IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 6, 2018, accepted November 21, 2018, date of publication December 4, 2018,
date of current version January 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2884777

Localization Performance Under Middle and Low Frequency Sound Source Based on Time Reversal Method in Enclosed Space

XIANGYANG ZENG, HUIYING MA, AND HAITAO WANG

School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China
Key Laboratory of Ocean Acoustics and Sensing, Ministry of Industry and Information Technology, Northwestern Polytechnical University, Xi'an 710072, China

Corresponding author: Xiangyang Zeng (zenggxy@nwpu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 11774291 and Grant 11604266.

ABSTRACT Multipath interference, reflection, sound absorption, and obstacles are important factors affect the sound source localization performance in enclosed space. Time reversal has the advantages of adaptive focusing, strong anti-reverberation capability, and the ability to overcome multipath effects, which are quite suitable for sound source localization in enclosed space due to its complex environment that many methods (e.g., beamforming, high-resolution spectral estimation, and the time delay of arrival) are not applicable. Therefore, this paper studies the localization performance under middle and low frequency sound sources cover single-frequency sinusoidal signal, bandwidth signal, and impulse signal through dual-channel matching based on time reversal method. When the frequency is below 1 kHz, the media's absorption is generally negligible. Therefore, this paper only studies frequencies within 1 kHz. In this paper, numerical simulations are used to verify the feasibility of dual-channel matching under middle and low frequency sound sources in ordinary enclosed space. Then, actual experiments are conducted in an ordinary office on the case of different sound sources. The numerical simulation and experiment results show that dual-channel matching method has satisfying localization effect. The localization effect of medium frequency signal is superior to low frequency signal. Bandwidth signals have better localization effect than single-frequency sinusoidal signals, and steady state sound source has better localization effect than impulse sound source. In general, dual-channel matching method has strong robustness and applicability for sound source localization in enclosed space.

INDEX TERMS Dual-microphone, enclosed space, middle and low frequency, sound source localization, time reversal.

I. INTRODUCTION

Sound source localization (SSL) using microphone arrays has been an active research topic since the early 1990 s [1]. It covers many practical applications, such as video conference, safety monitoring, smart home, aircraft cabin noise source localization, car phone system, and enclosed space sound tracking system. Based on these, demand on the enclosed space SSL is growing steadily. However, strong reverberation, acoustic attenuation, reflection, interference and diffraction and complex geometry structure contribute the difficulty on enclosed space SSL. Therefore, we need the localization method not only has high resolution, but also has strong robustness.

Existing SSL methods in enclosed space mainly include dual microphones based on binaural hearing mechanism and microphone array method. The microphone array method also covers the time delay of arrival (TDOA), beamforming (BF), high-resolution spectral estimation, and some updated algorithms based on those basic techniques combined with the particular array shapes. Those methods can get good spatial resolution in the large open environment [2]–[4]. The most common approach is TDOA [5]–[7]. TDOA has more applications in sound source location in enclosed space since it has low hardware demands and simple and quick algorithm, however, it cannot get the specific distance and merely can localize the direction of the sound source. In the reverberant environment, the signal received by the microphone not only has direct sound, but also reverberation sound, therefore, the localization performance is quite poor since this TDOA algorithm only considers the direct sound [8].

BF SSL technology has good performance in free field and low reverberation environment, however, its localization performance is quite suiTable unsuitable since the algorithm is largely affected by the array shape, and it also has poor localization performance in near-field [8]. High-resolution spectral estimation asks for smooth and time-invariant signal, however, most of the signals do not meet this requirement. Many hypothetical conditions are required if you take advantage of the high-resolution spectral estimation which is not suiTable unsuitable for practical applications [9]. Dual microphone method based on binaural hearing mechanism is mainly focused on theoretical study which is limited by localization precision [10]. There are many problems in the actual SSL in enclosed space, such as acoustic wave reflection, refraction and diffraction, the attenuation after encountering obstacles and wall surfaces, and the interference of background noise. From what has been discussed above, none of these methods are suiTable unsuitable for SSL in enclosed space. Therefore, we need to look for new suiTable unsuitable method which is time reversal. Existing SSL algorithms for small arrays still have two significant limitations: lack of range resolution, and accuracy degradation with increasing reverberation. Ribei *et al.* [11] take advantage of image method to conduct speaker localization with microphones arrays. Atmoko *et al.* estimate source direction and distances d from time of sound wave travel and distances at least three microphone sensors. They use different sound sources with different spectra, e.g., single frequency, multiple frequencies under 250 Hz, 500 Hz, 1 kHz and 2 kHz and male speech and different noise shapes. Much less than 1 for angle (direction) and distances are estimated at less than 4% inaccurate [12]. TDOA based methods with high sampling rate are used for 2 D high accuracy wideband SSL using three microphones [13]–[15]. Recently some papers were published which introduce 2 D SSL method using only two microphones in indoor cases using time delay estimation and interaural binaural level difference based methods simultaneously [16]–[18]. ILD based methods need to use source counting to find that one dominant source is active for high resolution localizing. Mandal *et al.* present Beep, an indoor location system that senses audible sound. Beep works with an accuracy of about 2 feet in more than 97% cases [19]. Liu *et al.* proposed a robust direction-of-arrival (DOA) estimation method for the sequential movement events of vehicles based on a small Micro-Electro-Mechanical System (MEMS) microphone array system. The simulations and experiments with frequency ranges from 0 to 4096 Hz and from 30 to 2500 Hz gained a good performance in emulating the DOA of a moving vehicle even in the case of severe wind interference [20]. Feng Guo et al introduced a sub-band direction-of-arrival (DOA) estimation method suitable for employment within an automatic bearing tracking system. The simulations and experiments with 3 KHz bandwidth also showed a good performance [21]. Antonello *et al.* proposes a novel method for joint source localization and dereverberation by interpolating the sound field using the measurements of a set of microphones and by solving an inverse problem that relies on a particular acoustic model. The localization result error is 4.5° [22]. Yingxiang Sun propose an sound source localization algorithm from 220 Hz to 3.4 kHz based on a probabilistic neural network, namely a generalized cross-correlation classification algorithm (GCA). Experimental results for adverse environments with high reverberation time T_{60} up to 600 ms and low SNR such as -10 dB show that the average azimuth angle error and elevation angle error by GCA are 4.6° and 3.1°, respectively [23]. Boaz *et al.* estimate the direction-of-arrival of a speaker in a room using direct path dominance test based on sound field directivity. Result demonstrate that the proposed method shows comparable performance to the original method in terms of robustness to reverberation and noise, and is about four times more computationally efficient for the given experiment [24].

Acoustic time reversal originated from the optical phase conjugation based on the transmission similarity of the sound waves and light waves [25]. It has been widely used in ultrasound, underwater and medical field. Fink developed the concept of time reversal mirror when he studied the phase conjugation of bandwidth signals in the field of ultrasound [26]. Acoustic time reversal (TRM) is made of a 1-D or 2-D array of N reversible transducers. Such a time reversal parallel procedure allows one to convert a divergent incident wave issued from a pointlike source into a convergent reflected wave, focusing on the source. Unlike the ordinary mirror that produces the virtual image of an acoustic object, the TRM produces the ''real'' image. The method is efficient even if there are weak inhomogeneities between the sound source and TRM. Fink also demonstrated the theory of the adaptive focusing of TRM. TR cavity was proposed by Fink in a non-dissipative medium and iterative TR technology in multi-target media. The results of ultrasound wave TR focus taking advantage of TRM array show that TRM can compensate the loss caused by multipath [26]–[29], Jackson *et al.* proved that TRM can be used to get the high-resolution focusing effect in underwater acoustic. TRM has advantages+ of slacking the influence caused by multipath effect and amend the inhomogeneity of the medium and the waveform distortion caused by acoustic interference [29]. One application of TRM technology on underwater acoustic field is underwater target detection due to its strong anti -reverberation capability. Seabed reverberation interference is particularly severe, therefore the time-reversal method is more suiTable unsuitable than other methods for target detection. TR theory and its first applications have been mainly developed in the work of Fink in the 1990 s [30], [31].

Whereas TR technique is widely spread, it has rarely been used in aeroacoustics up to now. SSL in enclosed space was first introduced by Draeger and Fink [32]. He studied one channel time reversal of elastic waves in a chaotic 2D-silicon

cavity. Numerical simulations illustrate the time reversal process, experiments carried out in silicon wafers show that it is possible to obtain an excellent temporal and spatial focusing quality'1. The sound source is a short pulse [33], [34]. Then Fink studied the time reversal of electromagnetic waves, and result shows that frequency bandwidth and the spectral correlations of the field within the cavity determine the focus on quality [35]. Yon *et al.* [36] took advantage of TRM to perform the localization experiments in an enclosed environment, and localization results show that TRM is effective in enclosed space sound localization. Rudolf Sprik conducted initial audio experiments in a reverberating room, and results showed that time reversal techniques is a powerful approach in acoustics to control multiple scattering of waves in complex systems [37]. Conti *et al.* [38] performed near-field time-reversal from the phase of the spatial spectrum of the field recorded on an array around the original probe source using an analytical continuation for the amplitude of the spatial spectrum. Following theory, $\lambda/20$ resolution is experimentally demonstrated with audible acoustic wave fields in the air. Stefan Catheline studied acoustic source localization model using in-skull reverberation and time reversal, and the authors anticipate that this general antenna like concept can be applied to many animals that use sound localization as well as to future design for microphone devices or sonars [39]. Padois *et al.* first considered monopolar sources that are either monochromatic or have a narrow or wide-band frequency content in a wind-tunnel flow. The source position estimation through TR method is well-achieved with an error inferior to the wavelength. An application to a dipolar sound source shows that this type of source is also very satisfactorily characterized. The acoustic source has a frequency range of 1.5-18 kHz [40]. The conceptual equivalence of Images Method was demonstrated through directly implementing the rigid-wall condition during TR for source localization/characterization according to the conventional beam-forming method, and results showed that it is highly comparable to those obtained using TR for the test-case of nonconvecting sources [41]. Lonzaga *et al.* [42] studied the time reversal for localization of sources of infrasound signals in a windy stratified atmosphere, and analyses show that the method can be used to substantially enhance estimates of the source back-azimuth and the source-to receiver distance.

Ma *et al.* [43] proposed a novel SSL method in enclosed space called dual-channel matching (DCM) based on advantages of overcoming multipath effects, adaptive focusing ability and anti-reverberation capability of TR method. In the second part, theory of dual-channel matching and TRM localization method is elaborated. Simulations are conducted to study the localization performance under different sound source conditions which cover single-frequency sinusoidal signals, impulse signal and bandwidth signals in the third part. The fourth part is the actual experimental verification.

II. THE THEORY OF TIME REVERSAL LOCALIZATION METHOD

A. THE THEORY OF TR METHOD

Fig 1 is a schematic diagram of the distribution of grids and microphones in a common room. Divide an interested area in the room into m grids with a grid space of d cm.

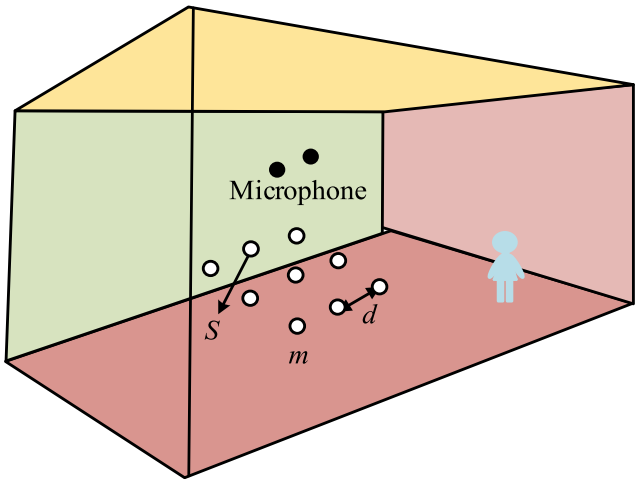


FIGURE 1. The distribution of the grids and dual-microphone.

Suppose the location of the center of the grid S to be the position of the sound source. In a common room, the propagation of acoustic waves can be regarded as a linear time-invariant system under the assumption of small amplitude acoustic waves. Ignoring the background noise, the frequency spectrum received by the dual-microphone emitted by the sound source s(t) located at S_i can be written as

$$F_{ij}(\omega) = S(\omega) \cdot H_{ij}\left(s, s_j, \omega\right) \tag{1}$$

where $S(\omega)$ is the frequency spectrum of the signal, $H_{ij}\left(s, s_j, \omega\right)$ $(i = 1, 2, \ldots n, j = \alpha, \beta)$ is the frequency response function of the actual sound channels. Time reversal in time domain corresponds to the phase conjugation in frequency domain. Therefore, the time reversed format of (1) is

$$F_{ij}^{*}(\omega) = H_{ij}^{*}\left(s, s_j, \omega\right) \cdot S^{*}(\omega) \tag{2}$$

Then, the TR signal is re-transmitted back through virtual channels:

$$F_{ij}^{\prime}(\omega) = H_{ij}^{*}(\omega) \cdot S^{*}(\omega) \cdot H_{ij}^{\prime}(\omega) \tag{3}$$

$H_{ij}^{\prime}(\omega)$ denotes the frequency response function of the virtual channel obtained by computation or measurement. The virtual channel will be identical with the actual sound channel without modeling error.

Conduct the cross-correlation of (3)

$$R(\omega) = F_{i\alpha}^{\prime}(\omega) \cdot F_{i\beta}^{\prime *}(\omega) \tag{4}$$

When the virtual channel is the same with the actual channel, formula (4) can be written as:

$$R(\omega) = H_{i\alpha}^{*}(\omega) \cdot H_{i\alpha}(\omega) \cdot H_{i\beta}(\omega) \cdot H_{i\beta}^{*}(\omega) \cdot S^{*}(\omega) \cdot S(\omega) \tag{5}$$

From the channel point, spatial gain reaches the maximum once the processing of TRM achieves channel matching, i.e., the main correlation peak of $R(\omega)$ is obtained when virtual channel and the actual channel is perfectly correlated.

B. THE THEORY OF DUAL CHANNEL MATCHING METHOD BASED ON TR

Since the sound source signal s(t) is unknown, formula (1) is handled as follows

$$X = \frac{F_{i\alpha}(\omega)}{F_{i\beta}(\omega)} = \frac{H_{i\alpha}(s, s_\alpha, \omega)}{H_{i\beta}(s, s_\beta, \omega)} \tag{6}$$

The division in frequency domain may result in drastic changes at some frequencies which will generate localization failing. The algorithm is improved by converting division to multiplication to overcome the defect

$$F_{i\alpha}(\omega) \cdot H_{i\beta}(s, s_\beta, \omega) = F_{i\beta}(\omega) \cdot H_{i\alpha}(s, s_\alpha, \omega) \tag{7}$$

Formula (7) indicates that the relationship of dual channel impulse responses from the sound source to microphones can be expressed by Fourier transform of signals received by microphones. Thus, sound source signal s(t) is unnecessary for localization as long as the signal received by the microphone is known.

Therefore, the SIR between every grid center to the fixed dual microphones can be measured in advance, then acquire channel response through Fourier transform

$$\boldsymbol{G}(f) = \begin{bmatrix} (G_{1\alpha}(f), G_{1\beta}(f)) & \cdots & (G_{m\alpha}(f), G_{m\beta}(f) \\ & \ddots & \vdots \\ & & (G_{n\alpha}(f), G_{n\beta}(f)) \end{bmatrix} \tag{8}$$

$$\Phi_i = F'_{r\alpha}(\omega) \cdot G_{i\beta}(\omega), \Psi_i = F'_{r\beta}(\omega) \cdot G_{i\alpha}(\omega) \tag{9}$$

The cross-correlation operation between Φ, Ψ is

$$\boldsymbol{R} = \text{corrcoef}(\boldsymbol{\Phi}, \boldsymbol{\Psi}) = \frac{\sum_f (\Phi_i - \bar{\Phi})(\Psi_i - \bar{\Psi})}{\sqrt{\sum_f (\Phi_i - \bar{\Phi})^2} \sqrt{\sum_f (\Psi_i - \bar{\Psi})^2}} \tag{10}$$

where corrcoef is the calculation of Pearson correlation coefficient, $\overline{\boldsymbol{\Phi}}$, $\overline{\boldsymbol{\Psi}}$, represent average values of $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$, respectively.

Calculate $\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_n$ based on $\boldsymbol{\Phi}_i, \boldsymbol{\Psi}_i$ through cross-correlation operation to find the maximum of correlation coefficient $\boldsymbol{R_i}$, where the location number i is defined as:

$$i = \arg\max_i \boldsymbol{R_i} \tag{11}$$

Then the location S_i of the sound source is deemed to be situated at the center of the grid which is corresponding to the maximum. So far an unknown SSL is accomplished.

The DCM method based on time reversal is a space matching method, which is different from the traditional localization estimation methods. Therefore, this paper proposes a simple indicator to describe the localization performance

$$R_c = \frac{R_{f-peak}}{R_{s-peak}} \tag{12}$$

R_{f-peak} is the first peak value of the correlation function, R_{s-peak} is the second peak value of the correlation function. Obviously, the larger the R_c, the more prominent the first peak value, the better the matching degree and the higher localization accuracy.

III. NUMERICAL SIMULATIONS

When the frequency is below 1 kHz, the media's absorption is generally negligible. Therefore, this paper only studies frequencies within 1 kHz. Performance of the proposed method under different sound sources which conclude single-frequency sinusoidal signal (125 Hz, 250 Hz, 500 Hz, 1000 Hz), bandwidth sinusoidal signal (125-250 Hz, 250-500 Hz, 500-1000 Hz, 125-1000 Hz) and square wave pulse signal are evaluated through dual-microphone simulations. These signals are generated by MATLAB simulation. This paper also studies the actual explosive sound.

A. NUMERICAL SIMULATIONS OF DUAL-CHANNEL MATCHING METHOD UNDER SINGLE FREQUENCY SINUSOIDAL SIGNALS

The spatial impulse response is obtained through simulation. Parameter settings are as follows: the room size is 1.2 m $*$ 1.4 m $*$ 1.6 m, the sound absorption coefficient of six walls are all set to 0.1. The sound source is located at (0.6 m, 0.7 m, 0.8 m). We also place two microphones at (0.4 m, 0.5 m, 0.8 m), (0.8 m, 0.5 m, 0.8 m). Divide 100 grids in the middle of the room. All grids are on a plane of 0.8 m high. The grid space is 5 cm. The localization effect of this method under the mid-low frequency sinusoidal signal (125 Hz, 250 Hz, 500 Hz, 1000 Hz) are studied. The SNR of all simulation experiments is set to 10 dB. The results are shown in Fig 2 and Fig 3. Table 1 shows R_c values at different frequencies.

TABLE 1. The average at R_c different frequencies.

f/Hz	125	250	500	1000
R_c	1.0910	1.2347	1.1130	1.2189

We can see that the dual-microphone localization method has satisfactory localization accuracy since the sound source position can be accurately found at four frequencies from Fig 2. We can see that the side lobe of the low-frequency sound is significantly higher than the high-frequency sound in terms of localization results from Fig 3. Overall, the localization performance of high-frequency sounds is superior to the low-frequency sound.

Since there is a total of 100 grids, that is, 100 positions, the result of one location may not represent results of all locations, we conduct 20 Monte Carlo experiments about 100 locations. The R_c values at different positions are shown

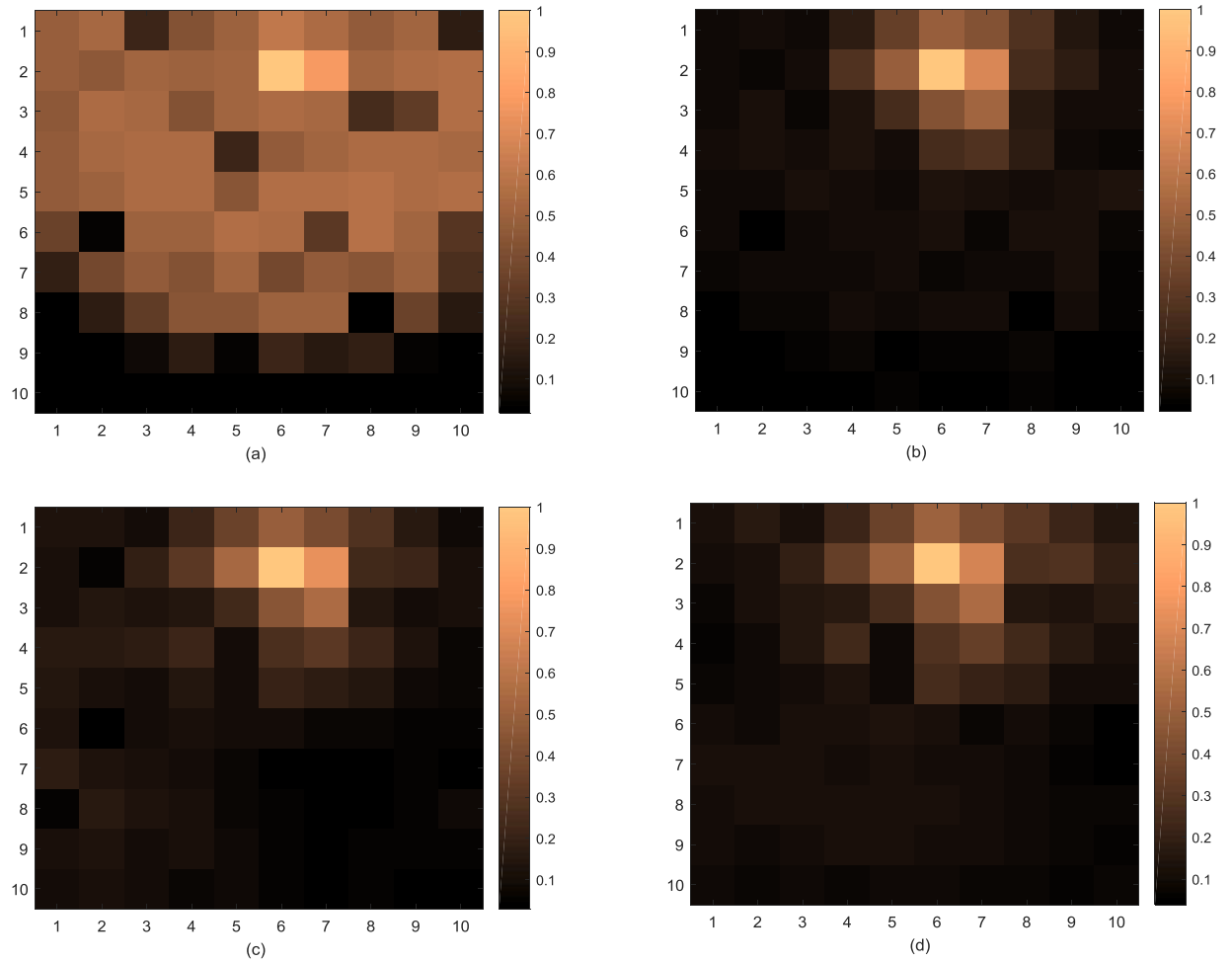


FIGURE 2. Localization results of single-frequency sinusoidal signal. (a) 125 Hz. (b) 250 Hz. (c) 500 Hz. (d) 1000 Hz.

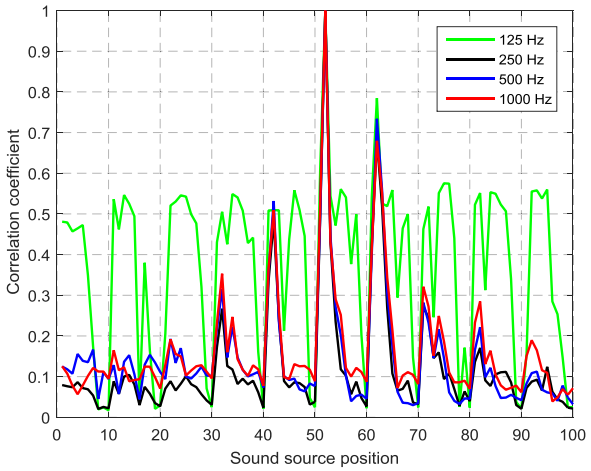


FIGURE 3. Correlation coefficient of different positions.

in Figure 4. The coordinates of the x-axis only represent 20 positions, not the specific position number. The average R_c at different frequencies is shown in Table 1.

We observed that R_c values at different positions are a little different, and only a few unique positions have significantly larger R_c values than other positions, but there is no regularity to follow in Fig 4 and Table 1. Overall, this method has better localization performance than the remaining three signals under the 250 Hz sinusoidal signal. In general, the localization performance is the worst when the sound source is a 125 Hz sinusoidal signal since the wavelength of the 125 Hz signal is long and scattering, diffraction and other phenomena cannot be ignored. We can see from Table 1 that the R_c value of 250 Hz sound source is the largest, which can be considered as the best localization performance. There is only a slight difference between the R_c values of the 1000 Hz and 500 Hz sound source, only 0.1059.

B. NUMERICAL SIMULATIONS OF DUAL-CHANNEL MATCHING METHOD UNDER BANDWIDTH SIGNALS

Room, grid, and microphone settings are the same with single frequency sinusoidal signal. The localization effect of this method under the mid-low frequency (125-250 Hz,

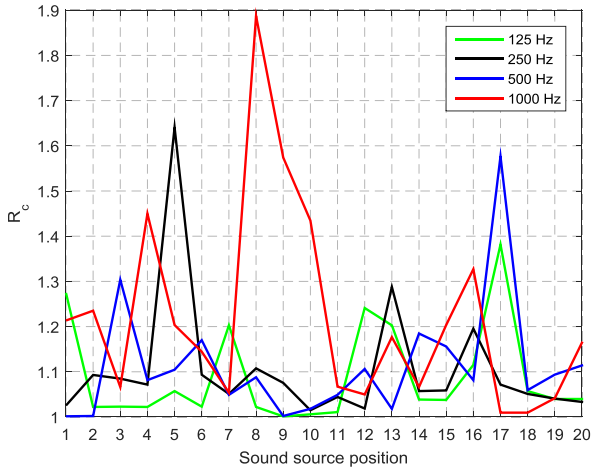


FIGURE 4. R_C values at different positions.

250-500 Hz, 500-1000 Hz, 125-1000 Hz) are studied. The results are shown in Fig 5. Fig 6 shows the correlation coefficient of different positions at bandwidth signals.

We can see from Fig 5 and Fig 6 that the dual-channel matching method has a better localization effect on the bandwidth signal, and the side lobe of the high-frequency signal is significantly lower than the low-frequency signal. For the localization performance of the dual-channel matching method, the high frequency signal is better than low frequency signal.

The R_c values at different positions are shown in Fig 7. The coordinates of the x-axis only represent 20 positions, not the specific position number. The average R_c at different frequencies is shown in Table 2.

TABLE 2. The average R_C at bandwidth signals.

f/Hz	125-250	250-500	500-1000	125-1000
R_c	1.1561	1.2991	1.2734	1.2679

Fig 7 and Table 2 show that the R_c values of the 500-1000 Hz sound source signal and the 125-1000 Hz sound source signal are not much different, which is only 0.0055. The 125-250 Hz sound source signal has the smallest R_c value and is the only one lower than 1.2. The 250-500 Hz sound source signal has the biggest R_c value. The remain frequency band have similar localization performance. Relatively speaking, localization performance of 125-250 Hz sound source is poor, however, it can also find the sound source position accurately. In general, the localization results R_c of the bandwidth signal vary slightly at different positions in the same frequency band, and there is no regularity. The localization result of the high frequency signal is better than that of the low frequency signal.

C. NUMERICAL SIMULATIONS OF DUAL-CHANNEL MATCHING METHOD UNDER IMPULSE SIGNALS

Room, grid, and microphone settings are the same with single frequency sinusoidal signal. We also take advantage of the square wave pulse signal and explosive sound to verify the performance of this method. Square wave pulse signals are generated by simulation at frequencies of 125 Hz, 250 Hz, 500 Hz, 1000 Hz. The result is shown in the Fig 8. Fig 9 shows the correlation coefficient of different positions.

Fig 8 and Fig 9 show that dual-channel matching method can accurately find the position of the square wave pulse sound source. The localization result of the explosion sound source signal is obviously better than the localization result of the square wave pulse sound source signal.

Fig 10 shows R_c values under 20 Monte Carlo experiments. The coordinates of the x-axis only represent 20 positions, not the specific position number. Table 3 shows the average R_c at bandwidth signals.

TABLE 3. The average R_C at impulse signals.

f/Hz	125	250	500	1000	Explosive sound
R_c	1.1498	1.1591	1.3605	1.6080	2.0078

Fig 10 and Table 3 indicate that the DCM method has better localization performance when the sound source is explosive sound. For the remaining four pulse sound sources, there is no regularity in the localization results. The average R_c of explosive sound is much bigger than 125 Hz square wave pulse, which up to 0.8580.

We can see from the above localization results that single frequency sine signals of 250 Hz, 500 Hz, bandwidth signal of 250-500 and explosive sound have relatively good localization performance. Therefore, this paper also compares the localization performance of the DCM method for single frequency sinusoidal signals (250 Hz and 500 Hz), bandwidth signal (250-500 Hz) and explosive sound. The comparison result is shown in Fig 11.

We can see from Fig 11 that the second peak is lowest at 250-500 Hz (0.5), the second peak at 500 Hz (0.67) is slightly lower than 250 Hz (0.65), and the second peak of explosive sound (0.7) is highest. The overall localization performance of 250-500 Hz bandwidth signal is better than 500 Hz single-frequency sinusoidal signal, poorer than 250 single-frequency sinusoidal signal. In general, the localization performance of bandwidth signal is better than that of single frequency sinusoidal signal, and the correlation coefficient of positions which are not the sound source position are not very high, so the dual-channel matching method has good localization performance.

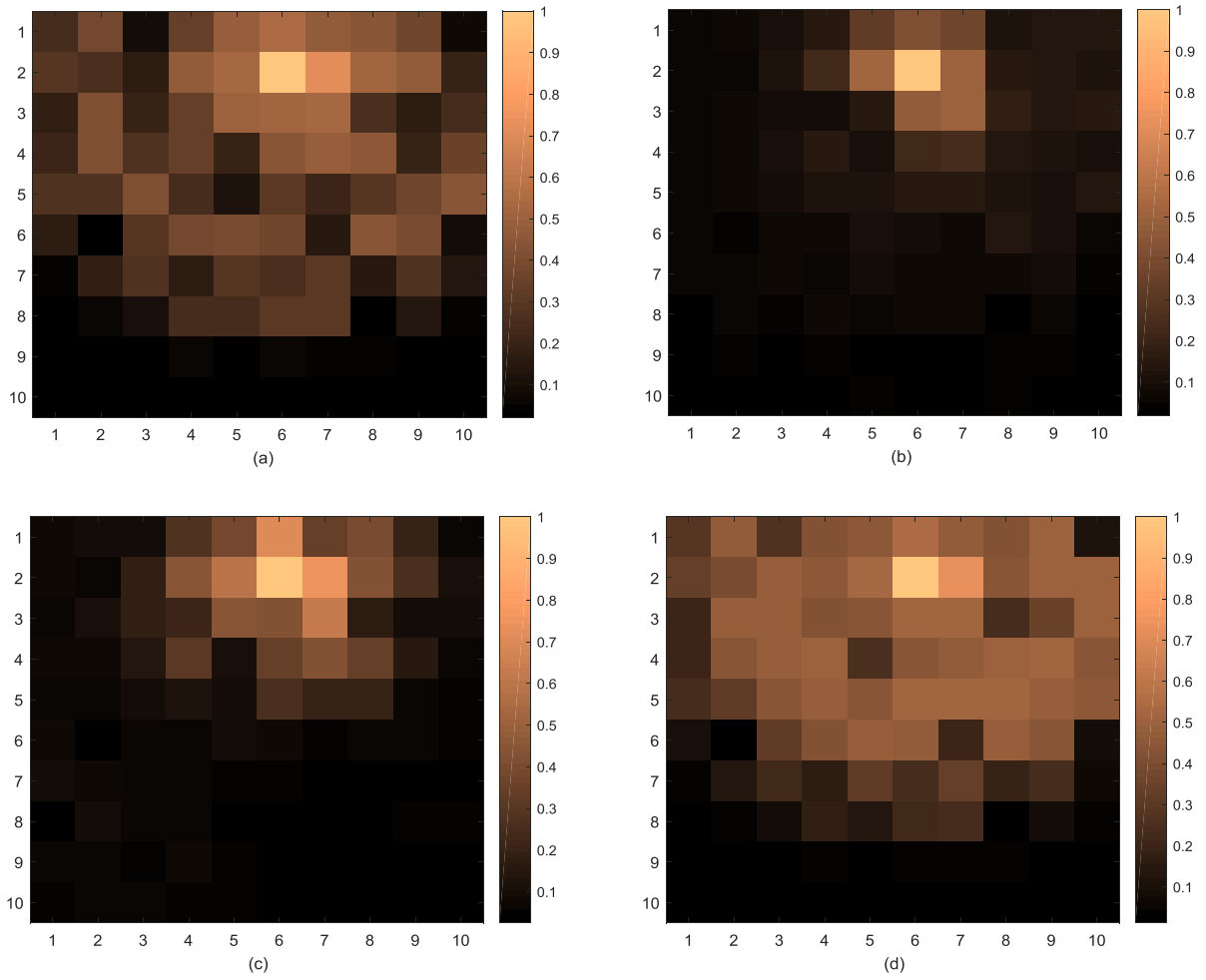


FIGURE 5. Localization results of bandwidth signal. (a) 125-250 Hz. (b) 250-500 Hz. (c) 500-1000 Hz. (d) 125-1000 Hz.

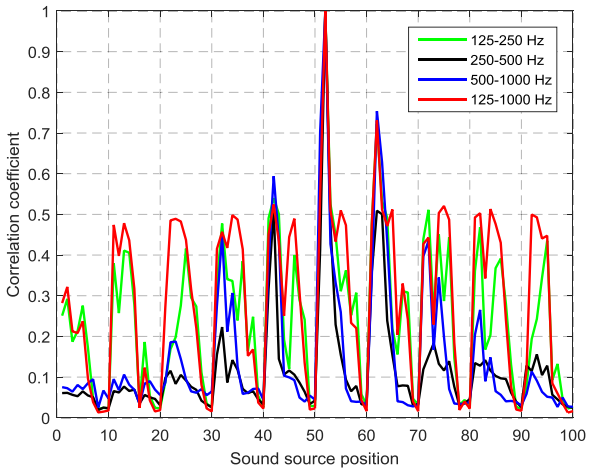


FIGURE 6. Correlation coefficient of different positions at bandwidth signals.

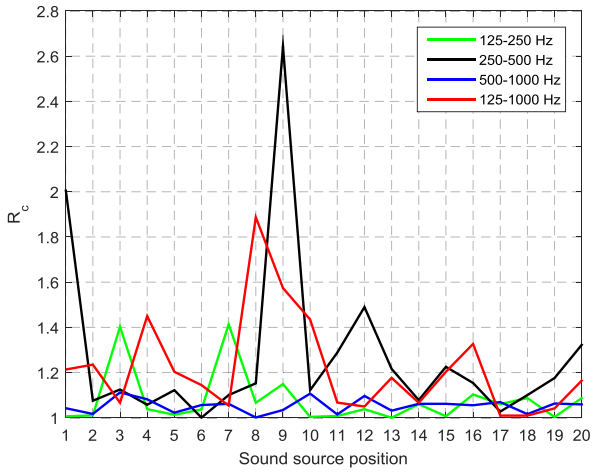


FIGURE 7. R_c values at different positions.

D. NUMERICAL SIMULATIONS OF HIGH-RESOLUTION SPECTRUM ESTIMATION

We also employ high-resolution spectrum estimation to conduct simulation experiments under 500 Hz sinusoidal signal and 250-500 Hz bandwidth signal for comparison. The room size is 3 m $*$ 4 m $*$ 5 m, the sound absorption coefficient is 0.3. The array includes 15 microphones, and the space of every two microphones is 10 cm. The sound source is

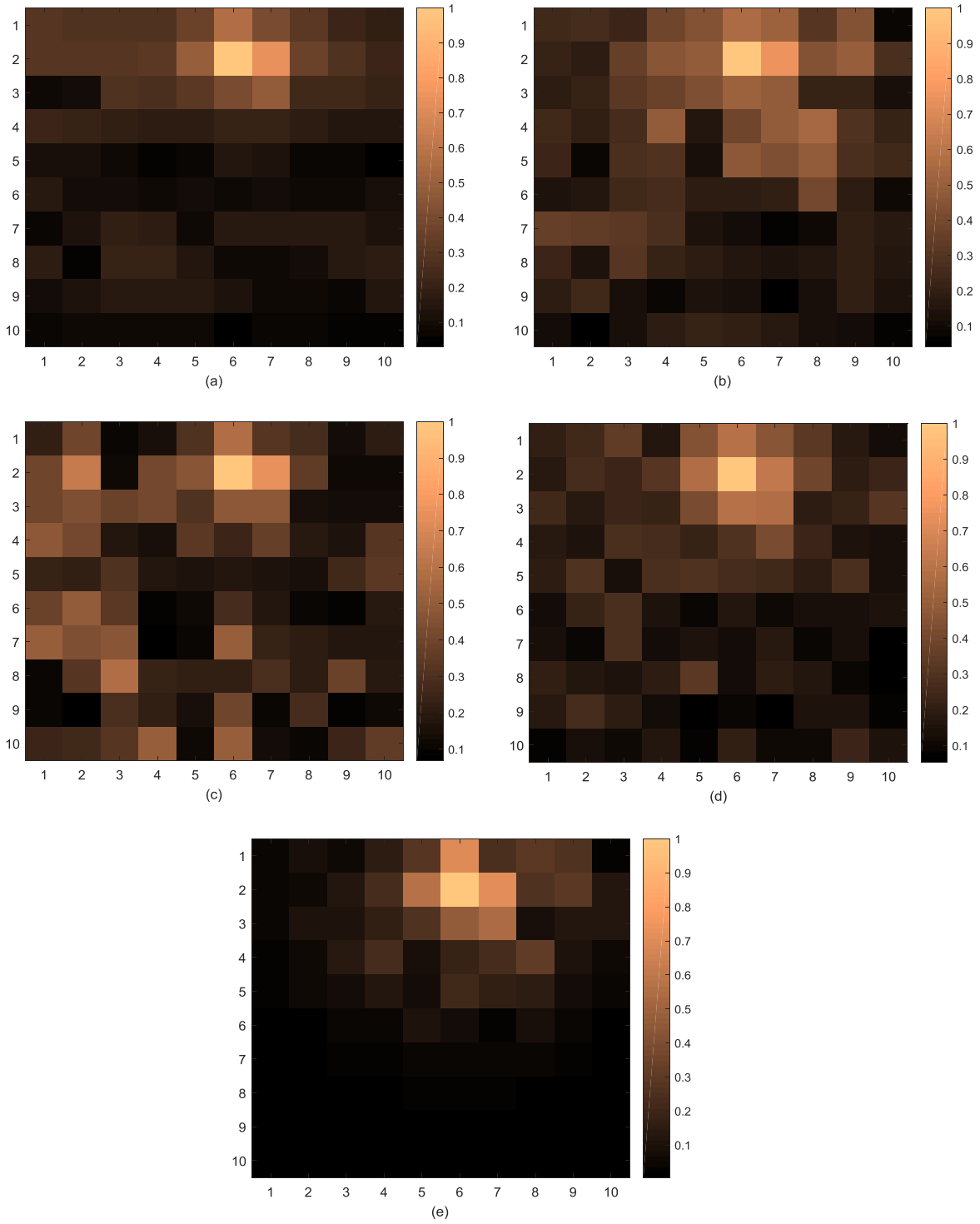


FIGURE 8. Localization results of dual channel matching for square wave pulse signal. (a) 125 Hz. (b) 250 Hz. (c) 500 Hz. (d) 1000 Hz. (e) Explosive sound.

located at 27 degrees and 1.6 m distance from reference array element. Single frequency sinusoidal signal and bandwidth signal localization results are shown in Fig 12. Green circle indicates the actual position of the sound source.

We can see from Fig 12 that side lobe value is so high that the position cannot be determined. Compared with DCM method, high-resolution spectrum estimation has poor localization performance.

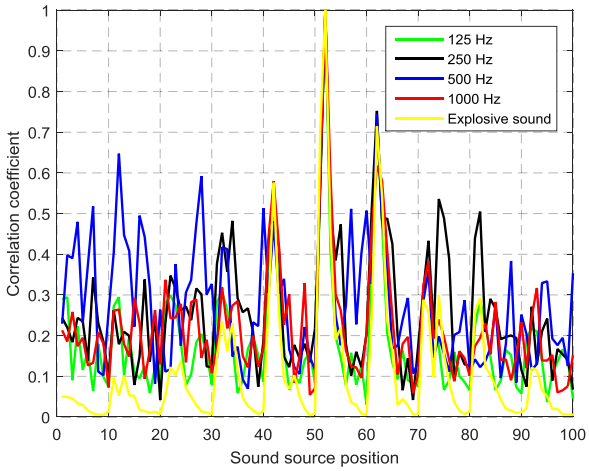


FIGURE 9. Correlation coefficient of different positions.

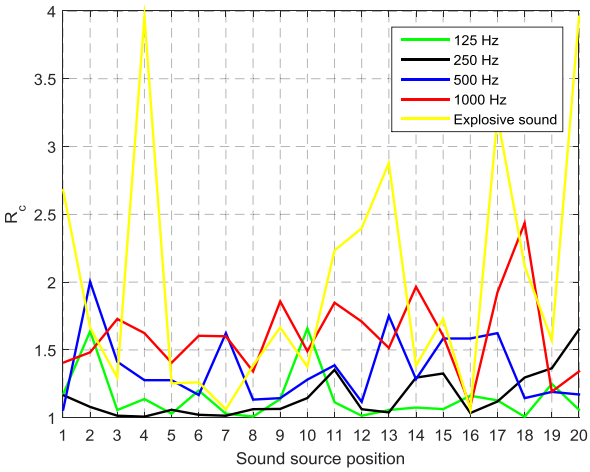


FIGURE 10. R_c values at different positions.

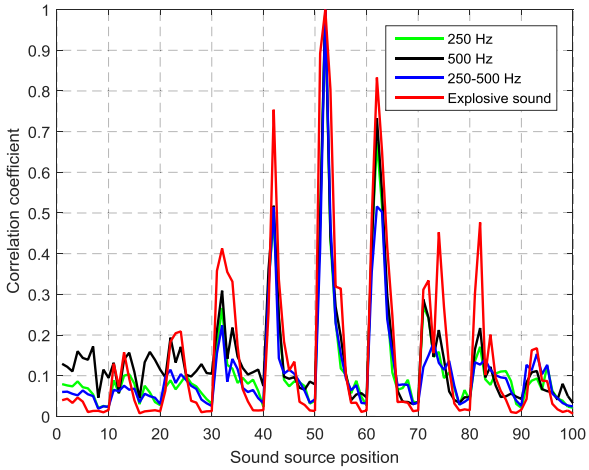


FIGURE 11. The localization result comparison of single frequency sinusoidal signal, bandwidth signal and explosive sound.

In general, the hypothetical simulation conditions are ideal since the complex environments such as obstacles and background noise in the room are ignored. Therefore, the simulation results will be better than the actual experimental results. The simulation result of high-resolution spectrum estimation is not satisfying, so actual experiments are not conducted to verify and compare.

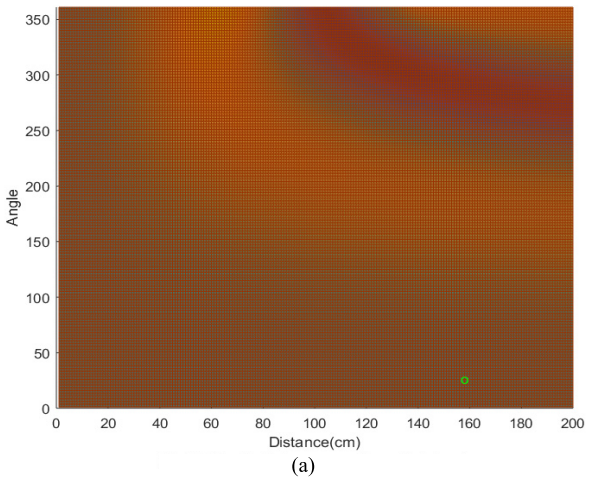


FIGURE 12. Single frequency sinusoidal signal and bandwidth signal localization results. (a) 500 Hz single frequency sinusodial signal. (b) 500-1000 Hz bandwidth signal.

IV. EXPERIMENTS BASED ON DUAL-CHANNEL MATCHING METHOD

This section describes the actual experiment process and results in detail. The spatial impulse response is obtained by acoustic field simulation software Dirac. The software is based on geometric acoustics that combines image method and ray tracing method to simulate the sound field. The spatial pulse response from the sound source to dual microphones measured by Dirac software is shown in Fig 13.

We can see from Fig 13 that the spatial impulse response of the sound source to dual microphones is different. Theoretically speaking, the spatial impulse responses of different points to the same point in a closed space are different. The farther the distance is, the more complex the environment is, and the greater the difference of spatial impulse response is. Therefore, the DCM method also has good localization performance in complex closed environment.

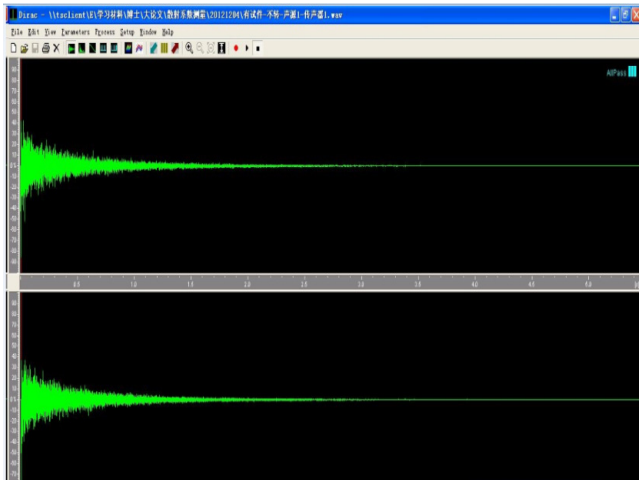


FIGURE 13. Spatial pulse response from the sound source to dual microphones.

We conduct experiments to study the dual-channel matching method in an ordinary office. The experimental system covers dual microphones, sound card, DIRAC software, a loudspeaker and a computer. The sound sources used for the tests are generated through simulation. Each sound source lasts 10 seconds. Specific area of the room is divided into $3 \times 5 = 15$ grids whose space is 30 cm. The sound source signal is emitted by a loudspeaker with a size of 10 cm$\times$15 cm$\times$20 cm. The layout of the room, experimental installation and grids is shown in Fig 14.

FIGURE 14. The experiment room.

We place the sound source on each grid and conduct the localization experiments according to the above

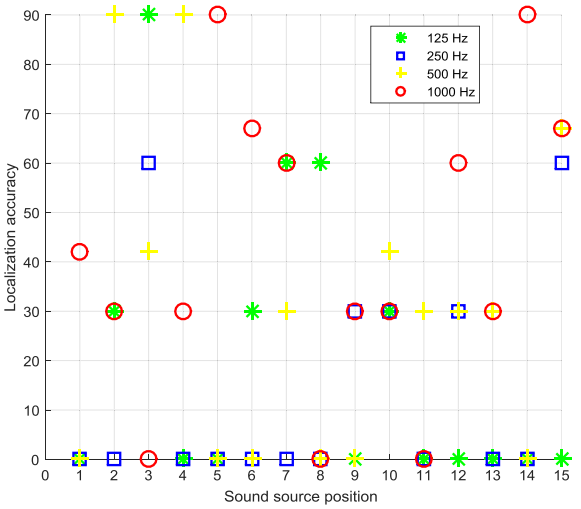


FIGURE 15. The accuracy of SSL under single-frequency sinusoidal signals.

TABLE 4. The average R_c at different frequencies.

f/Hz	125	250	500	1000
R_c	1.1189	1.1613	1.1549	1.1505

experimental procedures. The accuracy of SSL under single-frequency sinusoidal signals is shown in Figure 15. First, find the R_c of each grid, and then average the values of the 15 grids to get the final R_c. Table 4 shows the average R_c of all grids under single-frequency sinusoidal signals.

The localization accuracy of 0 cm means that the localization result is just the grid position where the sound source is located. When the localization accuracy is d cm, the localization result is d cm away from the actual position of the sound source. Seen from Fig 15, localization accuracy of 250 Hz and 500 Hz is better than 125 Hz and 1000 Hz. The sound source localization accuracy of 250 Hz is below 60 cm, and the other several sound sources are at 90 cm. Table 4 has the same conclusion. The side lobe of 250 Hz and 500 Hz is lower than 125 Hz and 1000 Hz, and the side lobe of 1000 Hz is lower than 125 Hz.

The accuracy of SSL under bandwidth signals is shown in Fig 16. The average R_c are shown in Table 5.

TABLE 5. The average R_c at bandwidth signals.

f/Hz	125-250	250-500	500-1000	125-1000
R_c	1.2864	1.2947	1.2890	1.1983

We can see from Fig 16 that as the localization accuracy decreases with the increase of frequency. Under 125-250 Hz, 13 positions have no errors in the localization result, and the localization accuracy of the remain-

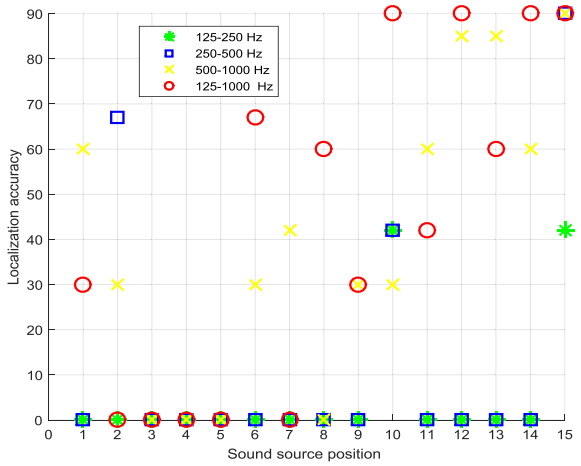


FIGURE 16. The accuracy of SSL under bandwidth signals.

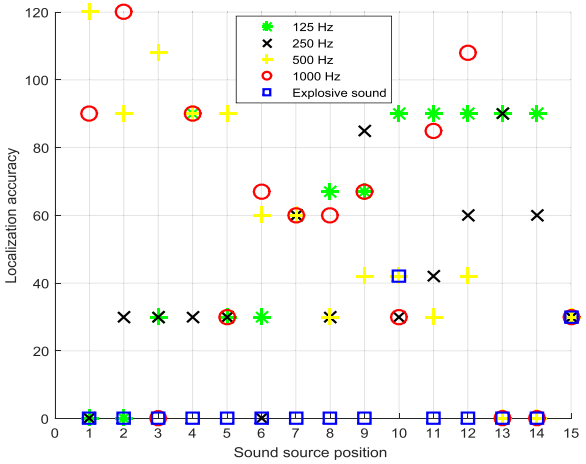


FIGURE 17. Localization accuracy of impulse sound source.

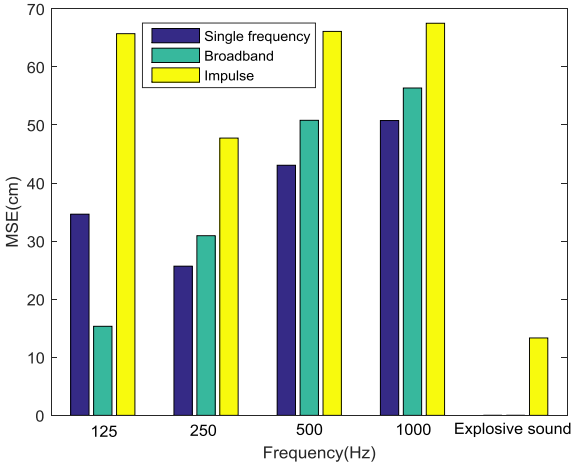


FIGURE 18. MSE of localization accuracy under different signals.

ing two positions is 42 cm. Under 125-1000 Hz, the lowest localization accuracy is 90 cm. The localization accuracy of 125-250 Hz and 250-500 Hz is significantly better than that of 500-1000 Hz and 125-1000 Hz. Seen from Table 5, the signal of 125-1000 Hz has highest side lobe and 250-500 Hz has lowest side lobe.

Fig 15, Fig 16, Table 4, and Table 5 show that localization results of single frequency sinusoidal signals are not very perfect. bandwidth signal has preferable localization result. On the whole, the bandwidth signal has a better localization result than single frequency sinusoidal signal, and high frequency signal has better localization result than low frequency signal.

In the previous section, we discussed the localization performance of steady-state sound source, and then we conduct experimental verification for the pulse sound source. The experimental environment remains unchanged. The experimental results are shown in Fig 17. The average R_c of all grids are shown in Table 6.

TABLE 6. The average R_c at impulse signals.

f/Hz	125	250	500	1000	Explosive sound
R_c	1.0456	1.0177	1.0935	1.0842	1.1841

We observed from Fig 17 that the localization accuracy of the square wave pulse is very low, and the localization accuracy of the explosive sound is basically the same as that of the bandwidth signal. Explosive sound also has lowest side lobe as can be seen from Table 6.

In addition, the mean square errors (MSE) of localization accuracy of single frequency sinusoidal signals bandwidth signals and impulse signals are computed and given in Fig 18. As shown in Fig 18, MSE of impulse sound sources are the biggest and are much bigger than the other two steady-state signals (single frequency sinusoidal signal and bandwidth signal). Single frequency sinusoidal signal only has the bigger MSE than bandwidth signal in 125 Hz, and the smaller MSE in the remain frequencies.

V. CONCLUSION

This paper first carries out numerical simulations on the localization performance of dual-channel matching method based on time reversal under middle and low frequency sound source (single frequency sinusoidal signal, bandwidth signal and impulse sound source) conditions. Then we verify the localization performance of the dual-channel matching method through actual tests in an ordinary office. For dual-channel matching method, the database is firstly established by measuring the actual impulse response of all grid points to dual-microphone, and then localization experiments are performed.

A series of experiments are presented to illustrate the applicability of dual-channel matching method on SSL. The localization results indicate that the medium frequency localization effect is better than the low frequency, the localization effect of the bandwidth signal is better than the single

frequency signal, and the localization effect of steady-state signals is better than impulse sound source; best localization accuracy up to 15 cm. In the single-frequency sinusoidal signals, the localization results of the 250 Hz and 500 Hz signals are better. In the bandwidth signals, the localization result of 250-500 Hz is better. In the impulse signals, the localization effect of the explosive sound is better.

In general, dual-channel matching method based on time reversal on SSL has masses a lot of advantage. The robustness is so strong that high localization precision and localization accuracy can be achieved in reverberation environment. This method also has the advantages of simple operation and low cost. Therefore, it is quite suiTable unsuitable for SSL in enclosed space.

REFERENCES

[1] M. S. Brandstein and H. F. Silverman, "A practical methodology for speech source localization with microphone arrays," *Comput. Speech Lang.*, vol. 11, no. 2, pp. 91–126, Apr. 1997.
[2] F. Huang, W. Sheng, and X. Ma, "Modified projection approach for robust adaptive array beamforming," *Signal Process.*, vol. 92, no. 7, pp. 1758–1763, Jul. 2012.
[3] M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos, "Conditional and unconditional Cramér–Rao bounds for near-field source localization," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2901–2907, May 2010.
[4] T. Yardibi, N. S. Zawodny, C. Bahr, F. Liu, L. N. Cattafesta, III, and J. Li, "Comparison of microphone array processing techniques for aeroacoustic measurements," *Int. J. Aeroacoust.*, vol. 9, no. 6, pp. 733–762, Jul. 2010.
[5] J. S. Hu, C. Y. Chan, C. K. Wang, M. T. Lee, and C. Y. Kuo, "Simultaneous localization of a mobile robot and multiple sound sources using a microphone array," *Adv. Robot.*, vol. 25, nos. 1–2, pp. 135–152, 2011.
[6] D. B. Ward, E. A. Lehmann, and R. C. Williamson, "Particle filtering algorithms for tracking an acoustic source in a reverberant environment," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 826–836, Nov. 2003.
[7] L. C. Mak and T. Furukawa, "Non-line-of-sight localization of a controlled sound source," in *Proc. IEEE/ASME Int. Conf. AIM*, Singapore, Jul. 2009, pp. 475–480.
[8] Q. Yuan, Q. Chen, and K. Sawaya, "Accurate DOA estimation using array antenna with arbitrary geometry," *IEEE Trans. Antennas Propag.*, vol. 53, no. 4, pp. 1352–1357, Apr. 2005.
[9] T. Lobos, Z. Leonowicz, J. Rezmer, and P. Schegner, "High-resolution spectrum-estimation methods for signal analysis in power systems," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 1, pp. 219–225, Feb. 2006.
[10] N. Roman and D. Wang, "Binaural tracking of multiple moving sources," *IEEE Trans. Audio, Speech, Language Process.*, vol. 16, no. 4, pp. 728–739, May 2008.
[11] F. Ribeiro, C. Zhang, D. A. Florencio, and D. E. Ba, "Using reverberation to improve range and elevation discrimination for small array sound source localization," *IEEE Trans. Audio, Speech, Language Process.*, vol. 18, no. 7, pp. 1781–1792, Sep. 2010.
[12] H. Atmoko, D. C. Tan, and B. Fazenda, "Accurate sound source localization in a reverberant environment using multiple acoustic sensors," *Meas. Sci. Technol.*, vol. 19, no. 2, p. 024003, Jan. 2008.
[13] M. S. Brandstein, J. E. Adcock, and H. F. Silverman, "A closed-form location estimator for use with room environment microphone arrays," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 1, pp. 45–50, Jan. 1997.
[14] E. Lleida, J. Fernandez, and E. Masgrau, "Robust continuous speech recognition system based on a microphone array," in *Proc. ICASS*, May 1998, pp. 241–244.
[15] P. Julian, A. G. Andreou, L. Riddle, S. Shamma, D. H. Goldberg, and G. Cauwenberghs, "A comparative study of sound localization algorithms for energy aware sensor network nodes," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 4, pp. 640–648, Apr. 2004.
[16] W. Cui, Z. Cao, and J. Wei, "Dual-microphone source location method in 2-D space," in *Proc. ICASSP*, Toulouse, France, May 2006, p. IV.
[17] N. Ono, S. Fukamachi, T. Nishimoto, and S. Sagayama, "Sound source localization by asymmetrically arrayed 2ch microphones on a sphere," in *Proc. MMSP*, Crete, Greece, Oct. 2007, pp. 56–59.
[18] Z. El Chami, A. Guerin, A. Pham, and C. Servière, "A phase-based dual microphone method to count and locate audio sources in reverberant rooms," in *Proc. WASPAA*, New Paltz, NY, USA, Oct. 2009, pp. 209–212.
[19] A. Mandal, C. V. Lopes, T. Givargis, A. Haghighat, R. Jurdak, and P. Baldi, "Beep: 3D indoor positioning using audible sound," in *Proc. CCNC*, Las Vegas, NV, USA, Jan. 2005, pp. 348–353.
[20] H. Liu, B. Li, X. Yuan, Q. Zhou, and J. Huang, "A robust real time direction-of-arrival estimation method for sequential movement events of vehicles," *Sensors*, vol. 18, no. 4, p. 992, Mar. 2018.
[21] F. Guo *et al.*, "Design of a direction-of-arrival estimation method used for an automatic bearing tracking system," *Sensors*, vol. 16, no. 7, p. 1145, Jul. 2016.
[22] N. Antonello, E. De Sena, M. Moonen, P. A. Naylor, and T. van Waterschoot, "Joint source localization and dereverberation by sound field interpolation using sparse regularization," in *Proc. ICASSP*, Calgary, AB, Canada, Apr. 2018, pp. 6892–6896.
[23] Y. Sun, J. Chen, C. Yuen, and S. Rahardja, "Indoor sound source localization with probabilistic neural network," *IEEE Trans. Ind. Electron.*, vol. 65, no. 8, pp. 6403–6413, Aug. 2018.
[24] B. Rafaely and K. Alhaiany, "Speaker localization using direct path dominance test based on sound field directivity," *Signal Process.*, vol. 143, pp. 42–47, Feb. 2018.
[25] I. Osamu, "An image reconstruction algorithm using phase conjugation for diffraction-limited imaging in an inhomogeneous medium," *J. Acoust. Soc. Amer.*, vol. 85, no. 4, pp. 1602–1606, Oct. 1988.
[26] C. Prada, F. Wu, and M. Fink, "The iterative time reversal mirror: A solution to self-focusing in the pulse echo mode," *J. Acoust. Soc. Amer.*, vol. 90, no. 2, pp. 1119–1129, Mar. 1991.
[27] D. Vigoureux and J.-L. Guyader, "A simplified time reversal method used to localize vibrations sources in a complex structure," *Appl. Acoust.*, vol. 73, no. 5, pp. 491–496, May 2012.
[28] M. Fink, "Time reversal of ultrasonic fields. I. Basic principles," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 39, no. 5, pp. 555–566, Sep. 1992.
[29] F. Wu, J.-L. Thomas, and M. Fink, "Time reversal of ultrasonic fields. II. Experimental results," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 39, no. 5, pp. 567–578, Sep. 1992.
[30] H. Zhang and S. Bing-Wen, "Time-reversal multi-user underwater acoustics communication in shallow water," *Appl. Acoust.*, vol. 28, no. 3, pp. 214–219, Mar. 2009.
[31] J. Kober, Z. Dvorakova, Z. Prevorovsky, and J. Krofta, "Time reversal transfer: Exploring the robustness of time reversed acoustics in media with geometry perturbations," *J. Acoust. Soc. Amer.*, vol. 138, no. 1, pp. 49–53, Jun. 2015.
[32] C. Draeger and M. Fink, "One-channel time reversal of elastic waves in a chaotic 2D-silicon cavity," *Phys. Rev. Lett.*, vol. 79, no. 63, pp. 407–410, Jul. 1997.
[33] C. Draeger and M. Fink, "One-channel time-reversal in chaotic cavities: Theoretical limits," *J. Acoust. Soc. Amer.*, vol. 105, no. 2, pp. 611–617, Oct. 1998.
[34] C. Draeger, J.-C. Aime, and M. Fink, "One-channel time-reversal in chaotic cavities: Experimental results," *J. Acoust. Soc. Amer.*, vol. 105, no. 2, pp. 618–625, Oct. 1999.
[35] G. Lerosey, J. de Rosny, A. Tourin, A. Derode, G. Montaldo, and M. Fink, "Time reversal of electromagnetic waves," *Phys. Rev. Lett.*, vol. 92, p. 193904, May 2004, doi: 10.1103/PhysRevLett.92.193904.
[36] S. Yon, M. Tanter, and M. Fink, "Sound focusing in rooms: The time-reversal approach," *J. Acoust. Soc. Amer.*, vol. 113, no. 3, pp. 1533–1543, Nov. 2002.
[37] R. Sprik, "Time reversed experiments with acoustics," *J. NAG NR*, vol. 174, pp. 1–10, 2005.
[38] S. G. Conti, P. Roux, and W. A. Kuperman, "Near-field time-reversal amplification," *J. Acoust. Soc. Amer.*, vol. 121, no. 6, pp. 3602–3606, Mar. 2007.
[39] S. Catheline and M. Fink, "Acoustic source localization model using in-skull reverberation and time reversal," *Appl. Phys. Lett.*, vol. 90, no. 6, p. 063902, Feb. 2007.
[40] T. Padois, C. Prax, V. Valeau, and D. Marx, "Experimental localization of an acoustic sound source in a wind-tunnel flow by using a numerical time-reversal technique," *J. Acoust. Soc. Amer.*, vol. 132, no. 4, p. 2397, 2012.

[41] A. Mimani, R. Porteous, and C. J. Doolan, "A simulation-based analysis of the effect of a reflecting surface on aeroacoustic time-reversal source characterization and comparison with beamforming," *Wave Motion*, vol. 70, pp. 65–89, Apr. 2017.
[42] J. B. Lonzaga, "Time reversal for localization of sources of infrasound signals in a windy stratified atmosphere," *J. Acoust. Soc. Amer.*, vol. 139, no. 6, pp. 3053–3062, May 2016.
[43] H. Ma, X. Zeng, and H. Wang, "A novel dual-channel matching method based on time reversal and its performance for sound source localization in enclosed space," *Acoust. Aust.*, vol. 44, no. 3, pp. 417–428, Dec. 2016.

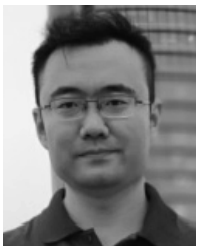
XIANGYANG ZENG was born in Yichang, Hubei, China, in 1974. He received the B.S., M.S., and Ph.D. degrees in engineering from Northwestern Polytechnical University (NPU), Xi'an, China, in 1997, 2000, and 2002, respectively.

From 2004 to 2005, he was a Visiting Scholar with the Department of Acoustics, Technical University of Denmark, Copenhagen, Denmark. His research interests include localization and recognition of noise sources, modeling of indoor and outdoor sound fields, and noise control.

He was a Professor and the Head of the Department of Environmental Engineering, NPU, China, from 2005 to 2018. He was a Communications Member of the European Acoustical Society.

HUIYING MA was born in Zhengzhou, Henan, China, in 1990. She received the B.S. degree in mining engineering from Chongqing University, Chongqing, China, in 2013, and the M.S. degree in environmental engineering from Northwestern Polytechnical University, Xi'an, China, in 2016, where she is currently pursuing the Ph.D. degree in acoustic.

Her research interests include the development of sound source localization in enclosed space and localization techniques using time reversal method, different sound sources, and enclosed space environment.



HAITAO WANG received the B.S. degree in environmental engineering, the M.S. degree in environmental science, and the Ph.D. degree in acoustics from Northwestern Polytechnical University, Xi'an, China, in 2008, 2010, and 2014, respectively. From 2004 to 2005, he was a Visiting Scholar with the Department of Acoustics, Technical University of Denmark, Copenhagen, Denmark. His research interest includes the numerical simulation of the sound field in the small enclosures, sound signal processing, noise control, localization and recognition of noise sources, modeling of indoor and outdoor sound fields, and noise control.

He was a Professor and the Head of the Department of Environmental Engineering, NPU, China, from 2005 to 2018. He was a Communications Member of the European Acoustical Society.