SPHERICAL LOUDSPEAKER ARRAY BEAMFORMING IN ENCLOSED SOUND FIELDS BY MIMO OPTIMIZATION

Hai Morgenstern and Boaz Rafaely

Department of Electrical and Computer Engineering
Ben-Gurion University of the Negev
Beer- Sheva 84105, Israel
haimorg@post.bgu.ac.il br@ee.bgu.ac.il

ABSTRACT

Spherical loudspeaker arrays have been recently studied for a variety of applications such as spatial sound reproduction and room acoustics. Array directivity is typically designed for free field, although loudspeaker arrays often operate in enclosures. Thererfore, current methods for directivity design may not be suitible. We present a new method for designing loudspeaker array directivity in enclosures, where the direct sound measured by a microphone array is emphasized compared to room reflections. The two spherical arrays form an acoustic multiple-input multiple-output (MIMO) system for which the loudspeaker array beamforming coefficients are designed based on the system transfer matrix. The proposed method uses the average output of the microhpne array, over all look directions to avoid signal cancellation due to room reflections. The performance of the proposed method is studied through a simulation example.

Index Terms— Room Acoustics, Beamforming, MIMO Systems, Spherical Arrays

1. INTRODUCTION

Spherical loudspeaker arrays, comprised of a set of loudspeaker units mounted around a surface of a sphere, have been recently studied for applications such as spatial sound reproduction [1], synthesis of radiation patterns for musical instruments [2, 3], local active control of sound [4], and the spatial analysis of sound fields within enclosures [5]. Driving each loudspeaker unit individually facilitates the production of directional beam patterns, capable of exciting desired directions under free-field conditions.

Methods for designing loudspeaker directivity, for the applications listed above, typically assume free-field conditions. However, loudspeaker arrays often operate in enclosures. Therefore, current methods for loudspeaker directivity design may not be suitable for enclosures due to the effect of room reflections. In this paper we propose a new approach for the design of loudspekaer directivity which considers the effect of room reflections. The method employs a spherical microphone array, where the response between each loudspeaker unit of the loudspeaker array and each microphone in the spherical microphone array is measured, forming an acoustic multiple-input multiple-output (MIMO) system. Loudspeaker directivity can now be optimized for the enclosure using adaptive beamforming, emphasizing the direct path and reducing the contribution of room reflections.

In a previous work, we proposed an adaptive algorithm for a similar problem, formulated in the time domain, which uses room impulse responses [6]. This algorithm suffers from two shortcomings. First, it requires recorded responses with high temporal resolution. Second, a closed form solution is given for the optimization of either source or receiver directivity; when optimizing directivities with regard to an acoustic MIMO system, the solution becomes suboptimal and is no longer given in closed form.

The method proposed here overcomes the shortcomings of current methods by considering the average output of the microphone array when rotated over all look directions, therefore capturing the entire sound field at the microphone position with high directional detail. As a result, the method requires knowledge of the transfer matrix for only a limited range of frequencies (supporting limited temporal resolution), has an optimal closed form solution, and can handle coherent room reflections.

The remainder of the paper is organized as follows. Following a brief presentation of background on spherical arrays and free-field beamforming, the formulation and physical interpretation of the new adaptive method are described in sections **5** and **6**, respectively. A simulation study illustrates the performance of the new method in comparison to standard beamforming.

2. BACKGROUND- THE SPHERICAL FOURIER TRANSFORM

This section presents the spherical Fourier transform and some of its properties. These will be used later in the paper for the derivation and interpretation of the new beamformer (BF) design method.

A spherical Fourier transform (SFT) represents a function on a sphere as a linear combination of spherical harmonics, a family of orthogonal functions that solve the angular component of the acoustic wave equation in spherical coordinates. Consider a function $f(\theta, \phi)$ which is square integrable on the unit sphere; then the SFT of f, denoted by f_{nm}, and the inverse transform are given by [7]:

$$f_{nm} = \int_{S^2} f(\theta, \phi) Y_n^{m^*}(\theta, \phi) \sin(\theta) \, d\theta d\phi, \qquad (1)$$

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{nm} Y_n^m(\theta, \phi), \qquad (2)$$

where "*" represents the complex conjugate and the integral covers the entire surface area of the unit sphere, denoted by S^2. $Y_n^m(\theta, \phi)$ represent the spherical harmonic functions, which are orthonormal and complete. The orthogonality of the spherical harmonic functions facilitates analytical analysis and design, and will be used later in the paper; it is given by [8]:

$$\int_{S^2} Y_n^{m^*}(\theta, \phi) Y_{n'}^{m'}(\theta, \phi) \sin(\theta) \, d\theta d\phi = \delta_{nn'} \delta_{mm'}, \qquad (3)$$

where $\delta_{nn'}$ is the Kronecker delta function.

3. SYSTEM MODEL

This section presents the formulation of two acoustics systems, a multiple-input single-output (MISO) system in free-field and a MIMO system in an enclosure. Both formulations are expanded in the spherical harmonic domain; i.e., in both systems the output of the array is calculated as a function of the SFT coefficients.

The far-field directivity pattern of a spherical loudspeaker array in free-field is formulated in [9] and is given by:

$$B(k, \theta, \phi) = \sum_{n=0}^{N_s} \sum_{m=-n}^{n} b_n(kr_s) w_{nm}(kr_0) s(k) Y_n^m(\theta, \phi), \qquad (4)$$

where N_s and r_s are the order and radius of the loudspeaker array, respectively, $b_n(kr_s)$ are coefficients dependant on the array radius and wavenumber k, and $w_{nm}(kr_0)$ is the SFT of the beamforming weighting function .The signal, $s(k)$, is assumed unity. The far-field directivity for a look direction, (θ_l, ϕ_l), and wavenumber k, denoted $B(k, \theta_l, \phi_l)$, is calculated using eq. (4) by substituting (θ_l, ϕ_l) in the spherical harmonic function.

Another system is constructed by placing the loudspeaker in an enclosure and measuring the sound pressure with a spherical microphone array with the order of N_r. The output of a far-field acoustic system in an enclosure is formulated in [6] and given in matrix notation:

$$\left[\gamma_{nm}^{(\mathrm{S})}\right]_{nm}^{\mathrm{H}} \left[H_{nm}^{n'm'}(k)\right]_{nm}^{n'm'} \left[\gamma_{nm}^{(\mathrm{R})}\right]_{nm} =: \boldsymbol{\gamma}_{\mathrm{S}}^{\mathrm{H}} \boldsymbol{H}(k) \boldsymbol{\gamma}_{\mathrm{R}}, \qquad (5)$$

where $\boldsymbol{\gamma}_{\mathrm{S}}$ and $\boldsymbol{\gamma}_{\mathrm{R}}$ are vectors of beamforming coefficients with dimensions of $(N_s+1)^2 \times 1$ and $(N_r+1)^2 \times 1$, for the loudspeaker and microphone arrays, respectively, $\boldsymbol{H}(K)$ is the MIMO transfer matrix for wavenumber k with dimension of $(N_s+1)^2 \times (N_r+1)^2$, and "$^{\mathrm{H}}$" is the Hermitian transpose. In **sec. 5**, the microphone array's directivity pattern is set to the maximum directivity index with a look direction steered over all directions. A maximum directivity beampattern with a look direction of (θ_l, ϕ_l) for the microphone array is achieved by setting $\left[\gamma_{nm}^{(\mathrm{R})}\right]_{nm} = Y_n^m(\theta_l, \phi_l)$ [9]. In vector notation, we denote the latter as $\boldsymbol{\gamma}_{\mathrm{R}} = \mathbf{y}(\boldsymbol{\theta}_l)$.

4. BEAMFORMING IN FREE-FIELD

In this section the maximum directivity index (DI) beamformer, a beamformer typically designed for free-field, is presented. In the next section the proposed adaptive beamformer, inspired by the maximum DI beamformer, will be developed.

The DI of a loudspeaker array quantifies the array's directivity [10], measuring the ratio between the peak and the average values of the squared directivity function. The directivity factor (DF) of an array is defined as [11]

$$\mathrm{DF} = \frac{|B(k, \theta_l, \phi_l)|^2}{\frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} |B(k, \theta, \phi)|^2 \sin(\theta) d\theta d\phi}, \qquad (6)$$

where $B(k, \theta_l, \phi_l)$ is defined in eq. (4). The DI is now computed by DI$= 10\log_{10}(\mathrm{DF})$. Maximizing the DI can be interpreted as maximizing array output for a single plane wave, relative to a spherically isotropic noise field.

An optimal loudspeaker array directivity pattern that maximizes the DI is axis-symmetric; for a look direction of (θ_l, ϕ_l), the SFT of the weighting function, w_{nm}, that achieves maximum directivity is given by $w_{nm} = \frac{Y_n^{m^*}(\theta_l, \phi_l)}{b_n(kr_s)}$ [9].

5. BEAMFORMING IN ENCLOSURES

In this section we expand the concept of the DI and modify it for acoustic MIMO systems. A new spatial measure is defined and loudspeaker beamforming coefficients are designed as to maximize it.

The directivity of the microphone array is defined first; the look direction is set to the direct path between the loudspeaker and the microphone arrays, and a directivity pattern that maximizes the standard DI is employed. Then, we define a new spatial measure, Q, as the microphone array output with its look direction pointing to the loudspeaker array, relative to the average of the microphone array output with its look direction steered over all directions. By maximizing the new spatial measure the beamforming coefficients of the loudspeaker array are designed to maximize the response at the look direction while attenuating significant reflections from other directions.

With $\boldsymbol{\gamma}_{\mathrm{R}} = \mathbf{y}(\theta_{\mathrm{R},0})$, where $\theta_{\mathrm{R},0}$ is the direction of the direct path, we aim to find $\boldsymbol{\gamma}_{\mathrm{S}}$ such that

$$\max_{\boldsymbol{\gamma}_{\mathrm{S}}} \frac{\boldsymbol{\gamma}_{\mathrm{S}}^{\mathrm{H}}\mathbf{H}(k)\mathbf{y}(\theta_{\mathrm{R},0})\mathbf{y}(\theta_{\mathrm{R},0})^{\mathrm{H}}\mathbf{H}(k)^{\mathrm{H}}\boldsymbol{\gamma}_{\mathrm{S}}}{\boldsymbol{\gamma}_{\mathrm{S}}^{\mathrm{H}}\left(\int_{\theta_r}\mathbf{H}(k)\mathbf{y}(\theta_r)\mathbf{y}(\theta_r)^{\mathrm{H}}\mathbf{H}(k)^{\mathrm{H}}d\theta_r\right)\boldsymbol{\gamma}_{\mathrm{S}}}. \tag{7}$$

The numerator represents the output of the micorphone array with the fixed look direction, and the denominator representss the mean of the microphone array output over all steering directions. This measure emphasizes the direct path and attenuates other refelction paths. This will be further discussed in the next section.

The similarity of the proposed BF to the maximum DI BF is now explained; the concept of the conventional DI is extended to MIMO systems, where instead of integrating the loudspeaker array far-field radiation in all directions, we integrate the output of the microphone array due to the sound field arriving from all directions.

Denoting $\mathbf{a} = \mathbf{H}(k)\mathbf{y}(\theta_{\mathrm{R},0})$ and
$\mathbf{B} = \left(\int_{\theta_{\mathrm{R}}}\mathbf{H}(k)\mathbf{y}(\theta_{\mathrm{R}})\mathbf{y}(\theta_{\mathrm{R}})^{\mathrm{H}}\mathbf{H}(k)^{\mathrm{H}}d\theta_{\mathrm{R}}\right)$, the problem is now formulated as a Rayleigh quotient:

$$\max_{\boldsymbol{\gamma}_{\mathrm{S}}} \frac{(\boldsymbol{\gamma}_{\mathrm{S}})^{\mathrm{H}}\mathbf{a}\mathbf{a}^{\mathrm{H}}(\boldsymbol{\gamma}_{\mathrm{S}})}{(\boldsymbol{\gamma}_{\mathrm{S}})^{\mathrm{H}}\mathbf{B}(\boldsymbol{\gamma}_{\mathrm{S}})}, \tag{8}$$

and the solution is given by [12]:

$$\boldsymbol{\gamma}_{\mathrm{S}} = \mathbf{B}^{-1}\mathbf{a}. \tag{9}$$

Using the orthogonality of the spherical harmonic functions as in eq. (3),

$$\begin{aligned}\mathbf{B} &= \int_{\theta_{\mathrm{R}}}\mathbf{H}(k)\mathbf{y}(\theta_{\mathrm{R}})\mathbf{y}(\theta_{\mathrm{R}})^{\mathrm{H}}\mathbf{H}(k)^{\mathrm{H}}d\theta_{\mathrm{R}} = \\ &= \mathbf{H}(k)\int_{\theta_{\mathrm{R}}}\mathbf{y}(\theta_{\mathrm{R}})\mathbf{y}(\theta_{\mathrm{R}})^{\mathrm{H}}d\theta_{\mathrm{R}}\mathbf{H}(k)^{\mathrm{H}} = \\ &= \mathbf{H}(k)\mathbf{I}\mathbf{H}(k)^{\mathrm{H}} = \mathbf{H}(k)\mathbf{H}(k)^{\mathrm{H}}.\end{aligned} \tag{10}$$

Finally, the solution is given by:

$$\boldsymbol{\gamma}_{\mathrm{S}}^{\mathrm{opt}} = (\mathbf{H}(k)\mathbf{H}(k)^{\mathrm{H}})^{-1}\mathbf{H}(k)\mathbf{y}(\theta_{\mathrm{R}}). \tag{11}$$

6. PHYSICAL INTERPRETATION

In order to provide a physical interpretation to the propsed method, we assume an ideal transfer function [6] composed of a collection of idealized reflections and given by:

$$H(k,\boldsymbol{\theta}_{\mathrm{R}},\boldsymbol{\theta}_{\mathrm{S}}) = \sum_i \mathrm{a}_i\, \mathrm{e}^{-jk\mathrm{R}_i}\, \delta(\boldsymbol{\theta}_{\mathrm{R}} - \theta_{\mathrm{R},i})\, \delta(\boldsymbol{\theta}_{\mathrm{S}} - \theta_{\mathrm{S},i}), \tag{12}$$

where $\{\mathrm{a}_i,\ \mathrm{R}_i,\ \theta_{\mathrm{R},i},\ \theta_{\mathrm{S},i}\}$ denote the amplitude, path length, and radiated and received directions, respectively, for the i^{th} refelction. Substituting a limited spatial resolution version of the idealized transfer function in eq. (12) to eq. (7) leads to:

$$\max_{\boldsymbol{\gamma}_{\mathrm{S}}} \frac{|\boldsymbol{\gamma}_{\mathrm{S}}\ \mathrm{a}_0\ \mathbf{y}(\theta_{\mathrm{S},0})|^2}{\sum_i |\boldsymbol{\gamma}_{\mathrm{S}}\ \mathrm{a}_i\ \mathbf{y}(\theta_{\mathrm{S},i})|^2}. \tag{13}$$

It is now clear that maximizing eq. (7) is equivalent to constraining the power for the direct sound while minimizing the power of all reflections.

7. SIMULATION STUDY

In order to study the performance of the new beamformer, a new spatial measure is defined. It compares the power for the direct path relative to the power of the reflected paths, as in eq. (13). The transfer matrix from eq. (5) can now be presented as:

$$\boldsymbol{H}(k) = \boldsymbol{H}_0(k) + \boldsymbol{H}_r(k) \tag{14}$$

where $\boldsymbol{H}_0(k)$ is the matrix for the direct path only, and $\boldsymbol{H}_r(k)$ is the transfer matrix for all reflections. The measure for performance analysis can now be written as:

$$\Psi = 10\log_{10}\left(\frac{|\boldsymbol{\gamma}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{H}_0(k)\boldsymbol{\gamma}_{\mathrm{R}}|^2}{|\boldsymbol{\gamma}_{\mathrm{S}}^{\mathrm{H}}\boldsymbol{H}_r(k)\boldsymbol{\gamma}_{\mathrm{R}}|^2}\right). \tag{15}$$

A room with dimensions of $2.8m \times 4.82m \times 5.71m$ was simulated using McRoomSIM [13]. The transfer matrix, as in eq. (12), relating order-limited spherical loudspeaker and microphone arrays, with $N_s = 3$ and $N_r = 9$, located at $(0.53m, 1.06m, 2m)$ and at $(2.21m, 3.67m, 4.18m)$, respectively, was constructed. A rigid loudspeaker array, comprised of 16 loudspeaker units with a radius of r_s =0.3m, and a rigid microphone array, comprised of 100 microphones with a radius of r_r =0.3m, were simulated.

The spatial measure, Ψ, was calculted for three different configurations. The first simulates an omnidirectional loudspeaker, the second simulates a spherical loudspeaker array with a maximum DI beam pattern (with the look direction set to the direction of the microphone array), and the third simulates the same loudspeaker array with the directivity optimized as in **sec. 5**. In all three cases, Ψ was calculated for an omnidirectional microphone, and k was chosen as to maintain $N_s = kr_s$.

	Ψ	WNG	DI
Omni	-30.30 dB	1.97 dB	0 dB
DI-FF	-20.11 dB	13.49 dB	12.0412 dB
DI-Room	-8.95 dB	11.40 dB	0.7931 dB

Table 1. Ψ, WNG, and DI for the different configurations.

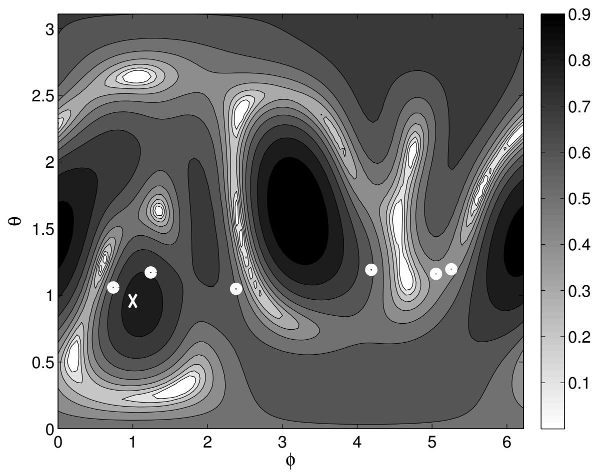

Fig. 1. DI-Room Beampattern. Cross and circles indicate direct and reflected sound respectively.

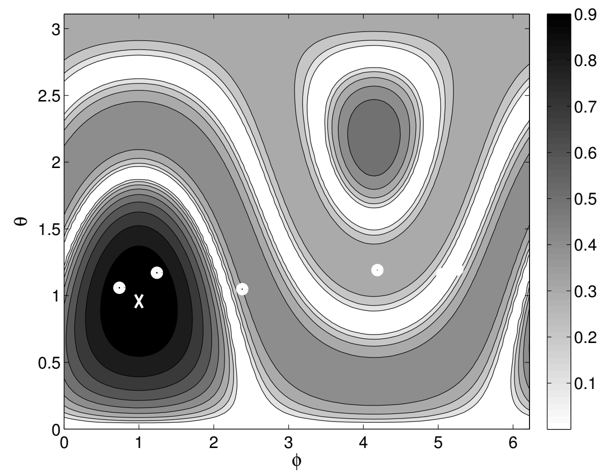

Fig. 2. DI-FF Beampattern. Cross and circles indicate direct and reflected sound respectively

Regularization was applied to the new beamformer, due to the matrix inversion in eq. (11). Table 1 shows the new measure, Ψ, for the three configurations, Omni, DI-FF, and the new method denoted as DI-Room. As can be seen, the new beamformer, DI-Room, achieves Ψ values that are higher by more than 10 dB relative to the values of the DI-FF, demonstrating its superiority. Robustness is also compared; to measure robustness, the white noise gain (WNG), as defined in [9], of each configuration is also presented. We can see that the performance of the DI-Room is very close (2 dB) to the DI-FF. For frequencies maintaing $N_s = kR_s$, the DI-FF achieves maximum WNG, implying the new beamformer is very robust. Table 1 also shows that the free-field DI of the DI-FF beamformer is the best. This is of no surprise, as this BF was designed under free-field conditions.

Figs. 1 and 2 demonstrates the beam pattern for DI-Room and DI-FF, respectively. The cross indicates the direction of the direct sound path and the circles indicate directions of the early reflections. Fig.. 1 shows the graphical interpretation of the optimization process; the main lobe of the beampattern points in the direction of the direct sound path while attenuating directions of early reflections.

8. CONCLUSIONS

In this paper a new algorithm for optimizing the directivity of a spherical loudspeaker array was presented. The algorithm is based on the spatial information acquired by the use of an acoustic MIMO system. It emphasizes the direct sound path while mitigating interfering paths of reflected sound. This facilitates efficient directional transmission and reception of narrowband signals within reverberant enclosures with applications in room acoustics. The algorithm can be modified to other disciplines; for example the method can be implemented in communications for transmission over multipath channels.

Generalizing the proposed algorithm for jointly optimizing both source and receiver directivities, and defining new spatial measures to quanitify their performance, is the topic of current research.

9. REFERENCES

[1] J. Ahrens and S. Spors, "An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 988–999, 2008.

[2] O. Warusfel and N. Misdariis, "Directivity synthesis with a 3d array of loudspeakers-application for stage performance," in *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01), Limerick, Ireland*, 2001, pp. 1–5.

[3] R. Avizienis, A. Freed, P. Kassakian, and D. Wessel, "A compact 120 independent element spherical loudspeaker array with programmable radiation patterns," in *Proceedings of the 120th Audio Engineering Society Convention*, 2012.

[4] B. Rafaely, "Spherical loudspeaker array for local active control of sound," *The Journal of the Acoustical Society of America*, vol. 125, pp. 3006, 2009.

[5] A.M. Pasqual, A. de Franca, J. Roberto, and P. Herzog, "Application of acoustic radiation modes in the directivity control by a spherical loudspeaker array," *Acta Acustica united with Acustica*, vol. 96, no. 1, pp. 32–42, 2010.

[6] H. Morgenstern, F. Zotter, and B. Rafaely, "Joint spherical beam forming for directional analysis of reflections in rooms," in *Acoustic 2012, Hong Kong, China.*

[7] J.R. Driscoll and D.M. Healy, "Computing fourier transforms and convolutions on the 2-sphere," *Advances in Applied Mathematics*, vol. 15, no. 2, pp. 202–250, 1994.

[8] G.B. Arfken, H.J. Weber, and L. Ruby, *Mathematical methods for physicists*, vol. 3, Academic press San Diego, 1985.

[9] B. Rafaely and D. Khaykin, "Optimal model-based beamforming and independent steering for spherical loudspeaker arrays," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 19, no. 7, pp. 2234–2238, 2011.

[10] J. Meyer and G. Elko, "Spherical microphone arrays for 3d sound recording," *audio signal processing for next-generation multimedia communication systems*, pp. 67–89, 2004.

[11] H.L. Van Trees and J. Wiley, *Optimum array processing*, Wiley-Interscience, 2002.

[12] G.H. Golub and C.F. Van Loan, *Matrix computations*, vol. 3, Johns Hopkins University Press, 1996.

[13] A. Wabnitz, N. Epain, C. Jin, and A. van Schaik, "Room acoustics simulation for multichannel microphone arrays," in *Proceedings of the International Symposium on Room Acoustics*, 2010.