HEMISPHERICAL MICROPHONE ARRAYS FOR SOUND CAPTURE AND BEAMFORMING

Zhiyun Li and Ramani Duraiswami

Perceptual Interfaces and Reality Lab, Computer Science and UMIACS
University of Maryland, College Park
zli@cs.umd.edu; ramani@umiacs.umd.edu

ABSTRACT

We design and demonstrate a hemispherical microphone array for spatial sound acquisition and beamforming. Our design makes use of the acoustic image principle. It is especially appropriate for a half 3D acoustic environment where all sound sources are constrained on one side of a rigid plane. It avoids the difficulties of building a full spherical microphone array yet keeps the advantage of achieving a direction-invariant beampattern. A special microphone layout is designed for simple implementation. We also propose an approach to effectively calibrate data-independent coefficients of the system. Simulation and experimental results are presented.

1. INTRODUCTION

Spherical microphone arrays are attracting increasing interest since they can capture a 3D soundfield and provide direction-invariant beampatterns in all directions [1][2]. In practice, using only a finite number of microphones, various layouts have been designed to optimally cancel the error caused by discreteness. The microphone positions can be either carefully selected to achieve optimal performances [3] or quite flexible with minimal performance compromise [4].

However, to physically build a full spherical microphone array on a rigid sphere is a challenging or impossible task. More importantly, in numerous real-world scenarios where sound sources are located in a constrained acoustic environment instead of a full 3D space, a full spherical array is either uneconomic or redundant. For example, in a conference room environment, all sound sources are usually above the table surface which forms a half 3D space. In this case, a hemispherical array may be a better choice because:

1. The table surface is usually rigid and inevitably creates acoustic images of the real sound sources, which validates the design of a hemispherical array.

2. Given a specified number of microphones and a sphere of given radius, a hemispherical array will have a denser microphone arrangement, thereby allowing for analysis of a wider frequency range. Even in an acoustic environment without image sources, using a hemispherical array mounted on a rigid plane to create images may be appropriate since it provides higher order beampatterns.
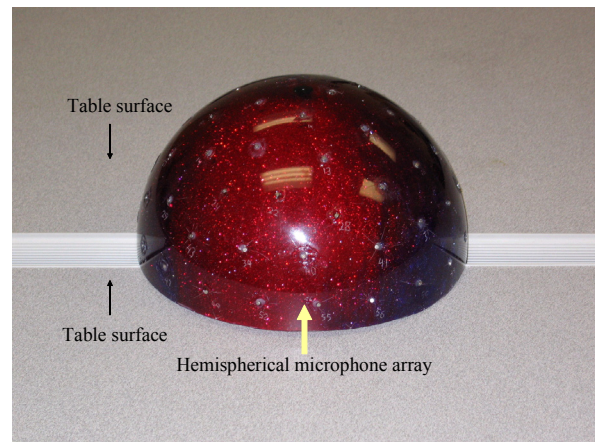
Figure 1: A hemispherical microphone array built on the surface of a half bowling ball. Its radius is 10.925cm.

3. A hemispherical array is easier to build and maintain, it can be mounted on a rigid surface such as table surface or wall, and wires can be conveniently placed to the microphones.

Fig. 1 shows our prototype of a hemispherical array with 64 microphones installed on the surface of a half bowling ball. This paper is organized into three parts. We first briefly review the theories of spherical beamformer and propose a methodology to design a hemispherical microphone array using the principle of images. Next, an effective calibration algorithm is proposed. Finally, our algorithms are demonstrated by simulation and experimental results using our prototype.

2. SPHERICAL BEAMFORMER

The basic principle of the spherical beamformer is to make use of the orthonormality of spherical harmonics to decompose the soundfield arriving at a spherical array. Then the orthogonal components of the soundfield are linearly combined to approximate a desired beampattern [1]. For a unit magnitude plane wave $\mathbf{k}$, incident from direction (θ_k, φ_k), the complex pressure field on the
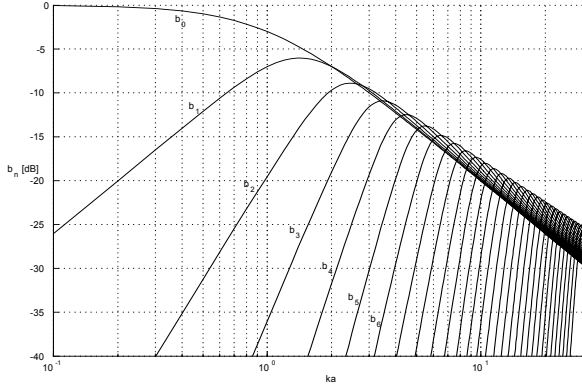
Figure 2: $b_n(ka)$ for orders from 0 to 30.

surface $(\theta_s, \varphi_s, r_s = a)$ of the rigid sphere is [5]:

$$p_t = 4\pi \sum_{n=0}^{\infty} i^n b_n(ka) \sum_{m=-n}^{n} Y_n^m(\theta_k, \varphi_k) Y_n^{m*}(\theta_s, \varphi_s), \tag{1}$$

$$b_n(ka) = j_n(ka) - \frac{j_n'(ka)}{h_n'(ka)} h_n(ka), \tag{2}$$

where j_n is the spherical Bessel function of order n, Y_n^m the spherical harmonics of order n and degree m, * denotes the complex conjugation, and h_n is the spherical Hankel function of the first kind. Some $b_n(ka)$ are plotted in Fig. 2.

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

$$W_{n'}^{m'}(\theta_s, \varphi_s, ka) = \frac{Y_{n'}^{m'}(\theta_s, \varphi_s)}{4\pi i^{n'} b_{n'}(ka)}, \tag{3}$$

then making use of the *orthonormality* of spherical harmonics:

$$\int_{\Omega_s} Y_n^{m*}(\theta_s, \varphi_s) Y_{n'}^{m'}(\theta_s, \varphi_s) d\Omega_s = \delta_{nn'}\delta_{mm'}, \tag{4}$$

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

$$P = \int_{\Omega_s} p_t W_{n'}^{m'}(\theta_s, \varphi_s, ka) d\Omega_s = Y_{n'}^{m'}(\theta_k, \varphi_k). \tag{5}$$

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 looking at the direction (θ_0, φ_0) can be modeled as a delta function:

$$F(\theta, \varphi) = \delta(\theta - \theta_0, \varphi - \varphi_0), \tag{6}$$

which can be expanded into an infinite series of spherical harmonics [6]:

$$F(\theta, \varphi) = 2\pi \sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_n^{m*}(\theta_0, \varphi_0) Y_n^m(\theta, \varphi). \tag{7}$$
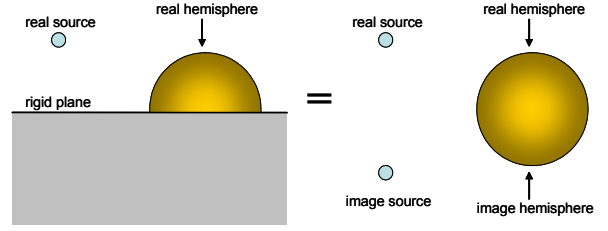

Figure 3: The hemispherical array with a rigid plane is equivalent to a spherical array in free space with real and image sources.

The weight at each point (θ_s, φ_s) to achieve this beampattern is:

$$w_s = \sum_{n=0}^{\infty} \frac{1}{2i^n b_n(ka)} \sum_{m=-n}^{n} Y_n^{m*}(\theta_0, \varphi_0) Y_n^m(\theta_s, \varphi_s). \tag{8}$$

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 pressure sensitive surface. For discrete arrays with finite number of microphones, the practical beampattern is a truncated version of (7) to some limited order N:

$$F_N(\theta, \varphi) = 2\pi \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^{m*}(\theta_0, \varphi_0) Y_n^m(\theta, \varphi). \tag{9}$$

3. DESIGN OF A HEMISPHERICAL MICROPHONE ARRAY

In this section, we will make use of the acoustic image principle to design a hemispherical microphone array.

3.1. Acoustic Image Principle

Suppose a sound source is placed on one side of a perfectly rigid plane, then in any point on the same side of the plane, the sound pressure is the combined result of unbounded irradiations of this sound source and the image sound source which lies on the other side of the plane symmetrically with respect to the plane. This is the *acoustic image principle* [7]. If we attach a rigid plane to the bottom of the hemispherical array (see Fig. 3), the pressure at each real microphone position (θ_s, φ_s) can be easily solved by removing the rigid plane and adding the image source at $(\tilde{\theta}_k = \pi - \theta_k, \varphi_k)$ and the image hemispherical array. In addition, the image microphone at $(\tilde{\theta}_s = \pi - \theta_s, \varphi_s)$ receives the same pressure as its counterpart. In other words, the rigid plane acts as an acoustic mirror. The solution on (θ_s, φ_s) is:

$$p_h = 4\pi \sum_{n=0}^{\infty} i^n b_n(ka) \sum_{m=-n}^{n} A_n^m(\theta_k, \varphi_k) Y_n^{m*}(\theta_s, \varphi_s), \tag{10}$$

$$A_n^m(\theta_k, \varphi_k) = Y_n^m(\theta_k, \varphi_k) + Y_n^m(\tilde{\theta}_k, \varphi_k). \tag{11}$$

3.2. A Symmetric and Uniform Layout

For a discrete hemispherical microphone array, the rigid plane creates a symmetric spherical layout of microphones with respect to
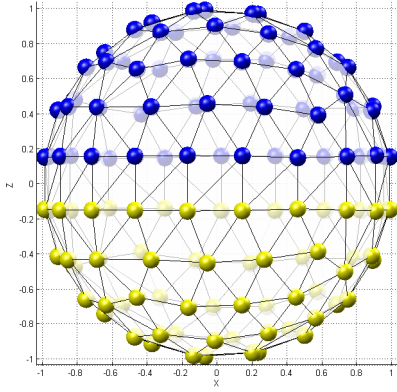
Figure 4: The symmetric and uniform layout of 128 nodes on a spherical surface. The blue (dark) nodes are for real microphones. The yellow (light) nodes are images.

Figure 5: Discrete orthonormality errors. Plot shows absolute values.

the plane. Intuitively, we desire this symmetric layout also be "uniform" over the whole spherical surface, especially in the neighborhood of the rigid plane. Although the spherical layout can be made more flexible at the cost of extra computation [4], we want a layout with minimal implementation overhead.

To find a "uniform" layout, we use the simulation described in [8]. It minimizes the potential energy of a distribution of movable electrons on a perfectly conducting sphere. It is obvious that if the simulation starts with a symmetric initial layout, the resulting layout in each iteration step is guaranteed symmetric, so is the final layout. We use this approach to obtain a 128-node layout that is both symmetric and uniform. Specifically, to make it repeatable, we start with Fliege's 64-node layout [8]. We flip it upside down and add it on the original nodes to create 128 nodes, then perform the simulation until all nodes are optimally separated. The resulting layout is shown in Fig. 4. No nodes are too close to the $z = 0$ plane, or the rigid plane, which also helps avoid technical construction difficulties.

3.3. Discrete Hemispherical Beamforming

For a hemispherical array with S microphones uniformly mounted at (θ_s, φ_s), $s = 1, 2, ..., S$, the image microphones are at $(\tilde{\theta}_s = \pi - \theta_s, \varphi_s)$. To adapt the spherical beamforming principle to the discrete case, the continuous integrals (4) are approximated by:

$$\frac{2\pi}{S} \sum_{s=1}^{S} Y_n^{m*}(\theta_s, \varphi_s) Y_{n'}^{m'}(\theta_s, \varphi_s) + Y_n^{m*}(\tilde{\theta}_s, \varphi_s) Y_{n'}^{m'}(\tilde{\theta}_s, \varphi_s)$$
$$\approx \delta_{nn'} \delta_{mm'}, \quad (12)$$

where $(n = 0, ..., N_{max}; m = -n, ..., n; n' = 0, ..., N; m' = -n', ..., n')$, N_{max} is the band limit of spatial frequency in terms of spherical harmonics orders, N is the order of beamforming. In general, more precise approximation can be achieved by using appropriate cubature weights [8]. For simplicity, we just use equal cubature weights without significantly affecting the results. To verify this, Fig. 5 shows the absolute errors of (12) using the layout in Fig. 4.

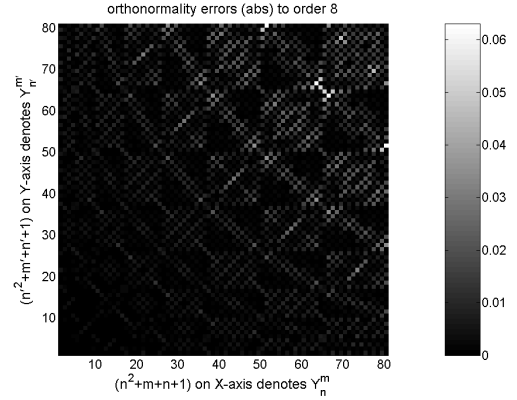Therefore, to approximate the regular beampattern of order N

as in (9), the weight for the s-th microphone is:

$$w_s = \sum_{n=0}^{N} \frac{\sum_{m=-n}^{n} Y_n^{m*}(\theta_0, \varphi_0) \left[Y_n^m(\theta_s, \varphi_s) + Y_n^m(\tilde{\theta}_s, \varphi_s) \right]}{2i^n b_n(ka)}.$$
$$(13)$$

4. EFFECTIVE CALIBRATION

In section 2, we derived the ideal beamformer from the ideal solution in (1). In practice, many factors will affect the complex pressure captured by the real-world hemispherical microphone array, such as table surface geometry and impedance, array placement, microphone positions and characteristics, etc. To achieve better results, the array has to be calibrated to match the theoretical solution.

According to our system settings and the beamforming algorithm, a complete calibration is unnecessary for our prototype. Instead, by examining (13), we propose a simple yet effective calibration. We notice that $b_n(ka)$ describes the theoretical strength of the order n expansion in (1). The actual $b_n(ka)$ captured by the real-world microphone array has to match the theoretical value in (13) so that they can cancel each other to synthesize the desired 3D beampatterns, such as (9). This is especially important for high order beamforming since high order components of captured soundfield are increasingly weak because of the convergence of (1). This can be clearly observed in Fig. 2. Another advantage to calibrate $b_n(ka)$ lies in its independence of the incident waves in calibration, if they are known in advance. In this case, the calibration can be made in one measurement.

We describe the actual soundfield captured by our hemispherical array as:

$$\bar{p}_h = 4\pi \sum_{n=0}^{N_{max}} i^n \bar{b}_n(ka) \sum_{m=-n}^{n} C_n^m(ka) Y_n^{m*}(\theta_s, \varphi_s) + \epsilon, \quad (14)$$

where $\bar{b}_n(ka)$ denotes the captured $b_n(ka)$. $C_n^m(ka)$ is the soundfield coefficient, which is known during calibration. For a single plane wave of unit amplitude incident from (θ_k, φ_k), in our hemispherical array setting, $C_n^m(ka) = A_n^m(\theta_k, \varphi_k)$ as in (11). ϵ is the residual error not included in $\bar{b}_n(ka)$. To calibrate $\bar{b}_n(ka)$ at order

n', we assign one component from the hemispherical beamforming weight in (13) for the s-th microphone:

$$w_s^{n'} = \sum_{m'=-n'}^{n'} Y_{n'}^{m'*}(\theta_0, \varphi_0) \left[Y_{n'}^{m'}(\theta_s, \varphi_s) + Y_{n'}^{m'}(\tilde{\theta}_s, \varphi_s) \right], \tag{15}$$

and the total output is:

$$\bar{P}^{n'} = \frac{2\pi}{S} \sum_{s=1}^{S} w_s^{n'} \bar{p}_h. \tag{16}$$

$$\approx 4\pi i^{n'} \bar{b}_{n'}(ka) \sum_{m'=-n'}^{n'} C_{n'}^{m'}(ka) Y_{n'}^{m'*}(\theta_0, \varphi_0),$$

We then have:

$$\bar{b}_{n'}(ka) = \frac{\bar{P}^{n'}}{4\pi i^{n'} \sum_{m'=-n'}^{n'} C_{n'}^{m'}(ka) Y_{n'}^{m'*}(\theta_0, \varphi_0)} + \varepsilon, \tag{17}$$

where ε contains ϵ and the orthonormality errors from (12). If ε is zero-mean Gaussian with respect to the beamforming direction (θ_0, φ_0), then we can estimate $\bar{b}_{n'}(ka)$ by averaging over every (θ_0, φ_0) with only one measurement if the calibration environment defined by $C_{n'}^{m'}(ka)$ can be precisely modeled.

5. SIMULATION AND EXPERIMENTAL RESULTS

We first show the simulation result. Suppose in free space, there are two plane waves of $4\,\text{kHz}$ incident from $(\pi/4, -\pi/2)$ and $(3\pi/4, -\pi/2)$, respectively. Using the 128-node spherical microphone array, we scan every 3D direction using the beamformer of order eight. For each direction, we plot the amplitude of the output from this beamformer as a 3D point with the amplitude value as the distance to the origin. The 3D scanning result is shown in Fig. 6(a). Please do not confuse this with the beampattern of the 128-node array. In fact, this scanning result is the sum of the *two* 3D beampatterns of order eight, steered respectively to $(\pi/4, -\pi/2)$ and its image $(3\pi/4, -\pi/2)$. If the beamformer is steered to $(\pi/4, -\pi/2)$, the sound from $(3\pi/4, -\pi/2)$ will be significantly suppressed. The plot is shown in linear scale for clearer separation of two sources.

In calibration, we play the same sound from a real sound source from $(\pi/4, -\pi/2)$. The hemispherical array is set up on the table surface as shown in Fig. 1. We then use the approach in section 4 to calibrate. The 3D scanning result using the calibrated beamformer is shown in Fig. 6(b). Ideally, it should be the same as Fig. 6(a), but it also detected some reflections from other surfaces of the room, especially from the back direction $(\pi/2, \pi/2)$.

To demonstrate this calibration is independent of the sound source locations, we move the real sound source to about $(\pi/3, -\pi/2)$. Fig. 6(c) shows the simulation result. In our experiment, we use the same calibrated beamformer to scan. The scanning result is shown in Fig. 6(d). As expected, the back reflection is stronger in this case.

6. CONCLUSIONS

We designed a hemispherical microphone array using the acoustic image principle. It can be seen as a symmetric spherical microphone array across the edge of real and image space, which is easy to build and install. It is especially appropriate in numerous scenarios where sound sources are constrained in a half 3D space. We
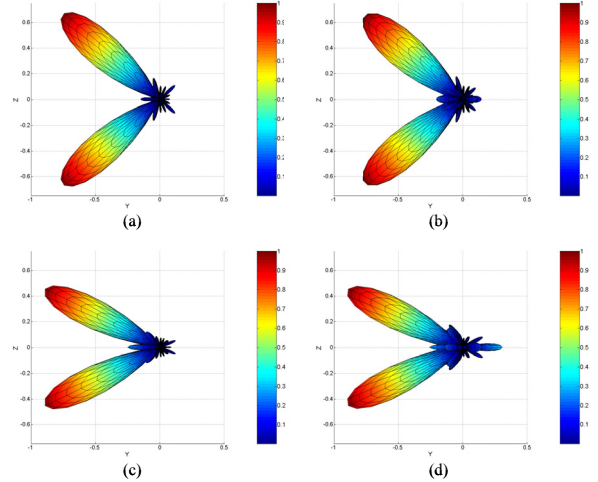

Figure 6: Simulation and Experimental results: (a) simulation of 3D scanning result with two sound sources, the beamformer is of order 8; (b) experimental result using the calibrated beamfomer of order 8; (c) simulation result after sound sources are moved; (d) experimental result using the same calibrated beamformer.

built an array using 64 microphones positioned on a hemispherical surface. The layout of real and image microphones is both symmetric and uniform, which leads to a simple implementation of hemispherical beamformer. An effective calibration method is proposed to extract the data-independent coefficients of the system from the captured soundfield. The simulation and experimental results demonstrate the effectiveness of our design. More comprehensive evaluation will be performed in the future work.

7. REFERENCES

[1] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," *ICASSP'02*, vol. 2, pp. 1781–1784, May 2002.

[2] T. D. Abhayapala and D. B. Ward, "Theory and design of high order sound field microphones using spherical microphone array," *ICASSP'02*, vol. 2, pp. 1949–1952, May 2002.

[3] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. on Speech and Audio Processing*, vol. 13, pp. 135–143, 2005.

[4] Z. Li and R. Duraiswami, "A robust and self-reconfigurable design of spherical microphone array for multi-resolution beamforming," *ICASSP'05*, vol. IV, pp. 1137–1140, Mar. 2005.

[5] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Elsevier Science, 2005. ISBN:0080443710.

[6] M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions*. U.S. Government Printing Office, 1964.

[7] P. Morse and K. Ingaard, *Theoretical Acoustics*. New York: McGraw Hill, 1968.

[8] J. Fliege and U. Maier, "The distribution of points on the sphere and corresponding cubature formulae," *IMA Journal on Numerical Analysis*, vol. 19, pp. 317–334, 1999.