Article

aeroacoustics

International Journal of Aeroacoustics
2022, Vol. 21(3-4) 98–114

DOI: 10.1177/1475472X22109371 1
journals.sagepub.com/home/jae
SAGE

A fast 3D-MUSIC method for near-field sound source localization based on the bat algorithm

C Yang, LL Sun, H Guo, YS Wang and Y Shao

Abstract

To improve the computation and real-time performances of the multiple signal classification (MUSIC) algorithm in 3D space, a fast sound source localization method based on the bat algorithm (BA) and the 3D-MUSIC, called BA-based 3D-MUSIC algorithm (3D-BMUSIC), is presented in this paper. 3D-BMUSIC greatly reduces the computation load by replacing the regular grid search with the BA. First, the near-field spherical wave model is established to obtain the spectral function of the 3D-MUSIC. Then, the spectral function is defined as the fitness function, which calculates the fitness value corresponding to each bat position. Finally, the global optimal bat position with the largest fitness value, as sound source localization, is obtained by successive iteration and sorting. The simulation and experiment show that 3D-BMUSIC accurately estimates the DOA and distance of near-field sources, and the root-mean-square error (RMSE) of 3D-BMUSIC is less than that of 3D-MUSIC. In addition, 3D-BMUSIC effectively reduces the computation time by approximately 96–98%. With shorter computation time and higher efficiency, 3D-BMUSIC promotes hardware implementation and is more suitable for high-precision localization of near-field sound sources.

Keywords

Near-field, Sound source localization, MUSIC, Bat algorithm, Uniform circular array

Date received: 25 March 2021; revised: 12 October 2021; accepted: 30 November 2021

Introduction

Sound source localization, as one of the key research directions in array signal processing,[1] is widely applied in many fields, such as sonar, wireless communications, and biomedical engineering. For multiple sound sources, the localization algorithms are divided into two categories: one is the super-

School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai, China

Corresponding author:
School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, No.333, Longteng Road, Songjiang District, Shanghai, 201620, P. R. China.
Email: hgsues@163.com

resolution spectrum estimation; the other is the beamforming. Over the past few decades, many 2D localization algorithms, such as multiple signal classification (MUSIC),[2] estimating signal parameter via rotational invariance techniques (ESPRIT),[3] maximum likelihood estimation (ML)[4] and various beamforming methods (BF), have been proposed for direction-of-arrival (DOA) estimation. The traditional algorithms are based on the far-field assumption that a signal travels as a plane wave and the difference in the amplitude attenuation of the received signal is ignored. Unlike the far-field DOA estimation, the near-field sound source localization needs to estimate the distance in three-dimensional space,[5] which makes the computation time-consuming.[6–11] The research of the super-resolution spectrum estimation mainly focuses on 2D space. However, in recent years, there is an increasing trend in localizing sound sources on 3D space by various beamforming methods, such as DAMAS, CLEAN, CLEAN-SC.

As a classical algorithm in super-resolution spatial spectrum estimation, the MUSIC algorithm has been improved for real-time performance in the past few decades. According to the Vandermode structure of the steering vector of a uniform linear array (ULA), the root-MUSIC algorithm is proposed to reduce the computation by constructing a root-finding polynomial instead of conducting a grid search. However, the shortcoming is obvious: the root-MUSIC is only applicable to ULA. To overcome this shortcoming, the AI, MST, and FD-root-MUSIC[12–14] algorithms are proposed by converting arbitrary arrays into virtual ULAs. Another problem is that the polynomial equations are very difficult to solve due to the high order of root-seeking polynomials constructed by these algorithms. It is claimed that converting complex domain computations of spatial spectrum estimation algorithms into real domains can reduce the computation complexity of DOA estimation.[15–19] Liu[18] proposed a fast algorithm for root-MUSIC with real-valued eigen-decomposition which requires a large number of snapshots and high SNR. In recent years, Yan and Ming applied compressive sensing theory in DOA estimation,[20–22] which effectively estimates the DOA by reconstructing the original signal with very few observations, but there is a limit to the amount of computation that can be reduced in this way. Masato[23] applied particle swarm optimization (PSO), a type of swarm intelligence optimization, to reduce the peak searching time. Similarly, Yang proposed a new bionic intelligent optimization algorithm that combines the advantages of PSO and the genetic algorithm (GA)[24] and named it the bat algorithm (BA).[25] There are very few studies on 3D positioning methods based on spectrum estimation.

The research of the 3D positioning method is mainly concentrated in the field of BF (such as DAMAS, CLEAN, CLEAN-SC, et al.). Brooks[26] investigated the exploitation of the DAMAS approach with several arrays of different sizes and designs, demonstrating that a large array is advantageous for extracting sound source location and strength in 3D application. Legg and Bradley[27] studied the amplitude and localization accuracy calculated by CLEAN-SC for 2D and 3D scanning surfaces with a spherical array and proved that 3D scanning surface provides higher localization accuracy for higher frequency sound sources. Geyer[28] presented the superiority of 3D beamforming in enhancing noise source distribution for standard 2D mapping. Sarradj[29–31] pointed out that the deconvolution approach is mandatory in the analysis of 3D beamforming and compared four different steering vectors in applications. Porteous[32] compared the performances of four beamforming algorithms for the localization of acoustic dipole sources in a 3D domain with two arrays placed orthogonally to each other. Thomas[33] compared Clean-SC, L1-GIBF, DAMAS, and SC-DAMAS in the case of 2D and 3D scan zones. Xenaki[34] proposed two Fourier-based deconvolution approaches to improve the quality of 3D beamforming with phased arrays. Ning[35] proposed the use of compressed sensing in three-dimensional localization. Chu[36–40] and Koutny[41] studied the three-dimensional deconvolution method based on the spherical array and compared the localization effect of various methods. Wang[42] proposed expectation-maximization (EM) and

Evidential EM (E2M) algorithm to estimate multiple sound sources in a statistical framework by maximum likelihood. The major contribution is the suppression of uncertain problems, such as uncertainty related to ill-known microphone locations and wavenumber. The 3D representation is very useful when analyzing the envelope surface of the car with a microphone array. Doebler[43] computed a 3D beamforming map of the surface enveloping the "Panamera S" body with three arrays and proved that CLEAN-SC is an efficient deconvolution method for this particular application. Dougherty[44] compared the performances of CLEAN-SC, DAMAS for the 3D localization of an engine noise source.

Based on the above discussions, the current research for MUSIC mainly focuses on DOA estimation (2D), while research on 3D-MUSIC is limited. Compared with BF-type methods, the research on MUSIC-type 3D positioning methods needs to be developed. The computation obstacle of 3D-MUSIC has not been effectively solved, and the optimization algorithm of 3D-MUSIC needs to be researched further. This paper proposes a 3D-BMUSIC algorithm for near-field sound source localization, which uses the BA instead of a grid searching in the 3D-MUSIC algorithm. First, the near-field sound source spherical wave signal model is established, the spectral function of independent variables such as the elevation, azimuth, and distance are derived. Second, the spectral function is taken as the fitness function, which is used to calculate the fitness value corresponding to each bat position. Finally, the global optimal bat position, as the localization of the near-field sound source, is obtained by iteration and sorting. The 3D-BMUSIC algorithm not only accurately estimates the localization of near-field sound sources but also has much lower computation complexity and excellent real-time performance. In the near-field sound source localization, its superiorities in terms of real-time and high precision are verified by simulations and experiments of multiple sources.

Signal model and 3D-MUSIC algorithm

Signal model

According to the distance between the sound source and the microphone array, the sound propagation model is divided into the near-field model and the far-field model.[45,46] For the near-field model, sound waves are transmitted in the form of spherical sound waves. The general near-field judgment formula is

$$L \leq \frac{2D^{\,2}}{\lambda} \tag{1}$$

where D is the array aperture, λ is the signal wavelength, and L is the distance between the sound source and the microphone array. The spherical wave model is more accurate and reliable than the far-field plane wave model when the sound source is in the near field. The array signal model based on spherical waves contains the phase and amplitude attenuation information of the sound source.

S narrow-band signals $s_1(t),s_2(t),\ldots,s_S(t)$ are measured by the microphone array. The signal received by the i th microphone can be expressed as

$$f_i(t) = \sum_{j=1}^{S} s_j\left(t - \tau_{ij}\right) + n_i(t),\ i = 1,2,\ldots,M, \quad j = 1,2,\ldots,S \tag{2}$$

where $n_i(t)$ is the Gaussian noise received by the i th microphone. τ_{ij} is the relative time delay factor of the j th signal received by the i th microphone.

The signal can be expressed in the complex envelope form

$$s_j(t) = u_j(t)e^{j(\omega_c t+\beta(t))} \tag{3}$$

$$s_j(t-\tau) = u_j(t-\tau)e^{j(\omega_c(t-\tau)+\beta(t-\tau))} \tag{4}$$

where $u(t)$ is the amplitude of the received signal, $\beta(t)$ is the phase of the received signal, and ω_c is the frequency of the received signal, $\omega_c = 2\pi f_c$.

For a narrowband signal, $\beta(t)$ can be obtained

$$\beta(t-\tau) \approx \beta(t) \tag{5}$$

In the spherical wave model, the amplitude of the sound pressure is inversely proportional to the propagation distance

$$\frac{u_j(t-\tau)}{u_j(t)} = \frac{r_{1j}}{r_{ij}} \tag{6}$$

Substituting equations (5) and (6) into equations (3) and (4) yields the signal in the complex envelope form

$$s_j(t-\tau) = \frac{r_{1j}}{r_{ij}} s_j(t)e^{-j\omega_c\tau_{ij}} \tag{7}$$

Therefore, the incoherent narrowband near-field signal model can be further written as

$$f_i(t) = \sum_{j=1}^{S} \frac{r_{1j}}{r_{ij}} e^{-j\omega_c\tau_{ij}} s_j(t) + n_i(t),\ i = 1,2,...,M,\quad j = 1,2,...,S \tag{8}$$

equation (2) can be expressed in matrix form as Figure 1

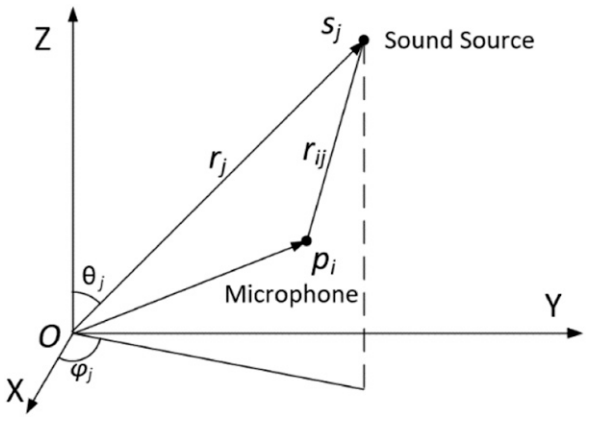

Figure 1. Microphone array receiving signal model.

$$\mathbf{F}(t) = [\,\mathbf{a}(r_1,\theta_1,\varphi_1), \quad \mathbf{a}(r_2,\theta_2,\varphi_2), \quad ,\ldots, \quad \mathbf{a}(r_S,\theta_S,\varphi_s)\,]\mathbf{S}(t) + \mathbf{N}(t) \;= \mathbf{AS} + \mathbf{N} \tag{9}$$

where

$$\mathbf{a}\big(r_j,\theta_j,\varphi_j\big) = \left[1, \frac{r_{1j}}{r_{2j}} e^{-jw_c\tau_{2j}}, \ldots, \frac{r_{1j}}{r_{Mj}} e^{-jw_c\tau_{Mj}}\right]^{\mathrm{T}} \tag{10}$$

denotes the M steering vector and τ_{ij} has the following form

$$\tau_{ij} = \big(\|s_j - p_i\| - \|s_j\|\big)/C \tag{11}$$

where the sound source localization vectors are s_1, s_2,…, s_S and the localization vectors of the microphone in the array are p_1, p_2,… p_M. c is the velocity of the sound wave, which is taken as 343 m/s at 20 degrees Celsius.

3D-MUSIC

$\mathbf{R}$, as the covariance matrix of $\mathbf{F}(t)$, is defined as

$$\mathbf{R} = E\big\{(\mathbf{AS} + \mathbf{N})(\mathbf{AS} + \mathbf{N})^{\mathrm{H}}\big\} = \mathbf{A}E\big\{\mathbf{SS}^{\mathrm{H}}\big\}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I} = \mathbf{AR}_{\mathrm{SS}}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I} \tag{12}$$

Where σ^2 is the Gaussian noise power. H represents the conjugate transpose of the matrix. Eigenvalue decomposition and ordering of the above formula can be obtained as follows

$$\mathbf{R}(t) = \mathbf{U}diag[\lambda_1,\ldots,\lambda_S,\lambda_{S+1},\ldots,\lambda_M]\mathbf{U}^{\mathrm{H}} \tag{13}$$

$\lambda_1,\ldots,\lambda_S,\lambda_{S+1},\ldots,\lambda_M$ are the eigenvalues of $\mathbf{R}(t)$. According to subspace theory, the matrix $\mathbf{U}$ is divided into two spaces: noise subspace ($\mathbf{U_N}$) and signal subspace.

$$\mathbf{A}^{\mathrm{H}}\mathbf{U}_{\mathbf{N}} = 0 \tag{14}$$

The spectral function is defined as

$$P_{MUSIC}(r,\theta,\varphi) = \frac{1}{\mathbf{a}^{\mathrm{H}}(r,\theta,\varphi)\mathbf{U}_{\mathrm{n}}\mathbf{U}_{\mathrm{n}}^{\mathrm{H}}\mathbf{a}(r,\theta,\varphi)} \tag{15}$$

r,θ, and φ are the distance, elevation and azimuth of the near-field sources, respectively. The 3D-MUSIC algorithm searches for the r,θ, and φ values by grid search, in which the computation load mainly depends on the grid spacing and search range, and the computation load formula can be expressed as

$$O\left(\frac{(r_{\max} - r_{\min})(\theta_{\max} - \theta_{\min})(\varphi_{\max} - \varphi_{\min})}{\Delta_r\Delta_\theta\Delta_\varphi}\right) \tag{16}$$

$(r_{max} - r_{min})$,$(\theta_{max} - \theta_{min})$ and $(\varphi_{max} - \varphi_{min})$ are the search range of the distance, elevation and azimuth, respectively; Δ_r,Δ_θ , and Δ_φ are the corresponding grid spacing.

3D-BMUSIC sound source localization estimation

Bat algorithm

BA is a random global optimization algorithm based on population. It was inspired by the echo localization behavior of bats during hunting. During the hunting process, bats first send out a sound signal, which is reflected when it encounters the surrounding objects. By analyzing the characteristics of the echo frequency, loudness, and pulse emission rate, bats can avoid obstacles and search for prey. The BA simulates the hunting process and maintains better coordination between global and local search, enabling more accurate judgments to be made on the localization of prey. The BA has the characteristics of a simple model, fast convergence, potential parallelism and distribution.

The process of bat algorithm optimization is as follows: in the whole search space, individual bats are randomly distributed as the basic unit of the bat algorithm. The optimal solution to the problem is the localization of the bats. All bats have a fitness value corresponding to the optimization problem. At the beginning of the algorithm, each bat senses its prey using its low pulse frequency and high loudness. When the prey is found (the current optimal solution), the pulse frequency increases and the loudness decreases. By comparing the fitness value of the inferior and superior position, the bats in the inferior position (smaller fitness value) gradually move toward the bats in the superior position (larger fitness value), after several searches, the bats will gather in the position of the prey (the optimal position). The optimization is to find the best parameter that maximizes a fitness function.

Optimization of 3D-MUSIC based on the BA

Grid search is the main factor that leads to the massive computation of the 3D-MUSIC algorithm. This paper introduces the BA into near-field sound source localization. As a solution to the optimization problem, the bat position represents the estimated localization of the sound sources. Primarily, set the bat population n, frequency range $[f_{min},f_{max}]$, loudness attenuation coefficient α, frequency increase coefficient γ, number of algorithm iterations $T_{\max}$, space dimension d, rate of the pulse emission rp_i, and loudness A_i.

The initialization formula of the sound source position is

$$x = x_{\min} + rand(0,1)(x_{\max} - x_{\min}) \tag{17}$$

where x_{max} and x_{min} are the upper and lower limits of the parameters, respectively.

The fitness function is defined as the spectral function of the 3D-MUSIC algorithm

$$fitness(x) = \frac{1}{\mathbf{a}^{\mathrm{H}}(r,\theta,\varphi)\mathbf{U}_{\mathrm{n}}\mathbf{U}_{\mathrm{n}}^{\mathrm{H}}\mathbf{a}(r,\theta,\varphi)} \tag{18}$$

where $\mathbf{U}_{\mathrm{n}}$ is the noise vector. The fitness value of the initial bat position is calculated and sorted to find the current optimal solution corresponding to the maximum fitness value. Record the optimal solution x^* of the iteration and its corresponding fitness value. Then, update the sound source localization as follow

$$f_i = f_{min} + (f_{max} - f_{min}) \times rand \tag{19}$$

$$v_i^t = v_i^{t-1} + \left(x_i^{t-1} - x^*\right) \times f_i \tag{20}$$

$$x_i^t = x_i^{t-1} + v_i^t \tag{21}$$

f_i is the pulse frequency at the current moment of the i th bat randomly assigned within the range $[f_{min}, f_{max}]$ of the pulse transmission frequency, rand $\in [0, 1]$; It is a factor affecting how quickly the bat changes its flight velocity to veer toward the current best solution; v_i^t, v_i^{t-1} are the corresponding flight speeds of the i th bat at time t and t-1; x_i^t, x_i^{t-1} are the positions at time t and t-1 of the i th bat, and x^* is the optimal individual at the moment.

If $rand_1 > rp_1^t$, select one solution from the set of the best bat positions and generate a new individual position through random flight

$$x_{new} = x_{old} + \varepsilon A^t, \quad \varepsilon \in [-1, 1] \tag{22}$$

x_{old} is a solution randomly selected from an existing set of optimal solutions, ε is a d-dimensional random vector, and A^t is the average loudness of all individual bats in the same period. The new individual position will then replace the current position of the i th bat.

If $rand_2 > A_i^t \& f(x_i^t) < f(x^*)$, save this new solution while increasing rp_i^t and decreasing A_i^t. Otherwise, the original position x_i^{t-1} remains unchanged

$$rp_i^{t+1} = rp_i^0[1 - \exp(\gamma t)] \tag{23}$$

$$A_i^{t+1} = \alpha A_i^t \tag{24}$$

rp_i^{t+1} is the pulse rate of the bat at time $t + 1$. A_i^t is the loudness at time t. α and γ are constants, with $\gamma > 0$ and $0 < \alpha < 1$.

Sort all sound source positions according to the adaptive function and find the current best sound source position x^*. If the maximum number of iterations is not reached, update the sound source localization. If the termination condition is satisfied, output the global optimal sound source position x^*.

For multi-sound source localization, another nested iterative searching is required. The basic idea is that iterative searching of multiple sound sources can be done by suppressing the relevant searching space. Figure 2 shows the flow chart of multiple sound source localization.

Simulation

In this section, the localization of incoherent sound sources is simulated. Based on the parameter L, the sound source is in the near field. The computation load, computation time and statistical performance of 3D-BMUSIC and 3D-MUSIC are compared. The coordinates of each microphone in the microphone array are shown in Figure 3, and the number of snapshots is 500. Other simulation settings are listed in Table 1.

Localization result and error analysis

Table 2 and Figure 4 show the ability of the 3D-BMUSIC algorithm for identifying near-field sound sources with SNR of 20 dB. The root means squared error (RMSE) of 3D-BMUSIC and 3D-MUSIC

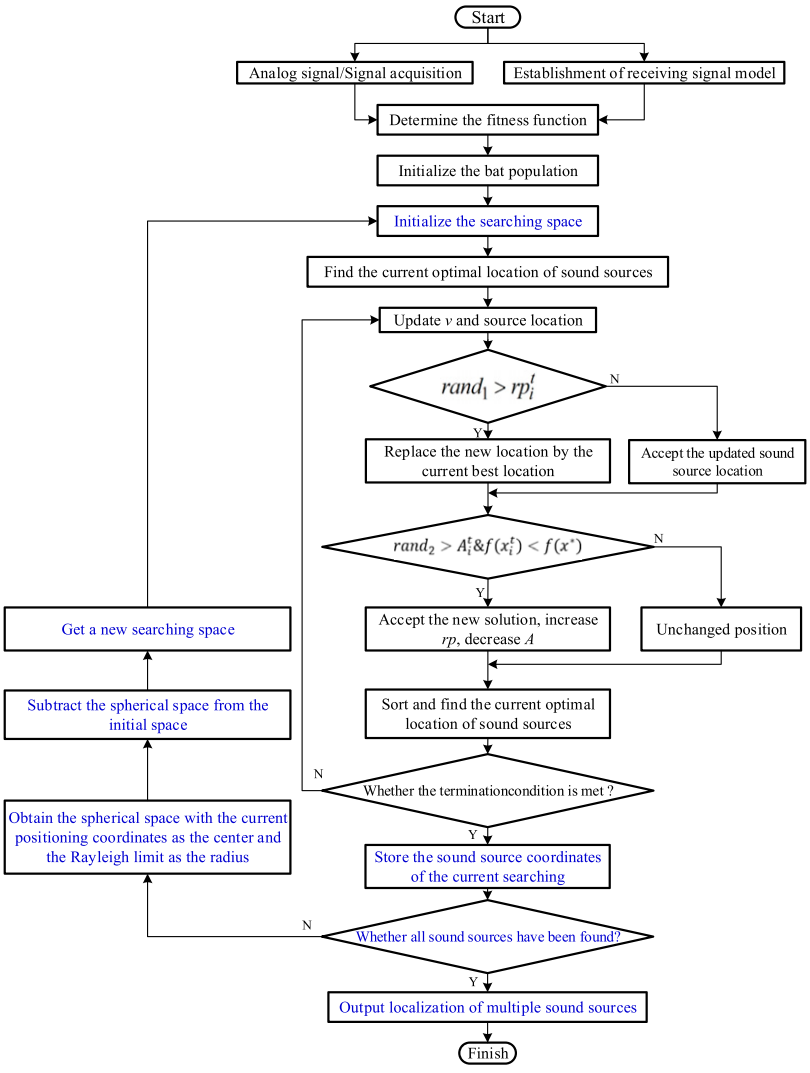

Figure 2. Flow chart of 3D-BMUSIC.

are compared in Figure 5; 200 Monte Carlo simulations were conducted to obtain the RMSE curves for error analysis.

Table 2 compares the actual source location and source localization estimated by 3D-BMUSIC. The sound source coordinates estimated by 3D-BMUSIC are very close to the actual location coordinates in Table 2. For convenience, Figure 4 shows the localization results in 2D and 3D spaces. The circle represents the actual location and the triangle represents the sound source localizations estimated by 3D-BMUSIC. Obviously, the estimated and the actual sound sources locations almost overlap, indicating that 3D-BMUSIC can accurately localize near-field sound sources.

Figure 5 shows the RMSE curves of 3D-BMUSIC and 3D-MUSIC with different SNRs. On the trend of the curve, as the SNR increases, the RMSE decreases significantly. The reason is that the

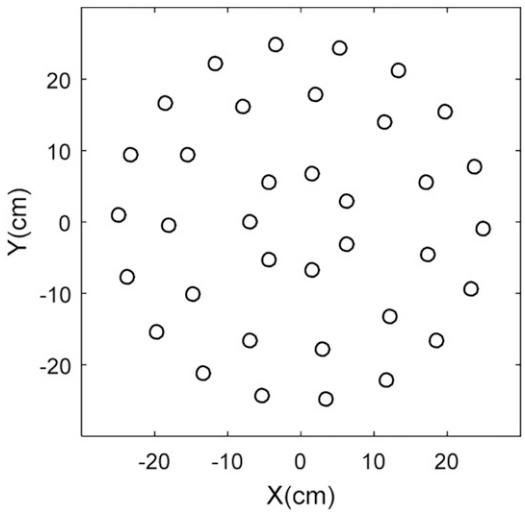

Figure 3. Layout of the microphone array.

Table 1. The simulation setting.

Name of parameter	Value
Coordinate system	Spherical
Number of microphones	M = 36
Source 1 localization	(75°, 77°, 60 cm)
Source 1	Narrowband (2000 Hz) random signal
Source 2 localization	(40°, 50°, 90 cm)
Source 2	Narrowband (2000 Hz) random signal
Sampling frequency	6144 Hz
Angle step in grid search of 3D-MUSIC	1°
Distance step in grid search of 3D-MUSIC	1 cm
Search range of elevation	1–90°
Search range of azimuth	1–360°
Search range of distance	1–100 cm
Bat population	n = 20
Frequency range	[0, 1]
Rate of pulse emission	rp_i = 0
Pulse emission rate coefficient	γ = 1
Number of algorithm iterations	$iter_{max}$ = 100
Spatial dimension	3D
Loudness	A_i = 0.6
Loudness attenuation coefficient	α = 0.6
Number of snapshots	500

Table 2. Actual position and localization result by 3D-BMUSIC.

	Actual position		Localization result	
	Spherical	Cartesian (cm)	Spherical	Cartesian (cm)
Source 1	(75°, 77°, 60 cm)	(13.04, 56.47, 15.53)	(74.99°, 77.03°, 60.02 cm)	(13.01, 56.49, 15.54)
Source 2	(40°, 50°, 90 cm)	(37.19, 44.32, 68.94)	(39.96°, 49.99°, 90.02 cm)	(37.17, 44.28, 69.00)

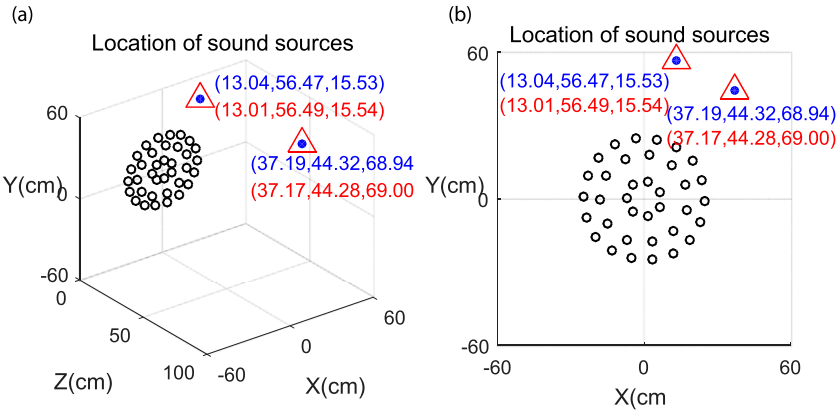

Figure 4. The localization of sound sources by 3D-BMUSIC: (a) spatial localization and (b) sound source projection on the XOY plane.

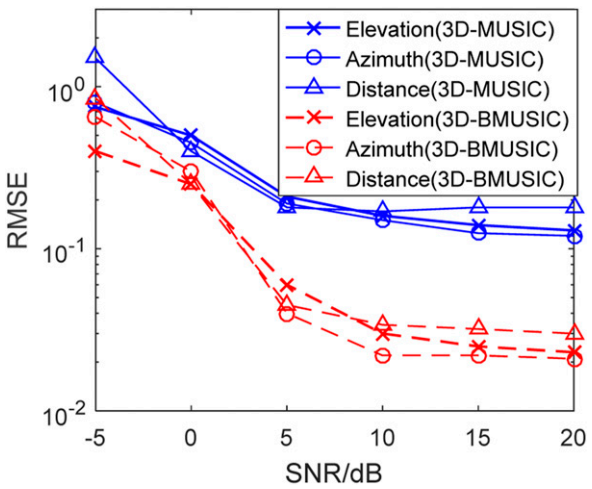

Figure 5. The RMSEs of 3D-BMUSIC and 3D-MUSIC.

noise and the signal subspace are completely separated, making the accuracy higher and the localization error smaller. On the other hand, the RMSE of 3D-BMUSIC is smaller than that of 3D-MUSIC as shown in Figure 5. Numerous simulations have shown that 3D-BMUSIC is excellent in the near-field sound source localization.

Convergence property and computation load

The convergence rate is the main indicator for evaluating algorithm efficiency. Figure 6 shows the iteration convergence curves of the elevation, azimuth and distance estimated by 3D-BMUSIC. Table 3 and Table 4 compare the computation load and time of 3D-BMUSIC and 3D-MUSIC. The computation time is the average of 200 runs of the algorithm on the Matlab2016b, and all parameters of 3D-MUSIC and 3D-BMUSIC are listed in Table 1.

As shown in Figure 6, the convergence of the iteration is very fast. The algorithm localizes the sound source well about 10 iterations and converges to a certain value within 85 iterations. The number of bat populations is one of the key parameters of the iteration. If the number of bats is large,

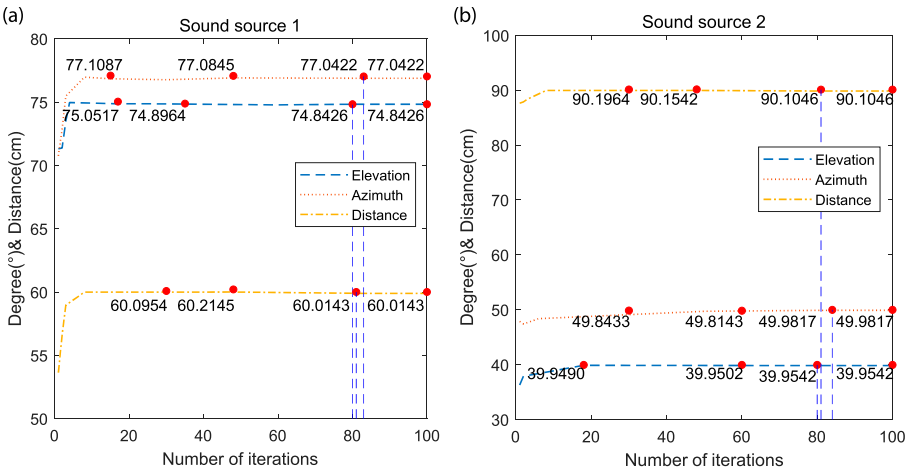

Figure 6. The convergence curves of 3D-BMUSIC in the simulation: (a) sound source 1 and (b) sound source 2.

Table 3. The computation time of 3D-BMUSIC and 3D-MUSIC.

3D-BMUSIC average time/s	3D-MUSIC average time/s	Average time ratio/%
0.67	35.65	1.90

Table 4. The computation load of 3D-BMUSIC and 3D-MUSIC.

Computation load of 3D-BMUSIC	Computation load of 3D-MUSIC
$O(2.0 \times 10^3)$	$O(3.24 \times 10^6)$

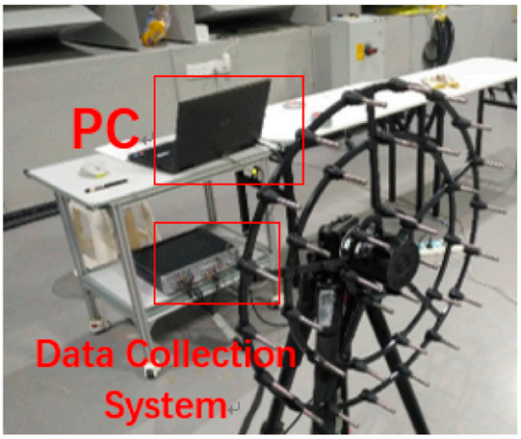

Figure 7. Experiment apparatus.

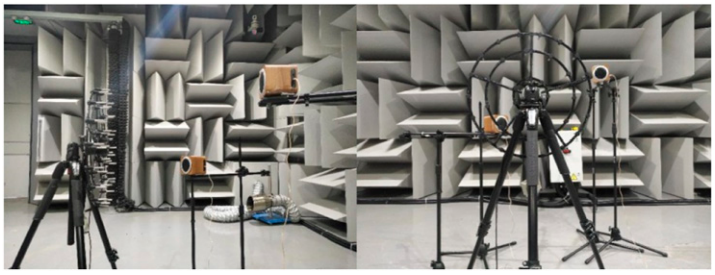

Figure 8. Spatial localization of sound sources.

Table 5. The experiment setting.

Name of parameter	Values
Number of microphones	M = 36
Sensitivity of microphone	64 mV/Pa
Source 1 localization	(28°,22°, 42 cm)
Source 1	Narrowband (2000 Hz) random signal
Source 2 localization	(48°,201°, 64 cm)
Source 2	Narrowband (2000 Hz) random signal
SNR	20 dB
Size of semi-anechoic chamber	9.8 m×8.6 m×3.5 m
Background noise	18 dB
Cutoff frequency	125 Hz
Measurement frequency range	1–22387 Hz
Number of snapshots	500

a more suitable position can be found in the first few iterations, so the convergence speed is faster. If the number of bats is small, the initial localization result deviates from the true position, so the number of iterations increases. There are two ways to terminate the iterations, one is to specify the maximum number of the iterations as used in the paper, and the other is to set the difference between the results of the two iterations. The number of iterations is mainly considered from two aspects: First of all, the parameter 100 is conservative, which can ensure high accuracy of localization; secondly, the number of iterations will increase for complex sound sources.

It can be seen from Table 3, the computation time of 3D-BMUSIC is only 1.9% of 3D-MUSIC. As shown in Table 4, the computation load of 3D-MUSIC $O(3.24 \times 10^6)$ is calculated by equation (16), while the computation load of 3D-BMUSIC is mainly determined by the maximum number of iterations and the population number, $O(iter_{max} \times \mathrm{n}) = O(2.0 \times 10^3)$. 3D-BMUSIC achieves high accuracy estimation with a lower computation load than that of the traditional 3D-MUSIC. The shorter computation time and faster convergence rate indicate that the real-time performance of 3D-BMUSIC is excellent and suits the practical applications better.

Experiment

The experiment consisted of data collection, data analysis and sound source localization. As shown in Figure 7, the sound pressure signals are sampled using the microphone array and the LMS-SCM-05 multichannel data acquisition system. The experiments are carried out in a semi-anechoic

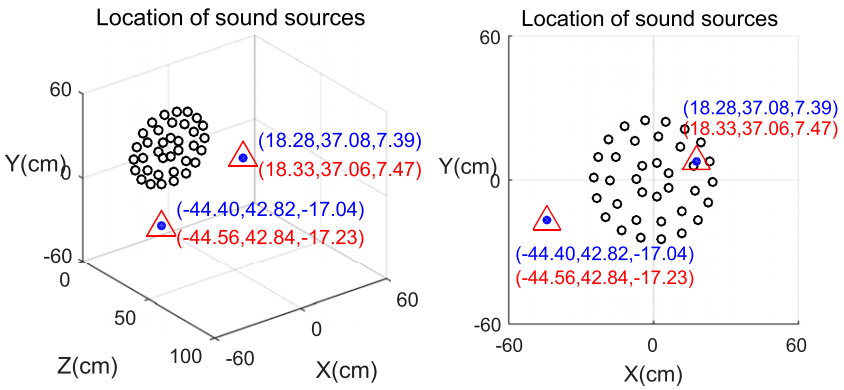

Figure 9. The localization of sound sources by 3D-BMUSIC: (a) spatial localization and (b) sound source projection on the XOY plane.

Table 6. Actual position and localization result by 3D-BMUSIC.

	Actual position		Localization result	
	Spherical	Cartesian (cm)	Spherical	Cartesian (cm)
Source 1	(28°, 22°, 42 cm)	(18.28, 37.08, 7.39)	(28.11°, 22.17°, 42.02 cm)	(18.33, 37.06, 7.47)
Source 2	(48°, 201°, 64 cm)	(−44.40, 42.82, −17.04)	(48.12°, 201.14°, 64.17 cm)	(−44.56, 42.84, −17.23)

chamber. The array is placed 50 cm above the ground. Figure 8 shows the side view and front view of the spatial position of the sound source in the experiment. Two speakers with a size of 11 cm×11 cm×13 cm are used to simulate sound sources. Compared with the distance to the array, the speaker can be considered as the compact sound source, and the center of the cone is defined as the position of the sound source. The basic parameters are shown in Table 5 and Table 1 mentioned in chapter 4.

Localization result and error Analysis

Table 6 lists the coordinates of the actual sound source position and localization results of 3D-BMUSIC in the experiment. The maximum distance deviation on the x-axis, y-axis and z-axis of the Cartesian coordinate system is 0.05 cm, 0.02 cm, and 0.19 cm, respectively. For simplicity, Figure 9 shows the positions of the actual sound sources, the sources identified by 3D-BMUSIC and the projection on the XOY plane in the Cartesian coordinate system. It can be found that 3D-BMUSIC still locates the near-field sound source accurately in the applications. Table 6

Convergence property and computation load

Figure 10 shows the iteration convergence curves of the elevation, azimuth and distance estimated by 3D-BMUSIC for two speakers. Compared with simulation, the iterative curve fluctuates slightly due to the measurement noise. When the number of iterations exceeds 80, the estimated values of the three parameters tend to be stable. The experiments prove the conclusion of the simulation:

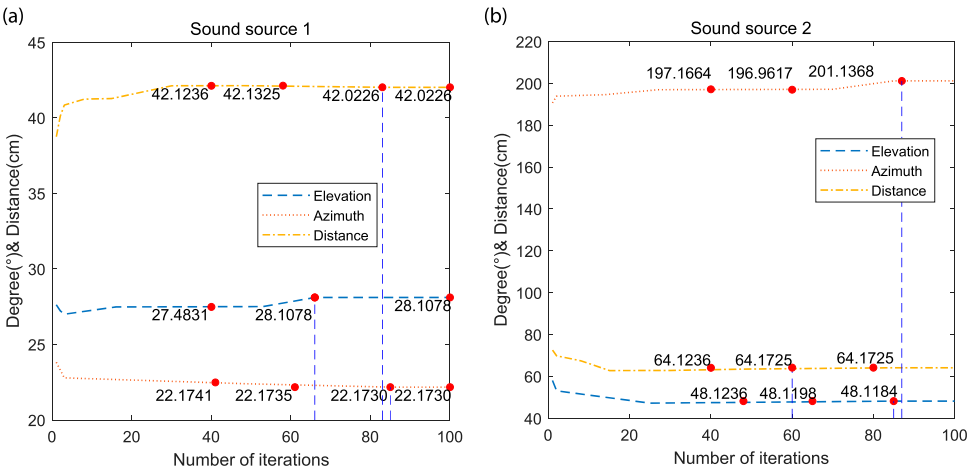

Figure 10. The convergence curves of 3D-BMUSIC in the experiment: (a) sound source 1 and (b) sound source 2.

Table 7. The time comparison of 3D-BMUSIC and 3D-MUSIC.

3D-BMUSIC average time/s	3D-MUSIC average time/s	Average time ratio/%
0.88	27.01	3.20

3D-BMUSIC can localize the spatial sound sources quickly and accurately. Table 7 compares the computation time of 3D-BMUSIC and 3D-MUSIC for two speakers. The computation time is the average of 20 groups of array data. It can be seen from Table 7, the computation time of 3D-BMUSIC is only 3.2% of 3D-MUSIC.

Conclusions

As a high-resolution algorithm, the study of MUSIC in 3D sound source localization is practical. The large computation load of the 3D-MUSIC limits its application in the near-field sound source localization. To improve the real-time performance of 3D-MUSIC, the BA-based 3D-MUSIC algorithm (3D-BMUSIC) is presented in this paper. 3D-BMUSIC treats grid search as an optimization problem, which achieves lower computation load and better localization accuracy by the grid search optimization based on the BA. The multiple sound sources are localized by successively iteration and sorting. Both simulation and experiment verify the excellent convergence property of 3D-BMUSIC, which localizes sound sources with higher precision and lower computation load than 3D-MUSIC. The maximum positioning deviation does not exceed 1 cm, and the computation time of 3D-BMUSIC is only 1%–4% of 3D-MUSIC. In application, 3D-BMUSIC is an accurate and efficient method for estimating the DOA and distance of sound sources in near-field by overcoming technical obstacles. Due to the lower computation time and higher accuracy, the 3D-BMUSIC can effectively improve the performance of the DSP-based sound source recognition system.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (51675324&52172371) and partly supported by the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning, and New Energy Vehicle Vibration and Noise Testing and Control Professional Technical Service Platform (18DZ2295900), China.

ORCID iDs

H Guo https://orcid.org/0000-0002-7874-7273
YS Wang https://orcid.org/0000-0001-8624-5268

References

1. Zheng Z, Fu M, Wang W-Q, et al. Mixed far-field and near-field source localization based on subarray cross-cumulant. *Signal Process* 2018; 150: 51–56.
2. Schmidt R. Multiple emitter location and signal parameter estimation. *IEEE Trans Antennas Propagation* 1986; 34: 276–280.
3. Roy R, Paulraj A and Kailath T. ESPRIT-estimation of signal parameters via rotational invariance techniques. 19th Asilomar Conference on Circuits. International Society for Optics and Photonics, 2002.
4. Marxim Rahula Bharathi B and Mohanty AR. Underwater sound source localization by EMD-based maximum likelihood method. *Acoust Aust* 2018; 46: 193–203.

5. Sun C and Ding J. *Application of the DOA algorithms in the uniform circular array antennas*. International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS). IEEE Computer Society, 2018.
6. Yan F-G, Shuai L, Wang J, et al. Real-valued root-MUSIC for DOA estimation with reduced-dimension EVD/SVD computation. *Signal Process* 2018; 152: 1–12.
7. Li JZ, Wang YD and Gang W. Signal reconstruction for near-field source localisation. *Signal Process* 2015; 9: 201–205.
8. Xie J, Tao H, Rao X, et al. Efficient method of passive localization for near-field noncircular sources. *IEEE Antennas Wireless Propagation Lett* 2015; 14: 1223–1226.
9. Wu Y, Ma L, Hou C, et al. Subspace-based method for joint range and DOA estimation of multiple near-field sources. *Signal Process* 2006; 86: 2129–2133.
10. Zhi W and Chia MY-W. Near-field source localization via symmetric subarrays. *IEEE Signal Process. Lett* 2007; 14: 409–412.
11. Shahimaeen A and Dehghani MJ. Two-dimensional DOA estimation for coherent signals using a novel covariance-like matrix. *Trans Emerg Telecommun Technol* 2019; 30: e3597.
12. Friedlander B. and Weiss A. J. Direction finding using spatial smoothing with interpolated arrays. *IEEE Trans Aerospace Electron Syst* 1992; 28: 574–587.
13. Belloni F., Richter A. and Koivunen V. DOA estimation via manifold separation for arbitrary array structures. *IEEE Trans Signal Process* 2007; 55: 4800–4810.
14. Rubsamen M. and Gershman A. B. Direction-of-arrival estimation for nonuniform sensor arrays: from manifold separation to Fourier domain MUSIC methods. *IEEE Trans Signal Process* 2009; 57: 588–599.
15. Keh-Chiarng Huarng KC and Chien-Chung Yeh CC. A unitary transformation method for angle-of-arrival estimation. *IEEE Trans Signal Process* 1991; 39: 975–977.
16. Zoltowski M. D., Kautz G. M. and Silverstein S. D. Beamspace Root-MUSIC. *IEEE Trans Signal Process* 1993; 41: 344–364.
17. Pesavento M., Gershman A. B. and Haardt M. Unitary root-MUSIC with a real-valued eigendecomposition: a theoretical and experimental performance study. *IEEE Trans Signal Process* 2000; 48: 1306–1314.
18. Liu C and Liao G. Fast algorithm for root-MUSIC with real-valued eigendecomposition. *Int Conf Radar2006* 2006; 10: 1–4.
19. Huang Q., Feng J., Huang J, et al. Two-stage MUSIC with reduced spectrum search for spherical arrays. *Digital Signal Process* 2020; 106: 102836.
20. Yan F, Jin M and Qiao X. Source localization based on symmetrical MUSIC and its statistical performance analysis. *Sci China Inf Sci* 2013; 56: 1–13.
21. Yan F, Jin M and Qiao X. Low-Complexity DOA Estimation Based on Compressed MUSIC and Its Performance Analysis. *IEEE Trans Signal Process* 2013; 61: 1915–1930.
22. Yan FG, Liu S, Jin M, et al. Fast DOA estimation based on MUSIC symmetrical compressed spectrum. *Syst Eng Electro* 2012; 34: 2198–2202.
23. Hirakawa M and Suyama K. Multiple sound source tracking by two microphones using PSO. *Int Symp Intell Signal Process Commun Syst* 2013; 20: 467–470.
24. Rodrigues DS, Delamaro ME, Corrêa CG, et al. Using genetic algorithms in test data generation. *ACM Comput Surv* 2018; 51: 1–23.
25. Yang X-S. A new metaheuristic bat-inspired algorithm. *Nat Inspired Coop Strateg Optimization (NICSO 2010)* 2010; 284: 65–74.
26. Brooks T and Humphreys W. Three-dimensional applications of DAMAS methodology for aeroacoustic noise source definition. In: *11th AIAA/CEAS Aeroacoustics Conference*. American Institute of Aeronautics and Astronautics; 2005.

27. Legg M and Bradley S. Comparison of CLEAN-SC for 2D and 3D scanning surfaces. In: *4th Berlin Beamforming Conference – 4th BeBeC*; 2012.
28. Geyer T, Sarradj E and Giesler J. Application of a beamforming technique to the measurement of airfoil leading edge noise. *Adv Acoust Vibration* 2012; 2012: 1–16.
29. Sarradj E. Three-dimensional acoustic source mapping. In: *4th Berlin Beamforming Conference – 4th BeBeC*; 2012.
30. Sarradj E. Three-dimensional acoustic source mapping with different beamforming steering vector formulations. *Adv Acoust Vib* 2012; 1–12.
31. Battista G, Chiariotti P, Herold G, et. al. Inverse methods for three-dimensional acoustic mapping with a single planar array. In: 7th Berlin Beamforming Conference – 7th BeBeC; 2018.
32. Porteous R., Prime Z., Doolan C. J, et al. Three-dimensional beamforming of dipolar aeroacoustic sources. *J Sound Vibration* 2015; 355: 117–134.
33. Padois T. and Berry A. Two and three-dimensional sound source localization with beamforming and several deconvolution techniques. *Acta Acustica United with Acustica* 2017; 103(3): 392–400.
34. Xenaki A., Jacobsen F. and Fernandez-Grande E. Improving the resolution of three-dimensional acoustic imaging with planar phased arrays. *J Sound Vibration* 2012; 331(8): 1939–1950.
35. Ning F., WeiQiu J.L.H., Qiu L, et al. Three-dimensional acoustic imaging with planar microphone arrays and compressive sensing. *J Sound Vibration* 2016; 380: 112–128.
36. Chu Z, Zhao S, Yang Y, et al. Deconvolution using CLEAN-SC for acoustic source identification with spherical microphone arrays. *J Sound Vib* 2018; 440: 161–173.
37. Yang Y., Chu Z., Shen L, et al. Fast Fourier-based deconvolution for three-dimensional acoustic source identification with solid spherical arrays. *Mech Syst Signal Process* 2018; 107: 183–201.
38. Yang Y., Chu Z., Shen L, et al. Functional delay and sum beamforming for three-dimensional acoustic source identification with solid spherical arrays. *J Sound Vibration* 2016; 373: 340–359.
39. Chu Z., Yang Y. and He Y. Deconvolution for three-dimensional acoustic source identification based on spherical harmonics beamforming. *J Sound Vibration* 2015; 344: 484–502.
40. Chu Z., Yang Y. and Shen L. Resolution and quantification accuracy enhancement of functional delay and sum beamforming for three-dimensional acoustic source identification with solid spherical arrays. *Mech Syst Signal Process* 2017; 88(5): 274–289.
41. Koutny A., Jiricek O., Thomas J.-H, et al. Source distance determination based on the spherical harmonics. *Mech Syst Signal Process* 2017; 85(2): 993–1004.
42. Wang X., Quost B., Chazot J.-D, et al. Estimation of multiple sound sources with data and model uncertainties using the EM and evidential EM algorithms. *Mech Syst Signal Process* 2016; 66-67: 159–177.
43. Doebler D., Ocker J. and Puhle C. On 3d beamforming in the wind tunnel. In: *6th Berlin Beamforming Conference – 6th BeBeC*; 2016.
44. Dougherty RP. Jet noise beamforming with several techniques. In: *3rd Berlin Beamforming Conference – 3rd BeBeC*; 2010.
45. Su X., Liu Z., Chen X, et al. Mixed incoherent far-field and near-field source localization under uniform circular array. *Sensors* 2018; 18: 1432–1444.
46. Tachikawa T., Yatabe K. and Oikawa Y. 3D sound source localization based on coherence-adjusted monopole dictionary and modified convex clustering. *Appl Acoust* 2018; 139: 267–281.