SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

December 2022, Vol. 65 222204:1–222204:12
https://doi.org/10.1007/s11432-022-3568-2

Prior area searching for energy-based sound source localization

Feng GAO[1,2], Yeyun CAI[1,2], Fang DENG[1,2*], Chengpu YU[1,2] & Jie CHEN[1,3]

[1]*Key Laboratory of Intelligent Control and Decision of Complex Systems, Beijing Institute of Technology, Beijing* 100081*, China;*
[2]*Beijing Institute of Technology Chongqing Innovation Center, Chongqing* 401120*, China;*
[3]*Guohao Academy, Tongji University, Shanghai* 200092*, China*

Received 26 January 2022/Revised 13 June 2022/Accepted 11 July 2022/Published online 15 November 2022

Abstract Traditional energy-based sound source localization methods have the problems of the large solution space and time-consuming calculation. Accordingly, this paper proposes to use the data collected by each acoustic sensor and their corresponding weights to adaptively initialize the prior area of a target. In this way, the potential existence range of the target is reduced and the location estimate can be determined in a small area. Specifically, we first determine the initial search point based on the current sound data and the set rules. Then, the prior location of the target is iteratively searched according to different sound energy circles' weights. Next, the prior area of the target is determined around the prior location. Finally, the precise location of the target is further traversed to minimize the objective function, which is constructed by the weighted nonlinear least squares location (WNLS) algorithm. A series of indoor experiments are performed. The results show that our method can effectively improve the positioning accuracy by approximately 13% and greatly reduce the calculation time.

Keywords prior area searching, sound source localization, adaptively initialization, iterative search, indoor environment

Citation Gao F, Cai Y Y, Deng F, et al. Prior area searching for energy-based sound source localization. Sci China Inf Sci, 2022, 65(12): 222204, https://doi.org/10.1007/s11432-022-3568-2

1 Introduction

Objects can be quickly located based on the sound they produce. Unlike visual signals [1], sound signals can be used to locate targets in all directions and at all hours of the day. At present, sound information and sound source localization techniques have been applied in various fields, such as target tracking [2–4], acoustic fingerprint recognition [5], medical health monitoring [6], and visual-audio joint analysis [7].

Nowadays, many sound source localization methods have been proposed, such as multi-signal classification (MUSIC) algorithm [8], beamforming-based localization algorithm [9], cross-correlation-based localization algorithm [10], and energy-based localization algorithm [11]. Among them, the MUSIC algorithm can perform frequency estimation and radio direction finding. It achieves a high localization performance with large parameters and data storage. An array of eight sound sensors was deployed in [12], and then the MUSIC algorithm was used to locate the sound source. A novel method was proposed in [13] to enhance the ability to distinguish noise signals from others for the MUSIC algorithm.

The beamforming-based localization algorithm [14] is widely used in the fields of unmanned aerial vehicles and robots. Source localization and enhancement were studied in [15] using the beamforming-based method. A robust adaptive beamforming-based method was proposed in [16] to estimate the steering vectors. In the cross-correlation localization algorithms, the time difference of arrival (TDOA)-based localization method was applied in [17] by estimating the time delay between two sensor nodes. The localization problem for a moving object was considered in [18]. The sound direction of a mobile robot interface was estimated in [19] with time delay information, and a cubic array was developed in [20] to locate the sound source by combining the TDOA method.

* Corresponding author (email: dengfang@bit.edu.cn)

Generally, energy-based sound source localization methods have the advantages of simple calculation form and low communication volume. Thus, they can be easily deployed on onboard platforms [21, 22]. Related energy-based methods can be classified into linear least squares (LLS) estimation, nonlinear least squares (NLS) estimation, and maximum likelihood (ML) estimation. Among them, the LLS algorithm eliminates the quadratic term of the sound source location by combining two sound energy circles, which reduces the computational complexity at the cost of losing accuracy. Moreover, LLS needs more sensors to locate targets in one measurement compared with NLS and ML. Objective functions built with a target's location in NLS and ML have complex nonlinear forms, which are difficult for computing analytical solutions. RV-SRP [23] realized the balance between accuracy and speed by deploying sparse volumetric grids. In [11], all feasible solutions were traversed to calculate the target location. A rough fine search method was adopted in [24, 25] to reduce the computational complexity. However, a large search step in the rough search stage may lead to the local minimum solution. In [26], a searching method based on the event region was proposed. This method limited the feasible solution of a target's location to the region near the sensor node, which received the largest sound energy. However, this method is highly dependent on the measurement data of a single sensor, so its positioning performance is sharply reduced when the sensor faults.

To improve the calculation speed and localization accuracy of traditional energy-based sound source localization methods, this paper proposes a prior area searching strategy, which adaptively initializes the prior area of a target. Then, the target's accurate position can be quickly found in the small prior area. In our work, the target's initial search point and prior area can be automatically determined. Moreover, there are few hyperparameters in our work. We perform a series of indoor sound source localization experiments on a rectangular sensor array. Experimental results demonstrate that compared with the traditional method, we can achieve a better localization accuracy with a higher computation speed using the proposed method. Moreover, the ablation studies further illustrate the effectiveness of our prior searching strategy under different conditions.

Some related studies have emerged in source-seeking areas, which attempt to design feedback controllers that guide movable onboard sensors to the source point in a limited time based on detected measurements. In this context, a guidance strategy for autonomous robots was proposed in [27], where robots were controlled to seek the odor source in an interactive way. The concentration measurement with a set of distributed robots was studied in [28] under limited or all-to-all communication conditions. Each robot could independently perform gradient estimation, which is the basis for its action. These studies can effectively locate the target by measuring the emitted signal strength in real time. Moreover, in the area of target localization based on the received signal strength, recurrent neural networks were used in [29] to solve the nonlinear inequality-constrained optimization problem defined on a graph. In [29], the blind sensor node localization was also successfully applied. By contrast, there are three differences between our work and the aforementioned methods. First, different from the iterative searching strategy, the global optimal position of a sound source target can be found in one measurement with our method. Second, the sensor array applied in our experiments is fixed in place. Third, compared with the sensor node networks in [29], targets can be located using fewer nodes (three or four) in our work.

The remainder of this paper is organized as follows. Section 2 models the sound source. Section 3 describes the sound source localization algorithm. Section 4 reports the prior area searching method. Section 5 provides a series of related experiments and results. The detailed analyses are also presented. Finally, Section 6 concludes this article.

2 Sound source modeling

In this section, we first present the definition of the near-field model of sound information and then establish the energy attenuation model of sound information.

2.1 Near-field model

According to the distance between a sound sensor array and target, the sound field model can be divided into the near-field model and the far-field model. The specific judgment condition is

$$r \leqslant \frac{2L^2}{\gamma}, \tag{1}$$

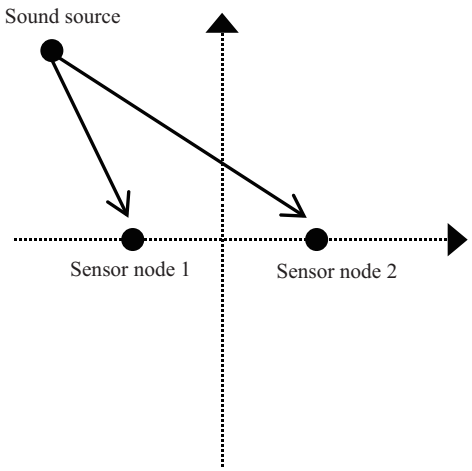

Figure 1 The diagram of the sound source's near-field model.

where r represents the distance from the sound source to the sound sensor array, L is the distance between sound sensor nodes, and γ is the wavelength of the sound source signal. As shown in Figure 1, we only consider the near-field model, which satisfies the condition (1). Because the sound source is in a short distance, the collected sound amplitude is different with different placements of sensor nodes.

2.2 Energy attenuation model

Generally, the signal produced by a sound source target will attenuate with the increase in distance. The sound amplitude detected by the node i at time t is

$$a_i(t) = \frac{a(t)}{\|r_i - r(t)\|} + v_i(t), \tag{2}$$

where $a_i(t)$ is the signal amplitude detected by the node i, $a(t)$ is the amplitude of the sound source's signal, and $v_i(t)$ represents noise and follows the normal distribution. r_i and $r(t)$ are the positions of the node i and sound source, respectively. The symbol $\|\cdot\|$ means calculating the Euclidean distance.

Non stationary sound signals can be considered stationary in a short time. Therefore, the energy E of a continuous signal $f(t)$ during a time interval T is

$$E = \int_{t_0}^{t_0+T} f^2(t) \cdot \mathrm{d}t. \tag{3}$$

The sound signal is sampled as a series of discrete values through the node i. Assuming that there are M sampling points in a time period T around time t, based on (3), the short-time energy $S_i(t)$ of a sound source at node i can be written as

$$S_i(t) = \frac{T}{M}\sum_{n=1}^{M} a_i^2(n) = \frac{T}{M}\sum_{n=1}^{M}\left[\frac{a(n)}{\|r_i - r(n)\|} + v_i(n)\right]^2. \tag{4}$$

Generally, noise is independent of the signal, so

$$\sum_{n=1}^{M} \frac{a(n)v_i(n)}{\|r_i - r(n)\|} = 0, \tag{5}$$

and according to (3), the energy $S(t)$ of the sound source is

$$S(t) = \frac{T}{M}\sum_{n=1}^{M} a^2(n). \tag{6}$$

Similarly, the energy $\varepsilon_i(t)$ of noise v_i can be written as

$$\varepsilon_i(t) = \frac{T}{M}\sum_{n=1}^{M} v_i^2(n). \tag{7}$$

By substituting (5)–(7) into (4), we can obtain

$$S_i(t) = \frac{S(t)}{\|r_i - r(n)\|^2} + \varepsilon_i(t). \tag{8}$$

In practice, the noise influence coefficient g_i is added, and the expression of (8) becomes

$$S_i(t) = \frac{g_i S(t)}{\|r_i - r(n)\|^2} + \varepsilon_i(t). \tag{9}$$

3 Sound source location reasoning

In this study, we use the weighted nonlinear least squares location (WNLS) algorithm [30] to solve the position of a sound source. Let k_{ij} be the ratio between the distances of nodes i and j to the sound source at time t. Then

$$k_{ij} = \frac{\|r_i - r(t)\|}{\|r_j - r(t)\|} = \left[\frac{S_i(t) - \varepsilon_i(t)}{S_j(t) - \varepsilon_j(t)}\right]^{-\frac{1}{2}}. \tag{10}$$

When the environmental signal-to-noise ratio is large, that is, $S(t) \gg \varepsilon(t)$, Eq. (10) can be simplified as

$$k_{ij} = \frac{\|r_i - r(t)\|}{\|r_j - r(t)\|} = \left[\frac{S_i(t)}{S_j(t)}\right]^{-\frac{1}{2}}, \tag{11}$$

where the left term can be rewritten as

$$k_{ij}^2 = \frac{\|r_i - r(t)\|^2}{\|r_j - r(t)\|^2}. \tag{12}$$

In actual measurements, the sound energies collected by two sensors can hardly be the same. Therefore, $k_{ij} > 0$ and $k_{ij} \neq 1$. Then, Eq. (12) can be transformed to

$$\|r(t) - C_{ij}\|^2 = \rho_{ij}^2, \tag{13}$$

where $C_{ij} = \frac{r_i - k_{ij}^2 r_j}{1 - k_{ij}^2}$ and $\rho_{ij} = |\frac{k_{ij}\|r_i - r_j\|}{1 - k_{ij}^2}|$. For sensors i and j, the sound source target is located on the circle with center of C_{ij} and radius of ρ_{ij}. This circle is called a sound energy circle in this paper.

Let the coordinates of the sound source target be $O(x, y)$, which should meet the following requirements:

$$\arg\min_{O(x,y)} \sum_{m=1}^{\frac{n(n-1)}{2}} |\beta_m(\|O - C_m\| - \rho_m)|, \tag{14}$$

where n is the number of sensors, C_m is the center of the sound energy circle, ρ_m is the radius of the sound energy circle, and β_m is the weight of a sound energy circle. Three circles can determine the approximate position of the sound source target. Therefore, the energy-based sound source location algorithm requires at least three sound sensor nodes in one measurement.

4 Prior area searching

Based on the above analyses, solving the sound source target's position is to search for the optimal solution according to the function (14). To improve the localization accuracy and searching speed, we propose a convex set-based algorithm for iteratively approaching the prior area of a target. In addition, we propose a method for finding the initial search point of the target.

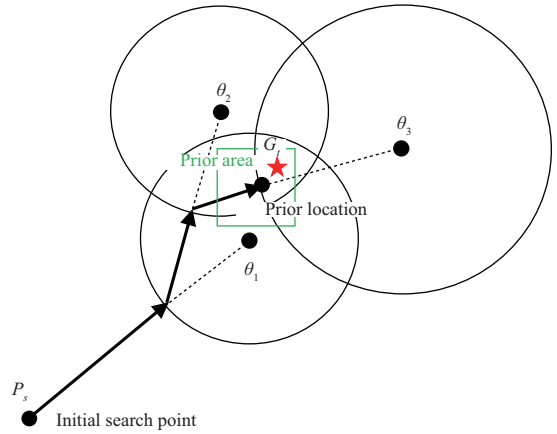

Figure 2 (Color online) The search process of finding the target's prior area, which is drawn with a green box.

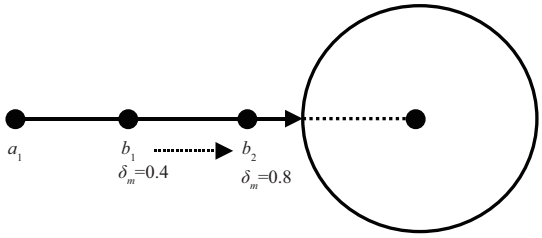

Figure 3 The influence of θ_m on the iterative process.

As shown in Figure 2, the convex set-based algorithm mainly includes two steps, which are simplified as follows.

Step 1: Set the initial search point P_s, which is used as the initial point to iteratively find the prior area of a target.

Step 2: Search for the optimal solution for the target's position in a small range.

Let θ_m be the sound energy circle determined by sensors i and j. In the first step, the coordinate of P_s is written as x_0, and then the next position x_1 is

$$x_1 = x_0 + \delta_m(P_{\theta_m}(x_0) - x_0), \tag{15}$$

where δ_m is the size of the iterative step. The symbol θ_m represents the sound energy circle, which is the closest to x_0. A line segment can be obtained by connecting x_0 and the center of θ_m, which intersects with θ_m at point $P_{\theta_m}(x_0)$.

Therefore, the iterative search process is

$$x_t = x_{t-1} + \delta_m(P_{\theta_m}(x_{t-1}) - x_{t-1}). \tag{16}$$

As shown in Figure 2, when there exist three sound energy circles θ_1, θ_2, and θ_3, the prior location of a target can be found through three iterations. In the second step, a small prior area is set around the prior location. The prior location is close to G_t, which is the target's actual position. In the small prior area, positions are divided with a certain step and traversed to locate the target.

A series of experiments in Section 5 show that prior area searching can effectively simplify the calculation and improve localization accuracy. The following two key factors affect the algorithm.

• **Value of δ_m.** As shown in Figure 3, the current position is point a_1, and the δ_m of position b_2 is larger than the δ_m of position b_1. The findings show that the larger the δ_m is, the closer the next position is to the sound energy circle. Therefore, δ_m represents the trust degree of the sound energy circle, which has the same meaning as β_m in (14). Thus, let $\delta_m = \beta_m$.

• **Location of P_s.** The selection of P_s affects the final positioning accuracy. A good P_s should be close to the sound source target's actual position. Here, we set P_s through the following steps.

Step 1: Select two sound sensor nodes i and j, which have the largest sum of the sound energy values detected at time t.

Step 2: According to (13), the sound energy circle determined by sensors i and j is C_{first}.

Step 3: Select the two sound sensor nodes p and q, which have the second largest sum of sound energy values detected at time t.

Step 4: According to (13), the sound energy circle determined by sensors p and q is C_{second}.

Step 5: Solve the intersection points of C_{first} and C_{second}. Three situations are considered.

(1) There are two intersections. Select any one of the two points as P_s.

(2) There is only one intersection. This point is chosen as P_s.

(3) There is no intersection. Draw a line segment by connecting the centers of two sound energy circles C_{first} and C_{second}. Then, the intersection of the line segment and C_{first} is chosen as P_s.

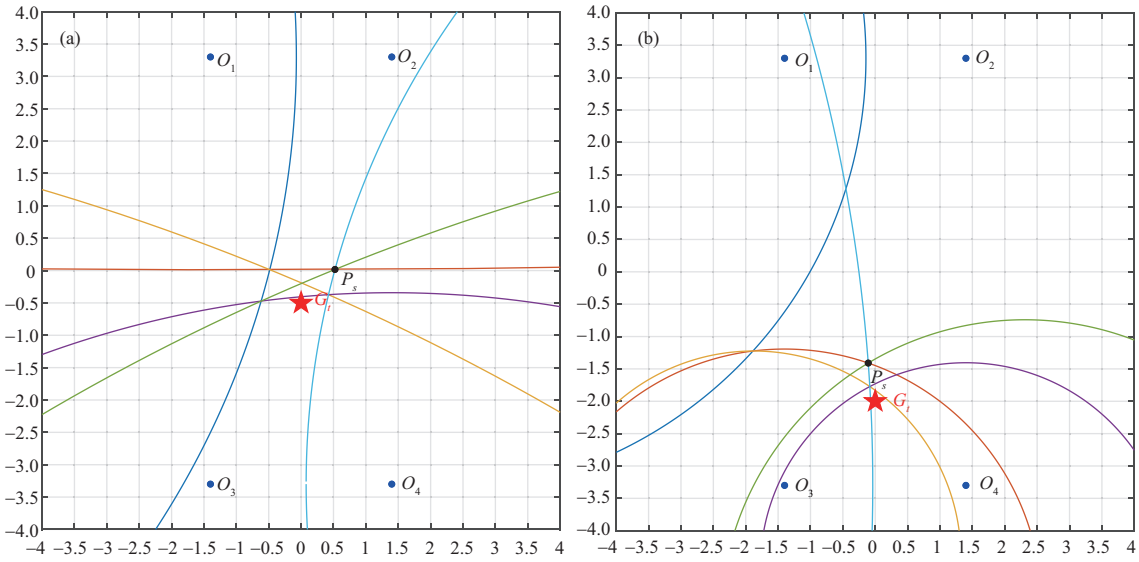

Figure 4 (Color online) The location of initial search point. O_1, O_2, O_3, and O_4 are four sound sensors placed in a square array. (a) The target is located at the position of (0, −0.5); (b) the target is located at the position of (0, −2.0).

The selected initial search points are shown in Figure 4, where Figures 4(a) and (b) are two situations with targets at different positions. Sound energy circles determined by four sound sensor nodes are drawn with different colors. It can be seen that P_s initialized by the above steps is close to the target's actual position G_t.

5 Experimental results

After the theoretical analyses, indoor sound source localization experiments and ablation studies are performed to verify the accuracy and speed of our algorithm.

5.1 Experimental configurations

A series of experiments are conducted in a quiet room. As shown in Figure 5, we use four acoustic sensor nodes, which are placed at O_1, O_2, O_3, and O_4 to receive sound signals. An Arduino board equipped with an automatic voltage regulator (AVR) microcontroller is applied as the main controller to receive signals, which are then sent to the computing platform for positioning targets. The coordinates of the four sensor nodes are $(-1.4, -3.3)$, $(-1.4, 3.3)$, $(1.4, 3.3)$, and $(1.4, -3.3)$. We use 13-bit analog-to-digital (AD) sampling, and the sampling frequency is 3 kHz. A car with horn is selected as the sound source target. We conduct 14 sets of experiments (red horns 1–14 in Figure 5), where the sound source target makes sound during a period of time. At each point, the sound information is measured nine times. Finally, the average localization error is calculated.

For comparison, we choose the traditional method, which solves (14) by iterating over all points in the sensor array's coverage area. The loss exponent [11] in (9) is a variable, which is related to many terms, such as temperature and wind speed. Here, we use the TDOA-based acoustic source localization method to measure the loss exponent, and its value is 2.08, which is close to 2. There are six sound energy circles that are produced by every two acoustic sensors. The weights of these sound energy circles are manually set to 0.9, 0.85, 0.75, 0.65, 0.65, and 0.65. The prior area mentioned in Section 4 is a square area with a length of 4 m. The step sizes in our proposed method and the traditional method are uniformly set to 0.1 m.

5.2 Indoor sound source localization experiments

The Euclidean distance between the estimated position and actual position is used as the localization error to measure the quality of different methods. During the experiments, random noise is added to the

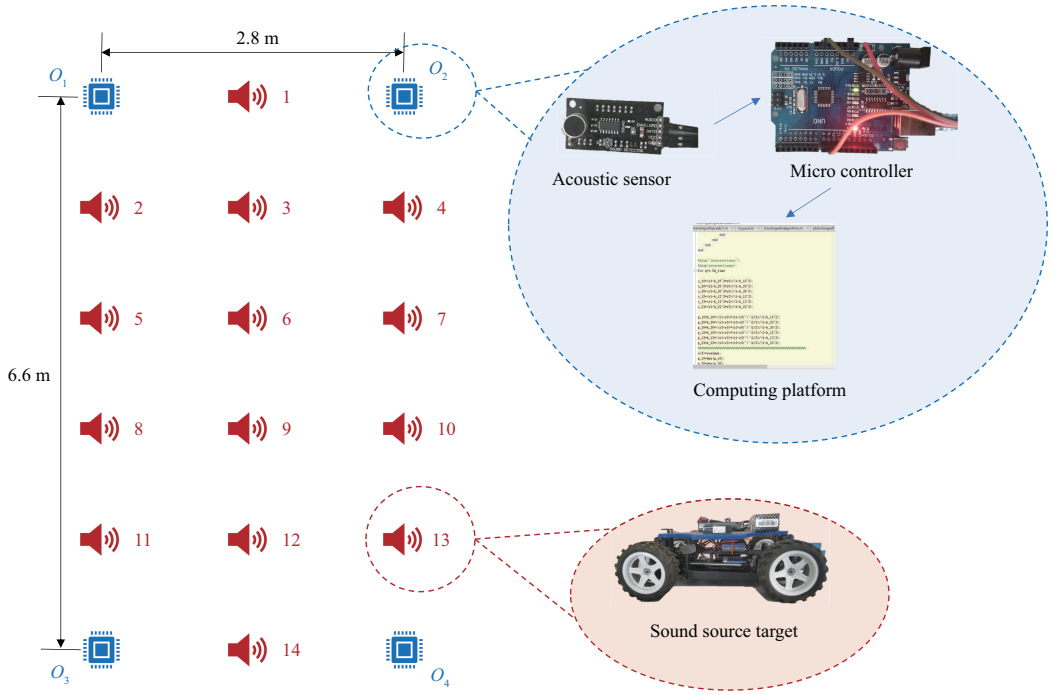

Figure 5 (Color online) The related indoor experimental configurations.

Table 1 Experimental results of proposed method and traditional method

Number	Target coordinate	Localization error of proposed method (m)				Localization error of traditional method (m)			
		SNR_{∞}	SNR_{30db}	SNR_{20db}	SNR_{10db}	SNR_{∞}	SNR_{30db}	SNR_{20db}	SNR_{10db}
1	(0, 3.3)	0.19	0.13	0.11	0.65	0.10	0.28	0.78	0.85
2	(−1.4, 2)	0.25	0.14	0.13	1.00	0.22	0.32	0.40	0.80
3	(0, 2)	0.33	0.35	0.05	1.32	0.36	0.36	0.36	1.30
4	(1.4, 2)	0.26	0.30	0.11	0.37	0.30	0.32	0.10	1.60
5	(−1.4, 0.5)	0.22	0.42	0.08	0.75	0.22	0.30	0.89	1.51
6	(0, 0.5)	0.26	0.26	0.45	0.74	0.20	0.22	0.32	1.00
7	(1.4, 0.5)	0.29	0.04	1.00	0.88	0.32	0.10	0.67	0.45
8	(−1.4, −0.5)	0.29	0.31	0.30	0.51	0.30	0.20	0.82	0.00
9	(0, −0.5)	0.35	0.39	0.91	0.45	0.30	0.40	0.78	1.43
10	(1.4, −0.5)	0.65	0.32	0.03	2.12	0.54	0.82	0.64	0.86
11	(−1.4, −2)	0.08	0.00	0.10	1.30	0.10	0.20	0.10	1.22
12	(0, −2)	0.59	0.54	0.63	1.32	0.61	0.58	0.71	1.00
13	(1.4, −2)	0.34	0.41	0.30	1.12	0.40	0.20	0.20	0.58
14	(0, −3.3)	0.80	1.01	0.63	0.00	0.80	0.81	0.51	1.24
	Average error	0.35	0.33	0.35	0.89	0.34	0.37	0.52	0.99

collected sound signal to verify the performances of our method and the traditional method under different signal-to-noise ratios (SNRs). The experimental results are shown in Table 1. Reasonably, the positioning errors of different methods decrease with the increase in SNR. Moreover, the average localization error of our method is approximately 0.48 m, whereas that of the traditional method is approximately 0.55 m. The results show that our accuracy improves by 13% compared to that of the traditional method. In addition, the advantage of our method becomes increasingly obvious as the SNR increases.

The processes of the sound source localization using the traditional method and our method are shown in Figures 6 and 7. Among them, P_r in Figure 6 represents the estimated position of a target using the traditional method. P_s and P_e in Figure 7 are the initial search point and final estimated position using

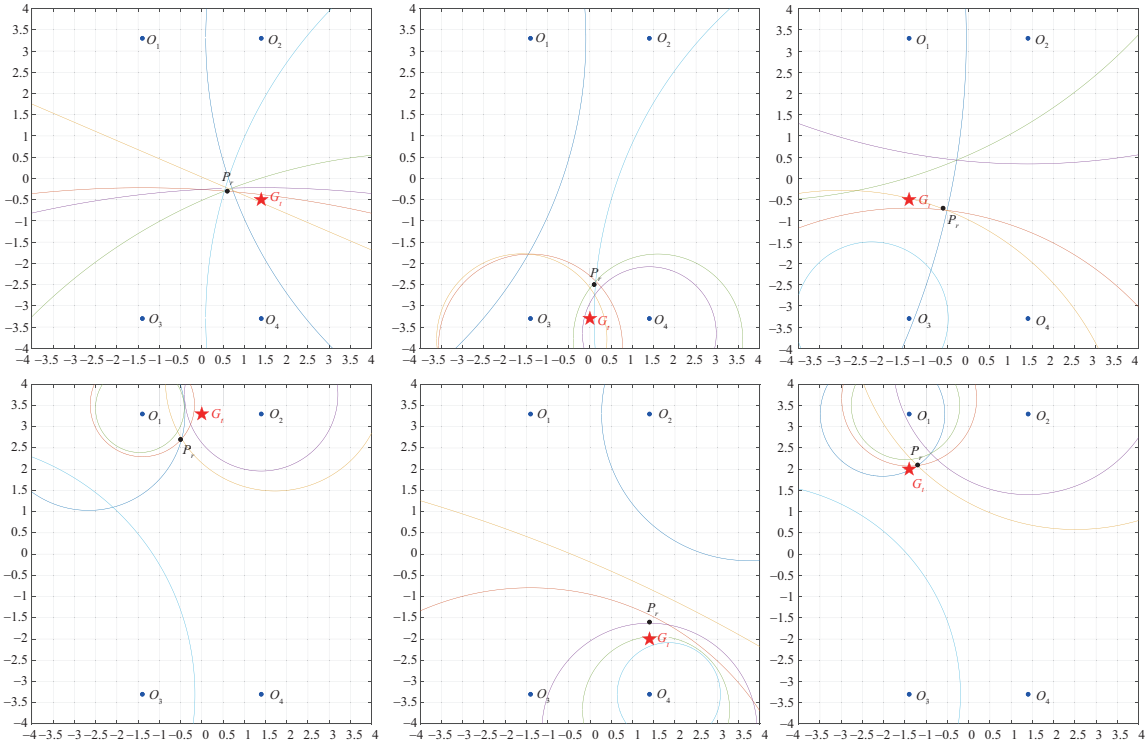

Figure 6 (Color online) The sound source localization processes with the traditional method in an indoor environment.

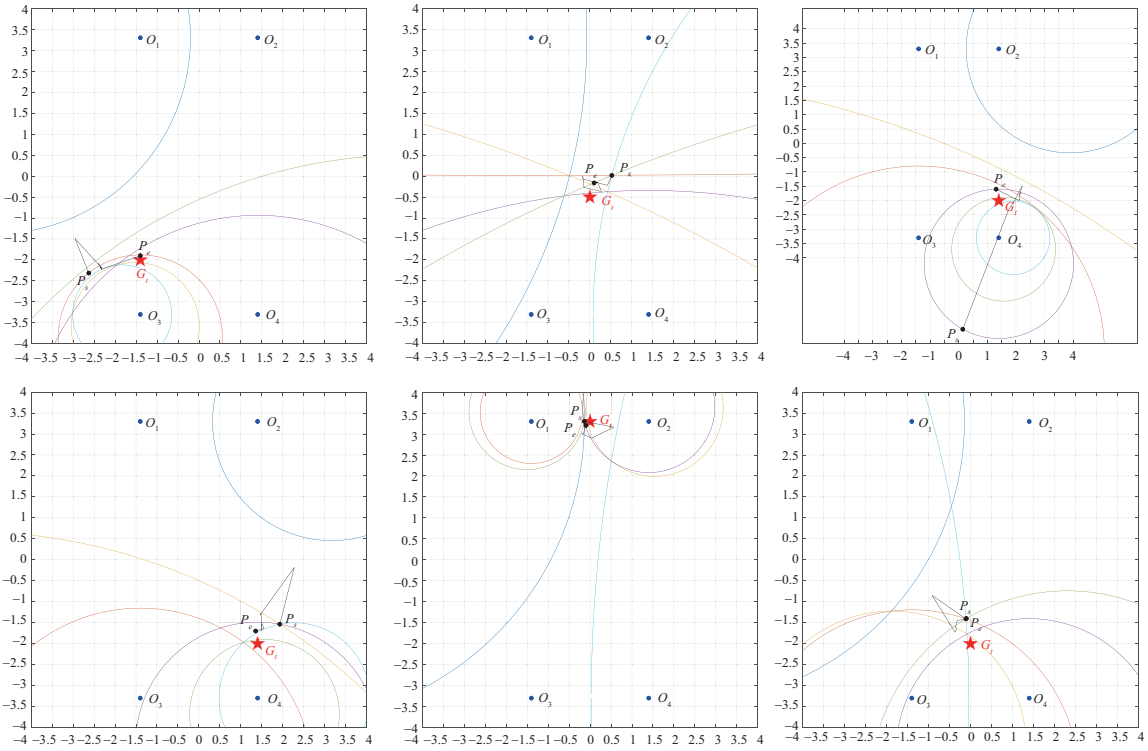

Figure 7 (Color online) The sound source localization processes with the proposed method in an indoor environment.

our method, respectively. In addition, the black line segments are the iterative processes from P_s to P_e. The results show that our method gradually approaches the target's ground-truth position from P_s to P_e, and the results of our method are more accurate than those of the traditional method.

Table 2 Experimental results with three acoustic sensors

Number	Target coordinate	Localization error of proposed method (m)				Localization error of traditional method (m)			
		SNR_{∞}	SNR_{30db}	SNR_{20db}	SNR_{10db}	SNR_{∞}	SNR_{30db}	SNR_{20db}	SNR_{10db}
1	(0, 3.3)	0.14	0.15	0.49	0.19	0.14	0.15	0.05	1.22
2	(−1.4, 2)	0.21	0.30	0.42	1.25	0.22	0.19	0.12	1.03
3	(0, 2)	0.32	0.34	0.33	0.17	0.31	0.30	0.23	1.46
4	(1.4, 2)	0.45	0.68	0.11	0.61	0.44	0.23	0.04	0.56
5	(−1.4, 0.5)	0.26	0.10	1.39	0.12	0.26	0.55	0.84	0.19
6	(0, 0.5)	0.21	0.25	0.60	2.16	0.21	0.30	1.32	1.42
7	(1.4, 0.5)	0.63	0.74	1.09	2.81	0.63	0.49	0.03	1.44
8	(−1.4, −0.5)	0.31	0.25	0.62	1.60	0.31	0.34	1.63	0.21
9	(0, −0.5)	0.70	1.16	0.21	0.89	0.70	0.64	1.43	0.27
10	(1.4, −0.5)	1.00	0.53	1.44	1.23	1.00	1.50	0.13	1.67
11	(−1.4, −2)	0.12	0.13	0.13	0.07	0.12	0.15	0.38	0.28
12	(0, −2)	1.62	1.43	0.77	0.69	1.61	1.54	1.37	2.01
13	(1.4, −2)	1.19	1.14	1.61	2.35	1.19	1.13	1.18	3.03
14	(0, −3.3)	2.02	2.05	1.24	2.41	2.02	2.02	2.38	2.64
	Average error	0.65	0.66	0.75	1.18	0.66	0.68	0.79	1.25

Table 3 Experimental results under SNR_{3db}

Number	Target coordinate	Localization error (m)	
		Our method	Traditional method
1	(0, 3.3)	1.38	1.33
2	(−1.4, 2)	0.52	1.50
3	(0, 2)	2.08	1.50
4	(1.4, 2)	1.20	0.55
5	(−1.4, 0.5)	2.11	0.93
6	(0, 0.5)	0.26	2.06
7	(1.4, 0.5)	2.81	3.54
8	(−1.4, −0.5)	1.93	2.85
9	(0, −0.5)	1.11	3.14
10	(1.4, −0.5)	2.14	1.80
11	(−1.4, −2)	0.06	2.85
12	(0, −2)	1.17	0.90
13	(1.4, −2)	0.13	2.31
14	(0, −3.3)	1.95	0.78
	Average error	1.35	1.86

5.3 Ablation studies

5.3.1 *Influence of sensor faults*

In practice, sensor faults may occur during measurements. In this part, we evaluate the performance of our method when one sensor faults. Experiments are conducted by removing the measurement data from one of the acoustic sensors (O_4). The experimental results are shown in Table 2. When a sensor faults, our method has a lower localization accuracy compared with those shown in Table 1. Moreover, major errors exist in the measurements 12, 13, and 14, where targets are located near O_4. This outcome is reasonable because sound energy circles are sparse near O_4. Furthermore, the average location errors of our method under different SNRs are smaller than those of the traditional method, which shows the effectiveness of our method.

5.3.2 *Influence of noisy area*

We conduct experiments under SNR_{3db} to clarify the performance of our method in a highly noisy area. The experimental results are shown in Table 3. The average positioning errors of our method and the traditional method in the highly noisy area are 1.35 m and 1.86 m, respectively. This result shows that our prior searching strategy brings an improvement of 0.51 m on the localization accuracy.

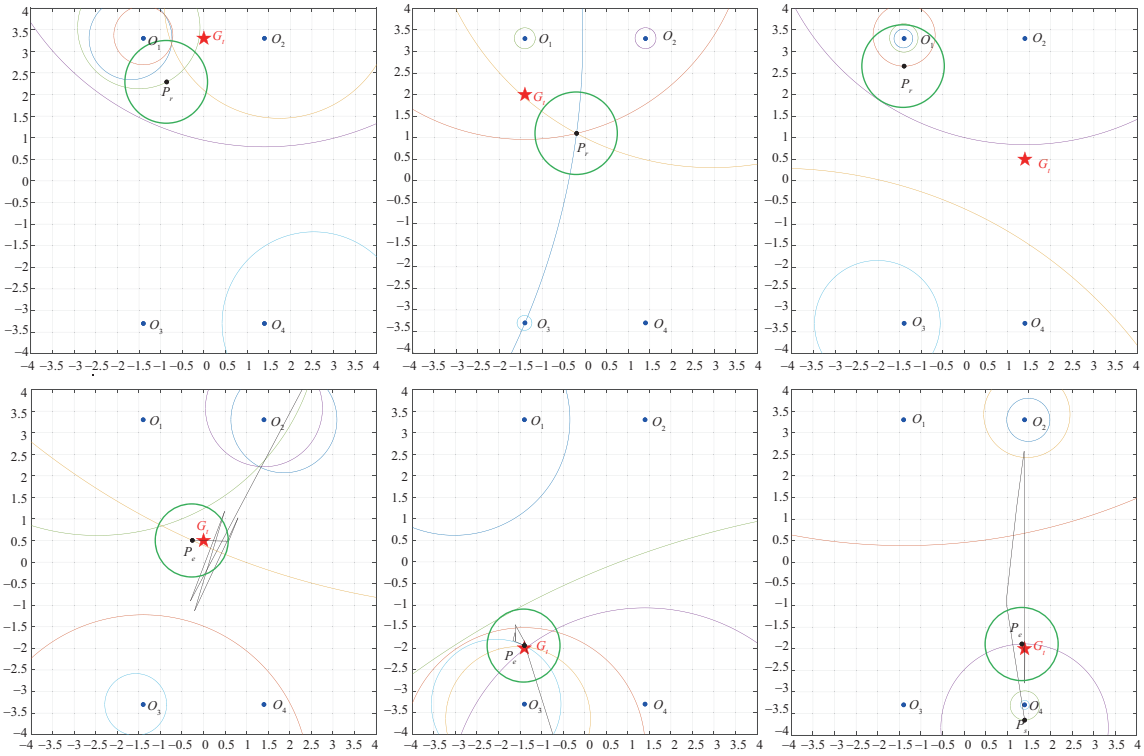

Figure 8 (Color online) Visualization results of different methods in a noisy area.

Table 4 Experimental results with different step sizes and prior areas

Method	Average localization error (m)				Mean error (m)
	SNR_{∞}	$\mathrm{SNR}_{30\mathrm{db}}$	$\mathrm{SNR}_{20\mathrm{db}}$	$\mathrm{SNR}_{10\mathrm{db}}$	
Tra-0.1	0.34	0.37	0.52	0.99	0.55
Tra-0.05	0.36	0.37	0.49	0.91	0.53
PAS-s	0.36	0.34	0.43	0.96	0.52
PAS-l	0.35	0.33	0.35	0.89	0.48

We further visualize the positioning results of the two methods in Figure 8. The green circles represent the confidence regions obtained by computing the standard deviations from Table 3 (0.86 m and 0.96 m for our method and the traditional method). In Figure 8, the first and second rows show the positioning results of the traditional method and our method, respectively. Compared with the traditional method, the positioning result of our method is more accurate, and the confidence region covers the target's ground-truth location more easily. One reason that explains this phenomenon may be that Eq. (14) contains many local minima under a noisy area, which confuses the traditional methods. Our method avoids this problem to some extent by setting the prior area to bound the potential existence area of a target in advance.

5.3.3 *Influence of step size and prior area*

We further analyze the influence of the step size and the size of prior area on the localization performance of different algorithms. Traditional methods with step sizes of 0.1 m and 0.05 m are named Tra-0.1 and Tra-0.05, respectively. Our proposed methods with prior area lengths of 2 m and 4 m are abbreviated as PAS-s and PAS-l, respectively. In addition, the step sizes in PAS-s and PAS-l are 0.1 m. The experimental results are shown in Table 4, where the mean localization errors of Tra-0.1 and Tra-0.05 are larger than those of PAS-s and PAS-l. Moreover, PAS-l has the best accuracy of 0.48 m.

The traditional method can achieve a more accurate positioning result with a smaller step size and larger search area but with a longer calculation time. For the traditional method, when the search area is 10 m $\times$ 10 m and the step size is 0.05 m, the average positioning time is 0.42 s. When the search area is 20 m $\times$ 20 m and the step size is 0.05 m, the average positioning time is 1.11 s. In this context,

our proposed method does not need to manually set the search area, and the average positioning time is 0.33 s.

6 Conclusion

This paper proposes a method to adaptively determine the prior area of a sound source target to reduce the solution space and calculation time. First, the initial search point is determined according to the received signals. Then, the prior location of the target is iteratively found through a limited number of steps. Next, a small prior area is set around the target's prior location. Finally, the precise location of the target is traversed in the small prior area using the WNLS method. A series of experiments are carried out with a rectangular array of acoustic sensors. The extensive experimental results and analyses show the superiority of our method compared with the traditional method. In the future, we will focus on designing a sound source localization method under practical conditions, including highly noisy areas and data loss.

Acknowledgements This work was supported in part by National Science Fund for Distinguished Young Scholars of China (Grant No. 62025301), Basic Science Center Project (Grant No. 62088101), and Key Program of National Natural Science Foundation of China (Grant No. 61933002).

References

1 Yin S, Zhou B, Yang M J, et al. Localizing object parts in 3D from a single image. Sci China Inf Sci, 2019, 62: 074101
2 Usman M, Keyrouz F, Diepold K. Real time humanoid sound source localization and tracking in a highly reverberant environment. In: Proceedings of the 9th International Conference on Signal Processing, Beijing, 2008. 2661–2664
3 Wang Q H, Ivanov T, Aarabi P. Acoustic robot navigation using distributed microphone arrays. Inf Fusion, 2004, 5: 131–140
4 Faraji M M, Shouraki S B, Iranmehr E, et al. Sound source localization in wide-range outdoor environment using distributed sensor network. IEEE Sens J, 2020, 20: 2234–2246
5 Yue X H, Deng F, Xu Y. Multidimensional zero-crossing interval points: a low sampling rate acoustic fingerprint recognition method. Sci China Inf Sci, 2019, 62: 019202
6 Wang F L, Wu D, Jin P, et al. A flexible skin-mounted wireless acoustic device for bowel sounds monitoring and evaluation. Sci China Inf Sci, 2019, 62: 202402
7 Makantasis K, Liapis A, Yannakakis G N. The pixels and sounds of emotion: general-purpose representations of arousal in games. IEEE Trans Affective Comput, 2021. doi: 10.1109/TAFFC.2021.3060877
8 Hao H-K, Liang H-M, Liu Y-W. Particle methods for realtime sound source localization based on the multiple signal classification algorithm. In: Proceedings of International Conference on Intelligent Green Building and Smart Grid, Taipei, 2014. 1–5
9 Gao S, Liu R, Wu X, et al. Eigen beam based sound source localization algorithms evaluation on a non-spherical microphone array. In: Proceedings of IEEE 2nd International Conference on Information Communication and Signal Processing, Weihai, 2019. 185–189
10 Do H T. Robust cross-correlation-based methods for sound-source localization and separation using a large-aperture microphone array. Dissertation for Ph.D. Degree. Providence: Brown University, 2011
11 Deng F, Guan S, Yue X, et al. Energy-based sound source localization with low power consumption in wireless sensor networks. IEEE Trans Ind Electron, 2017, 64: 4894–4902
12 Asono F, Asoh H, Matsui T. Sound source localization and signal separation for office robot JiJo-2. In: Proceedings of International Conference on Multisensor Fusion and Integration for Intelligent Systems, Taipei, 1999. 243–248
13 Nakamura K, Nakadai K, Asano F, et al. Intelligent sound source localization and its application to multimodal human tracking. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, 2011. 143–148
14 da Silva B, Segers L, Braeken A, et al. Runtime reconfigurable beamforming architecture for real-time sound-source localization. In: Proceedings of the 26th International Conference on Field Programmable Logic and Applications, Lausanne, 2016. 1–4
15 Salvati D, Drioli C, Ferrin G, et al. Beamforming-based acoustic source localization and enhancement for multirotor UAVs. In: Proceedings of the 26th European Signal Processing Conference, Rome, 2018. 987–991
16 Sun S, Ye Z. Robust adaptive beamforming based on a method for steering vector estimation and interference covariance matrix reconstruction. Signal Processing, 2021, 182: 107939
17 Li J Z, Guo F C, Jiang W L. Source localization and calibration using TDOA and FDOA measurements in the presence of sensor location uncertainty. Sci China Inf Sci, 2014, 57: 042315
18 Ho K C, Lu X, Kovavisaruch L. Source localization using TDOA and FDOA measurements in the presence of receiver location errors: analysis and solution. IEEE Trans Signal Process, 2007, 55: 684–696
19 Kim C-T, Choi T-Y, Choi B, et al. Robust estimation of sound direction for robot interface. In: Proceedings of IEEE International Conference on Robotics and Automation, Pasadena, 2008. 3475–3480
20 Badali A, Valin J-M, Michaud F, et al. Evaluating real-time audio localization algorithms for artificial audition in robotics. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, 2009. 2033–2038
21 Horiike D, Scheibler R, Kinoshita Y, et al. Energy-based multiple source localization with blinkies. In: Proceedings of Asia Pacific Signal and Information Processing Association Annual Summit and Conference, Auckland, 2020. 443–448
22 Tang R, Zuo Y, Liu W, et al. Efficient energy-based orthogonal matching pursuit algorithm for multiple sound source localization with unknown source count. Meas Sci Technol, 2022, 33: 045018
23 Lima M V S, Martins W A, Nunes L O, et al. A volumetric SRP with refinement step for sound source localization. IEEE Signal Process Lett, 2014, 22: 1098–1102

24 Sheng X H, Hu Y-H. Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. IEEE Trans Signal Process, 2004, 53: 44–53
25 Hafezi S, Moore A H, Naylor P A. 3D acoustic source localization in the spherical harmonic domain based on optimized grid search. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Shanghai, 2016. 415–419
26 You Y, Yoo J, Cha H. Event region for effective distributed acoustic source localization in wireless sensor networks. In: Proceedings of IEEE Wireless Communications and Networking Conference, Hong Kong, 2007. 2762–2767
27 Jiang X, Li S, Luo B, et al. Source exploration for an under-actuated system: a control-theoretic paradigm. IEEE Trans Contr Syst Technol, 2019, 28: 1100–1107
28 Li S, Guo Y. Distributed source seeking by cooperative robots: all-to-all and limited communications. In: Proceedings of IEEE International Conference on Robotics and Automation, Saint Paul, 2012. 1107–1112
29 Li S, Wang Z, Li Y. Using Laplacian eigenmap as heuristic information to solve nonlinear constraints defined on a graph and its application in distributed range-free localization of wireless sensor networks. Neural Process Lett, 2013, 37: 411–424
30 Saric Z M, Kukolj D D, Teslic N D. Acoustic source localization in wireless sensor network. Circ Syst Signal Process, 2010, 29: 837–856