

Article

Arbitrary Microphone Array Optimization Method Based on TDOA for Specific Localization Scenarios

Haitao Liu ^{1,*†‡}, Thia Kirubarajan ^{2‡} and Qian Xiao ¹

¹ School of Mechatronics and Vehicle Engineering, East China Jiaotong University, Nanchang 330013, China; jxralph@ecjtu.edu.cn

² Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada; kiruba@mcmaster.ca

* Correspondence: 2860@ecjtu.edu.cn; Tel.: +86-0791-8704-6135

† Building 13, south district, 808 Shuanggang east street, Changbei economic development zone, Nanchang 330013, Jiangxi province, China.

‡ These authors contributed equally to this work.

Received: 16 August 2019; Accepted: 4 October 2019; Published: 7 October 2019

Abstract: Various microphone array geometries (e.g., linear, circular, square, cubic, spherical, etc.) have been used to improve the positioning accuracy of sound source localization. However, whether these array structures are optimal for various specific localization scenarios is still a subject of debate. This paper addresses a microphone array optimization method for sound source localization based on TDOA (time difference of arrival). The geometric structure of the microphone array is established in parametric form. A triangulation method with TDOA was used to build the spatial sound source location model, which consists of a group of nonlinear multivariate equations. Through reasonable transformation, the nonlinear multivariate equations can be converted to a group of linear equations that can be approximately solved by the weighted least square method. Then, an optimization model based on particle swarm optimization (PSO) algorithm was constructed to optimize the geometric parameters of the microphone array under different localization scenarios combined with the spatial sound source localization model. In the optimization model, a reasonable fitness evaluation function is established which can comprehensively consider the positioning accuracy and robustness of the microphone array. In order to verify the array optimization method, two specific localization scenarios and two array optimization strategies for each localization scenario were constructed. The optimal array structure parameters were obtained through numerical iteration simulation. The localization performance of the optimal array structures obtained by the method proposed in this paper was compared with the optimal structures proposed in the literature as well as with random array structures. The simulation results show that the optimized array structure gave better positioning accuracy and robustness under both specific localization scenarios. The optimization model proposed could solve the problem of array geometric structure design based on TDOA and could achieve the customization of microphone array structures under different specific localization scenarios.

Keywords: source localization; TDOA; array optimization; PSO

1. Introduction

In the past two decades, microphone array technology has consistently been a hot research field. Microphone arrays are mainly used for sound source localization and identification, and have been an important practical technology with many valuable applications, such as noise source localization [1,2], target sound source tracking [3], teleconferencing systems [4,5], intelligent robots [6–8], and so on.

In microphone array technology, there are three main methods for sound source localization, namely, beamforming, acoustic holography, and time difference of arrival (TDOA). The beamforming

method applies delay-and-sum to signals from an array of microphones, and in the direction of the source, a beam peak will form to locate the sound sources [9]. Acoustic holography reconstructs the acoustic fields to locate the sound sources by solving the inverse propagation problems [10]. Beamforming and acoustic holography methods usually involve planar microphone arrays and calculation points located on a surface at a certain distance with respect to the array, which provides a poor resolution in the direction perpendicular to the array. In recent years, beamforming with several deconvolution techniques [11] and inverse methods with additional issues [12,13] have been proposed to construct volumetric sound source imaging, which can give the exact three-dimensional (3D) coordinates of sound sources. The method based on TDOA, virtually a triangulation method, locates the sound source using geometric relationships between microphones and sound sources, which can give the spatial position of sound sources with reasonable accuracy using a small number of sensors [14]. TDOA methods have been widely used for real-time sound source localization [15,16]. Moreover, in some sound source localization scenarios, such as simple sound source tracking, TDOA methods show better application prospects.

Different numbers of microphones and different kinds of array structures are used in these three methods. In general, the number of microphones used in beamforming and acoustic holography is much larger than in the TDOA method because the number of microphones has a significant influence on the reconstruction accuracy of the sound source mapping [17]. Nevertheless, the number of microphones required in the TDOA method is much smaller. In theory, only four microphones are needed to locate a sound source in three-dimensional space. For example, Wu and Zhu [15] used only four microphones to locate arbitrarily time-dependent acoustic sources in a free three-dimensional space in real-time. In addition, the number of microphones is not the decisive factor of its location accuracy. The array structure is another main factor that relates to the accuracy of source localization for these three methods.

Many kinds of microphone array structures are applied in sound source localization, which can mainly be divided into three categories: 1-dimensional, 2-dimensional, and 3-dimensional. The 1-dimensional array structure group mainly comprises binaural arrays [18] and linear arrays [19]. 2-dimensional array structures include square [20], cross [21], spiral [22,23], and circular geometries [24]. 3-dimensional arrays mainly include cubic [25], pyramidal [25], hemispherical [26], and spherical [27] geometries. Relevant scholars have analyzed the performance of various arrays, indicating that each kind of array structure is only suitable for specific localization algorithms and scenarios. There is no array structure that can achieve good localization performance under any kinds of scenarios and algorithms. For example, the two- or three- dimensional localization accuracy of a randomly distributed array will vary widely with respect to the relative position of the sound source [28]. Therefore, the optimization of the microphone array structure becomes an important research point. Wang and Bei [29] proposed an optimization method based on acoustic holography theory to optimize the microphone array coordinates on a fixed cross X-type array structure, and the main side lobe ratio and the main lobe area were selected as the optimization objective function. Kodrasi et al. [30] adopted different heuristic optimization approaches and an exhaustive search approach to optimize the microphone positions for an arbitrary planar array based on the beamforming method, and Kodrasi's methods found near-optimal configurations. Recently, Yan and Ma [31], Sarradj [32], Bjelić et al. [33], Teng and Lv [34], and Le Courtois et al. [35] also proposed new methods for planar array optimization based on the beamforming method, and compared the array performance under different localization scenarios. In the optimization procedure, the main-lobe width and side-lobe level are generally selected as the optimization objective function. Padois et al. [36] proposed a spherical microphone array with polyhedral discretization and compared it with a spherical array with a slightly different geometry based on the beamforming method. The results showed that the polyhedral discretization array could obtain better positioning accuracy. In 2019, Padois et al. [37] carried out further research on array geometry optimization based on time-domain beamforming. They proposed an optimal spherical microphone array geometry using a nonlinear optimization.

Numerical and experimental results showed that the optimized geometry improved the sound source mapping.

From the above, it can be seen that a great deal of research work has been done in the field of microphone array optimization for sound source localization. However, these array optimization methods are mainly based on the beamforming and acoustic holography methods. The optimization procedure is also usually based on existing array structures, such as cross array [29], circle array [31], spiral array [32], irregular planar array [33,34], spherical array [38], and so on. As such, certain constraints for the array structure have been introduced to the optimization. A pre-constrained array structure may lead to a local optimum, which may not be suitable for certain specific localization scenarios. In addition, besides the array optimization based on beamforming and acoustic holography, the research on array optimization based on TDOA is relatively rare. Zietlow et al. [39] established a simulation model based on TDOA to compare the source positioning accuracy of different microphone arrangements. The microphone arrays consist of eight microphones in three different arrangements, namely cube, twisted cube, and random. These array arrangements were fixed, and no optimization was performed for the array structures. Hu et al. [40] proposed an analytical method based on TDOA to optimize microphone array structure, which could guarantee that the sound source localization had the same performance in all directions for omni-directional estimation. However, the optimal result led to a set of nonlinear equations which could not give deterministic analytical solutions. With additional constraints, only a particular solution in a regular polyhedron form can be obtained. Further, only five kinds of array structures with a specified number of microphones belong to the solution of regular polyhedron form, including the tetrahedron (5 microphones), the hexahedron (9 microphones), the octahedron (7 microphones), the dodecahedron (21 microphones), and the icosahedron (13 microphones). The limited solution of array structures restricts the practical application of Hu's method. Meanwhile, due to some constraints in modeling and solving the method, these five kinds of array structures may not give the best positioning results under some specific localization scenarios, such as the scenarios with the asymmetrical distribution of sound sources. Therefore, more in-depth research needs to be carried out in the field of array structure optimization for sound source localization based on TDOA.

This paper is devoted to an arbitrary microphone arrays optimization method for sound source localization based on TDOA. The method proposed is a numerical approach based on the particle swarm optimization (PSO) algorithm, which can optimize the array structure of an arbitrary number of microphones under any specific localization scenarios without prior array structure information. Examples of localization scenarios were constructed to obtain the optimal array structures through the proposed method. Additionally the optimal array structures were compared with the array structures proposed by Hu et al. [40] as well as random array structures under the constructed specific scenarios.

This article makes four main contributions. First, a numerical approach of microphone array optimization based on the PSO algorithm for the TDOA method is proposed. Second, the proposed model can perform array structure optimization with an arbitrary number of microphones, and no prior array structure information is introduced into the optimization procedures, which is likely to obtain the more optimal solutions. Third, the array optimization model has general applicability, which can effectively solve the problem of microphone arrangements in sound source localization under different specific localization scenarios. Fourth, the fitness evaluation function constructed in the optimization model can give good consideration to the accuracy and robustness of sound source localization based on TDOA. The two specific localization scenarios established here verify the proposed optimization method.

In the following sections, the optimization model is introduced in detail, and the localization performance is compared with the array structures proposed by Hu et al. as well as with random array structures. Section 2 introduces the construction of the TDOA-based sound source localization model for an arbitrary microphone array, as well as the solution for the localization model. The numerical optimization model based on PSO for an arbitrary array structure is presented in Section 3, in addition

to the optimization procedure. Simulations were performed and their results are discussed in Section 4, followed by conclusions in Section 5.

2. Construction of Localization Model Based on TDOA

The sound source localization model is the basis of the array structure optimization. The TDOA method was used to locate the sound source. Therefore, the time difference and the spatial geometric relationship between the array and the sound source were used to establish the localization model.

2.1. Geometric Structure Parameterization for Arbitrary Microphone Array

In order to optimize the microphone array, the geometric structure of the array needs to be parameterized first. Because the sound source localization method in this paper is based on TDOA, a reference microphone is needed in the microphone array. For convenience, the coordinate of the reference microphone \mathbf{M}_0 is set as $(0, 0, 0)$. Then, the other microphones' coordinates can be expressed by the radial distance l_i , the azimuth angle α_i , and the elevation angle β_i in three-dimensional space, as shown as Figure 1.

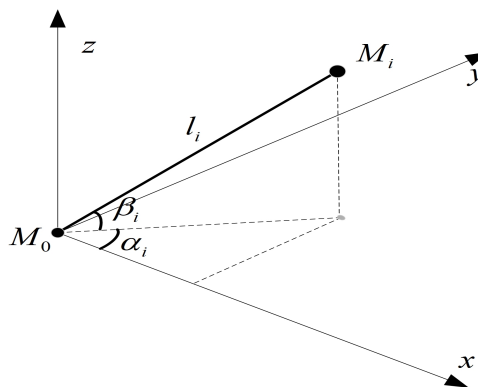


Figure 1. Diagram of the microphones' coordinates.

The coordinates of the other microphones are $\mathbf{M}_i(l_i \cos(\beta_i) \cos(\alpha_i), l_i \cos(\beta_i) \sin(\alpha_i), l_i \sin(\beta_i))$, where $i = 1, 2, 3, \dots, N_m$, N_m stands for the number of microphones except for the reference microphone. Therefore, the optimal parameter of the microphone array is

$$\mathbf{M}_p = \begin{pmatrix} l_1 & \alpha_1 & \beta_1 \\ l_2 & \alpha_2 & \beta_2 \\ \vdots & \vdots & \vdots \\ l_i & \alpha_i & \beta_i \end{pmatrix} \quad i = 1, 2, \dots, N_m. \quad (1)$$

The constraint of the azimuth angle α_i is $[0^\circ, 360^\circ]$ and for the elevation angle it is β_i is $[-90^\circ, 90^\circ]$. The range of radial distance l_i is related to the size of microphone array and the frequency band of sound source, as well as the requirements for actual positioning scenarios. Normally, the lower limit of l_i is the diameter of microphone d_m , and the upper limit of l_i is $c/(2f_m)$, where c is the speed of sound and f_m is the main periodic frequency present in the sources. Then, the set of the optimal search space can be described as

$$\mathbb{O}\mathbb{S} = \{ \mathbf{M}_p | \mathbf{M}_p = \begin{pmatrix} l_1 & \alpha_1 & \beta_1 \\ \vdots & \vdots & \vdots \\ l_i & \alpha_i & \beta_i \end{pmatrix}, l_i \in [d_m, c/2f_m], \alpha_i, \beta_i \in [0^\circ, 360^\circ], i = 1, 2, \dots, N_m \}. \quad (2)$$

2.2. Spatial Source Localization Model Based on TDOA

The spatial source location model was constructed based on the TDOA method, which is a triangulation method. Suppose the coordinate of the sound source is $\mathbf{S} = (x, y, z)$. The mathematical description is shown as follows:

$$\begin{cases} r_0 = \sqrt{(x_0 - x)^2 + (y_0 - y)^2 + (z_0 - z)^2} = \|\mathbf{M}_0 - \mathbf{S}\|, \\ r_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} = \|\mathbf{M}_i - \mathbf{S}\|, \end{cases} \quad (3)$$

where r_0 is the distance between the sound source and the reference microphone \mathbf{M}_0 , and r_i is the distance between the sound source and the other microphones \mathbf{M}_i . $i = 1, 2, \dots, N_m$.

By constructing the time arrival difference from the sound source to the reference microphone and the other microphones, the spatial source location model can be obtained. The model is shown as follows:

$$r_{i,0} = \|\mathbf{M}_i - \mathbf{S}\| - \|\mathbf{M}_0 - \mathbf{S}\| = c\tau_{i,0}, \quad (4)$$

where $\tau_{i,0}$ is the sound arrival time difference between \mathbf{M}_i and \mathbf{M}_0 . $\tau_{i,0}$ can be estimated by the method of cross correlation.

Suppose that $u_i(t)$ and $u_0(t)$ are the acoustic signals acquired by microphones \mathbf{M}_i and \mathbf{M}_0 separately. The cross-correlation function between the two signals is

$$R_{i,0}(\tau) = \int_{-\infty}^{\infty} u_i(t)u_0(t + \tau)dt, \quad (5)$$

$$\tau_{i,0} = \arg \max[R_{i,0}(\tau)]. \quad (6)$$

The spatial source location model consists of a group of nonlinear multivariate equations, which is difficult to solve. An alternative method is to transform the model into a set of linear equations. Spatial distance satisfies the relationship shown in Equation (7):

$$r_i^2 = (r_{i,0} + r_0)^2. \quad (7)$$

Then, Equation (4) can be rewritten as:

$$2r_{i,0}r_0 + 2x_{i,0}x + 2y_{i,0}y + 2z_{i,0}z = K_i - K_0 - r_{i,0}^2, \quad (8)$$

where $K_i = x_i^2 + y_i^2 + z_i^2$, $K_0 = x_0^2 + y_0^2 + z_0^2$, $x_{i,0} = x_i - x_0$, $y_{i,0} = y_i - y_0$, $z_{i,0} = z_i - z_0$.

Equation (8) is a group of linear equations, which can be written in matrix form:

$$\mathbf{A}\mathbf{X} = \mathbf{B} \quad (9)$$

$$\text{where } \mathbf{X} = [x, y, z]^T, \mathbf{A} = - \begin{pmatrix} x_{i,0} & y_{i,0} & z_{i,0} \\ \vdots & \vdots & \vdots \\ x_{N,0} & y_{N,0} & z_{N,0} \end{pmatrix}, \mathbf{B} = \begin{pmatrix} r_{i,0} \\ \vdots \\ r_{N,0} \end{pmatrix} \cdot r_0 + 1/2 \begin{pmatrix} r_{i,0}^2 - K_i + K_0 \\ \vdots \\ r_{N,0}^2 - K_N + K_0 \end{pmatrix}.$$

2.3. Solution for Spatial Source Localization model

When the number of microphones is $N_m = 3$, the spatial source location model (Equation (9)) can be solved directly, as represented in [41]. However, the direct solution method may produce two different answers, which leads to localization ambiguity. Meanwhile, the accuracy and robustness of source localization are not very good under the condition of $N_m = 3$. Adding redundant sensors can effectively improve the performance of source localization. When there are more than four microphones ($N_m \geq 4$), the system is overdetermined, as the number of measurements is greater than the number of

unknowns. The LS (least-square) method can be used to solve the overdetermined linear equations. Chan and Ho [42] proposed an alternative solution algorithm in closed-form, valid for both distant and close sources, which used twice-weighted LS to give the localization results. Chan's method gives an explicit solution with reasonable accuracy and is non-iterative with low computational complexity. Therefore, Chan's method is more suitable for the optimization calculation of acoustic array structure in this paper.

2.3.1. The First Weighted Least-Square Solution Process

In order to solve the source localization model by least-square method, Equation (9) should be rewritten to construct an error vector. Because of noise in the TDOA estimation, the error vector can be derived as:

$$\boldsymbol{\psi} = \mathbf{h} - \mathbf{G}_a \mathbf{z}_a^0, \quad (10)$$

where $\mathbf{z}_a = [x, y, z, r_0]^T$ is the unknown vector. $\mathbf{h} = 1/2 \begin{pmatrix} r_{1,0}^2 - K_1 + K_0 \\ \vdots \\ r_{N,0}^2 - K_N + K_0 \end{pmatrix}$. $\mathbf{G}_a = \begin{pmatrix} x_{1,0} & y_{1,0} & z_{1,0} & r_{1,0} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N,0} & y_{N,0} & z_{N,0} & r_{i,0} \end{pmatrix}$. $(\cdot)^2$ stands for the expectations of variables without noise.

Suppose that the noise of TDOA estimation is n_i . $\tau_{i,0} = \tau_{i,0}^0 + n_{i,0}$. $r_{i,0} = r_{i,0}^0 + cn_{i,0}$, and $r_i^0 = r_{i,0}^0 + r_0^0$. Then, $\boldsymbol{\psi}$ can be expressed as

$$\boldsymbol{\psi} = c\mathbf{R}\mathbf{n} + 0.5c^2\mathbf{n} \odot \mathbf{n}, \quad (11)$$

where $\mathbf{R} = \text{diag}\{r_1^0, r_2^0, \dots, r_N^0\}$, $\mathbf{n} = [n_{1,0}, n_{2,0}, \dots, n_{N,0}]^T$. The symbol \odot stands for the Schur product. The noise vector \mathbf{n} approximately obeys a Gaussian normal distribution.

In practice, the condition $r_i^0 \gg r_{i,0} = cn_{i,0}$ is usually satisfied. Therefore, the second term on the right hand side of Equation (11) can be ignored. Then, the covariance matrix of $\boldsymbol{\psi}$ can be given as:

$$\boldsymbol{\Psi} = \mathbf{E}[\boldsymbol{\psi}\boldsymbol{\psi}^T] = c^2\mathbf{R}\mathbf{Q}\mathbf{R}, \quad (12)$$

where $\mathbf{Q} = \mathbf{E}(\mathbf{n}\mathbf{n}^T) = \text{Cov}(\mathbf{n})$. Then, the first weighted LS method is used to solve Equation (10):

$$\mathbf{z}_a = (\mathbf{G}_a^T \boldsymbol{\Psi}^{-1} \mathbf{G}_a)^{-1} \mathbf{G}_a^T \boldsymbol{\Psi}^{-1} \mathbf{h}. \quad (13)$$

When the source is far from the array, each r_i^0 is close to r_0^0 , so $\mathbf{R} \approx r_0^0 \mathbf{I}$. Then, an approximate solution of Equation (13) is

$$\mathbf{z}_a \approx (\mathbf{G}_a^T \mathbf{Q}^{-1} \mathbf{G}_a)^{-1} \mathbf{G}_a^T \mathbf{Q}^{-1} \mathbf{h}. \quad (14)$$

When the source is close to the array, Equation (14) can be firstly used to obtain an initial solution to estimate \mathbf{R} , which can be substituted into Equations (12) and (13) to get a more accurate result.

2.3.2. The Second Weighted Least-Square Solution Process

The above solution of \mathbf{z}_a assumes that x , y , and r_0 are independent. However, r_0 is related to the source location. In order to incorporate this relationship to give an improved estimate, \mathbf{z}_a can be expressed as

$$z_{a,1} = x^0 + e_1, z_{a,2} = y^0 + e_2, z_{a,3} = z^0 + e_3, z_{a,4} = r_0^0 + e_4, \quad (15)$$

where e_1, e_2, e_3 , and e_4 are estimation errors of \mathbf{z}_a . (x^0, y^0, z^0) are the coordinates of the real sources. Then, a new error vector $\boldsymbol{\psi}'$ can be obtained as:

$$\psi' = \mathbf{h}' - \mathbf{G}'_a \mathbf{z}'_a, \quad (16)$$

$$\text{where } \mathbf{h}' = 1/2 \begin{pmatrix} (z_{a,1} - x_0)^2 \\ (z_{a,2} - y_0)^2 \\ (z_{a,3} - z_0)^2 \\ z_{a,4}^2 \end{pmatrix}, \mathbf{G}'_a = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \mathbf{z}'_a = \begin{pmatrix} (x - x_1)^2 \\ (y - y_1)^2 \\ (z - z_1)^2 \end{pmatrix}.$$

Substitute Equation (15) into Equation (16):

$$\psi' = \begin{pmatrix} 2(x^0 - x_0)e_1 + e_1^2 \\ 2(y^0 - y_0)e_2 + e_2^2 \\ 2(z^0 - z_0)e_3 + e_3^2 \\ 2r_0^0 e_4 + e_4^2 \end{pmatrix} \approx \begin{pmatrix} 2(x^0 - x_0)e_1 \\ 2(y^0 - y_0)e_2 \\ 2(x^0 - z_0)e_3 \\ 2r_0^0 e_4 + e_4^2 \end{pmatrix}. \quad (17)$$

The covariance matrix of ψ' is

$$\Psi' = \mathbf{E}[\psi' \psi'^T] = 4\mathbf{R}' \text{cov}(\mathbf{z}_a) \mathbf{R}', \quad (18)$$

where $\mathbf{R}' = \text{diag}\{x^0 - x_0, y^0 - y_0, z^0 - z_0, r_0^0\}$. Then, the second weighted LS method is used to solve Equation (16):

$$\mathbf{z}'_a = (\mathbf{G}'_a{}^T \Psi'^{-1} \mathbf{G}'_a)^{-1} \mathbf{G}'_a{}^T \Psi'^{-1} \mathbf{h}'. \quad (19)$$

The matrix Ψ' is not known since it contains the true values. However, \mathbf{R}' can be approximated by using the values in \mathbf{z}_a . If the source is far away, then the covariance matrix of \mathbf{z}_a can be represented as:

$$\text{cov}(\mathbf{z}_a) \approx c^2 r_0^2 (\mathbf{G}_a^T \mathbf{Q}^{-1} \mathbf{G}_a)^{-1}. \quad (20)$$

Then, Equation (19) reduces to:

$$\mathbf{z}'_a \approx (\mathbf{G}'_a{}^T \mathbf{R}'^{-1} \mathbf{G}'_a \mathbf{Q}^{-1} \mathbf{G}_a \mathbf{R}'^{-1} \mathbf{G}'_a)^{-1} (\mathbf{G}'_a{}^T \mathbf{R}'^{-1} \mathbf{G}'_a \mathbf{Q}^{-1} \mathbf{G}_a \mathbf{R}'^{-1}) \mathbf{h}'. \quad (21)$$

The final sound source position is estimated as:

$$\mathbf{z}_p = \sqrt{\mathbf{z}'_a} + [x_0, y_0, z_0]^T \quad (22)$$

or

$$\mathbf{z}_p = -\sqrt{\mathbf{z}'_a} + [x_0, y_0, z_0]^T. \quad (23)$$

3. Numerical Optimization Method for Array Structures

Given a certain number of microphones, there are infinite spatial geometric structures for microphone arrays. Nevertheless, in various practical scenarios, it is necessary to find the optimal array structure to effectively reduce the positioning error of sound sources in the target area. Because the microphone array consists of multiple microphones and each microphone's coordinates have three independent variables, the structure optimization of the microphone array in this paper is a multidimensional optimization problem.

The evolutionary algorithm is a global optimization method with high robustness and broad applicability. Unlike classic optimization methods such as gradient descent and quasi-Newton methods, the gradient of the problem being optimized is not required for the evolutionary algorithm. Meanwhile, the evolutionary algorithm makes few or no assumptions about the optimization problem and has great advantages in the application of unsupervised, complex multidimensional problems that cannot be solved using traditional deterministic algorithms [43]. The genetic algorithm (GA) and particle swarm optimization (PSO) algorithm are evolutionary algorithms. GA searches for the optimal solution by

imitating the mechanism of selection and inheritance in nature. The selection of crossover rate and mutation rate in GA seriously affects the quality of the solution, and the selection mostly depends on experience. Additionally, GA is very slow and difficult to converge for high-dimensional problems. Particle swarm optimization (PSO) is a metaheuristic global optimization algorithm, and the inner workings of the PSO make sufficient use of probabilistic transition rules to search very large spaces of candidate solutions in parallel [44]. Compared with GA, PSO has the advantage of simplicity, easy implementation, and few parameters requiring adjustment. PSO does not have genetic operations such as crossover and mutation. Instead, it determines the search based on its speed. Another essential feature of PSO is that particles have memories. The full search and update process of PSO follows the current optimal solution. Compared with GA, PSO may converge to an optimal solution more quickly. For the optimization problem of the microphone array structure in this paper, the gradient of the optimization objective function is difficult to derive. Additionally, the structure optimization of an array with many microphones is a high-dimensional optimization problem. These factors make PSO an effective method to solve the optimization problem in this paper. Therefore, an optimization model based on PSO was constructed to optimize the geometric parameter of the microphone array under different localization scenarios.

3.1. Optimization Model Based on PSO

A swarm of particles which traverse a multidimensional search space are employed in the PSO algorithm to find optima. Each particle is a potential solution and is influenced by the experiences of other particles, as well as its own experiences. Let \mathbf{p}_j be the position in the search space of the j -th particle, and the number of particles is set as N_p . Then, a swarm of particles can be expressed as:

$$\mathbf{P}_s = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{N_p}], \quad (24)$$

where each particle can be denoted as

$$\mathbf{p}_j = \begin{pmatrix} l_1 & \alpha_1 & \beta_1 \\ l_2 & \alpha_2 & \beta_2 \\ \vdots & \vdots & \vdots \\ l_{N_m} & \alpha_{N_m} & \beta_{N_m} \end{pmatrix} \quad j = 1, 2, \dots, N_p. \quad (25)$$

A new fitness evaluation function for the array structure optimization is constructed by the mean squared error (MSE) and the variance (VAR) of the localization results, which can comprehensively consider the localization accuracy and robustness. The fitness function is shown as follows:

$$f(\mathbf{p}_j) = \phi_w \text{MSE}(\mathbf{z}_p) + (1 - \phi_w) \text{VAR}(\mathbf{z}_p), \quad (26)$$

where ϕ_w is the weight value, $\phi_w \in [0, 1]$. \mathbf{z}_p is the final estimated sound source position. $\text{MSE}(\mathbf{z}_p)$ is the mean squared error of the localization results, which can be defined as

$$\text{MSE}(\mathbf{z}_p) = \frac{1}{N_s} \sum_i^{N_s} \left\| \mathbf{z}_{pi} - \mathbf{z}_{pi}^0 \right\|^2, \quad (27)$$

where \mathbf{z}_{pi}^0 is the coordinate of the real source. N_s is the number of sources involved in the optimization. $\text{VAR}(\mathbf{z}_p)$ is the variance of the localization results, which can be defined as:

$$\text{VAR}(\mathbf{z}_p) = \frac{1}{N_s} \sum_i^{N_s} \left(\left\| \mathbf{z}_{pi} - \mathbf{z}_{pi}^0 \right\| - \overline{\left\| \mathbf{z}_{pi} - \mathbf{z}_{pi}^0 \right\|} \right)^2, \quad (28)$$

where $\overline{\left\| \mathbf{z}_p - \mathbf{z}_p^0 \right\|} = \frac{1}{N_s} \sum_i^{N_s} \left\| \mathbf{z}_{pi} - \mathbf{z}_{pi}^0 \right\|$.

In Equation (26), the mean squared error can be used to judge the accuracy of sound source location results, and the variance can be used to judge the robustness of sound source localization results. The weight ratio between them can be adjusted according to the requirement of localization scenarios.

Then, an optimization problem (minimization) is defined as:

$$\exists \mathbf{p}'_j \in \mathbb{OS} \subseteq \mathbb{R}^d \Rightarrow \forall \mathbf{p}_j \in \mathbb{OS}, f(\mathbf{p}'_j) \leq f(\mathbf{p}_j) \quad j = 1, 2, \dots, N_p, \quad (29)$$

where \mathbb{R}^d is the real number field in d -dimensional space.

The PSO algorithm is used to solve this optimization problem. To seek the optimal solution, each particle moves in the direction of its previously best (\mathbf{p}_{best}) position and the global best (\mathbf{g}_{best}) position in the swarm. The expression of \mathbf{p}_{best} is

$$\mathbf{p}_{best}(j, k) = \arg \min [f(\mathbf{p}_j(k))] \quad j \in \{1, 2, \dots, N_p\}; k = 1, \dots, I_t, \quad (30)$$

and the expression of \mathbf{g}_{best} is

$$\mathbf{g}_{best}(k) = \arg \min [f(\mathbf{p}_j(k))] \quad j = 1, 2, \dots, N_p; k = 1, \dots, I_t, \quad (31)$$

where k denotes the current iteration number, and I_t denotes the maximum iteration number.

The velocity \mathbf{V} and position \mathbf{p} of particles are updated by the following equations:

$$\mathbf{V}_j(k+1) = w\mathbf{V}_j(k) + c_1 \text{rand}(\cdot)(\mathbf{p}_{best}(j, k) - \mathbf{p}_j(k)) + c_2 \text{rand}(\cdot)(\mathbf{g}_{best}(j, k) - \mathbf{p}_j(k)), \quad (32)$$

$$\mathbf{p}_j(k+1) = \mathbf{p}_j(k) + \mathbf{V}_j(k+1), \quad (33)$$

where \mathbf{V} stands for the migration velocity of particles, which is common to be set as a boundary to limit particles flying out of the search space. $\text{rand}(\cdot)$ are uniformly distributed random variables within range $[0, 1]$. c_1 and c_2 stand for learning factors, which are positive constant parameters. w is the inertia weight used to balance the global exploration and local exploitation. Shi [45] suggested a solution to determine the inertia weight:

$$w_t = w_{max} - \frac{w_{max} - w_{min}}{I_t} k, \quad (34)$$

where w_{max} and w_{min} are maximum and minimum weight, respectively.

3.2. PSO Optimization Procedure

The optimization procedure for the acoustic array is summarized as follows:

Step 1. Initialize PSO parameters including the number of particles N_p , the learning factors c_1 and c_2 , inertia weights w_{max} and w_{min} , and the total iteration number I_t .

Step 2. Initialize the particles' positions with a random distribution $p_j(0)$, and the parameters of the each particle ($j = 1, 2, \dots, N_p$) do not go beyond the boundaries of the search space.

Step 3. Initialize $p_{best}(j, 0) = p_j(0)$; $\mathbf{g}_{best}(0) = \text{argmin} f(p_j(0))$.

Step 4. Update particles' velocity and position through Equations (30) and (31), and maintain $p_j(k+1) \in \mathbb{OS}$.

Step 5. If $f(p_j(k)) < f(p_{best}(j, k))$, update the best known particle position $p_{best}(j, k) = p_j(k)$; if $f(p_{best}(j, k)) < f(\mathbf{g}_{best}(k))$, update the global best position $\mathbf{g}_{best}(k) = p_{best}(j, k)$.

Step 6. Judge the termination criteria: $f(\mathbf{g}_{best}(k)) \leq \delta$ (δ is presented as threshold) or the iteration number reaches the maximum I_t with the fitness function converging steadily. If not, repeat Steps 4 and 5. Otherwise, go to Step 7.

Step 7. Output the $\mathbf{g}_{best}(k)$ that stands for the best optimized result.

The flow chart of the optimization procedure is shown in Figure 2.

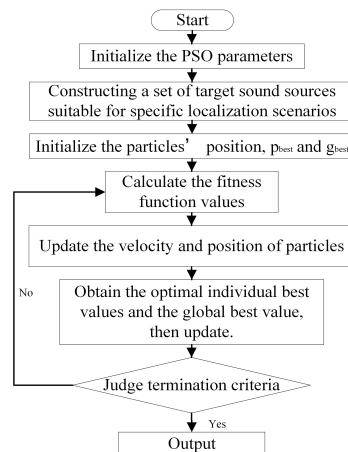


Figure 2. Flow chart for microphone array optimization. PSO: particle swarm optimization.

4. Simulation and Analysis

In order to verify the effectiveness of the method proposed in this paper, two kinds of sound source localization scenarios were constructed for microphone array optimization. One scenario was a ring-shaped sound source distribution, and the other was a cuboid sound source distribution. These two scenarios represent some specific sound source localization scenarios in practical applications, such as surround sound sources localization and road traffic flow noise sources tracking. For each specific localization scenario, two strategies of structure optimization were adopted to generate two kinds of optimal structures. In addition, the regular polyhedron microphone array structure proposed by Hu et al. as well as random array structures were used as a comparative study of the performance of sound source localization. The model established in Sections 1 and 2 was edited to code and run on the Matlab platform.

4.1. Scenario I—Ring-Shaped Sound Sources Distribution

In scenario I, sound sources were distributed in a cyclic annular band, here referred to as the ring-shaped sound source distribution. The distribution was controlled by Equation (35), as follows:

$$C_{SI} = [R_{SI} \sin(2\pi\theta_{SI}/360), R_{SI} \cos(2\pi\theta_{SI}/360), h_{SI}] \quad (35)$$

where C_{SI} is the coordinates of the sound source. R_{SI} is the radius of the cyclic annular band. θ_{SI} is the azimuth angle of the source. h_{SI} is the height of the source. In scenario I, $R_{SI} \in [6 \text{ m}, 6.5 \text{ m}]$, $\theta_{SI} \in [0^\circ, 360^\circ]$, $h_{SI} \in [0 \text{ m}, 0.5 \text{ m}]$.

The source distribution for the array structure optimization under scenario I is shown in Figure 3. The microphone array was located in the center of the ring, in which the reference microphone was located at the origin of the coordinates, and the location of other positioning microphones was obtained by optimization calculation.

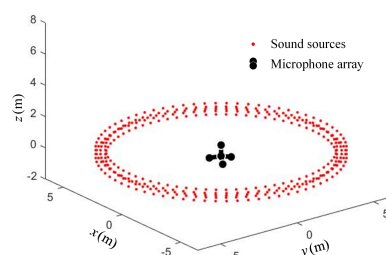


Figure 3. Scenario I—ring-shaped sound source distribution.

For the sound source localization based on TDOA, the time difference estimation error is the main influencing factor for localization accuracy. To facilitate the optimization and verification of microphone array structures, the TDOAs were directly obtained by calculating the relative position relationship between the sound sources and the array microphones. A noise component was added to the TDOAs, which was used to represent measurement noise in actual applications.

$$r_{i,0} = c\tau_{i,0} + \eta_{i,0}, \quad (36)$$

where $\eta_{i,0}$ is the time delay estimation noise component. $\eta_{i,0}$ is assumed to be a mutually independent, zero-mean stationary Gaussian random process, and the standard deviation of $\eta_{i,0}$ is σ . In this simulation, σ was set as 0.01.

Five microphones were selected to optimize the microphone array structures, which were compared with the tetrahedral structure array proposed by Hu et al. as well as a random array. One microphone in the array was chosen as the reference microphone, and the coordinates were set as $(0,0,0)$. The other microphones' coordinates were set as $M_i = (l_i \cos(\beta_i) \cos(\alpha_i), l_i \cos(\beta_i) \sin(\alpha_i), l_i \sin(\beta_i))$. The array parameters (l_i, α_i, β_i) were set as the properties of each particle. The parameters of the PSO model were initialized. The learning factors c_1 and c_2 were all set to 1.5. The maximum weight w_{max} was set to 0.8. The minimum weight w_{min} was set to 0.4. The weight value ϕ_w of the fitness function was set to 0.5, which means that the localization accuracy and robustness were equally considered.

In this simulation, two optimization strategies were applied to search for the optimal array structure. For the first kind of array optimization (Opt-array I), the distances l_i between M_i and M_0 were set to the same length 0.7 m, which is comparable with the tetrahedral structure array proposed by Hu et al. and the random array. Then, the azimuth angle α_i and the elevation angle β_i are the main geometric parameters to be optimized for the microphone array. Therefore, for Opt-array I, the dimension of the particles in the PSO algorithm was eight, since the number of other microphones used in scenario I was four. The constraints of the optimization space were $\alpha_i \in [0^\circ, 360^\circ]$ and $\beta_i \in [-90^\circ, 90^\circ]$. The number of particles N was set to 250. For the second kind of array optimization (Opt-array II), the radial distance l_i was not predefined. The azimuth angle α_i , the elevation angle β_i , as well as the radial distances l_i were all used for the array geometric parameters to be optimized. For convenience, the radial distance l_i was limited between 0.2 m and 0.8 m, considering the general array size for sound source localization. Then, the dimension of the particles for Opt-array II was twelve, and the constraints of the optimization space were $\alpha_i \in [0^\circ, 360^\circ]$, $\beta_i \in [-90^\circ, 90^\circ]$, and $l_i \in [0.2 \text{ m}, 0.8 \text{ m}]$. Given that the optimizing search space is much larger than in Opt-array I, the number of particles N was set to 450 for Opt-array II.

Then, the optimization model based on PSO ran on the Matlab platform to obtain the optimal array structure under scenario I. The fitness evolution curves of Opt-array I and the Opt-array II are shown in Figure 4.

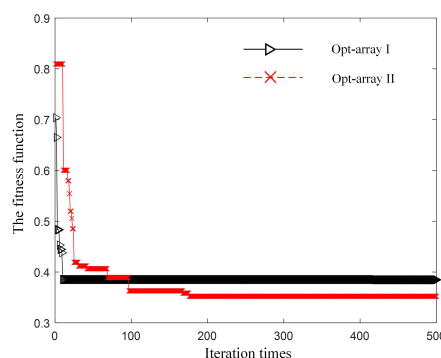


Figure 4. The fitness evolution curve under scenario I for the two kinds of optimal array.

It can be seen from Figure 4 that the fitness evolution curve tended to be stable after 200 steps of iteration, which indicates that the optimization process was basically convergent. The optimization results are deemed to be the optimal array structures under scenario I.

For the manufacturing of a microphone array, an angle precision of 5° is achievable. Therefore, the angle values of the optimized array were all rounded each 5° . The microphone coordinates and geometric parameters of the optimal arrays, the tetrahedral array proposed by Hu et al., and a random array are listed in Table 1.

Table 1. The microphone coordinates of the arrays under scenario I.

Items	Opt-Array I	Opt-Array II	Tetrahedral Array	Random Array
	$M_0(0,0,0)$	$M_0(0,0,0)$	$M_0(0,0,0)$	$M_0(0,0,0)$
Coordinates	$M_1(-0.09, -0.04, 0.69)$	$M_1(0.10, 0.07, 0.79)$	$M_1(0, 0, 0.7)$	$M_1(0.56, 0.32, 0.38)$
	$M_2(-0.08, -0.1, -0.69)$	$M_2(0.24, -0.49, -0.28)$	$M_2(0.66, 0, -0.24)$	$M_2(0, -0.54, 0.54)$
	$M_3(0.27, -0.16, 0.63)$	$M_3(0.01, 0.04, -0.79)$	$M_3(-0.33, 0.57, -0.24)$	$M_3(-0.58, -0.21, -0.29)$
	$M_4(-0.02, 0.18, -0.68)$	$M_4(-0.22, 0.02, 0.77)$	$M_4(-0.33, -0.57, -0.24)$	$M_4(0.06, -0.11, -0.71)$
Parameters	$l_i = 0.7$ m	$l_1 = 0.75$ m	$l_i = 0.7$ m	$l_1 = 0.75$ m
		$l_2 = 0.73$ m		$l_2 = 0.77$ m
		$l_3 = 0.77$ m		$l_3 = 0.68$ m
		$l_4 = 0.75$ m		$l_4 = 0.72$ m
Geometric Array	$\alpha_1 = 25^\circ; \beta_1 = 100^\circ$	$\alpha_1 = 35^\circ; \beta_1 = 80^\circ$	$\alpha_1 = 0^\circ; \beta_1 = 90^\circ$	$\alpha_1 = 30^\circ; \beta_1 = 30^\circ$
	$\alpha_2 = 230^\circ; \beta_2 = -80^\circ$	$\alpha_2 = 295^\circ; \beta_2 = -25^\circ$	$\alpha_2 = 0^\circ; \beta_2 = -20^\circ$	$\alpha_2 = 270^\circ; \beta_2 = 45^\circ$
	$\alpha_3 = 330^\circ; \beta_3 = 65^\circ$	$\alpha_3 = 75^\circ; \beta_3 = -85^\circ$	$\alpha_3 = 120^\circ; \beta_3 = -20^\circ$	$\alpha_3 = 200^\circ; \beta_3 = -25^\circ$
	$\alpha_4 = 95^\circ; \beta_4 = -75^\circ$	$\alpha_4 = 175^\circ; \beta_4 = 75^\circ$	$\alpha_4 = 240^\circ; \beta_4 = -20^\circ$	$\alpha_4 = 300^\circ; \beta_4 = -80^\circ$

It can be seen from Table 1 that the array geometric parameters between Opt-array I and the Opt-array II were different. The average radial distances of Opt-array II were larger than those of Opt-array I. At the same time, there were small differences among the four radial distances of Opt-array II. The array structures of Opt-array I and Opt-array II are shown in Figure 5.

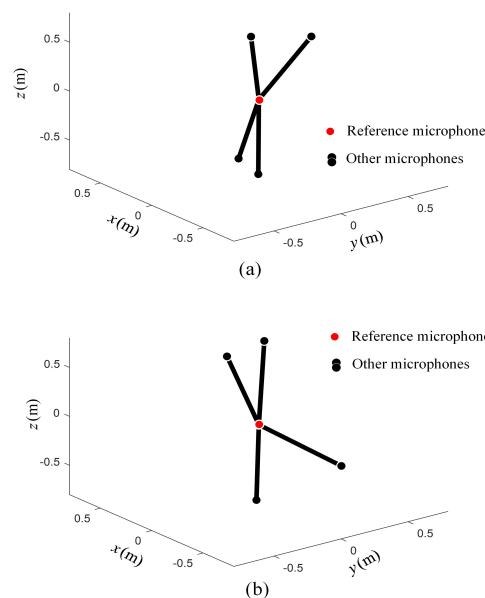


Figure 5. The optimized array structure under scenario I. (a) The structure of Opt-array I; (b) The structure of Opt-array II.

In order to verify the performance of the optimal arrays obtained in the simulation, a scenario with sound sources randomly distributed in the cyclic annular band was constructed, as shown in Figure 6.

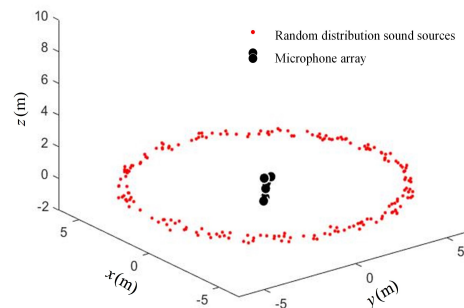


Figure 6. Random distribution of sound sources in the cyclic annular band.

As shown in Figure 6, 200 sources were randomly distributed in the cyclic annular band of scenario I. Opt-array I, opt-array II, the tetrahedral structure array, and the random array were used to locate these sources. The distances between the located sources and the corresponding real sources were counted to measure the positioning accuracy and robustness. Meanwhile, in order to analyze the influence of input noise on array positioning performance, Gaussian random noises $\eta_{i,0}$ with four different standard deviations were added to the time delay estimation, namely, $\sigma = 0.005$, $\sigma = 0.01$, $\sigma = 0.015$, $\sigma = 0.018$. The statistical chart is shown in Figure 7.

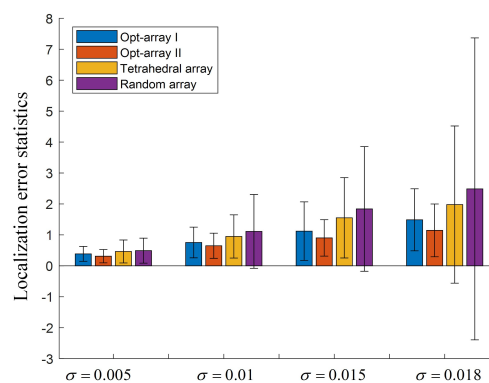


Figure 7. The localization error statistics for the arrays under scenario I.

In Figure 7, the height of the rectangular bar stands for mean localization error, and the length of the line bar presents the standard deviation of the localization error. It can be seen from Figure 7 that the mean values and the standard deviations of the localization error were enlarged with the increase of the input noise component of the time delay estimation. Under the same input noise amplitude, the mean value and the standard deviation of Opt-array I and Opt-array II were all much lower than that of the tetrahedral array and the random array. The bigger the input noise, the more significant the gap. This means that the optimized arrays by the proposed method could improve the accuracy and robustness of the sound source localization based on TDOA. The results illustrate the effectiveness of the array structure optimization method proposed in this paper.

Under four different input noise levels, the mean value and standard deviation of the localization error for the random array were all much larger than that for Opt-array I, Opt-array II, and the tetrahedral array, which illustrates that array optimization—whether the method of this paper or Hu’s method—produced a positive effect. The random arrays may achieve excellent positioning performance, but the possibility is tiny. Also, the mean value and standard deviation of the localization

error for Opt-array II were lower than for Opt-array I. Considering that there were differences among the four radial distances l_i of Opt-array II after array structure optimization, the optimization of the radial distance l_i contributes to promoting the positioning performance of the microphone array besides the optimization of the azimuth angle α_i and the elevation angle β_i .

4.2. Scenario II—Cuboid-Shaped Sound Sources Distribution

In scenario II, the sound sources were distributed in a cuboid space band, here referred to as the cuboid-shaped sound source distribution. The cuboid space band is $15\text{ m} \times 6\text{ m} \times 3\text{ m}$. The microphone array was located on one side of the cuboid distribution. The location of the reference microphone coincided with the origin of the coordinate system. The constructed scenario II is shown in Figure 8.

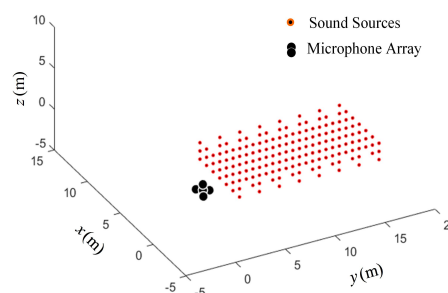


Figure 8. Scenario II—cuboid-shaped sound source distribution.

In the simulation of scenario II, five microphones were also selected to optimize the microphone array structure, which was compared with the tetrahedral structure array proposed by Hu et al. and a random array. The noise component $\eta_{i,0}$ with zero-mean Gaussian normal distribution was introduced into the time delay estimation. The standard deviation σ of the noise was the same as in scenario I, namely, 0.01. The parameters of the PSO model were set to be the same as in scenario I.

The fitness evolution curve of Opt-array I and Opt-array II under scenario II are shown in Figure 9.

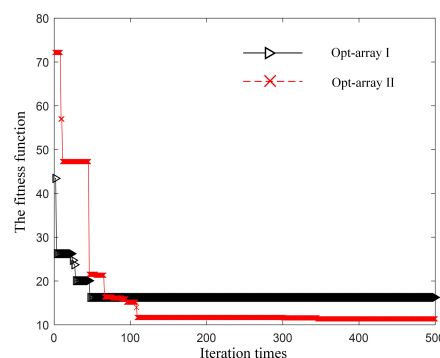


Figure 9. The fitness evolution curve under scenario II for the two kinds of optimal array.

It can be seen from Figure 9 that the fitness evolution curve tended to be stable after 150 steps of iteration. At the beginning of the iteration, the fitness function value of Opt-array II was higher than that of Opt-array I. Nevertheless, after numbers of iterative calculation, the fitness function value of Opt-array II was lower than that of Opt-array I when the iterations approached convergence, which means that the optimized structure of Opt-array II may have better localization performance than Opt-array I. Also, the angle values of the optimized array were all rounded each 5° . The microphone coordinates of the optimal arrays and the tetrahedral array proposed by Hu et al. and the random array are listed in Table 2.

Table 2. The microphone coordinates of the arrays under scenario II.

Items	Opt-Array I	Opt-Array II	Tetrahedral Array	Random Array
Coordinates	$M_0(0,0,0)$	$M_0(0,0,0)$	$M_0(0,0,0)$	$M_0(0,0,0)$
	$M_1(0.29, -0.34, 0.54)$	$M_1(-0.26, -0.22, -0.58)$	$M_1(0,0,0.7)$	$M_1(0.56, 0.32, 0.38)$
	$M_2(-0.14, 0.32, 0.61)$	$M_2(0.17, -0.57, -0.50)$	$M_2(0.66, 0, -0.24)$	$M_2(0, -0.54, 0.54)$
	$M_3(-0.06, 0.01, -0.70)$	$M_3(0.13, 0.16, 0.78)$	$M_3(-0.33, 0.57, -0.24)$	$M_3(-0.58, -0.21, -0.29)$
	$M_4(0.68, 0.18, 0)$	$M_4(-0.12, 0.67, 0.40)$	$M_4(-0.33, -0.57, -0.24)$	$M_4(0.06, -0.11, -0.71)$
Parameters	$l_i = 0.7$ m	$l_1 = 0.73$ m	$l_i = 0.7$ m	$l_1 = 0.75$ m
		$l_2 = 0.65$ m		$l_2 = 0.77$ m
		$l_3 = 0.80$ m		$l_3 = 0.68$ m
		$l_4 = 0.69$ m		$l_4 = 0.72$ m
Geometric Array	$\alpha_1 = 130^\circ; \beta_1 = 130^\circ$	$\alpha_1 = 220^\circ; \beta_1 = -60^\circ$	$\alpha_1 = 0^\circ; \beta_1 = 90^\circ$	$\alpha_1 = 30^\circ; \beta_1 = 30^\circ$
	$\alpha_2 = 115^\circ; \beta_2 = 60^\circ$	$\alpha_2 = 285^\circ; \beta_2 = -40^\circ$	$\alpha_2 = 0^\circ; \beta_2 = -20^\circ$	$\alpha_2 = 270^\circ; \beta_2 = 45^\circ$
	$\alpha_3 = 175^\circ; \beta_3 = -85^\circ$	$\alpha_3 = 50^\circ; \beta_3 = 75^\circ$	$\alpha_3 = 120^\circ; \beta_3 = -20^\circ$	$\alpha_3 = 200^\circ; \beta_3 = -25^\circ$
	$\alpha_4 = 15^\circ; \beta_4 = 0^\circ$	$\alpha_4 = 100^\circ; \beta_4 = 30^\circ$	$\alpha_4 = 240^\circ; \beta_4 = -20^\circ$	$\alpha_4 = 300^\circ; \beta_4 = -80^\circ$

It can be seen from Table 2 that the array geometric parameters between Opt-array I and Opt-array II under scenario II were different. The difference of radial distances l_i of Opt-array II under scenario II was much larger than that under scenario I. The array structures of Opt-array I and Opt-array II are shown in Figure 10.

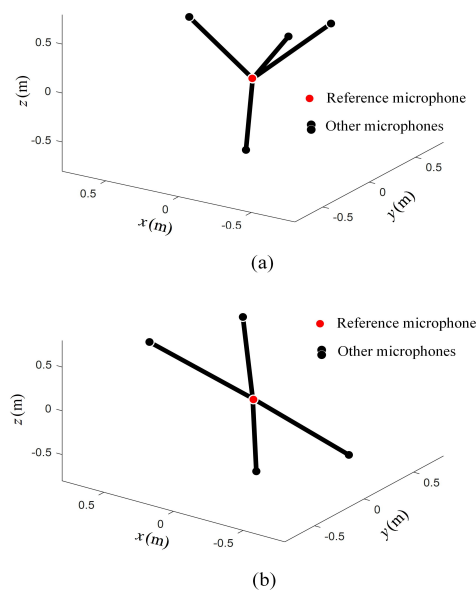


Figure 10. The optimized array structure under scenario II. (a) The structure of Opt-array I; (b) The structure of Opt-array II.

In order to verify the performance of the arrays, the scenario of randomly distributed sound sources in the cuboid space band was constructed, as shown in Figure 11.

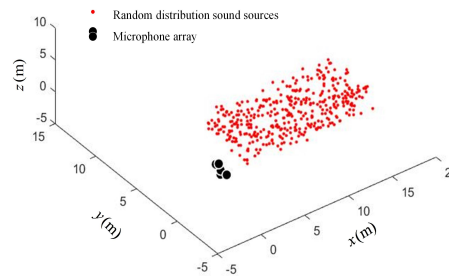


Figure 11. The random distribution of sound sources in the cuboid space band.

In Figure 11, 400 sources are randomly distributed in the cuboid space band. The Opt-array I, the Opt-array II, the tetrahedral structure array, and the random array are used to locate these sources. The Gaussian random noise $\eta_{i,0}$ with five different standard deviations are added to the time delay estimation, namely $\sigma = 0.002$, $\sigma = 0.005$, $\sigma = 0.008$, $\sigma = 0.01$, $\sigma = 0.012$. The statistics of the distances between the located sources and the corresponding real sources are drawn in Figure 12.

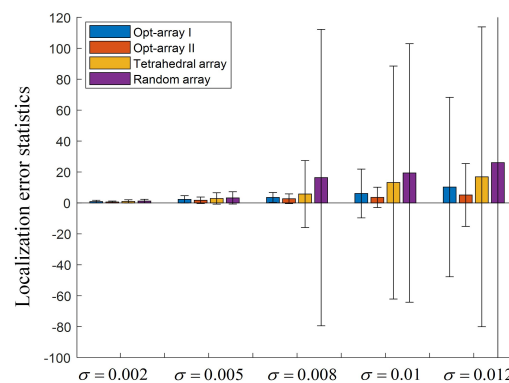


Figure 12. The localization error statistics for the arrays under scenario II.

It can be seen from Figure 12 that the mean values and the standard deviations of the localization error were enlarged with the increase of the input noise component of the time delay estimation. Under the same input noise amplitude, the mean value and the standard deviation of Opt-array I and Opt-array II were much lower than those of the tetrahedral array proposed by Hu et al. and the random array, and the gap increased rapidly with the increase of input noise. The optimized arrays by the proposed optimization method could improve the accuracy and robustness of the sound source localization based on TDOA.

Under five different input noise levels, the mean value and standard deviation of the localization error for the random array were much larger than that for Opt-array I, Opt-array II, and the tetrahedral array, especially when the input noise component was large, which illustrates that array optimization produced a positive effect. Random arrays have little chance of achieving excellent positioning performance under specific localization scenarios. Also, the mean value and standard deviation of the localization error for Opt-array II were lower than those for Opt-array I, which indicates that the optimization of the radial distance l_i contributed to promoting the positioning performance of the microphone array besides the optimization of the azimuth angle α_i and the elevation angle β_i . Moreover, the localization error reduction of Opt-array II under scenario II was more significant than that under scenario I when the input noise component was large. Considering that the difference of radial distance l_i of Opt-array II under scenario II was much larger than that under scenario I, the radial distance under scenario II was a more significant factor of the array structure optimization than under scenario I.

In addition, compared with scenario I, the mean values and the standard deviations of scenario II were much larger. The standard deviation rose sharply with the increase of input noise. The main reason for this is that the location area and the size of the sound sources in scenario II were much larger than in scenario I, and the sound sources were asymmetrically distributed. For scenario II, increasing the number of array microphones may help to reduce positioning errors and improve positioning robustness. Therefore, another optimization case was applied in scenario II, which is that seven microphones were chosen for the array structure optimization. The octahedron structure array proposed by Hu et al. [40] and a random array with seven microphones were used for comparative study.

Two kinds of optimization strategies were also used in the simulation. In the first kind of array optimization (Opt-array-7mic I), the radial distances l_i between M_i and M_0 were set to the same length of 0.7 m. For the second kind of array optimization (Opt-array-7mic II), the radial distances l_i , the azimuth angle α_i , and the elevation angle β_i were all used as the array geometric parameters to be optimized. The constraints and the initial parameters of the optimization model were set to be the same as in the case of the five microphone array optimization. Given that the optimizing search space was much larger than the array optimization with five microphones, the number of particles N for Opt-array-7mic I and Opt-array-7mic II were set to 400 and 650, respectively.

The optimal array structures were obtained after running the optimization model on the Matlab platform under scenario II. The geometric parameters of the optimal arrays, the octahedron array, and the random array are listed in Table 3. Also, the angles of the optimized array were all rounded 5° . The array structures of Opt-array-7mic I and Opt-array-7mic II are shown in Figure 13.

Table 3. The geometric parameters of the arrays with 7 microphones under the scenario II.

Items	Opt-Array-7mic I	Opt-Array-7mic II	Octahedron Array	Random Array
Coordinates	$M_0(0, 0, 0)$	$M_0(0, 0, 0)$	$M_0(0, 0, 0)$	$M_0(0, 0, 0)$
	$M_1(0.10, -0.07, 0.69)$	$M_1(0.33, -0.03, -0.72)$	$M_1(0, 0, 0.7)$	$M_1(0.56, 0.32, 0.38)$
	$M_2(-0.08, -0.16, -0.68)$	$M_2(-0.17, 0.47, 0.60)$	$M_2(-0.495, -0.495, 0)$	$M_2(0, -0.54, 0.54)$
	$M_3(0, 0.50, 0.50)$	$M_3(0.23, 0.62, 0.38)$	$M_3(0.495, -0.495, 0)$	$M_3(-0.58, -0.21, -0.29)$
	$M_4(0.14, -0.38, -0.57)$	$M_4(-0.09, -0.50, -0.61)$	$M_4(0.495, 0.495, 0)$	$M_4(0.06, -0.11, -0.71)$
	$M_5(0.08, -0.09, 0.69)$	$M_5(-0.05, -0.05, -0.75)$	$M_5(-0.495, 0.495, 0)$	$M_5(0.79, 0, 0)$
	$M_6(-0.18, 0.23, -0.63)$	$M_6(-0.07, -0.19, 0.75)$	$M_6(0, 0, -0.7)$	$M_6(-0.18, 0.30, 0.61)$
Parameters	$l_i = 0.7$ m	$l_1 = 0.79$ m	$l_i = 0.7$ m	$l_1 = 0.75$ m
		$l_2 = 0.78$ m		$l_2 = 0.77$ m
		$l_3 = 0.76$ m		$l_3 = 0.68$ m
		$l_4 = 0.79$ m		$l_4 = 0.72$ m
		$l_5 = 0.75$ m		$l_5 = 0.79$ m
		$l_6 = 0.78$ m		$l_6 = 0.70$ m
Geometric Array	$\alpha_1 = 325^\circ; \beta_1 = 80^\circ$	$\alpha_1 = 355^\circ; \beta_1 = -65^\circ$	$\alpha_1 = 0^\circ; \beta_1 = 90^\circ$	$\alpha_1 = 30^\circ; \beta_1 = 30^\circ$
	$\alpha_2 = 245^\circ; \beta_2 = -75^\circ$	$\alpha_2 = 110^\circ; \beta_2 = 50^\circ$	$\alpha_2 = 225^\circ; \beta_2 = 0^\circ$	$\alpha_2 = 270^\circ; \beta_2 = 45^\circ$
	$\alpha_3 = 90^\circ; \beta_3 = 45^\circ$	$\alpha_3 = 70^\circ; \beta_3 = 30^\circ$	$\alpha_3 = 315^\circ; \beta_3 = 0^\circ$	$\alpha_3 = 200^\circ; \beta_3 = -25^\circ$
	$\alpha_4 = 290^\circ; \beta_4 = -55^\circ$	$\alpha_4 = 260^\circ; \beta_4 = -50^\circ$	$\alpha_4 = 45^\circ; \beta_4 = 0^\circ$	$\alpha_4 = 300^\circ; \beta_4 = -80^\circ$
	$\alpha_5 = 310^\circ; \beta_5 = 80^\circ$	$\alpha_5 = 225^\circ; \beta_5 = -85^\circ$	$\alpha_5 = 135^\circ; \beta_5 = 0^\circ$	$\alpha_5 = 0^\circ; \beta_5 = 0^\circ$
	$\alpha_6 = 130^\circ; \beta_6 = -65^\circ$	$\alpha_6 = 250^\circ; \beta_6 = 75^\circ$	$\alpha_6 = 0^\circ; \beta_6 = -90^\circ$	$\alpha_6 = 120^\circ; \beta_6 = 60^\circ$

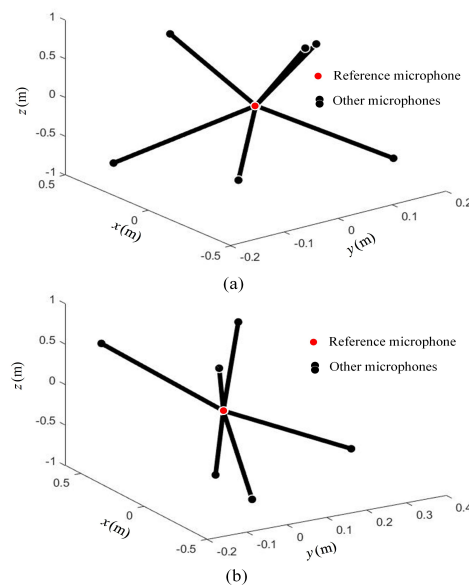


Figure 13. The optimized array structure with seven microphones under scenario II. (a) The structure of Opt-array-7mic I; (b) The structure of Opt-array-7mic II.

It can be seen from Table 3 and Figure 13 that the array structure between Opt-array-7mic I and Opt-array-7mic II were different. The difference of the radial distances l_i of Opt-array-7mic II was much smaller than that of Opt-array II under scenario II.

In order to verify the performance of the optimal arrays, the scenario of randomly distributed sound sources in the cuboid space band was constructed, similar to Figure 11. Gaussian random noises $\eta_{i,0}$ with five standard deviations were also added to the time delay estimation, namely, $\sigma = 0.002, \sigma = 0.005, \sigma = 0.008, \sigma = 0.01, \sigma = 0.012$. The statistics of the distances between the located sources and the corresponding real sources are drawn in Figure 14.

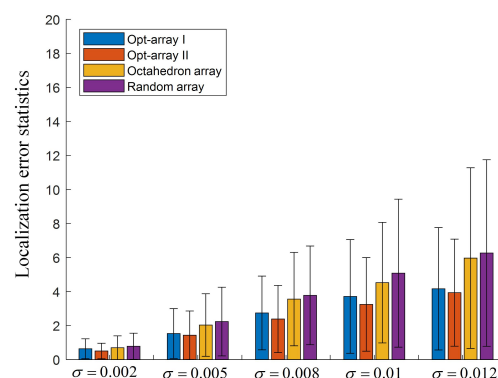


Figure 14. The localization error statistics for the arrays with seven microphones under scenario II.

Figure 14 shows that the mean values and standard deviations of localization error for Opt-array-7mic I and Opt-array-7mic II were lower than for the octahedron array proposed by Hu et al. and the random array, which illustrates the effectiveness of the proposed array optimization method. Comparing Figures 12 and 14, it can be seen that the mean values and the standard deviations of localization error for Opt-array-7mic I and Opt-array-7mic II were lower than for Opt-array-I and Opt-array II under scenario II. Considering that the standard deviations of the optimal arrays with seven microphones were significantly lower than those of the optimal arrays with five microphones when the input noise component was large, optimal array structures with more microphones could

significantly improve the robustness of the source localization based on TDOA. In addition, the mean values and the standard deviations of localization error for the octahedron array and the random array with seven microphones were also much lower than that for the tetrahedral array and the random array with five microphones, which demonstrates that increasing the number of microphones can greatly improve the positioning accuracy and robustness of the array based on TDOA.

5. Conclusions

This paper proposed a method of microphone array optimization for sound source localization based on TDOA under specific localization scenarios, which can be applied to the optimization of arbitrary array structure without prior information. For any number of microphones, a more optimal array structure can be given under any localization scenario. The proposed method is a numerical approach based on the particle swarm optimization algorithm. The mean squared error and the variance of the localization results combined with a weight value are used to construct the fitness function of the optimization model, which can consider both positioning accuracy and robustness. The geometric structure of the microphone array was established in parametric form, which is assigned as particle attributes and substituted into the optimization model to obtain the more optimal results. Two specific localization scenarios were constructed to optimize the array structures. For both specific scenarios, two kinds of array optimization strategies were utilized to obtain two optimal array structures. The optimized array structures were compared with the regular polyhedron structure array under different input noise amplitude.

For scenario I, the mean value and the standard deviation of the localization error for Opt-array I and Opt-array II were much lower than for the tetrahedral array and the random array, and the higher the input noise, the more significant the gap. Under four different input noise levels, the mean value and standard deviation of the localization error for the random array were the largest, and those of Opt-array II were the smallest. The results indicate that the array optimization produced a positive effect, and the optimization of the radial distance l_i contributed to promoting the positioning performance of the microphone array under scenario I.

For scenario II, the mean value and the standard deviation of Opt-array I and Opt-array II were also much lower than those of the tetrahedral array and the random array. Under five different input noise levels, the mean value and standard deviation of the localization error for the random array were the largest, and those of Opt-array II were the smallest. The array optimization and the optimization of the radial distance l_i all showed a positive effect on the positioning performance of the microphone array under scenario II. Moreover, the localization error reduction of Opt-array II under scenario II was more significant than that under scenario I. Considering that the difference of the radial distance of Opt-array II under scenario II was much larger than that under scenario I, the radial distance under scenario II was a more significant factor of the array structure optimization than that under scenario I. Under scenario II, the mean value and standard deviation of the optimal array were much higher than those of the optimal array under scenario I. The array with seven microphones was introduced into the optimization under scenario II, compared with the octahedron array and a random array. The results show that under five different input noise levels, the mean value and standard deviation of the localization error for Opt-array-7mic II were the smallest, and those for the random array were the largest. The mean value and standard deviation of the optimal array with seven microphones were lower than those of the optimal array with five microphones; especially, the standard deviation of the optimal array was significantly lower. This indicates that an optimal array structure with more microphones can significantly improve the robustness of source localization based on TDOA.

For both specific localization scenarios, the comparison results show that the localization accuracy and robustness of the optimized array structures were better than those of the regular polyhedron array structures proposed by Hu et al. and random array structures, which illustrates the effectiveness of the proposed array structure optimization method. The random arrays may achieve excellent positioning performance, but the likelihood is small. The optimization of the radial distance l_i contributed to

promoting the positioning performance of the microphone array besides the optimization of the azimuth angle α_i ; and the elevation angle β_i , particularly for scenario II. In the future, the efficiency of the optimization algorithm can be further studied, as well as the correlation between the positioning performance of the array and the array geometric parameters.

Author Contributions: Conceptualization, H.L. and Q.X.; methodology, H.L.; software, H.L.; validation, H.L.; formal analysis, H.L.; investigation, H.L.; resources, H.L. and T.K.; data curation, H.L.; writing—original draft preparation, H.L.; writing—review and editing, T.K.; visualization, H.L.; supervision, T.K.; project administration, H.L.; funding acquisition, H.L.

Funding: This work is jointly funded by the National Natural Science Foundation of China (NSFC) (Grant No. 51765017), the National Foundation from China Scholarship Council (Grant No. 201808360093), and the Research Foundation of Transportation Department of Jiangxi Province, China (Grant No. 2015D0062).

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

Abbreviations

The following abbreviations are used in this manuscript:

TDOA Time Difference of Arrival
PSO Particle Swarm Optimization

References

- Alameda-Pineda, X.; Horaud, R. A geometric approach to sound source localization from time-delay estimates. *IEEE Trans. Audio Speech Lang. Process.* **2014**, *22*, 1082–1095. [[CrossRef](#)]
- Padois, T.; Berry, A. Application of acoustic imaging techniques on snowmobile pass-by noise. *J. Acoust. Soc. Am.* **2017**, *141*, EL134–EL139. [[CrossRef](#)] [[PubMed](#)]
- Okuyama, F.; Takayama, J.Y.; Ohyama, S.; Kobayashi, A. A study on determination of a sound wave propagation direction for tracing a sound source. In Proceedings of the 41st SICE Annual Conference. SICE 2002, Osaka, Japan, 5–7 August 2002; Volume 2, pp. 1102–1104. [[CrossRef](#)]
- Huang, G.; Chen, J.; Benesty, J. Direction-of-arrival estimation of passive acoustic sources in reverberant environments based on the Householder transformation. *J. Acoust. Soc. Am.* **2015**, *138*, 3053–3060. [[CrossRef](#)] [[PubMed](#)]
- Jacob, F.; Haeb-umbach, R. On the bias of direction of arrival estimation using linear microphone arrays. In Proceedings of the Speech Communication; 12th ITG Symposium, Paderborn, Germany, 5–7 October 2016.
- Argentieri, S.; Danès, P.; Souères, P. A survey on sound source localization in robotics: From binaural to array processing methods. *Comput. Speech Lang.* **2015**, *34*, 87–112. [[CrossRef](#)]
- Tourbabin, V.; Rafaely, B. Direction of arrival estimation using microphone array processing for moving humanoid robots. *IEEE Trans. Audio Speech Lang. Process.* **2015**, *23*, 2046–2058. [[CrossRef](#)]
- Rascon, C.; Meza, I. Localization of sound sources in robotics: A review. *Rob. Autom. Syst.* **2017**, *96*, 184–210. [[CrossRef](#)]
- Chiariotti, P.; Martarelli, M.; Castellini, P. Acoustic beamforming for noise source localization—Reviews, methodology and applications. *Mech. Syst. Sig. Process.* **2019**, *120*, 422–448. [[CrossRef](#)]
- Hoffmann, F.M.; Fazi, F.M.; Williams, E.G.; Fontana, S. A general radiation model for sound fields and nearfield acoustical holography in wedge propagation spaces. *J. Acoust. Soc. Am.* **2017**, *142*, 1249–1260. [[CrossRef](#)]
- Padois, T.; Berry, A. Two and three-dimensional sound source localization with beamforming and several deconvolution techniques. *Acta Acust. United Acust.* **2017**, *103*, 392–400. [[CrossRef](#)]
- Battista, G.; Chiariotti, P.; Herold, G.; Sarradj, E.; Castellini, P. Inverse methods for three-dimensional acoustic mapping with a single planar array. In Proceedings of the 7th Berlin Beamforming Conference, Berlin, Germany, 5–6 March 2018.
- Ning, F.; Wei, J.; Qiu, L.; Shi, H.; Li, X. Three-dimensional acoustic imaging with planar microphone arrays and compressive sensing. *J. Sound Vib.* **2016**, *380*, 112–128. [[CrossRef](#)]
- Ferguson, B.G.; Criswick, L.G.; Lo, K.W. Locating far-field impulsive sound sources in air by triangulation. *J. Acoust. Soc. Am.* **2002**, *111*, 104–116. [[CrossRef](#)] [[PubMed](#)]

15. Wu, S.F.; Zhu, N. Locating arbitrarily time-dependent sound sources in three dimensional space in real time. *J. Acoust. Soc. Am.* **2010**, *128*, 728–739. [[CrossRef](#)] [[PubMed](#)]
16. Zhu, N.; Wu, S.F. Sound source localization in three-dimensional space in real time with redundancy checks. *J. Comput. Acoust.* **2012**, *20*, 1–16. [[CrossRef](#)]
17. Miao, F.; Yang, D.; Wen, J.; Lian, X. Moving sound source localization based on triangulation method. *J. Sound Vib.* **2016**, *385*, 93–103. [[CrossRef](#)]
18. Zhong, X.; Sun, L.; Yost, W. Active binaural localization of multiple sound sources. *Rob. Autom. Syst.* **2016**, *85*, 83–92. [[CrossRef](#)]
19. Klein, U.; Vö, T.Q. Direction-of-arrival estimation using a microphone array with the multichannel cross-correlation method. In Proceedings of the 2012 IEEE International Symposium on Signal Processing and Information Technology, Ho-Chi-Minh, Vietnam, 12–15 December 2012. [[CrossRef](#)]
20. Luo, R.C.; Huang, C.H.; Huang, C.Y. Search and track power charge docking station based on sound source for autonomous mobile robot applications. In Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 18–22 October 2010. [[CrossRef](#)]
21. Zheng, S.; Hao, P.; Lian, X.; Li, K. Time-domain transfer path analysis of multiple moving noise sources. *Noise Control Eng. J.* **2011**, *59*, 541–548. [[CrossRef](#)]
22. Sarradj, E. Optimal planar microphone array arrangements. The 41st Deutsche Jahrestagung für Akustik (DAGA) conference, Nuremberg, Germany, 16–19 March 2015; pp. 220–223.
23. Prime, Z.; Doolan, C. A comparison of popular beamforming arrays. In Proceedings of the Acoustics 2013-Victor Harbor, Victor Harbor, Australia, 17–20 November 2013.
24. Karbasi, A.; Sugiyama, A. A new DOA estimation method using a circular microphone array. In Proceedings of the 15th European Signal Processing Conference, Poznan, Poland, 3–7 September 2007; pp. 778–782.
25. Valin, J.M.; Michaud, F.; Hadjou, B.; Rouat, J. Localization of Simultaneous Moving Sound Sources for Mobile Robot Using a Frequency-Domain Steered Beamformer Approach. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), New Orleans, LA, USA, 26 April–1 May 2004; pp. 1033–1038. [[CrossRef](#)]
26. Ohata, T.; Nakamura, K.; Nagamine, A.; Mizumoto, T.; Ishizaki, T.; Kojima, R.; Sugiyama, O.; Nakadai, K. Outdoor sound source detection using a quadcopter with microphone array. *J. Rob. Mechatron.* **2017**, *29*, 177–187. [[CrossRef](#)]
27. Nikunen, J.; Virtanen, T. Time-difference of arrival model for spherical microphone arrays and application to direction of arrival estimation. In Proceedings of the 25th European Signal Processing Conference, Kos, Greece, 28 August–2 September 2017; pp. 1255–1259. [[CrossRef](#)]
28. Martinson, E.; Apker, T.; Bugajska, M. Optimizing a reconfigurable robotic microphone array. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011; pp. 125–130. [[CrossRef](#)]
29. Wang, R.; Bei, S. Optimization of fixed microphone array in high speed train noises identification based on far-field acoustic holography. *Adv. Acoust. Vibr.* **2017**, *2017*, 1–11. [[CrossRef](#)]
30. Kodrasi, I.; Rohdenburg, T.; Doclo, S. Microphone position optimization for planar superdirective beamforming. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; pp. 109–112.
31. Yan, L.; Ma, W. Arrangements of phased microphone arrays for acoustic source localization based on deconvolution algorithms. *J. Phys. Conf. Ser.* **2018**, *1065*, 1–4. [[CrossRef](#)]
32. Sarradj, E. A Generic Approach To Synthesize Optimal Array Microphone Arrangements. In Proceedings of the 6th Berlin Beamforming Conference, Berlin, Germany, 29 February–1 March 2016; pp. 1–12.
33. Bjelić, M.; Stanojević, M.; Šumarac Pavlović, D.; Mijić, M. Microphone array geometry optimization for traffic noise analysis. *J. Acoust. Soc. Am.* **2017**, *141*, 3101–3104. [[CrossRef](#)]
34. Teng, P.; Lv, J. The optimization design of microphone array layout for wideband noise sources. In Proceedings of the 22nd International Congress on Acoustics, Buenos Aires, Argentina, 5–9 September 2016, pp. 1–9.
35. Le Courtois, F.; Thomas, J.H.; Poisson, F.; Pascal, J.C. Genetic optimisation of a plane array geometry for beamforming. Application to source localisation in a high speed train. *J. Sound Vib.* **2016**, *371*, 78–93. [[CrossRef](#)]

36. Padois, T.; Sgard, F.; Doutres, O.; Berry, A. Acoustic source localization using a polyhedral microphone array and an improved generalized cross-correlation technique. *J. Sound Vib.* **2017**, *386*, 82–99. [[CrossRef](#)]
37. Padois, T.; Doutres, O.; Sgard, F.; Berry, A. Optimization of a spherical microphone array geometry for localizing acoustic sources using the generalized cross-correlation technique. *Mech. Syst. Sig. Process.* **2019**, *132*, 546–559. [[CrossRef](#)]
38. Chardon, G.; Kreuzer, W.; Noisternig, M. Design of a robust open spherical microphone array. In Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 6860–6864.
39. Zietlow, T.; Hussein, H.; Kowerko, D. Acoustic source localization in home environments—the effect of microphone array geometry. In Proceedings of the 28th Conference on Electronic Speech Signal Processing, Saarbrücken, Germany, 15–17 March 2017; pp. 219–226.
40. Hu, J.S.; Tsai, C.M.; Chan, C.Y.; Chang, Y.J. Geometrical arrangement of microphone array for accuracy enhancement in sound source localization. In Proceedings of the 8th Asian Control Conference (ASCC), Kaohsiung, Taiwan, 15–18 May 2011.
41. Fang, B. Simple solutions for hyperbolic and related position fixes. *IEEE Trans. Aerosp. Electron. Syst.* **1990**, *26*, 748–753. [[CrossRef](#)]
42. Chan, Y.; Ho, K. A simple and efficient estimator for hyperbolic location. *IEEE Trans. Signal Process.* **1994**, *42*, 1905–1915. [[CrossRef](#)]
43. Sengupta, S.; Basak, S.; Peters, R.; Sengupta, S.; Basak, S.; Peters, R.A. Particle swarm optimization: A survey of historical and recent developments with hybridization perspectives. *Mach. Learn. Knowl. Extr.* **2018**, *1*, 157–191. [[CrossRef](#)]
44. Zhang, Y.; Wang, S.; Ji, G. A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications. *Math. Prob. Eng.* **2015**, *2015*, 1–38. [[CrossRef](#)]
45. Shi, Y.; Eberhart, R.C. Parameter selection in particle swarm optimization. In Proceedings of the 7th International Conference on Evolutionary Programming, San Diego, CA, USA, 25–27 March 1998; pp. 591–600.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).