Using Uniform Microphone Arrays to Design Sparse Microphone Arrays with Frequency-Independent Beam Pattern

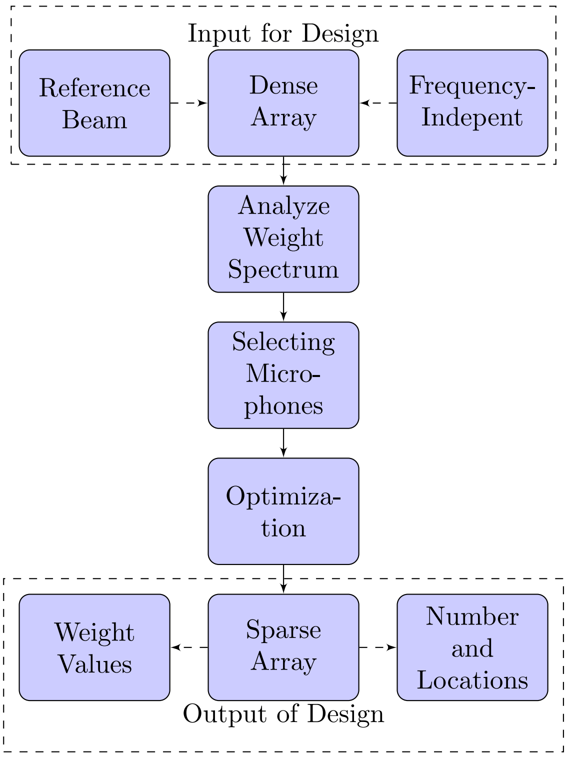
Phan Le Son[1]

[1] *Technische Universität Kaiserslautern, Kaiserslautern, Germany, Email:phan@eit.uni-kl.de*

Introduction

Frequency-Independent (FI) beamforming is a technique used to obtain signals with a wide range of frequencies. It maintains signal integrity and spatial selectivity over frequencies. This is important for audio signals where the bandwidth of signals is several octaves. The design methodology for a Dense and Uniform Array (DUA) with FI beamforming is known [1]. If one has a good design for a DUA with FI beam pattern, one might be able to remove unimportant microphones in the DUA to obtain a Sparse Array (SA). Most studies focus on optimization methods [2, 5, 6] to determine the position of important microphones, but they do not use the information from the DUA. In some circumstances, it is not only difficult to set up the optimization with good parameters, but it is also difficult to solve the optimization with subject to sparse solutions. In this paper, a method is provided to design a SA that takes information from a DUA as computational input (Figure 1).

Figure 1: Flowchart for Sparse Array design.



First, the Coordinate Transformation (CT) method in [1] is used to design a DUA with FI beam pattern. The microphone gains (weights of spatial filter) of the DUA on different frequencies, which are achieved by CT, constitute a weight matrix. Principal Component Analysis (PCA) [3] is applied to reduce the size of the weight matrix because this matrix does not have a full-rank. After compressing the weight matrix, we find the rows with superior energy in the compressed matrix corresponding to the critical microphones that need to be kept in the SA. Next, we apply K-mean clustering algorithm to categorize the remaining microphones in the DUA into different groups [2]. For each group, microphones closest to the its centroid are taken as the group's representative. The representative microphones are added to the SA. The number of microphones in the SA is relative to the number of groups. Moreover, as the number of groups increases, the total distance from the points to the centroid points decreases. This decrease is correlated with the feasibility of optimization in the next step. Once the positions of the microphones in the SA are identified, optimization problems need to be solved to find the microphone gains at every frequency.

Signal Model

In the far-field signal, the wave is planar. The plane of array consolidates with x-y plane in the Cartesian coordinate. Beam pattern in an interested bandwidth, $\forall\omega \in \Omega$:

$$b(\phi, \theta, \omega) = w^H(\omega)d(\phi, \theta, \omega)$$

where the superscript H denotes the conjugates-transpose operator, $w(\omega)$ is a weight vector that contains the complex value of the spatial filter at a frequency ω and $d(\phi, \theta, \omega)$ is a steering vector at direction defined by azimuth and elevation angle (ϕ, θ).
From M microphones with equidistance d_H in the DUA, we select a set s_K contains $K(K << M)$ microphones with indices, $i_k \in [1, .., M], k = 1, ..., K$.
Beam pattern at ω is formed by the microphones in the set s_K:

$$b_S(\phi, \theta, \omega) = w_S^H(\omega)d(\phi, \theta, \omega)$$

Where w_S, has K non-zero elements and $M - K$ zero elements, is $M \times 1$ weight vector of the SA.
Suppose a reference beam pattern is given, then it is possible to design a DUA with M sensors, equidistance d_H and weight vector $w(\omega)$ by the CT method [1] to assure that $b(\phi, \theta, \Omega)$ is almost independent of Ω.
In this paper, a method to find the set of active sensors in a DUA $s_K = [i_1, i_2, ..., i_K]$ is presented so that $b_S(\phi, \theta, \Omega)$ is close with $b(\phi, \theta, \Omega)$.

Design Method for Sparse Array

In recent studies, optimization methods are used to seek s_K and w_S either together [5] or separately [2]. The set-up of optimization problems normally does not have a general rule for using the characteristics of the reference beam pattern, which indeed decide the possibility of the sparseness of an array. With this concern in mind, a three-step method is proposed to design a SA: Analysis, Selecting and Optimization. This method finds s_K and w_S separately. However, in the Analysis step, the optimizations is not used to find microphones' positions, instead a DUA is designed and analyzed. In the Selecting step, the weight matrix of the DUA is used as a basis to select a set s_K. Finally, in the last step, we use optimization method to find the weight vector ω_S. The optimization efforts to find the weight vector is negligible compared with that to find sparse solutions.

Analysis

Let's design a DUA with FI beam pattern $\forall \omega \in \Omega$. We reuse the design constraints of uniform array from [1] to find parameters for a DUA: N, d_H.

$$\begin{cases} \frac{2\pi c}{N d_H} \leq \omega \leq \frac{\pi c(N-2)}{2N d_H}, & \text{N is even} \\ \frac{2\pi c}{N d_H} \leq \omega \leq \frac{\pi c(N-1)}{2N d_H}, & \text{N is odd} \end{cases} \quad (1)$$

Where c is sound speed, N is the number of microphones in one axis (vertical or horizontal) of planar array. Applying the CT method for every $\omega_j \in \Omega$ [1]:

- Step 1: define a reference beam pattern $b_{ref}(\phi, \theta)$.
- Step 2: with radius $R = \frac{\omega_j N d_H}{2\pi c}$, presenting the reference beam pattern to a gain function $b_R(\phi, \theta)$ in the Spherical coordinate.
- Step 3: the gain function in the Cartesian coordinate $b_{\omega_j}(u, v)$ is achieved by transforming $b_R(\phi, \theta)$ in the Spherical coordinate to the Cartesian coordinate.
- Step 4: apply Inverse Fourier Transform of $b_{\omega_j}(u, v)$ to achieve $w(\omega_j)$.

After J iteratives for J frequencies in Ω, a weight matrix is obtained:

$$W_A = [w(\omega_1), w(\omega_2), ..., w(\omega_J)]$$

The weight matrix achieves from this step could help

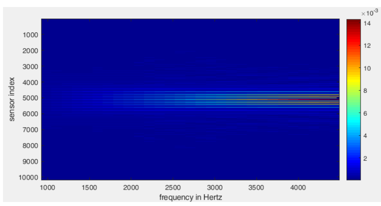


Figure 2: The weight matrix spectrum of a DUA.

us to analyze the possibility of reducing the number of microphones. Figure 2 presents a weight spectrum W_A of a DUA: consists of 101×101 microphones which are rearranged into vertical of the matrix. If there are a few rows in the weight matrix having energy that dominates the other rows, the reference beam pattern has a high potential to reduce a lot of microphones in the DUA. If the energy spectrum of W_A spreads out over the row and column, there is less chance to reduce the number of microphones in the DUA. Therefore, using optimization methods in this step may be inefficient in general because we do not know the good parameters for optimizations without analyzing the weight matrix of the DUA.

Selecting

Before choosing important microphones, we could apply the dimensional reduction algorithm for the weight matrix to save the computation time of the selecting algorithm. If W_A is a rank-deficient matrix with decaying singular-values, using PCA is a good option for dimensional reduction in terms of keeping the essential information.

The input of PCA algorithm is the weight matrix W_A of size $M \times J$. Let matrix $\overline{W}_A$ be the central version of W_A:

$$\overline{W}_A(:, j) = W_A(:, j) - mean(W_A(:, j)), \forall j \in J$$

Where $mean(v)$ is the mean value of vector v.

The output of PCA algorithm is a dimensional reduction matrix:

$$W_R = \overline{W}_A U$$

Where U, is a matrix with size $J \times L$, contains L eingenvectors corresponding to L largest eigenvalues, $L << J$. The column dimension L of the matrix U is a set according to the following criteria: $i^{th}(i \leq L)$ first eigenvalues of $R_A = W_A^H W_A$ in decreasing order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L$ are selected which satisfy [2]:$0.6 < \alpha < 1$.

$$\alpha = \frac{\sum_{j=1}^{L} \lambda_j}{\sum_{j=1}^{J} \lambda_j}$$

Dimensional reduction of weight matrix W_R is showed in

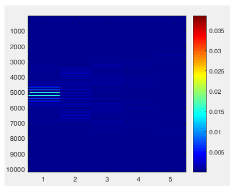


Figure 3: Dimensional reduction of weigh spectrum.

Figure 3. The compressed matrix needs to be analyzed in order to find the critical microphones. The selecting process follows 2 criteria: energy contribution in the weight spectrum and well-support for Optimization step.

Let s_{K_1} be a subset containing indices of K_1 microphones in the DUA which have the strongest energy span over row of W_R. Subset $\overline{s}_{K_1}$ contains the microphones in the DUA excluding microphones in s_{K_1}.

Let a spanning energy for a microphone be a sum of that microphone's energy over frequencies,

$$\epsilon_i = \|W_R(i, :)\|_2^2, i = 1, ..., M$$

$$\epsilon_{max} = max(\epsilon_i), i = 1, ..., M$$

The set of microphones having spanning energy is greater than a threshold,

$$s_{K_1} = find\_index(\epsilon_i > \beta \epsilon_{max}), i = 1, ..., M.$$

Remaining microphones,

$$\overline{s}_{K_1} = find\_index(\epsilon_i \leq \beta \epsilon_{max}), i = 1, ..., M.$$

Where $0 < \beta < 1$ is an arbitration energy factor for the SA. If β is close to 0, then SA is close to DUA. It means more microphones from the DUA are taken. If β is close to 1, the SA only takes a few most important microphones from the DUA.
From subset $\overline{s}_{K_1}$, we apply a clustering algorithm over the rows of $W_R(\overline{s}_{K_1}, :)$ to categorize microphones in $\overline{s}_{K_1}$ into K_2 groups. Each group contains the microphones having similar characteristics (similar pattern of energy distribution over frequencies). K-means clustering algorithm [7] is used to minimize the within-cluster sums of point-to-centroid distances. The sums of point-to-centroid distances could present the gap between the SA and the DUA in terms of FI beam pattern. However, mathematical formula of the relation between the sums of point-to-centroid distances and the distortions of FI beamforming in the SA is not covered in this paper.
K-means algorithm is used to divide $(M - K_1)$ microphones into K_2 groups. The bigger K_2 is relative to the smaller sum of distances to centroids, which means the difference between the beam pattern of SA and the beam pattern of DUA is smaller. Next, for every group, one or several representative microphones which are closest to the centroid of its group are chosen to constitute the subset s_{K_2}.
Finally, the set s_K is the union of subset s_{K_1} and subset s_{K_2},

$$s_K = s_{K_1} \cup s_{K_2}$$

Optimization

We uniform discretized angle spaces and introduce P directions $\rho_i = (\phi_i, \theta_i) \in \Theta$ that cover the entire direction of the beam pattern. We select U out of P directions that cover the main-lobe region Θ_m, and let K_i^m be a set containing these directions at a single frequency ω_i,

$$K_i^m = \{(\rho_1, \omega_i), (\rho_2, \omega_i), ..., (\rho_U, \omega_i)\}$$

Where superscript m stands for main-lobe. Similarly, we define K_i^s containing $P-U$ directions that cover the side-lobe region Θ_s,

$$K_i^s = \{(\rho_{U+1}, \omega_i), (\rho_{U+2}, \omega_i), ..., (\rho_P, \omega_i)\}$$

Where superscript s stands for side-lobe. We define Steering Matrix over angle spaces,

$$D_\Theta(\Theta, \omega_i) = [d(\rho_1, \omega_i), d(\rho_2, \omega_i), ..., d(\rho_P, \omega_i)]$$

Reference beam pattern for main-lobe region Θ_m: b_d^m
Reference beam pattern for side-lobe region Θ_s: b_d^s
Main-lode constraints:

$$\mathcal{C}_1 : \left\| b_d^m - w_S^H(\omega_i) D_\Theta(K_i^m) \right\|_2 \leq \epsilon 1(\omega_i), \forall \omega_i \in \Omega.$$

Side-lode constraints:

$$\mathcal{C}_2 : \left\| b_d^s - w_S^H(\omega_i) D_\Theta(K_i^s) \right\|_2 \leq \epsilon 2(\omega_i), \forall \omega_i \in \Omega.$$

Where $\epsilon 1(\omega_i), \epsilon 2(\omega_i)$ are positive parameters. As we mentioned in the Analysis step, 'the sums of point-to-centroid distances' is correlated with $\epsilon 1(\omega_i), \epsilon 2(\omega_i)$. The smaller 'the sums of point-to-centroid distances' is, the smaller $\epsilon 1(\omega_i), \epsilon 2(\omega_i)$ could be set.
Distortionless response constraint for looking direction (ϕ_0, θ_0):

$$\mathcal{C}_3 : w_S^H(\omega_i) d(\phi_0, \theta_0, \omega_i) = 1, \forall \omega_i \in \Omega.$$

The optimization could find the vector $w_S(\omega_i)$, which may contain the elements with a large number. In such a scenario, the array is sensitive with the white noise. Therefore, the WNG constraint is needed:

$$\mathcal{C}_4 : w_S^H(\omega_i) w_S(\omega_i) \leq \gamma(\omega_i), \forall \omega_i \in \Omega.$$

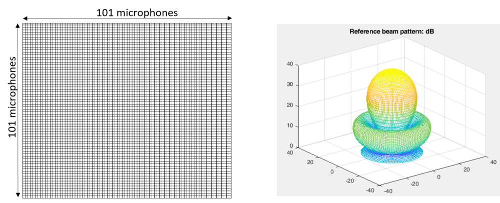
A reasonable choice is to minimize the white noise's power $\mathcal{C}_4$ with subject to the remaining constraints $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ [2].

$$\begin{aligned} &\underset{w_S(\omega_i)}{\text{minimize}} \quad : w_S^H(\omega_i) w_S(\omega_i) \\ &\text{subject to} \\ &\left\| b_d^m - w_S^H(\omega_i) D_\Theta(K_i^m) \right\|_2 \leq \epsilon 1(\omega_i) \\ &\left\| b_d^s - w_S^H(\omega_i) D_\Theta(K_i^s) \right\|_2 \leq \epsilon 2(\omega_i) \\ &w_S^H(\omega_i) d(\phi_0, \theta_0, \omega_i) = 1 \end{aligned} \tag{2}$$

Numerical Simulation

An example for planar microphone array are given.

Step 1:Analysis



(a) (b)

Figure 4: (a) DUA and (b) Reference beam pattern.

Define a reference beam pattern (Figure 4(b)):

$$b(\phi, \theta) = \begin{cases} \left|\frac{\sin(\alpha\pi\theta)}{\alpha\pi\theta}\right|, & \theta > 0, \alpha \text{ is a constant,} \\ 1, & \theta = 0. \end{cases}$$

From (1), we could find a configuration for the DUA: $N = 101$ microphones, $d_H = 0.01$ m, $c = 340$ m/s (Figure 4(a)) and the possible frequency range ($\frac{c}{N d_H} \leq f \leq \frac{c(N-1)}{2 d_H N}$): 337 Hz $\leq f \leq$ 16832 Hz.
The frequency range: $\Omega = \{1Khz, \ldots, 4.5Khz\}$ is selected.
Applying the CT method, we yield the weight matrix and its dimensional reduction is presented in Figure 2 and Figure 3.

Step 2:Selecting

$\beta = 0.75$ is set then $|s_{K_1}| = 21$ microphones. The remaining microphones in the DUA are divided into 15 groups by K-means clustering algorithm. The iteration of K-means clustering is 500 times for this example to find the acceptable solution. Finally, we have $|s_{K_2}| = 60$ microphones (four representative microphones for each group) and the total microphone in the SA is 81, Figure 5.

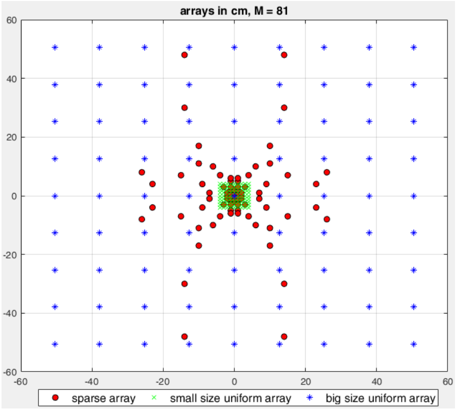


Figure 5: Layout for planar arrays: sparse array ('circle' red), SUA ('x' green) and BUA ('*' blue).

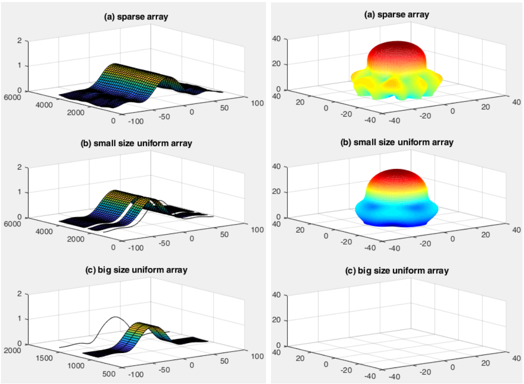


Figure 6: (left) cross-cut of beam pattern versus frequency, (right) beam pattern at 2Khz

Step 3:Optimization

Main-lobe constraint: $\epsilon 1(\omega_i) = 0.002U, \quad \forall \omega_i \in \Omega.$
Side-lobe constraint: $\epsilon 2(\omega_i) = 0.006(P - U), \quad \forall \omega_i \in \Omega.$
Where $P = 180, U = 60$.
The same constraints in (2) apply for SA, SUA and BUA. The synthesized FI beam patterns versus frequencies are presented in the left in Figure 6 and the beam pattern at 2Khz are presented in the right in Figure 6. The optimization of the SA always give the feasible solution while the optimization for SUA and BUA are infeasible for some frequencies (the blanks in the graphs in Figure 6 and Figure 7). Figure 7 depicts the WNGs versus frequencies.

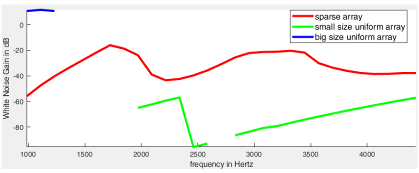


Figure 7: White Noise Gains over frequencies.

Note that it is possible to increase the WNG for the SA either by increasing the number of microphones or reducing the hardness $\epsilon 1(\omega_i), \epsilon 2(\omega_i)$ of constraints $\mathcal{C}_1, \mathcal{C}_2$ for FI beam pattern.
Beam Pattern Error (BPE) indicates the difference between the real beam pattern b_r and the desired beam pattern b_d (the comparison of arrays' BPEs are given in Table 1),

$$BPE = |b_r - b_d|.$$

Table 1: Beam Pattern Error average

Array	Frequency		
	1Khz	2Khz	4Khz
SA	0.0328	0.0305	0.0315
SUA	Infeasible	0.0422	0.0331
BUA	0.0319	Infeasible	Infeasible

Conclusion

A uniform array with near distance of microphones ensures FI beam pattern at high frequency but is worse at the WNG index, while a uniform array with far distance of microphones ensures FI beam pattern at low frequency and is good at the WNG index. In case of wide range frequencies, the uniform array needs to be rather dense and large size to cover both near distance microphones and far distance microphones, therefore the number of microphones needs to be huge. Thus, the sparse array is an optimal solution because it reduces the number of microphones, ensures the FI beam pattern and increases the WNG. Our design method for the sparse arrays uses the information from the uniform array to seek the locations of microphones, which is efficient not only in performance but also in computing time.

References

[1] Phan Le Son, *Uniform Array with Broadband Beamforming for Arbitrary Beam Patterns*, arXiv:2003.00991 [eess.SP],2020

[2] Yaakov Buchris, Alon Amar, Jacob Benesty, and Israel Cohen, *Incoherent Synthesis of Sparse Arrays for Frequency-Invariant Beamforming*, IEEE/ACM Transaction on Audio, Speech, and Language Processing, vol. 27, no. 3, March 2019

[3] I. Jolliffe, *Principal Component Analysis*, Springer Series in Statistics, New York, NY, USA:Springer, 1986.

[4] M. Branstein, D. B. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*, Berlin, Germany:Springer-Verlag, 2001.

[5] Y. Liu, L. Zhang, L. Ye, Z. Nie, Q. H. Liu, *Synthesis of sparse arrays with frequency-invariant-focused beam patterns under accurate sidelobe control by iterative second-order cone programming*, IEEE Trans. Antennas Propag., vol. 63, no. 12, Dec. 2015.

[6] M. B. Hawes, W. Liu, *Sparse array design for wideband beamforming with reduced complexity in tapped delay-lines*, IEEE/ACM Trans. Audio Speech Lang. Process., vol. 22, no. 8, Aug. 2014.

[7] J. A. Hartigan, M. A. Wong, *Algorithm as 136: A k-means clustering algorithm*, J. Roy. Statist. Soc. C (Appl. Statist.), vol. 28, no. 1, 1979.