A ROBUST AND SELF-RECONFIGURABLE DESIGN OF SPHERICAL MICROPHONE ARRAY FOR MULTI-RESOLUTION BEAMFORMING

Zhiyun Li, Ramani Duraiswami

ABSTRACT

We describe a robust and self-reconfigurable design of a spherical microphone array for beamforming. Our approach achieves a multi-resolution spherical beamformer with performance that is either optimal in the approximation of desired beampattern or is optimal in the directivity achieved, both robustly. Our implementation converges to the optimal performances quickly while exactly satisfying the specified frequency response and robustness constraint in each iteration step without accumulated round-off errors. The advantage of this design lies in its robustness and self-reconfiguration in microphone array reorganization, such as microphone failure, which is highly desirable in online maintenance and anti-terrorism. Design examples and simulation results are presented.

1. INTRODUCTION

Spherical microphone arrays are recently becoming the subject of some study as they allow omnidirectional sampling of the 3D soundfield, and may find applications in multi-resolution soundfield capture and recreation [4]. In [5], a modal beamformer design in orthogonal beam-space was presented. In [3], we proposed a preliminary extension to allow relatively flexible microphone placements with minimal performance compromise. Our main contributions in this paper are: 1) we balance the trade-off between accuracy and robustness to allow even more flexible layouts with optimal performances; 2) we design a self-reconfigurable implementation to make the beamformer robust to microphone reorganization; 3) it seamlessly achieves multi-resolution beampatterns, either regular beampatterns or optimal directivity, both robustly.

The rest of this paper is organized into four sections. In section 2, we present the basic principle of spherical beamformer. In section 3, we formulate the beamformer for discrete array into a finite linear system for specified beamforming direction. The solution optimally approximates the desired beampattern in *least mean square* (LMS) sense. In section 4, we optimize the accuracy of approximation under robustness constraint. To allow efficient implementation, we rewrite this constrained optimization problem into an ellipsoidal form under a linear and a spherical constraints. This naturally leads to a self-reconfigurable design in the form similar to [1] and [2], but with different inputs and optimization goals. Obviously, our implementation inherits their advantages, such as absence of round-off error accumulation, exact satisfaction of constraints, etc. Design examples and simulations will be presented in section 5.

This work was partially supported by NSF Award 0205271. PIRL, UMIACS, Univ. of MD at College Park. Email: zli@cs.umd.edu; ramani@umiacs.umd.edu

2. BACKGROUND

The basic idea 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 [5].

For a unit magnitude plane wave with wavenumber k, incident from direction (θ_k, φ_k), the complex pressure field on the surface $(\theta_s, \varphi_s, r_s = a)$ of the rigid sphere is [6]:

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

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

where a is the radius of the sphere, j_n is the spherical Bessel function of order n, Y_n^m is the spherical harmonics of order n and degree m. $*$ denotes the complex conjugation. h_n is the spherical Hankel function of the first kind.

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

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

Then making use of orthonormality of spherical harmonics:

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

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

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

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 function $F(\theta, \varphi)$ can be expanded in terms of 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:

$$F(\theta, \varphi) = \delta(\theta - \theta_0, \varphi - \varphi_0), \quad (6)$$

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

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

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

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

For an ideal continuous microphone array, the spherical beamformer can be steered into any 3D directions *digitally* with the same beampattern.

For discrete arrays with finite number of microphones, the practical beampattern is a truncated version of (7) to some limited order N [5]:

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

3. DISCRETE SPHERICAL BEAMFORMER AS FINITE LINEAR SYSTEM

To achieve a regular beampattern of order N (9), a discrete spherical beamformer with S microphones can be formulated as a finite linear system:

$$\begin{aligned} \mathbf{AW} &= c_N \mathbf{B}_N, \quad (10) \\ \mathbf{dW} &= 1, \quad (11) \end{aligned}$$

where (10) defines the beampattern, and (11) the frequency response to the sound from the beamforming direction. Without loss of generality, here we consider an all-pass filter. In (10), $\mathbf{A}$ are the coefficients of the spherical harmonics expansion of the soundfield in (1):

$$\mathbf{A} = [\ \mathbf{A}_1 \quad \mathbf{A}_2 \quad \cdots \quad \mathbf{A}_S\ ], \quad (12)$$

$$\mathbf{A}_s = \begin{bmatrix} i^0 b_0(ka) Y_0^{0*}(\theta_s, \varphi_s) \\ i^1 b_1(ka) Y_1^{-1*}(\theta_s, \varphi_s) \\ \cdots \\ i^N b_N(ka) Y_N^{N*}(\theta_s, \varphi_s) \\ i^{(N+1)} b_{(N+1)}(ka) Y_{(N+1)}^{-(N+1)*}(\theta_s, \varphi_s) \\ \cdots \\ i^{N_{eff}} b_{N_{eff}}(ka) Y_{N_{eff}}^{N_{eff}*}(\theta_s, \varphi_s) \end{bmatrix}, \quad (13)$$
$$(s = 1, ..., S.)$$

N_{eff} is the maximum order with significant amplitude in the expansion (1). $\mathbf{W}$ is the vector of complex weights to be assigned to each microphone at (θ_s, φ_s, a):

$$\mathbf{W} = \begin{bmatrix} W(ka, \theta_0, \varphi_0, \theta_1, \varphi_1) \\ W(ka, \theta_0, \varphi_0, \theta_2, \varphi_2) \\ \cdots \\ W(ka, \theta_0, \varphi_0, \theta_S, \varphi_S) \end{bmatrix}. \quad (14)$$

$\mathbf{B}_N$ is the vector of coefficients of the beampattern of order N steered to (θ_0, φ_0) in (9):

$$\mathbf{B}_N = \begin{bmatrix} Y_0^{0*}(\theta_0, \varphi_0) \\ Y_1^{-1*}(\theta_0, \varphi_0) \\ \cdots \\ Y_N^{N*}(\theta_0, \varphi_0) \\ 0 \\ \cdots \\ 0 \end{bmatrix}. \quad (15)$$

Common constants are omitted in (13) and (15). In (11), $\mathbf{d}$ is the row vector of the complex pressure at each microphone position produced by a plane wave of unit magnitude from the desired beamforming direction (θ_0, φ_0):

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \cdots \\ d_S \end{bmatrix}^T = \begin{bmatrix} p_t(\theta_0, \varphi_0, \theta_1, \varphi_1) \\ p_t(\theta_0, \varphi_0, \theta_2, \varphi_2) \\ \cdots \\ p_t(\theta_0, \varphi_0, \theta_S, \varphi_S) \end{bmatrix}^T. \quad (16)$$

In (10), c_N is a normalizing coefficient to satisfy the all-pass frequency response (11). The LMS solution of (10) is:

$$\mathbf{W} = \left[(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \right] c_N \mathbf{B}_N. \quad (17)$$

Then c_N can be determined using (11). If we assume (10) has small residues, from (9), the *a priori* estimate of c_N is:

$$c_N \approx \frac{1}{2\pi \|\mathbf{B}_N\|_2^2}. \quad (18)$$

According to the spherical harmonic addition theorem, c_N is independent of (θ_0, φ_0) and can be simplified easily:

$$c_N \approx \frac{1}{\sum_{n=0}^{N} \frac{2n+1}{2} P_n(\cos 0)} = \frac{2}{(N+1)^2}. \quad (19)$$

4. ROBUST AND SELF-RECONFIGURABLE IMPLEMENTATION

To design a spherical beamformer with finite microphones under *white noise gain* (WNG) constraint of δ^2:

$$|\mathbf{dW}|^2/(\mathbf{W}^H\mathbf{W}) \geq \delta^2, \tag{20}$$

yet optimally approximate the desired beampattern of order N as (9), we need to minimize:

$$\min_{\mathbf{W}} \|\mathbf{AW} - c_N\mathbf{B}_N\|_2^2, \tag{21}$$

subject to:

$$\begin{aligned} \mathbf{dW} &= 1, &(22)\\ \mathbf{W}^H\mathbf{W} &\leq \delta^{-2}. &(23) \end{aligned}$$

This optimization can be numerically solved by some black-box software packages, such as MATLAB function `fmincon`, etc. Another way is to use Tikhonov regularization. The solution then becomes:

$$\mathbf{W} = \left[(\mathbf{A}^H\mathbf{A}+\lambda^2\mathbf{I})^{-1}\mathbf{A}^H\right] c_N\mathbf{B}_N, \tag{24}$$

where λ is the regularization parameter, which unfortunately is not directly related to the WNG constraint. A trial-and-error strategy can be used in implementation.

The most straightforward way to implement this system is to precompute all the weights for each pre-defined 3D direction and store them in a lookup table. This method, however, is inefficient because of the obvious trade-off between spatial resolution and memory. In addition, the resulted beamformer is not robust to microphone failure. It can be shown that the failure of even one microphone may significantly damage the beampattern.

In this section, we reformulate our problem so that we can parallel the methods in [1] and [2] to design a self-reconfigurable implementation which automatically and robustly converges to the desired beampattern of specified order in any steering directions. We rewrite the object function into an ellipsoidal form:

$$\min_{\mathbf{W}} \|\mathbf{A}\times\mathbf{W} - c_N\mathbf{B}_N\|_2^2 = \min_{\tilde{\mathbf{W}}} \tilde{\mathbf{W}}^H\mathbf{R}\tilde{\mathbf{W}}, \tag{25}$$

subject to:

$$\begin{aligned} \mathbf{C}^H\tilde{\mathbf{W}} &= \mathbf{g}, &(26)\\ \tilde{\mathbf{W}}^H\tilde{\mathbf{W}} &\leq \delta^{-2}+1, &(27) \end{aligned}$$

where

$$\tilde{\mathbf{W}} = \begin{bmatrix} \mathbf{W} \\ W_0 \end{bmatrix}, \ \mathbf{R} = \begin{bmatrix} \mathbf{A}^H \\ c_N\mathbf{B}_N^H \end{bmatrix} \begin{bmatrix} \mathbf{A} \\ c_N\mathbf{B}_N \end{bmatrix}^T,$$

$$\mathbf{C} = \begin{bmatrix} d_1^* & 0 \\ d_2^* & 0 \\ \vdots & \vdots \\ d_S^* & 0 \\ 0 & 1 \end{bmatrix}, \ \mathbf{g} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

We know $W_0 = -1$ from (25), however, we include it as an extra variable into $\tilde{\mathbf{W}}$ and its actual value is automatically determined by the constraint (26) in the process of optimization.

To solve this optimization, we first decompose $\tilde{\mathbf{W}}$ into its orthogonal components:

$$\tilde{\mathbf{W}} = \mathbf{W}_c + \mathbf{V}, \tag{28}$$

$$\mathbf{W}_c = \mathbf{C}[\mathbf{C}^H\mathbf{C}]^{-1}\mathbf{g}. \tag{29}$$

$\mathbf{W}_c$ is the LMS solution to satisfy the linear constraint (26). The residue is expected to be zero since usually (26) is a highly under-determined system. Substituting (29) into (27), we have:

$$\mathbf{V}^H\mathbf{V} \leq \delta^{-2} + 1 - \mathbf{g}^H[\mathbf{C}^H\mathbf{C}]^{-1}\mathbf{g} = b^2. \tag{30}$$

Thus, the WNG constraint becomes a spherical constraint on $\mathbf{V}$. Since $\mathbf{R}\tilde{\mathbf{W}}(t)$ is the gradient of the object function (25) at step t, the tentative update vector is:

$$\tilde{\mathbf{V}}(t+1) = \tilde{\mathbf{P}}_c[\mathbf{V}(t) - \mu\mathbf{R}\tilde{\mathbf{W}}(t)], \tag{31}$$

$\mathbf{V}(t)$ is the scaled projection of $\tilde{\mathbf{V}}(t)$ into the sphere surface of radius b:

$$\mathbf{V}(t) = \begin{cases} \tilde{\mathbf{V}}(t) & \text{for } \left|\tilde{\mathbf{V}}\right|^2 \leq b^2 \\ b\frac{\tilde{\mathbf{V}}(t)}{\left|\tilde{\mathbf{V}}(t)\right|} & \text{for } \left|\tilde{\mathbf{V}}\right|^2 > b^2 \end{cases}, \tag{32}$$

μ is the step size, and $\tilde{\mathbf{P}}_c$ is the null space of $\mathbf{C}^H$:

$$\tilde{\mathbf{P}}_c = \mathbf{I} - \mathbf{C}[\mathbf{C}^H\mathbf{C}]^{-1}\mathbf{C}^H. \tag{33}$$

Update the weights:

$$\tilde{\mathbf{W}}(t+1) = \mathbf{W}_c + \begin{cases} \tilde{\mathbf{V}}(t+1) & \text{for } \left|\tilde{\mathbf{V}}\right|^2 \leq b^2 \\ b\frac{\tilde{\mathbf{V}}(t+1)}{\left|\tilde{\mathbf{V}}(t+1)\right|} & \text{for } \left|\tilde{\mathbf{V}}\right|^2 > b^2 \end{cases}. \tag{34}$$

We set the initial guess as:

$$\begin{aligned} \tilde{\mathbf{W}}(0) &= \mathbf{R}^{-1}\mathbf{C}[\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}]^{-1}\mathbf{g}, &(35)\\ \tilde{\mathbf{V}}(0) &= \tilde{\mathbf{W}}(0) - \mathbf{W}_c, &(36) \end{aligned}$$

which is equivalent to the solution we get in section 3. If the resulting WNG is within constraint, the iteration will stay with this solution, otherwise, it will start the constrained optimization process, both automatically. At each step, the constraints (26) and (27) are satisfied exactly. In addition, similar to the methods in [1] and [2], round-off errors don't accumulate. This iteration is independent of the actual signal processing rate, so it may be implemented more efficiently as a parallel unit with other processors.

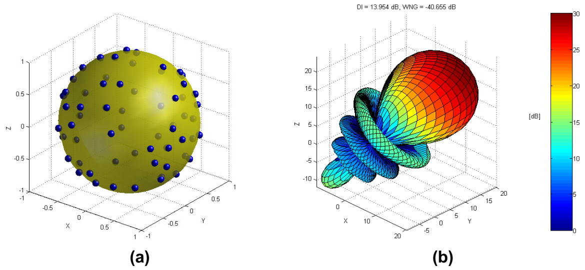

Fig. 1. (a) The random layout of 64 microphones on a sphere of radius 10cm. (b) The unconstrained beampattern of order four at 1KHz (WNG≈-41dB).

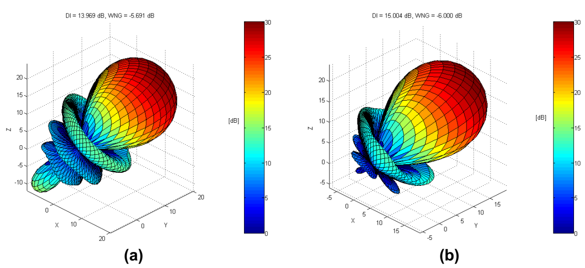

Fig. 2. Constrained optimizations. (a) Optimal beampattern of order four under the WNG constraint for the same setup in Fig. 1. (b) Optimal directivity under the WNG constraint.

5. SIMULATIONS

As a typical example, we use a random layout of 64 microphone on a spherical surface as shown in Fig. 1(a). Using the solution (17), the beamformer of order four is shown in Fig. 1(b), which is unrobust if we require a minimum WNG of -6dB. Fig. 2(a) shows the optimal approximations of the regular beampattern of order four under the WNG constraint. There is minimal difference between the beampatterns in Fig. 2(a) and Fig. 1(b). The comparisons of residues are shown in Fig. 3.

If we desire optimal directivity, we can approximate the ideal beampattern as (7). In practice, we just need to approximate an order high enough[1], such as order 7 in this case. Fig. 2(b) shows the resulted beampattern. We use *directivity index* (DI) to evaluate the directivity. Fig. 4 shows the iteration process using unit step size. The system converges quickly. It also demonstrates our implementation can robustly reconfigure itself in microphone reorganization.

6. CONCLUSIONS

This paper describes a robust and self-reconfigurable design of spherical microphone arrays for beamforming. Our design achieves optimal performances with multi-resolution beampatterns. Design examples and simulation results are presented.

[1] It can be pre-determined strictly for given spherical microphone array. This will appear in the extended version of this paper.

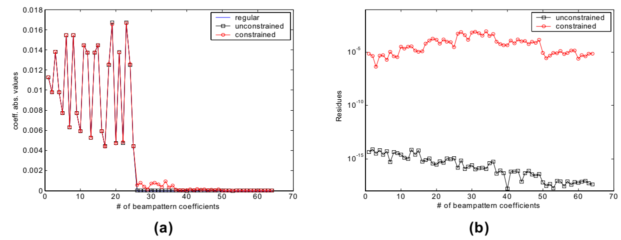

Fig. 3. (a) Comparison of beampattern coefficients with $c_4\mathbf{B}_4$. (b) Residue comparison between Fig. 2(a) and Fig. 1(b). Both plots show the absolute values.

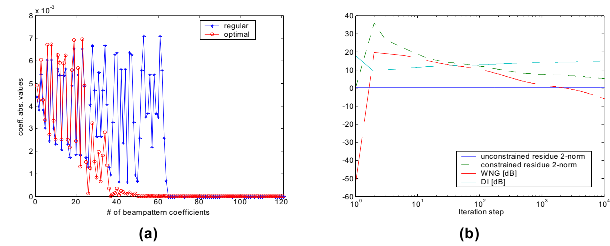

Fig. 4. Optimal directivity under WNG constraint. (a) The comparison of resulted beampattern and regular beampattern of order 7. (b)The iteration process.

7. REFERENCES

[1] H. Cox, R. M. Zeskind, and M. M. Owen. Robust adaptive beamforming. *IEEE Trans. on ASSP*, 35(10):1365–1376, Oct. 1987.

[2] O. L. Frost. An algorithm for linearly constrained adaptive array processing. *Proc. of the IEEE*, 60(8):926–935, Aug. 1972.

[3] Z. Li, R. Duraiswami, E. Grassi, and L. S. Davis. Flexible layout and optimal cancellation of the orthonormality error for spherical microphone arrays. In *ICASSP'04*, pages IV41–44, May 17-21 2004.

[4] Z. Li, R. Duraiswami, and N. A. Gumerov. Capture and recreation of higher order 3d sound fields via reciprocity. In *ICAD'04*, Sydney, Australia, July 6-9 2004.

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

[6] E. G. Williams. *Fourier Acoustics*. Academic Press, San Diego, 1999.