VPS VIBRATIONS IN PHYSICAL SYSTEMS

DOI: 10.21008/j.0860-6897.2021.2.04
Vibrations in Physical Systems, 32(2):2021204, 2021

A Simulation of Thinning of Microphone Array in Near-field Broadband Beamformers

Agnieszka WIELGUS[1], Bogusław SZLACHETKO[1]

Corresponding author: Agnieszka WIELGUS, email: agnieszka.wielgus@pwr.edu.pl

[1] Department of Acoustic, Multimedia and Signal Processing, Wroclaw University of Science and Technology, Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland

Abstract This paper is devoted to the problem of designing an optimal microphone matrix. We define a criterion function where the performance of our matrix should be as close as possible to the desired one based on L_2 norm. In the classical approach, increasing a size of the matrix is used to improve the system performance. However, in many cases it is not a good solution. In this paper we propose a solution based on thinning technique. We work with rectangular, equispaced microphone matrix and using metaheuristic approach called simulated annealing we optimise the set of active microphones (we switch off some of the microphones from the regular matrix). For illustrations, few numerical examples are solved. Comparing to the classical approach we show that thinning microphone matrix can significantly improve system performance.

Keywords: acoustics, beamforming systems, microphone array, thinning technique

1. Introduction

Beamforming is one of the most important issues in many real-life problems: wireless communication, radars, sonars or matrix antennas and microphones [1, 2]. This term is applicable to both radiation and reception of energy. Many systems designed to receive spatially propagating signals often meet the presence of some interference signals [3]. If the frequency of interfering signal and our desired coincide, it is impossible to use classical temporal filtering. However, if they come from different points in space, signals can be separated using method called spatial filtering. For this purpose a microphone array is used. It acts as a spatial filter that consolidates the acoustic signals received by individual microphones.

The classical approach to the spatial filtering problem is based on a one of the standard microphone layouts (e.g. horizontal, vertical). Each microphone is treated as a filter (FIR or IIR), which coefficients should be determined so that the system output meets the desired criterion, e.g. signal maximisation or minimisation of noise level or its elimination. Such an approach is well known from literature and many different approaches have been developed to solve it [4-6].

In [7] authors presented some assumptions and methods that allow simultaneous optimisation of microphone placement, along with the determination of filter parameters. In addition the presented approach shows that using unequally spaced arrays can provide better results than classical microphone arrays. Since both the placement vector and the filter coefficient vector interact and have influence on system efficiency it cannot be concluded easily that one placement vector is better than another placement vector because for each placement we need to consider the corresponding filter coefficient vector. To isolate the effect of the placement vector, they inhibit the effect of the filter coefficients by considering an infinite length filter. Such an approach was next used in [8,9] to provide some efficient algorithms for microphone placement problem.

In this paper we propose a novel approach to the beamforming problem, based on a "thinning" technique. Such an approach allow us to use the same big matrix for different position of a speaker and different environment. We start with a large, regular microphone matrix and using a metaheuristic we thin this matrix out to get a better system response. In [10] authors considered a problem of sparse beamformer design in case of L_1 norm and based on properties of this norm they are able to reduce the size of microphone matrix. However, to the best of our knowledge, there is no research devoted to the criterion based on L_2 norm.

The paper is organised as follows: section 2 describes the problem, the proposed solution is presented in section 3, while section 4 contains an experimental analysis of the proposed approach. The work ends with a short summary and future research ideas (section 5).

2. Problem formulation

There is given a rectangular microphone array with size **N** ($n \times m$ microphones) with centre located in r_c. Signals from microphones are sampled synchronously, and then digital signals are directed to the inputs of the FIR filters of order L (each array element is an L-tap finite impulse response filter). At a given time a chosen number of microphones (1, 2, …., N) can be active, while others are inactive.

The desired system response $G_d(r, L, f)$ is also defined, where r is the location of the sound source and f is a frequency. The location of the sound source is not constant and can change (e.g. speaker moves around the room). Therefore the desired system response depends on the actual speaker's position. For the N-element microphone array the transfer function of the active i-th microphone in the near field is a function of:

$$A_i(r, f) = \frac{1}{\| r - r_i \|} e^{\frac{-j2\pi f \| r - r_i \|}{c}}, \tag{1}$$

where c is a speed of sound in the air and r_i is a location of the i-th microphone.
According to [7], the frequency responses of these FIR filters are:

$$H_i(h, f, L) = h_i^T d_0(f) \text{ for } i = 1, \dots, N, \text{ where } h_i = [h_i(0), h_i(1), \dots, h_i(L-1)] h_i \in R^L \text{ and}$$
$$d_0(f) = \left[1, e^{\frac{-j2\pi f}{f_S}}, \dots, e^{\frac{-j2\pi f (L-1)}{f_S}}\right]. \tag{2}$$

It is important to notice that the information about group delay is included in the transfer function.

For the given microphone number and their placement (we only consider active microphones) a system response can be found by solving the following equation:

$$G(r, f) = \sum_{i=1}^{N} H_i(h, L, f) A_i(r, f) = A^T(r, f) H(h, f, L). \tag{3}$$

The problem we consider in this paper is to design the microphone array (i.e. to determine which of the microphones should be active for a given position of a speaker and to calculate the FIR filter coefficients) so that the current output the beamformer is as close as possible to the desired one in the sense of L_2 norm. The cost function for a given speaker position r is defined as follows:

$$E(h) = \frac{1}{|\Omega|} \int_{\Omega} \rho(r, f) |A^T(r, f) H(h, f, L) - G_d(r, L, f)|^2 dr df, \tag{4}$$

where $\rho(r,f)$ is a positive weighting function, while Ω defines spatial-frequency domain which consist of passband Ω_P and stopband Ω_S regions, i.e. $\Omega = \Omega_P \cup \Omega_S$.

The criterion function depends on the filter coefficients and the set of chosen microphones. However, with each filter length, a different set of coefficients is connected. In [7, 8] authors noticed that the optimal cost function value will not increase as filter length increases to infinity and based on it, they defined the system performance limit. Their experiments shows that with increasing the filter length the minimum cost function values approach the performance limit quickly. If the filter length is fixed and sufficiently long, the problem can be formulated as finding the filter coefficients together with the placement vector simultaneously.

Let $\lambda = (r_1, r_2, \dots, r_N)$be the vector of microphone locations and $\lambda_a = (r_{a1}, r_{a2}, \dots, r_{aM})$, where $M \leqslant N$is a vector of active microphone locations. We define the problem as:

$$\min_{h\in\mathbb{R}^{ML}} E(h,\lambda_a,r) \text{ where } E(h,\lambda_a,r)=\frac{1}{|\Omega|}\int_\Omega \rho(r,f)\left|A^T(r,f,\lambda_a)H(h,f,L)-G_d(\lambda_a,r,L,f)\right|^2 drdf \tag{5}$$

and $\boldsymbol{\lambda}_a$ is decision variable for current $\boldsymbol{r}$ (position of a speaker).
For any given $\boldsymbol{\lambda}_a$ function (8) is convex, however non-convex with respect to the placement vector λ_a.

3. Algorithms

The size of the solution space is 2^N (exponential), where N is a number of all microphones, thus, it is hard to check all possible solutions in a reasonable time. Therefore, we propose a solution based on a metaheuristic algorithm called a simulated annealing (SA), a well known local search algorithm [11]. This simple in implementation algorithm is a powerful approach to solve a complex problems.

The name of the algorithm refers to the thermodynamic cooling process in which the crystalline substance is heated and then slowly cooled up to the reaching of regular crystal structure. During each iteration a new solution is generated from the neighbourhood of the current solution. These two solutions are compared and a solution with a better criterion value is always accepted, while a solution with a worse value of the objective function can be accepted with a certain probability. This probability decreases with the number of iterations and its value depends on the current value of parameter called temperature.

For the considered problem SA algorithm starts with a criterion value calculated for the case in which all microphones are active. Next, a random solution $\boldsymbol{\lambda}_a$ is generated (a random microphones are active). In each iteration a new solution $\boldsymbol{\lambda}_{anew}$ is generated from a neighbourhood of the current solution $\boldsymbol{\lambda}_a$. As a neighbourhood of $\boldsymbol{\lambda}_a$ we define all solutions we can receive after switching an activeness of two random microphones. In addition during each iteration we negate current status of two random microphones. The SA algorithm is presented below (see Fig. 1).

Algorithm 1: Simulated annealing

Define objective function $\min_{h\in R^{M,L}} E(\mathbf{h},\lambda_a,r)$
Calculate criterion for full active matrix λ_f, set T, max_{it} and γ, $E_{best}=E(\lambda_f)$
Generate random initial solution λ_a, $E_{curr}=E(\lambda_a)$
while ($iter < max_{it}$) or (stop criterion) **do**
  Choose λ_{anew} by a random switch of two microphone activeness and negate activeness of two random microphones
  Assign $\lambda_a=\lambda'_{anew}$ with probability $P(T,\lambda_a,\lambda_{anew})=\min\left\{1,\exp\frac{-(E(\lambda_{anew})-E(\lambda_a))}{T}\right\}$
  if $E(\lambda_{anew})<E_{best}$ **then**
    $\lambda_{best}=\lambda_{anew}$
  end
  $T=\frac{T}{1+\gamma T}$
end
Show the best solution λ_{best}

Fig. 1. SA algorithm.

4. Numerical experiments

This section is devoted to the results of the numerical experiments. Note that there are no benchmark instances for this problem. All codes are implemented in MATLAB platform and run on PC with Intel(R) Core i7 CPU with 2.5 GHz. As in [7] the desired response function is specified over a region that would fit into a room with a single speaker and some listeners. It includes the frequency range of human voice, however we cut the upper frequency to 1.5 kHz. It still maintains speech recognition. Since we should allow for the delay of the speech to reach the microphones, the desired response function in the passband region is defined as:

$$G_d(\lambda,r,f)=e^{-j2\pi f\left(\frac{\|r-r_c\|}{c}+\frac{L-1}{2}T\right)}, \tag{6}$$

where $r_c=\sum_{i=1}^{M} r_i/M$ denotes the centre position of the placement variable $\boldsymbol{\lambda}_a$ and c = 340.9 m/s is the sound speed in the air. The desired response is depicted in Fig. 2. The sampling rate is set as 8 kHz, and the maximum frequency is chosen as 4 kHz. In addition it is assumed that the minimum distance parameter between two different microphone elements ε_d can not be smaller than 0.015^2 m^2. The weighting function is chosen as $\rho(r, f) = 1$. According to [7] we set maximum filter length as L=40.

The placement configuration problem is considered in two dimensions. The microphone array can consists of N=16, 25, 36, 49 elements. The passband region is defined as follows:

$$\Omega_p = (r, f) \colon \{(r, f) \colon 0.5 \text{ kHz} \leqslant f \leqslant 1.5 \text{ kHz}, -0.4 \text{ m} \leqslant x \leqslant 0.4 \text{ m}, y = 0 \text{ m}\} \tag{7}$$

and the stopband Ω_s is a sum of the following three parts:

$$\begin{gathered} (r, f) \colon \{(r, f) \colon 2 \text{ kHz} \leqslant f \leqslant 4 \text{ kHz}, -0.4 \text{ m} \leqslant x \leqslant 0.4 \text{ m}, y = 0 \text{ m}\}, \\ \{(r, f) \colon 0.5 \text{ kHz} \leqslant f \leqslant 1.5 \text{ kHz}, 1.8 \text{ m} \leqslant | x | \leqslant 3.0 \text{ m}, y = 0 \text{ m}\}, \\ \{(r, f) \colon 2 \text{ kHz} \leqslant f \leqslant 4 \text{ kHz}, 1.8 \text{ m} \leqslant | x | \leqslant 3.0 \text{ m}, y = 0 \text{ m}\} \end{gathered} \tag{8}$$

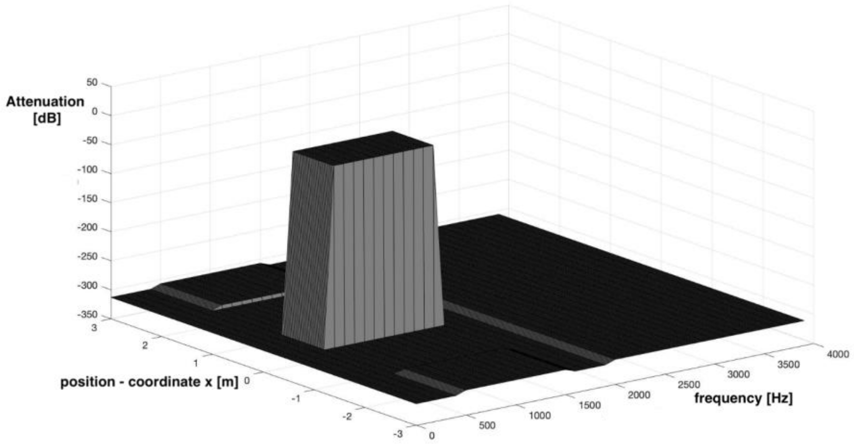

Fig. 2. Desired response of the system – no dumping in passband region.

The regions of the locations for the microphones are chosen as:

$$\begin{gathered} \Lambda_1 = \{(x, y) \colon -0.1 \text{ m} \leqslant x \leqslant 0.1 \text{ m}, 0.9\ m \leqslant y \leqslant 1.5 \text{ m}\} \\ \Lambda_2 = \{(x, y) \colon -0.15 \text{ m} \leqslant x \leqslant 0.15 \text{ m}, 0.9 \text{ m} \leqslant y \leqslant 1.5 \text{ m}\} \\ \Lambda_3 = \{(x, y) \colon -0.2 \text{ m} \leqslant x \leqslant 0.2 \text{ m}, 0.9 \text{ m} \leqslant y \leqslant 1.5 \text{ m}\} \end{gathered} \tag{9}$$

Both passband and stopband are discretised, the frequency points are taken every 0.1 kHz and the spatial points are taken every 0.02m. In each location Λ_1, Λ_2, Λ_3 the microphones at the beginning are equispaced and fulfil all region. In [7] the authors claim that for 5 microphones and the considered passband and stopband regions and desired response function the optimal cost function value is -52.76 dB. The optimised placement vector is: λ^*={(0, 1.4975), (0,1.4736), (0,1.4139), (0, 1.2585), (0,0.9)}. We treat this solution as reference solution, and with our settings of calculations we get cost function value equal to -40.6072 dB for this placement.

The results of our numerical experiments are gathered in Tab. 1. We present the initial value of cost function (for all microphones active), best found criterion value with the number of active microphones and mean criterion value with the mean number of microphones. The parameters of SA algorithm were as follows: T=100000 (initial temperature), max_{it}=1000 (number of iterations), γ=0.1 (cooling ratio).

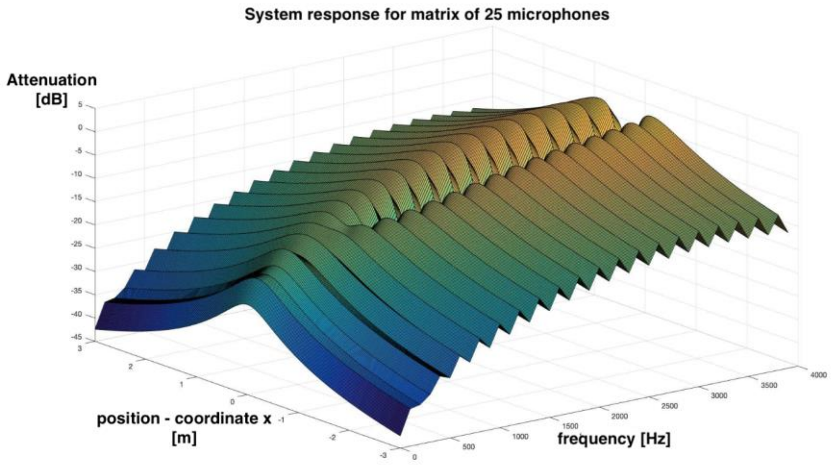

Fig. 3. Initial magnitude of the actual response for matrix 5x5 Λ_1, criterion value -16.8800 dB.

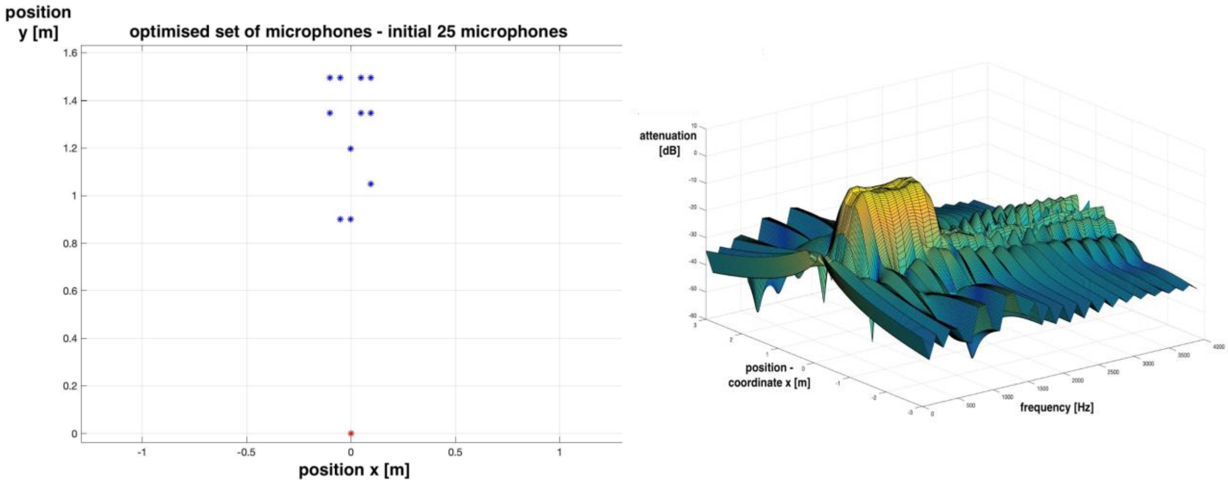

Fig. 4. Optimised microphone matrix and magnitude of the actual response for this matrix, Λ_1, initial matrix size 5x5 microphones, initial criterion value -16.8800 dB, optimised criterion value -41.4011 dB.

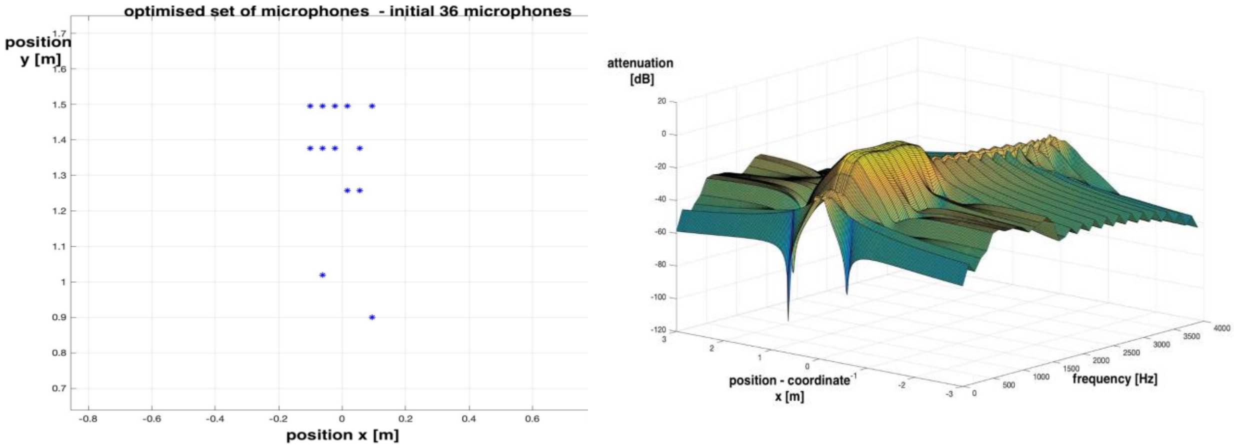

Fig. 5. Optimised microphone matrix and magnitude of the actual response for this matrix, Λ_1, initial matrix size 6x6 microphones, initial criterion value -20.5185 dB, optimised criterion value -41.4477 dB.

In Figure 3 a magnitude of the actual response for all 25 microphone is depicted (cross-section in *y*=1). In Figure 4 we can see the best found solution. It can be seen that removing some microphones can provide a significant improvement in system efficiency. In Figure 5 we can see an optimised set of microphones and the magnitude of the actual response for this optimised placement. The initial size of matrix was 36 (6 rows and 6 columns). In both cases thinning of microphone matrix provides better solutions.

5. Conclusions

In this paper we show that increasing the size of microphone matrix to get better results is not the correct approach. Having a large-sized microphone matrix we can use thinning techniques to get a significantly better objective function. Even the proposed, simple in implementation approach, allow us to improve the quality of our system. The future research will focus on verifying the theoretical results during experiments, developing and verifying more sophisticated stochastic optimisation techniques like particle swarm optimisation or modern hybrid algorithms. The second approach can be an attempt to construct an exact algorithm based on a branch and bound technique.

Tab. 1. Cost function values.

Λ	Matrix size	Initial c.v. [dB]	Best c.v. [dB]	Best m.nbr	Mean c.v [dB]	Mean m.nbr
Λ_1	4x4	-24.7404	-41.9219	7	-39.1719	11
	5x5	-16.8800	-41.4011	11	-35.0871	11
	6x6	-20.5195	-41.4474	13	-31.2801	16
	7x7	-25.7158	-40.3300	8	-31.0219	16
Λ_2	4x4	-26.3237	-42.4611	6	-40.7075	7
	5x5	-17.8042	-41.6706	11	-39.6546	15
	6x6	-19.0891	-31.5189	16	-30.7452	20
	7x7	-23.7878	-31.1964	18	-30.5550	24
Λ_3	4x4	-26.0102	-42.7924	6	-40.0542	9
	5x5	-24.1060	-43.5116	8	-41.4291	12
	6x6	-25.5757	-42.3267	10	-36.9096	16
	7x7	-27.1484	-34.0625	21	-31.7225	22

Additional information

The authors declare no competing financial interests.

References

1. D. H. Johnson, D. E. Dudgeon. Array Signal Processing: Concepts and Techniques. New York, NY, USA: Simon & Schuster, Inc., 1992.
2. W. Liu, S. Weiss. Design of frequency invariant beamformers for broadband arrays. IEEE Transactions on Signal Processing, 56(2): 855–860, 2008.
3. B. D. Van Veen, K. M. Buckley. Beamforming: A Versatile Approach to Spatial Filtering. IEEE ASSP Magazine, 5:4-24, 1998.
4. B. K. Lau, Y. H. Leung, K. L. Teo, V. Steeram. Minimax filters for microphone arrays. IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, 46(12):1522–1524, 1999.
5. R. A. Kennedy, D. B. Ward, T. D. Abhayapala. Nearfield beamforming using radial reciprocity. IEEE Transactions on Signal Processing, 47(1):33–40, 1999.
6. S. Nordebo, I. Claesson, S. Nordholm. Weighted chebyshev approximation for the design of broadband beamformers using quadratic programming. IEEE Signal Processing Letters, 1(7):103–105 1994.
7. Z. G. Feng, K. F. C. Yiu, S. E. Nordholm. Placement design of microphone arrays in near-field broadband beamformers. IEEE Transactions on Signal Processing, 60(3):1195–1204, 2012. DOI: 10.1109/TSP.2011.2178491
8. Z. Li, K. F. C. Yiu, and Z. Feng. A hybrid descent method with genetic algorithm for microphone array placement design. Applied Soft Computing, 13(3):1486-1490, 2013. DOI: 10.1016/j.asoc.2012.02.027
9. Z. G. Feng, K. F. C. Yiu, S. E. Nordholm. Performance Limit of Broadband Beamformer Designs in Space and Frequency. J. Optim. Theory Appl., 164:316–341, 2015. DOI: 10.1007/s10957-014-0543-5
10. M. Gao , K. F. C. Yiu, S. E. Nordholm. On the Sparse Beamformer Design. Sensors, 18(10):3536, 2018. DOI:10.3390/s18103536
11. S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi. Optimization by simulated annealing. Science, 220(4598): 671-680, 1983.