

Article Optimal Microphone Array Placement Design Using the Bayesian Optimization Method

Yuhan Zhang ^{1,2}, Zhibao Li ² and Ka Fai Cedric Yiu ^{1,*}

- ¹ Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China; yhmath.zhang@gmail.com
- ² School of Mathematics and Statistics, Central South University, Changsha 410083, China; zblimath@csu.edu.cn
- * Correspondence: cedric.yiu@polyu.edu.hk; Tel.: +852-34008981

Abstract: In addition to the filter coefficients, the location of the microphone array is a crucial factor in improving the overall performance of a beamformer. The optimal microphone array placement can considerably enhance speech quality. However, the optimization problem with microphone configuration variables is non-convex and highly non-linear. Heuristic algorithms that are frequently employed take a long time and have a chance of missing the optimal microphone array placement design. We extend the Bayesian optimization method to solve the microphone array configuration design problem. The proposed Bayesian optimization method does not depend on gradient and Hessian approximations and makes use of all the information available from prior evaluations. Furthermore, Gaussian process regression and acquisition functions make up the Bayesian optimization method. The objective function is given a prior probabilistic model through Gaussian process regression, which exploits this model while integrating out uncertainty. The acquisition function is adopted to decide the next placement point based upon the incumbent optimum with the posterior distribution. Numerical experiments have demonstrated that the Bayesian optimization method could find a similar or better microphone array placement compared with the hybrid descent method and computational time is significantly reduced. Our proposed method is at least four times faster than the hybrid descent method to find the optimal microphone array configuration from the numerical results.

Keywords: Bayesian optimization; beamformer design; microphone placement; Gaussian process regression; acquisition function

1. Introduction

Beamforming techniques can effectively obtain the sound of interest via spatial filtering to reduce interference and ambient noise from a mixed signal received by a set of microphone arrays. They are now widely used in the fields of wireless communications, hearing aids, and speech recognition [1-4]. Many techniques currently exist for solving the filter coefficients to achieve speech enhancement under specific conditions; for example, the linearly constrained minimum variance (LCMV) beamformer presented in [5] minimizes the power of the background noise, and dereverberation and interference suppressing are employed as constraints. It is worth noting that the length of the filters and number of microphones also greatly affect the beamformer's performance. When filters reach a certain length, the performance limit enters stagnation and is far from satisfactory; in contrast, as the number of microphones is increased, the desired directivity pattern can be achieved under some circumstances [6]. In addition, the design of the microphone array's location has a big influence on how well the beamformer works. Regular microphone array placements are always chosen [7], but it was discovered that the optimized microphone array placement within specific dimensions and areas significantly increased the overall performance compared to the regular placement in [8,9].

Many optimization problems and algorithms have been established to solve the microphone array configuration issue. The array-thinning technique [10,11] carefully selects



Citation: Zhang, Y.; Li, Z.; Yiu, K.F.C. Optimal Microphone Array Placement Design Using the Bayesian Optimization Method. *Sensors* **2024**, 24, 2434. https://doi.org/ 10.3390/s24082434

Academic Editor: Iren E. Kuznetsova

Received: 1 March 2024 Revised: 2 April 2024 Accepted: 8 April 2024 Published: 10 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). the location of the microphone array while using fewer microphones to preserve the prior performance in the one-dimensional situation. Several studies have employed heuristic methods to identify the global optimal solution since the problem of optimization is nonconvex and nonlinear. These include evolutionary programming [12], genetic algorithm (GA) [13–16], simulated annealing algorithm [17], pattern-search algorithm [18], and differential evolution [19]. However, these methods are often very time-consuming and have the risk of missing the optimal solution. In [20], an approach based on compressive sensing is described for building wideband sparse microphone arrays. The Taguchi method makes an effort to perform systematic experiments based upon orthogonal arrays to analyze the microphone design after pre-selecting multiple possible positions for the microphone elements [21]. By making the filter length sufficiently long, a nonlinear optimization problem on filter coefficients and microphone array placement using the l_2 norm is presented in [8]. This problem is then reduced to one where the placement variable is the only decision variable in [9], and a hybrid descent method incorporating a genetic algorithm is provided to obtain a more general solution.

Many methods mentioned above based on heuristic algorithms tend to be very inefficient. However, a Bayesian optimization method makes good use of the prior information from previous iterations and provides a posterior probability distribution to describe potential microphone array locations. The computational efficiency can be greatly improved. Bayesian optimization [22,23] is mainly for independent variables over continuous domains. It is widely applied in machine learning [24], the design of mechanical systems and materials [25–27], and the development of pharmaceuticals [28,29] because it is capable of tackling optimization problems with complicated objective functions.

This paper aims to extend the Bayesian optimization method to solve the microphone array configuration problem to improve the computational efficiency. Given that the nonconvexity and non-linearity of the microphone array configuration optimization issue and the objective function form a dual integral, using heuristic methods is inefficient. The Bayesian optimization method can make full use of previously evaluated information by employing Gaussian progress (GP) regression to proxy the objective function of the microphone array design problem, while the acquisition function directs iterations to the point with the highest probability of being the minimum value.

In this study, we applied the Bayesian optimization method to find the optimal microphone array placement since the optimization problem with respect to the microphone array variables is non-convex and takes a long time to compute. GP regression can be applied to approximate the objective function and obtain the posterior distributions for the rest of the points in the feasible domain. The next sampling point should be smaller than the current minimum with greater probability and improvement by optimizing the acquisition function (obtained from the posterior probability distribution). As more configuration samples are evaluated, the posterior distribution is continually updated. This can bring the new feasible solution closer to the optimal locations for the microphone array. This method can efficiently achieve more excellent performance than the hybrid descent method with higher computational efficiency.

Our contributions can be summarized as the following:

- Considering that the microphone array design problem is non-convex and non-linear, the Bayesian optimization method is extended to solve the microphone array placement design problem;
- 2. GP regression is used to surrogate objective function in the microphone array placement design optimization problem, while different acquisition function strategies are applied;
- 3. Numerical experiments demonstrate that the proposed Bayesian optimization method could produce the same or better performance with shorter computational time compared with the hybrid descent method [9].

2. Problem Formulation

Assume that the signal received by each element of an array with N microphones is processed by a finite impulse response (FIR) filter and that L is the length of filter. The transfer function from the sound source to the *i*th microphone can be given as

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

the microphone has been fixed in r_i , i = 1, ..., M, the location of the sound signal is determined by s, its frequency is determined by f, and c denotes the speed of sound in air. These FIR filter frequency responses are as follows when the signals are sampled synchronously at a rate of f_s per second:

$$W_i(\boldsymbol{w}_i, f, L) = \boldsymbol{w}_i^T \boldsymbol{d}_0(f, L),$$

where $w_i = [w_i(0), w_i(1), \dots, w_i(L-1)]^T$ indicates the coefficients of the *i*th FIR filter, and the vector $d_0(f, L)$ is stated as

$$d_0(f,L) = [1, e^{\frac{-j2\pi f}{f_s}}, \dots, e^{\frac{-j2\pi f}{f_s}(L-1)}]^T,$$

in which $(\cdot)^T$ denotes the matrix transpose.

Figure 1 illustrates the structure of the microphone array. The actual response by the beamformer could be constructed as follows based on the *i*th frequency response and the transfer function to the *i*th element microphone:

$$G(\boldsymbol{\lambda}, \boldsymbol{s}, \boldsymbol{w}, L) = \sum_{i=1}^{M} A_i(\boldsymbol{r}_i, \boldsymbol{s}, f) W_i(\boldsymbol{w}_i, f, L),$$
(1)

where $\lambda = \{r_1, r_2, \dots, r_M\}$ is the set of microphone array locations and $w = \{w_1, w_2, \dots, w_M\}$ denotes the coefficients of all FIR filters.



Figure 1. The structure of a microphone array with N microphones.

As has been proven in Lemma 1 in [8], infinite-length filters and frequency-response functions have an equal relationship according to the infinite-length technique. In addition, the infinite-length technique has been applied in [8,9] to provide a beamforming output that is independent of the filter length L:

$$\tilde{G}(\boldsymbol{\lambda}, \boldsymbol{s}, \tilde{\boldsymbol{w}}, f) = \sum_{i=1}^{M} A_i(\boldsymbol{r}_i, \boldsymbol{s}, f) \tilde{W}_i(\tilde{\boldsymbol{w}}_i, f),$$

where $\tilde{w}_i \in \tilde{\Gamma}$, $\tilde{\Gamma} = {\tilde{u}(f) + j\tilde{v}(f) : \tilde{u}(f) \text{ and } \tilde{v}(f) \text{ are continuous and absolutely integrable, and the right-hand and left-hand derivatives exist, <math>\tilde{v}(0) = 0$, $\tilde{v}(fs/2) = 0$ }.

Given that the desired response is $G_d(\lambda, s, f)$ and that the l_2 norm is frequently employed as a measure of the error between $\tilde{G}(\lambda, s, \tilde{w}, f)$ and $G_d(\lambda, s, f)$, the objective function with regard to the coefficients of beamformer \tilde{w} and the microphone configuration variables λ are

$$F(\boldsymbol{\lambda}, \boldsymbol{\tilde{w}}) = \frac{1}{|\Omega|} \int_{\Omega} \rho(\boldsymbol{\lambda}, f) |\tilde{G}(\boldsymbol{\lambda}, \boldsymbol{s}, \boldsymbol{\tilde{w}}, f) - G_d(\boldsymbol{\lambda}, \boldsymbol{s}, f)|^2 d\boldsymbol{s} df,$$

where Ω is a predefined spatial–frequency domain. $\rho(\lambda, f)$ is a positive weighting function. The domain Ω is frequently composed of passband area Ω_p and stopband area Ω_s . The following optimization problem determines a set of microphone array placements λ and a set of beamformer coefficients \tilde{w} that minimize the error:

$$\min_{\tilde{w}\in\tilde{\Gamma}^{N},\lambda\in\Lambda}F(\lambda,\tilde{w})$$
s.t. $||r_{i}-r_{j}||^{2}\geq\bar{\varepsilon}_{d},i\neq j,$
(2)

A indicates the possible area for the microphone array. $\bar{\varepsilon}_d$ is the square of the minimum distance between two independent microphone elements, and restrictions $||r_i - r_j||^2 \ge \bar{\varepsilon}_d$, $i \neq j$ ensure that microphone elements work efficiently at a minimum distance from each other.

It is challenging to solve the non-convex optimization problem in Equation (2) as a whole since it consists of two separate kinds of variables. If the microphone array configuration is determined, the beamformer coefficient design is reduced to a convex optimization problem. Therefore, the optimization problem in (2) might be rewritten as

$$\begin{array}{l} \min_{\boldsymbol{\lambda} \in \Lambda} \quad F(\boldsymbol{\lambda}, \tilde{\boldsymbol{w}}^*) \\ s.t. \quad \|\boldsymbol{r}_i - \boldsymbol{r}_j\|^2 \ge \bar{\varepsilon}_d, i \neq j. \end{array} \tag{3}$$

The optimum beamformer coefficients under the specified array placements are \tilde{w}^* . However, the only decision variables λ nested inside $A_i(r_i, s, f)$, $G_d(\lambda, s, f)$, and the objective function in (3) are non-convex with regard to λ . Although the hybrid descent method proposed in [9] can find the optimal set of microphone array placements, its evaluation of the next microphone array location λ requires considerable time and results in an inefficient algorithm; a Bayesian optimization method is introduced to improve computational efficiency and to find a better microphone array configuration.

3. Bayesian Optimization Method

The non-convex optimization problem in (3) for the location variables λ is an extremely difficult to compute integral objective function. But, in the Bayesian optimization method, the multiple integral function in (2) might be demonstrated with a GP model. Moreover, the Bayesian optimization method, as has been proven by [23], can finally converge to the global optimal solution. A detailed description of the Bayesian optimization method for solving microphone array configuration problems is presented.

3.1. GP Regression

GP [30,31] has been considered as a good way to model loss functions in Bayesian statistical methods and has been applied in classification [32], face recognition [33], and neural networks [34]. Suppose that a finite collection of *n* different placements $\lambda_{1:n}$ is selected and that the objective function values and noisy observations are denoted by the variables $F(\lambda_{1:n})$ and $\bar{F}(\lambda_{1:n})$, respectively. In GP regression, it is assumed that $F(\lambda_{1:n})$ will follow the a priori GP distribution and observation error $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, resulting in observations $\bar{F}(\lambda) = F(\lambda) + \varepsilon$. Let $\mathbb{D}_n = \{(\lambda_i, \bar{F}(\lambda_i))\}_{i=1}^n$ denote the group of observations

$$\bar{F}(\lambda_{1:n}) \sim \mathcal{N}(m_0(\lambda_{1:n}), \Sigma_0(\lambda_{1:n}, \lambda_{1:n})), \tag{4}$$

where

$$\lambda_{1:n} = [\lambda_1, \lambda_2, \dots, \lambda_n],$$

$$\bar{F}(\lambda_{1:n}) = [\bar{F}(\lambda_1), \bar{F}(\lambda_2), \dots, \bar{F}(\lambda_n)],$$

$$m_0(\lambda_{1:n}) = [m_0(\lambda_1), m_0(\lambda_2), \dots, m_0(\lambda_n)],$$

$$\Sigma_0(\lambda_{1:n}, \lambda_{1:n}) = \begin{bmatrix} \mathbb{Q}(\lambda_1, \lambda_1) & \cdots & \mathbb{Q}(\lambda_1, \lambda_n) \\ \vdots & \ddots & \cdots \\ \mathbb{Q}(\lambda_n, \lambda_1) & \cdots & \mathbb{Q}(\lambda_n, \lambda_n) \end{bmatrix} + \sigma^2 I,$$
(5)

Equation (5) is from [22]. $\mathcal{N}(m_0(\cdot), \Sigma_0(\cdot))$ denotes a Gaussian prior distribution with $m_0(\cdot) : \mathbb{R}^{3 \times 1} \mapsto \mathbb{R}$ as the prior mean and $\Sigma_0(\cdot) \in \mathbb{R}^{n \times n}$ as the covariance matrix. $\mathbb{Q}(\lambda, \hat{\lambda})$ is a kernel to measure the correlation of λ and $\hat{\lambda}$. Generally, the squared exponential kernel

$$\mathbb{Q}(\lambda,\hat{\lambda}) = \sigma_f^2 e^{\left(-\frac{\|\lambda-\hat{\lambda}\|^2}{2l^2}\right)}$$
(6)

is used. Equation (6) is from [32]. Notice that, the closer the two points, the bigger the value of the function, and that, the further away, the smaller the value of the function. This property shows that squared exponential functions are suitable to characterize the similarity between different microphone configurations.

3.2. Choosing Prior Hyperparameters

In the covariance matrix $\Sigma_0(\lambda_{1:n}, \lambda_{1:n})$, prior hyperparameters $\theta =: \{\sigma_f, l, \sigma\}$ must be chosen in accordance with the provided observations samples \mathbb{D}_n . Maximum likelihood estimation (MLE) is frequently used in probability and statistics to fit the GP model [35]. The distribution under these previous hyperparameters is known to us:

$$\overline{F}(\boldsymbol{\lambda}_{1:n})|\boldsymbol{\theta} \sim \mathcal{N}(m_0(\boldsymbol{\lambda}_{1:n}), \Sigma_0(\boldsymbol{\lambda}_{1:n}, \boldsymbol{\lambda}_{1:n})),$$

where we modify the notation in (4) to show that it depends on θ . The log-likelihood function can be easily obtained:

$$\log P(\bar{F}(\lambda_{1:n})|\theta) = -\frac{1}{2}\log|\Sigma_0| - \frac{1}{2}\bar{F}^T(\lambda_{1:n})\Sigma_0\bar{F}(\lambda_{1:n}) - \frac{n}{2}\log(2\pi).$$
(7)

Equation (7) is from [23]. Subsequently, the maximum likelihood function is employed to determine the prior hyperparameters:

$$\hat{\theta} = \arg\max_{\theta} \log P(\bar{F}(\lambda_{1:n})|\theta).$$
(8)

3.3. Acquisition Function

Bayes' rules predict that the random variable $\bar{F}(\bar{\lambda})$ is normally distributed. The posterior mean and variance function are as follows:

$$\bar{F}(\bar{\lambda})|\bar{F}(\lambda_{1:n}) \sim N(m(\bar{\lambda}), \sigma^2(\bar{\lambda})), \tag{9}$$

where

$$m(\bar{\lambda}) = \Sigma_0(\bar{\lambda}, \lambda_{1:n}) \Sigma_0(\lambda_{1:n}, \lambda_{1:n})^{-1} (\bar{F}(\lambda_{1:n}) - m_0(\lambda_{1:n})) + m_0(\bar{\lambda}),$$

$$\sigma_n^2(\bar{\lambda}) = \Sigma_0(\bar{\lambda}, \bar{\lambda}) - \Sigma_0(\bar{\lambda}, \lambda_{1:n}) \Sigma_0(\lambda_{1:n}, \lambda_{1:n})^{-1} \Sigma_0(\lambda_{1:n}, \bar{\lambda}),$$
(10)

Equation (10) is from [22]. The input data $\overline{F}(\lambda_{1:n})$ and prior $m_0(\overline{\lambda})$ are averaged jointly to obtain the posterior mean $m(\overline{\lambda})$, whose weight is dependent on the kernel. The data provide more information; it should be shown that the posterior variance is always less than the previous variance.

The objective function's prediction and uncertainty are represented by the posterior mean $m(\bar{\lambda})$ and variance $\sigma_n^2(\bar{\lambda})$, calculated at each point $\bar{\lambda}$ in (10). The acquisition function has the responsibility of directing the pursuit of the optimum by these posterior functions.

To locate the new sample placement, conventional improvement-based and optimistic acquisition methods are introduced.

The earliest acquisition function makes the next placement candidate superior to the optimal incumbent $\bar{F}_n^* = \min_{m \le n} \bar{F}(\lambda_{1:m})$ by maximizing the probability of improvement (PI) [36]:

į

$$\Lambda_{n+1} = \arg\max_{\bar{\lambda}} PI(\bar{\lambda}),\tag{11}$$

where

$$PI(\bar{\lambda}) := P(\bar{F}(\bar{\lambda}) < \bar{F}_n^*)$$
$$= \Phi(\frac{m(\bar{\lambda}) - \bar{F}_n^*}{\sigma_n(\bar{\lambda})}).$$

The standard normal cumulative distribution function is denoted by $\Phi(\cdot)$. The posterior distribution of $\overline{F}(\overline{\lambda})$ is expressed as in (9). The new point λ_{n+1} is intended to have a high probability of being larger than the optimal incumbent \overline{F}_n^* , which will miss the point with larger gain but lower certainty.

Expected improvement (EI) [26,37] considers both the probability of improvement and the quantity of improvement. Suppose that we compute one of the remaining points λ_{n+1} and the corresponding values $\bar{F}(\lambda_{n+1})$ in the subsequent iterations; the optimal function value is either \bar{F}_n^* or $\bar{F}(\lambda_{n+1})$. If this quantity $[\bar{F}_n^* - \bar{F}(\lambda_{n+1})]$ is positive, the improvement in the best observed point is $[\bar{F}_n^* - \bar{F}(\lambda_{n+1})]$; if not, it is 0. This improvement could be expressed more succinctly as $[\bar{F}_n^* - \bar{F}(\lambda_{n+1})]^+$, where $a^+ := \max(a; 0)$ represents the positive part.

Since $\bar{F}(\lambda_{n+1})$ is unknown, we can maximize the expected value of the improvement to make both this improvement $[\bar{F}_n^* - \bar{F}(\lambda_{n+1})]^+$ and the possibility $P(\bar{F}(\lambda_{n+1}) < \bar{F}_n^*)$ large in the next point λ_{n+1} :

$$\lambda_{n+1} = \arg\max_{\bar{\lambda}} EI_n(\bar{\lambda}),\tag{12}$$

where

$$EI_{n}(\bar{\lambda}) := \mathbb{E}_{n}\left[\left[\bar{F}_{n}^{*} - \bar{F}(\bar{\lambda})\right]^{+}\right] \\ = \begin{cases} \sigma_{n}(\bar{\lambda})\phi(\frac{\Delta_{n}(\bar{\lambda})}{\sigma_{n}(\bar{\lambda})}) + \Delta_{n}(\bar{\lambda})\Phi(\frac{\Delta_{n}(\bar{\lambda})}{\sigma_{n}(\bar{\lambda})}), & \text{if } \sigma_{n}(\bar{r}) > 0, \\ 0, & \text{if } \sigma_{n}(\bar{\lambda}) = 0, \end{cases}$$
(13)

denotes the expectation provided under the posterior distribution given observations \mathbb{D}_n . Equation (13) is from [38]. If $\Delta_n(\bar{\lambda}) = m(\bar{\lambda}) - \bar{F}_n^*$ is the expected difference between the mean of the new point \bar{r} and the previous best \bar{F}_n^* , then $\mathbb{E}_n[[\bar{F}_n^* - \bar{F}(\bar{r})]^+]$ is the expected value of improvement. The standard normal probability density function is denoted by $\phi(\cdot)$.

The lower confidence bound (LCB) [39] strategy is widely applied in the field of multiarmed bandit [40]. Since the remaining points obey a Gaussian distribution $N(m(\bar{\lambda}), \sigma^2(\bar{\lambda}))$ and we want to find the minimum value of the objective functions in (3), the confidence lower bound can be expressed as

$$LCB(\bar{\lambda}) := m(\bar{\lambda}) - \beta \sigma(\bar{\lambda}).$$

To balance the mean and variance, the hyperparameter β is employed. Choose the next sampling point by minimizing $LCB(\bar{\lambda})$:

$$\lambda_{n+1} = \arg\min_{\bar{\lambda}} LCB(\bar{\lambda}). \tag{14}$$

A formal statement of Bayesian optimization to solve the microphone array placement location problem based on the design broadband beamformer is given in Algorithm 1.

Algorithm 1 Bayesian optimization method for microphone array placement design

Initial step. Select sensor location sample $r_1, r_2, ..., r_n$, calculate objective function $F(r_1), F(r_1), ..., F(r_n)$, prior mean value $m(\lambda) = 0$ and set t = n. **S1.** Choose prior hyperparameters $\theta =: \{\sigma_f, l, \sigma\}$ by MLE in (8) and $\overline{F}(r_1), \overline{F}(r_2), ..., \overline{F}(r_t)$ can be obtained.

S2. The resulting prior distribution on $\overline{F}(\lambda_1), \overline{F}(\lambda_2), \ldots, \overline{F}(\lambda_t)$ is

$$\overline{F}(\boldsymbol{\lambda}_{1:t}) \sim \mathcal{N}(m_0(\boldsymbol{\lambda}_{1:t}), \Sigma_0(\boldsymbol{\lambda}_{1:t}, \boldsymbol{\lambda}_{1:t})).$$

S3. Find the current optimal array placement r^* corresponding to the $\bar{F}_t^* = \min_{s \le t} \bar{F}(r_{1:t})$.

S4. Choose r_{t+1} as the next sample point by finding the optimal value of optimization problem (11), (12) or (14) using conditional distribution (9).

S5. Add $\{r_{t+1}, \overline{F}(r_{t+1})\}$ to the known sensor location sample, set t = t + 1. **S6.** Repeat step 1,2 3,4 and 5 until convergence.

4. Numerical Experiment

To demonstrate the algorithm's performance, the microphone array placement design issue in different dimensions is provided. Convex optimization subproblems are solved using the quadprog function in Matlab, and Bayesian optimization pocket GpyOpt is employed in Python 3.7. All codes are performed on a laptop with Intel(R), Core(TM) i5 CPU, and 2.42 GHz.

In the following example, the desired response function is defined throughout an area that would be suitable for a hands-free or multimedia mobile phone application in the passband area:

$$G_d(s,\lambda,f) = e^{-j2\pi f\left(\frac{\|s-r_{\ell}\|}{c} + \frac{L-1}{2}T\right)},\tag{15}$$

where $r_c = \sum_{i=1}^{M} r_i$ is the center position of all placement variables λ and c = 340.9 m/s is the speed of sound in air. Equation (15) is from [6]. We put $G_d(\lambda, s, f) = 0$ in the stopband to remove the interference and background noise. The minimum distance between two distinct elements is $\bar{\varepsilon}_d = 0.015^2$, $\rho(s, f) = 1$, and $f_s = 8$ khz; maximum frequency is selected as 4 kHz. A performance limit, which is the logarithmic value of observations $\bar{F}(\lambda^*)$ under the optimal microphone array design λ^* , is provided to represent the differences among various microphone array configurations:

$$PLIM =: 10 \log F(\lambda^*).$$

4.1. The 2D Microphone Array Placement Design Problem

A two-dimensional microphone array configuration problem is considered firstly. Both the passband and the stopband are specified on the plane z = 0, where the speaker is located. The microphones are located on plane z = 1 (see Figure 2). Discussion and comparison of microphone arrays containing nine elements follow.

These are the specific region definitions:

$$\Omega_p = \{(s, f) | \| (x, y) \| \le 0.4 \,\mathrm{m}, z = 0 \,\mathrm{m}, 0.5 \,\mathrm{kHz} \le f \le 1.5 \,\mathrm{kHz} \},\$$

and

$$\begin{aligned} \Omega_{s} &= \{ (s, f) | \| (x, y) \| \leq 0.4 \text{ m}, z = 0 \text{ m}, 2.0 \text{ kHz} \leq f \leq 4 \text{ kHz} \} \\ &\cup \{ (s, f) | 1.8 \text{ m} \leq \| (x, y) \| \leq 3.0 \text{ m}, z = 0 \text{ m}, 0.5 \text{ kHz} \leq f \leq 1.5 \text{ kHz} \} \\ &\cup \{ (s, f) | 1.8 \text{ m} \leq \| (x, y) \| \leq 3.0 \text{ m}, z = 0 \text{ m}, 2.0 \text{ kHz} \leq f \leq 4.0 \text{ kHz} \}, \end{aligned}$$

and the placement feasible region is

$$\Lambda = \{ \lambda | |x| \le 1.5 \,\mathrm{m}, |y| \le 1.5 \,\mathrm{m}, z = 1 \,\mathrm{m} \}.$$



The discretization of $\Omega = \Omega_p \cup \Omega_s$ is applied: 60 points are generated for each frequency domain area and 0.2 m for each spatial domain region.

Figure 2. The setup of the 2D microphone array placement design problem.

The performance and CPU time (measured in seconds) for the proposed Bayesian optimization method and the hybrid descent method applied to microphone arrays with nine elements are displayed in Table 1. In this table, the Bayesian optimization method consists mainly of a GP and different acquisition functions. It is evident that the proposed algorithm could considerably boost the speed of computing while achieving the same broadband beamformer performance as the hybrid descent approach [9]. The proposed algorithm reaches the optimal performance in only 1518 s, which is more than four times faster than the hybrid descent method. Moreover, an optimal set of microphone arrays can effectively improve the performance of the beamformer compared with linear placement. In Table 1, we also show the average stopband gain G_s at f = 1400 Hz and filter length L = 50. They illustrate that the noise in the stopband region is better suppressed with an optimal set of microphone array configurations compared with linear placement.

	GP-PI	GP-EI	GP-LCB	GA-Gradient *	Linear
CPU time (s)	1255	1518	1375	6407	-
PLIM (dB)	-38.5996	-39.1635	-38.1337	-39.1635	-19.4577
G_s (dB)	-43.5372	-45.6333	-42.0707	-45.6333	-25.0743

Table 1. Summary of beamformer performance with different array placement design.

* GA-gradient stands for hybrid descent method.

By applying the proposed algorithm and hybrid descent method, the optimal placement r^* is shown in Figure 3 below. Because the optimization problem in (3) is non-convex, as can be seen from these figures, the microphone array configuration can vary significantly. Figure 4 shows the beamformer's performance for the optimal microphone array configuration r^* in the (x, y)-plane at 1400 Hz and in the (x, f)-plane at y = 0 for a filter of finite length L = 50 to illustrate the impact of the beamformer, where *G* denotes the beamformer's output (Equation (1)) over the whole targeted region.



Figure 3. Optimal microphone array placement in the 2D plane with different methods, the blue stars represent the position of each microphone. (a) Bayesian optimization method based on GP regression and PI acquisition function. (b) Bayesian optimization method based on GP regression and EI acquisition function. (c) Bayesian optimization method based on GP regression and LCB acquisition function. (d) Hybrid descent method based on GA and gradient descent algorithm.



Figure 4. Performance under the optimal 2D configuration. (a) Beamformer output response in the (x, y)-plane at 1400 Hz with filter length L = 50. (b) Beamformer output response in the (x, f)-plane at y = 0 with filter length L = 50.

It is worth noting that a failure of one of the microphones in the optimal array configuration will not invalidate the beamformer, but it will degrade the performance. The stopband gain G_s is -45.6333 dB in the optimal microphone array. If a microphone in the array fails randomly, stopband gain G_s reduces to -31.7916 dB.

4.2. The 3D Microphone Array Placement Problem

In the next example, a 3D microphone array placement design problem is considered. The microphone elements are selected in a solid that is 0.5 m away from the desired cubic region for beamforming, as shown in Figure 5 below.

These are the specific region definitions:

$$\Omega_p = \{(s, f) | \| (x, y) \| \le 0.4 \text{ m}, 1.5 \text{ m} \le |z| \le 2 \text{ m}, 0.5 \text{ kHz} \le f \le 1.5 \text{ kHz} \},\$$

and

$$\begin{split} \Omega_{\rm s} &= \{(s,f) | \| (x,y) \| \leq 0.4 \text{ m}, 1.5 \text{ m} \leq |z| \leq 2 \text{ m}, 2.0 \text{ kHz} \leq f \leq 4 \text{ kHz} \} \\ &\cup \{(s,f) | \| (x,y) \| \geq 1.8 \text{ m}, |x|, |y| \leq 3.0 \text{ m}, 1.5 \text{ m} \leq |z| \leq 2 \text{ m}, 0.5 \text{ kHz} \leq f \leq 1.5 \text{ kHz} \} \\ &\cup \{(s,f) | 1.8 \text{ m} \leq \| (x,y) \| \leq 3.0 \text{ m}, 1.5 \text{ m} \leq |z| \leq 2 \text{ m}, 2.0 \text{ kHz} \leq f \leq 4.0 \text{ kHz} \}, \end{split}$$

and the placement feasible region is

$$\Lambda = \{\lambda | |x| \le 1.5 \text{ m}, |y| \le 1.5 \text{ m}, 1.5 \text{ m} \le |z| \le 2 \text{ m}\}.$$



Figure 5. The setup of the 3D microphone array placement design problem.

The frequency and spatial domain areas are divided in a way that is similar to the 2D microphone array situation. Table 2 displays that the computational efficiency of the proposed algorithm can be increased by almost five times and that the performance of the beamformer is better than the hybrid descent method. Moreover, an optimal set of microphone arrays can effectively improve the performance of the beamformer compared with linear placement. The average stopband gain G_s at 1400 Hz and z = 1.6 m in Table 2 demonstrates that the noise at the stopband is suppressed well with the optimal array placement.

Table 2. Summary of beamformer performance with different array placement designs.

	GP-PI	GP-EI	GP-LCB	GA-Gradient *	Linear		
CPU time(s)	1718	1597	1535	8340	-		
PLIM(dB)	-36.7066	-36.7088	-36.7082	-36.6520	-16.9972		
$G_s(dB)$	-61.2713	-62.7589	-63.1721	-61.7405	-27.4568		
* CA gradient stands for hybrid descent method							

* GA-gradient stands for hybrid descent method.

We primarily display in Figure 6 where the microphone arrays are located. It can be seen that the microphone placements are scattered evenly over the middle of the rectangle's feasible domain, with a higher density on the side that is the furthest from the sound source.

Figure 7 shows the beamformer's performance for the ideal set of microphone arrays in the (x, y)-plane at 1400 Hz, z = 1.6 m and in the (x, z)-plane at 1400 Hz, y = 0 for a filter of finite length L = 50.

The stopband gain G_S that can be achieved with an optimal microphone array placement is -63.1721 dB. If an element in the microphone array fails, the stopband gain G_S would reduce to -51.5854 dB without being completely ineffective.

In fact, we design 2D and 3D examples based on different scenarios. The 2D case is also a sub-example of the 3D one. In the 2D case, the microphone array is placed at the same level, but the microphones can be positioned at different levels in the 3D case. The 3D case is a bit more complex, and solving the optimization problem demands more computational time. In addition, better suppression of noise in the stopband is achieved in the 3D case.



Figure 6. Optimal microphone array placement in the 3D plane with different methods, the blue stars represent the position of each microphone. (**a**) Bayesian optimization method based on GP regression and PI acquisition function. (**b**) Bayesian optimization method based on GP regression and EI acquisition function. (**c**) Bayesian optimization method based on GP regression and LCB acquisition function. (**d**) Hybrid descent method based on GA and gradient descent algorithm.



Figure 7. Performance under the optimal 3D configuration. (a) Beamformer output response in the (x, y)-plane at 1400 Hz, z = 1.6 m, with filter length L = 50. (b) Beamformer output response in the (x, z)-plane at 1400 Hz, y = 0, with filter length L = 50.

5. Conclusions

In this paper, the Bayesian optimization method has been employed to solve the microphone array configuration design problem to enhance beamformer performance. Since the configuration design problem is non-convex and highly non-linear and the objective function is time-consuming to calculate, GP has been used as a surrogate function to approximate the objective function. The acquisition function guided the iterations toward the optimal set of microphone array placements in the sense of probability, and different acquisition functions were used for comparison. Numerical experiments have demonstrated that the proposed Bayesian optimization method finds similar or better microphone array configuration more efficiently. The proposed Bayesian optimization method is at least four times faster than the hybrid descent method to find the optimal placement from the numerical results. Therefore, the method is a competitive approach to design microphone placements when short time is required. As a future extension, it is interesting to consider alternative probabilistic agent models in Bayesian optimization to approximate the complicated objective function. Also, more advanced versions of acquisition functions could be considered. Optimal microphone arrays can be realized in many commercial products such as receivers in smart home systems and multi-function classrooms.

Author Contributions: Conceptualization, Z.L. and K.F.C.Y.; methodology, Z.L. and K.F.C.Y.; software, Y.Z.; validation, Y.Z., Z.L. and K.F.C.Y.; formal analysis, Y.Z.; investigation, Y.Z.; resources, Z.L.; data curation, Y.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, K.F.C.Y. and Z.L.; visualization, Y.Z.; supervision, Z.L. and K.F.C.Y.; project administration, Z.L. and K.F.C.Y.; funding acquisition, Z.L. and K.F.C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This paper is supported by RGC Grant 15203923 and PolyU Grant (1-WZ0E, 4-ZZPT), the Natural Science Foundation of China (12271526), the Natural Science Foundation of Hunan Province (2022JJ30675), and the Natural Science Foundation of Changsha (kq2202068).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- 1. Winters, J.H. Smart antennas for wireless systems. *IEEE Pers. Commun.* **1998**, *5*, 23–27. [CrossRef]
- 2. Benesty, J.; Chen, J.; Huang, Y. Microphone Array Signal Processing; Springer: Berlin, Germany, 2008.
- 3. Gannot, S.; Burshtein, D.; Weinstein, E. Signal enhancement using beamforming and nonstationarity with applications to speech. *IEEE Trans. Signal Process.* **2001**, *49*, 1614–1626. [CrossRef]
- 4. Brandstein, M.; Ward, D. Microphone Arrays: Signal Processing Techniques and Applications; Springer: Berlin, Germany, 2001.
- 5. Frost, O.L. An algorithm for linearly constrained adaptive array processing. Proc. IEEE 1972, 60, 926–935. [CrossRef]
- Feng, Z.G.; Yiu, K.F.C.; Nordholm, S.E. Performance limit of broadband beamformer designs in space and frequency. J. Optim. Theory Appl. 2015, 164, 316–341. [CrossRef]
- Khalid, L.; Nordholm, S.E.; Dam, H.H. Design study on microphone arrays. In Proceedings of the 2015 IEEE International Conference on Digital Signal Processing (DSP), Singapore, 21–24 July 2015; pp. 1171–1175.
- Feng, Z.G.; Yiu, K.F.C.; Nordholm, S.E. Placement design of microphone arrays in near-field broadband beamformers. *IEEE Trans. Signal Process.* 2011, 60, 1195–1204. [CrossRef]
- 9. Li, Z.B.; Yiu, K.F.C.; Feng, Z.G. A hybrid descent method with genetic algorithm for microphone array placement design. *Appl. Soft Comput.* **2013**, *13*, 1486–1490. [CrossRef]
- Oliveri, G.; Donelli, M.; Massa, A. Linear array thinning exploiting almost difference sets. *IEEE Trans. Antennas Propag.* 2009, 57, 3800–3812. [CrossRef]
- 11. Rocca, P.; Haupt, R.L. Dynamic array thinning for adaptive interference cancellation. In Proceedings of the Fourth European Conference on Antennas and Propagation, Barcelona, Spain, 12–16 April 2010; pp. 1–3.
- 12. Fernandez-Delgado, M.; Rodriguez-Gonzalez, J.; Iglesias, R.; Barro, S.; Ares-Pena, F. Fast array thinning using global optimization methods. *J. Electromagn. Waves Appl.* 2010, 24, 2259–2271. [CrossRef]
- 13. Chen, K.; Yun, X.; He, Z.; Han, C. Synthesis of sparse planar arrays using modified real genetic algorithm. *IEEE Trans. Antennas Propag.* **2007**, *55*, 1067–1073. [CrossRef]
- 14. Khatami, I.; Jamalabadi, M.Y.A. Optimal design of microphone array in a planar circular configuration by genetic algorithm enhanced beamforming. *J. Therm. Anal. Calorim.* **2021**, *145*, 1817–1825. [CrossRef]
- Macho-Pedroso, R.; Domingo-Perez, F.; Velasco, J.; Losada-Gutierrez, C.; Macias-Guarasa, J. Optimal microphone placement for indoor acoustic localization using evolutionary optimization. In Proceedings of the 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Alcala de Henares, Spain, 4–7 October 2016; pp. 1–8.
- 16. Yu, J.; Donohue, K.D. Optimal irregular microphone distributions with enhanced beamforming performance in immersive environments. *J. Acoust. Soc. Am.* **2013**, *134*, 2066–2077. [CrossRef]
- 17. Trucco, A. Weighting and thinning wide-band arrays by simulated annealing. Ultrasonics 2002, 40, 485–489. [CrossRef] [PubMed]
- 18. Razavi, A.; Forooraghi, K. Thinned arrays using pattern search algorithms. *Prog. Electromagn. Res.* 2008, 78, 61–71. [CrossRef]
- 19. Malgoezar, A.; Snellen, M.; Sijtsma, P.; Simons, D. Improving beamforming by optimization of acoustic array microphone positions. In Proceedings of the 6th Berlin Beamforming Conference, Berlin, Germany, 29 February–1 March 2016; p. 5.
- Hawes, M.B.; Liu, W. Sparse microphone array design for wideband beamforming. In Proceedings of the 2013 18th International Conference on Digital Signal Processing (DSP), Fira, Santorini, Greece, 1–3 July 2013; pp. 1–5.
- Chan, K.Y.; Yiu, C.K.F.; Nordholm, S. Microphone configuration for beamformer design using the Taguchi method. *Measurement* 2017, 96, 58–66. [CrossRef]
- 22. Frazier, P.I. A tutorial on Bayesian optimization. arXiv 2018, arXiv:1807.02811.
- 23. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; Freitas, N.D. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* 2015, 104, 148–175. [CrossRef]
- 24. Snoek, J.; Larochelle, H.; Adams, R.P. Practical bayesian optimization of machine learning algorithms. *Proc. Adv. Neural Inf. Process. Syst.* **2012**, *25*, 2960–2968.
- 25. Mockus, J.B.; Mockus, L.J. Bayesian approach to global optimization and application to multiobjective and constrained problems. *J. Optim. Theory Appl.* **1991**, *70*, 157–172. [CrossRef]
- 26. Jones, D.R.; Schonlau, M.; Welch, W.J. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* **1998**, 13, 455–492. [CrossRef]
- 27. Forrester, A.; Sobester, A.; Keane, A. Engineering Design via Surrogate Modelling: A Practical Guide; John Wiley & Sons: New York, NY, USA, 2008.
- 28. Frazier, P.I.; Wang, J. Bayesian optimization for materials design. Proc. Inf. Sci. Mater. Discov. Des. 2015, 225, 45–75.
- 29. Packwood, D. Bayesian Optimization for Materials Science; Springer: Berlin, Germany, 2017.
- 30. Williams, C.; Rasmussen, C. Gaussian Processes for Machine Learning; MIT Press: Cambridge, MA, USA, 2006; Volume 2.
- 31. Mockus, J. Application of Bayesian approach to numerical methods of global and stochastic optimization. *J. Glob. Optim.* **1994**, 4, 347–365. [CrossRef]
- 32. Schulz, E.; Speekenbrink, M.; Krause, A. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *J. Math. Psychol.* **2018**, *85*, 1–16. [CrossRef]
- Lu, C.C.; Tang, X.O. Surpassing human-level face verification performance on LFW with GaussianFace. In Proceedings of the 29th AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015; Volume 29, pp. 1–9.
- 34. Neal, R.M. Bayesian Learning for Neural Networks; Springer Science & Business Media: New York, NY, USA, 2012.

- 36. Kushner, H.J. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *J. Fluids Eng.* **1964**, *86*, 97–106. [CrossRef]
- 37. Mockus, J.; Tiesis, V.; Zilinskas, A. The application of Bayesian methods for seeking the extremum. *Towards Glob. Optim.* **1978**, 2, 117–129.
- Brochu, E.; Cora, V.M.; De, F.N. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv* 2010, arXiv:1012.2599.
- 39. Srinivas, N.; Krause, A.; Kakade, S.M.; Seeger, M. Gaussian process optimization in the bandit setting: No regret and experimental design. In Proceedings of the 27th International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010; pp. 1015–1022.
- 40. Lai, T.L.; Robbins, H. Asymptotically efficient adaptive allocation rules. Adv. Appl. Math. 1985, 6, 4–22. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.