, \dots, (N + 1)^2; n' = 0, 1, \dots, N; m' = -n', \dots, n')$. So that

$$\frac{4\pi}{S} \sum_{s=1}^S C_{n'}^{m'}(\theta_s, \varphi_s) Y_n^{m*}(\theta_s, \varphi_s) Y_{n'}^{m'}(\theta_s, \varphi_s) = \delta_{nn'} \delta_{mm'} \quad (20)$$

$$\begin{aligned} (S &= (N + 1)^2; s = 1, 2, \dots, S; \\ n &= 0, 1, \dots, N; m = -n, \dots, n; \\ n' &= 0, 1, \dots, N; m' = -n', \dots, n') \end{aligned}$$

We can simplify (20) further by defining:

$$\mathbf{D}_{n'}^{m'} = \begin{bmatrix} C_{n'}^{m'}(\theta_1, \varphi_1) Y_{n'}^{m'}(\theta_1, \varphi_1) \\ C_{n'}^{m'}(\theta_2, \varphi_2) Y_{n'}^{m'}(\theta_2, \varphi_2) \\ \dots \\ C_{n'}^{m'}(\theta_S, \varphi_S) Y_{n'}^{m'}(\theta_S, \varphi_S) \end{bmatrix} \quad (21)$$

and

$$\mathbf{Y} = \begin{bmatrix} Y_0^{0*}(\theta_1, \varphi_1) & Y_0^{0*}(\theta_2, \varphi_2) & \dots & Y_0^{0*}(\theta_S, \varphi_S) \\ Y_1^{-1*}(\theta_1, \varphi_1) & Y_1^{-1*}(\theta_2, \varphi_2) & \dots & Y_1^{-1*}(\theta_S, \varphi_S) \\ \dots & \dots & \dots & \dots \\ Y_N^{N*}(\theta_1, \varphi_1) & Y_N^{N*}(\theta_2, \varphi_2) & \dots & Y_N^{N*}(\theta_S, \varphi_S) \end{bmatrix}$$

So we have:

$$\frac{4\pi}{S} \mathbf{Y} \cdot \mathbf{D}_{n'}^{m'} = [\delta_{0n'} \delta_{0m'}, \delta_{1n'} \delta_{(-1)m'}, \dots, \delta_{Nn'} \delta_{Nm'}]' \quad (22)$$

To have a stable solution for $\mathbf{D}_{n'}^{m'}$, we have to find S points on the spherical surface such that the condition number of \mathbf{Y} is reasonably small. We use an iterative approach to search for these optimal points. First, we distribute the S points on the spherical surface randomly, then we can pick a random moving direction for each point. We also can use some

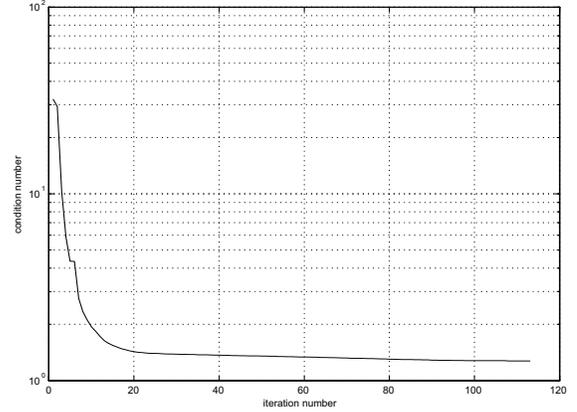


Fig. 2. Iteratively search for the optimal nodes which minimize the condition number of \mathbf{Y}

heuristic criteria to decide the directions. Then we move all points a step forward and check the condition number of \mathbf{Y} . If we get smaller condition number, then move forward again, if not, choose another directions. We do this iteratively until the condition number converges. A typical search process is shown in Fig. 2.

For each n' and m' , (22) is a linear system with $(N + 1)^2$ constraints and $(N + 1)^2$ unknowns. So we can solve $\mathbf{D}_{n'}^{m'}$ exactly and uniquely. Now the orthonormality holds to order N exactly for $(N + 1)^2$ arbitrarily positioned microphones.

Once we have solved (22), we can use the results in the real system implementation. We just need to replace the original weight (5) on each microphone to the modified weight

$$W_{n'}^{m'}(\theta_s, \varphi_s, ka) = \frac{D_{n'}^{m'}(\theta_s, \varphi_s)}{4\pi i^{n'} b_{n'}(ka)} \quad (23)$$

5. FLEXIBLE DESIGN EXAMPLE

Our algorithm can be used to design flexible microphone layout. Suppose we want to build a spherical array with radius 10cm using 64 microphones. The microphone positions are restricted in the area $(\pi/4 \leq \theta \leq \pi, -\pi \leq \varphi \leq \pi)$. We use our iterative algorithm to compute the 64 nodes and 64×64 weights.

Fig. 3 is the resulting layout for 64 microphones. Fig. 4 is the three dimensional beampattern of order four at 1.5kHz.

6. CONCLUSIONS

This paper describes a flexible and optimal design of spherical microphone arrays. We analyzed the effects of discrete

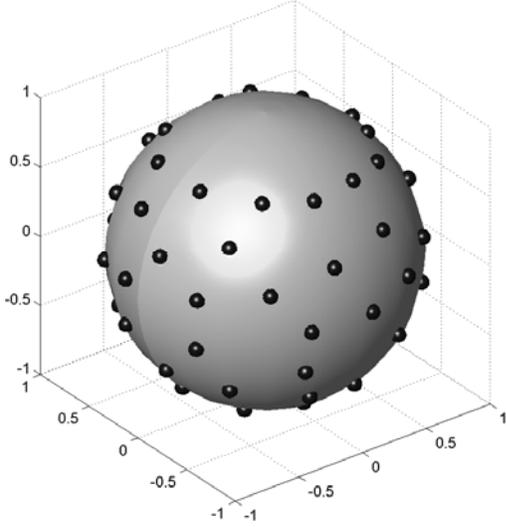


Fig. 3. Optimal layout of 64 microphones on the surface defined as $(\pi/4 \leq \theta \leq \pi, -\pi \leq \varphi \leq \pi)$

orthonormality error and proposed a linear constraint approach to get exact orthonormalities. We then used an iterative method to search for the optimal microphone layout with respect to the condition number. We demonstrated the effectiveness of our algorithms by performing a practical design.

7. REFERENCES

- [1] Earl G. Williams, *Fourier Acoustics*, Academic Press, San Diego, 1999
- [2] Jens Meyer and Gary Elko, “A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield”, *IEEE Proc. ICASSP’02*, vol.2, pp. 1781-1784, May, 2002
- [3] T. D. Abhayapala and D. B. Ward, “Theory and design of high order sound field microphones using spherical microphone array”, *IEEE Proc. ICASSP’02*, vol.2, pp. 1949-1952, May, 2002
- [4] Jörg Fliege and Ulrike Maier: “A Two-Stage Approach for Computing Cubature Formulae for the Sphere”. *Ergebnisberichte Angewandte Mathematik, No. 139T*. Fachbereich Mathematik, Universität Dortmund, 44221 Dortmund, Germany. September 1996.
- [5] D. Healy, D. Rockmore, and S. Moor, “An FFT for the 2-sphere and applications”, *IEEE Proc. ICASSP’96*, vol.3, pp. 1323-1326, May, 1996

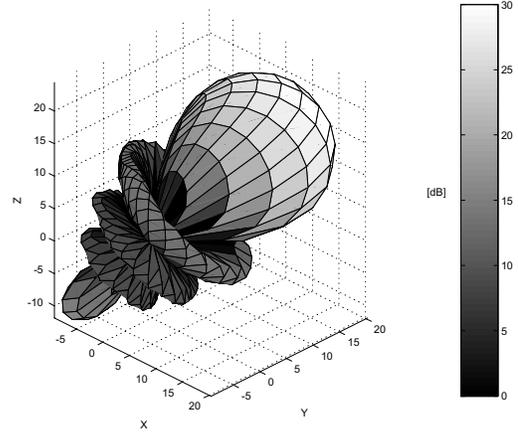


Fig. 4. Beampattern of order 4 at 1.5kHz, steering at $(\pi/4, \pi/4)$.

- [6] P.K. Morse, K.U. Ingaard. *Theoretical Acoustics*, McGraw Hill, New York, 1968.

A. PROOF OF THE ORTHONORMALITY ERROR

Assume we can use S microphones to satisfy the following orthonormalities:

$$\begin{cases} \frac{4\pi}{S} \sum_{s=1}^S Y_2^{0*}(\theta_s, \varphi_s) Y_0^0(\theta_s, \varphi_s) = 0 \\ \frac{4\pi}{S} \sum_{s=1}^S Y_3^{0*}(\theta_s, \varphi_s) Y_1^0(\theta_s, \varphi_s) = 0 \end{cases} \quad (24)$$

We have:

$$\begin{cases} \sum_{s=1}^S \cos^2 \theta_s = \frac{1}{3} \\ \sum_{s=1}^S \cos^4 \theta_s = \frac{1}{5} \end{cases} \quad (25)$$

Now we check the following orthonormality:

$$\begin{aligned} & \frac{4\pi}{S} \sum_{s=1}^S Y_2^{0*}(\theta_s, \varphi_s) Y_2^0(\theta_s, \varphi_s) = 1 \\ \Leftrightarrow & \frac{4\pi}{S} \left(\frac{1}{4} \sqrt{\frac{5}{\pi}} \right)^2 \left(9 \sum_{s=1}^S \cos^4 \theta_s - 6 \sum_{s=1}^S \cos^2 \theta_s + S \right) = 1 \\ & \text{(use (25)) } \Leftrightarrow S = 1 \end{aligned} \quad (26)$$

It is obvious that (26) and (25) are contradictory.