IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 24, 2019, accepted September 24, 2019, date of publication October 11, 2019, date of current version November 15, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2947097

SRP-PHAR Combined Velocity Scanning for Locating the Shallow Underground Acoustic Source

PENGFEI NIE[1,2], BIN LIU[1,2], PING CHEN[1,2], KUN LI[2], AND YAN HAN[1,2]
[1]Shanxi Key Laboratory of Information Survey and Processing, North University of China, Taiyuan 030051, China
[2]School of Information and Communication Engineering, North University of China, Taiyuan 030051, China

Corresponding author: Bin Liu (liubin@nuc.edu.cn)

This work was supported by the National Defense Science and Technology Key Laboratory of Transient Impact Technology under Grant 614260603030817.

ABSTRACT Shallow underground acoustic source localization is a component of near-field source localization, which is involved in numerous application fields. The positioning accuracy is mainly limited by the accuracy of time of arrival (TOA)/time difference of arrival (TDOA) extraction and velocity extraction from noisy data. The steered response power with phase transform (SRP-PHAT) is one of the most robustness and high-precision acoustic source localization approaches, which avoids extracting the TDOA in advance. But SRP-PHAT is constrained for only using under known velocity. Furthermore, it is barely possible for shallow underground sound source localization to easily obtain high-quality velocity models. This paper proposes an improved SRP-PHAT with unknown velocity (SRP-PHAT-UNVEL), which avoids extracting the TDOA and velocity in advance. SRP-PHAT-UNVEL matches the calculated time delay curve with the actual time delay curve by scanning of the candidate spatial position and the candidate velocity simultaneously, so as to maximize the output energy to fulfill the positioning. However, SRP-PHAT-UNVEL has larger computational complexity as it proceeds with the optimization of space and velocity. Since the spatial position and velocity affect the shape and the curvature of the calculated delay curve respectively, these are two relatively independent processes. Therefore, the simultaneous optimization of space and velocity can be replaced by a two-stage optimization to improve the efficiency and accuracy of SRP-PHAT-UNVEL. Spatial optimization is equivalent to the optimization of SRP-PHAT, and the spatial optimization of the bat algorithm has faster convergence rate and higher location precision than traditional methods. Velocity optimization can be achieved by the common linear search technique since the function of velocity and energy is an ideal convex function. Simulation experiment results show that the proposed method is insensitive to noise, which can achieve high accuracy of the acoustic source position and velocity simultaneously. With grouping the measured data, the proposed method can further improve the robustness and accuracy by fusing the grouping location results with principal component analysis.

INDEX TERMS Shallow underground acoustic source location, steered response power with phase transform, bat algorithm, principal component analysis.

I. INTRODUCTION

Underground acoustic source localization has become a popular research topic in the earthquake and microseismic fields, which form part of far-field scenarios, as the source and sensor arrays are far apart. However, we are interested in locating the near-field source in many cases [1]. Moreover, there exist a variety of practical problems in the areas of border security [2], core safety monitoring, water conservancy, and performance evaluation of earth-penetrating weapons, which require the localization of an acoustic source positioned less than several hundred meters from the sensor, i.e. the applications of shallow underground acoustic source localization. Two challenging issues arise in locating shallow subsurface acoustic sources. First, the underground medium is not uniform and is unknown,

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Feng.

which presents difficulties in most model-driven localization methods. Second, because shallow underground acoustic sources are often distributed at depths of less than several hundred meters, higher positioning accuracy is required than for deep underground acoustic source. This is the engineering requirement of the related application field of shallow acoustic source localization.

Acoustic source localization methods can be broadly classified into two categories: indirect and direct methods. The time difference of arrival (TDOA) and time of arrival (TOA) methods are two commonly used indirect positioning approaches. The TDOA-based method first calculates the time differences between varying sensor pairs. Many techniques can be applied to estimate the TOAs [3]–[7], among which the generalized cross-correlation (GCC) function [3] is the most classical method. In the second step, TDOAs are combined with knowledge of the sensor positions and velocities to generate a nonlinear equation set. The sound source estimation position is obtained by solving the nonlinear equations, which can be solved by the maximum likelihood solution [8], [9] and closed-form solution [8], [10]–[12]. The TOA-based method is a commonly used positioning approach for natural earthquakes and microseisms, in which the objective function is constructed based on the TOAs, velocity information, and sensor position information, while sound source position estimation is achieved by minimizing the objective function. The differences among the various TOA-based methods lie in the construction, processing, and optimization of the objective function. The Geiger [13] positioning method and its modified method [14]–[16] belong to the typical TOA-based methods.

Both the TDOA and TOA-based methods only use time information, which is part of the collected signal information. The error of the indirect method consists of time information extraction and position estimation, and any large component of the two parts will results in overall positioning errors. Meanwhile, indirect methods are sensitive to noise, and require accurate time as well as reliable velocity information.

The direct method mainly includes the subspace decomposition method [17]–[20] and beamforming method [21]–[23]. The MUSIC method [17] is a classical subspace decomposition method that decomposes data into signal and noise subspaces by observing the eigenvalue decomposition of the signal covariance matrix. Positioning is achieved using the orthogonality of the signal and noise subspaces. The MUSIC algorithm and other subspace decomposition methods [18]–[20] require accurate velocity information. The steered response power with phase transform (SRP-PHAT) method [23]–[26] is one of the most effective beamforming-based positioning method. Under noise and reverberation conditions, SRP-PHAT is one of the most efficient and robust localization methods, which performs global optimization using all available information [23], [27]. SRP-PHAT is implemented in two stages. First, the GCC function between the signals acquired by each microphone pair is acquired, and then the source location is searched over a grid of space. The second stage is the most computationally demanding, as the high localization accuracy implies using dense grids [23]. Since, the SRP space has many local extremes [23], [27]–[29], the traditional optimization algorithms (e.g. gradient descent) are not applicable to SRP-PHAT, which need to be optimized to reduce the computational expense. Various improved algorithms have been proposed, such as stochastic region contraction (SRC) [27], coarse-to-fine region contraction (CFRC) [28], stochastic particle filtering (SPF) [29], and other methods [30], [31]. These optimization methods belong to region contraction methods, at the expense of accuracy and robustness. Moreover, the convergence velocity is slow, resulting in low computational efficiency when the spatial search area is relatively large. The bat algorithm [32]–[35] is a heuristic search algorithm and provides an effective method for searching for global optimal solutions. The bat algorithm is an iteration-based optimization technique, which is initialized into a set of random solutions and then searches for the optimal solution by means of iteration. Around the optimal solution, a new local solution is generated by random flight, which strengthens the local search. Compared with other algorithms, the bat algorithm is far superior in terms of accuracy and effectiveness [34], [35], and few parameters need to be adjusted. The bat algorithm may be a suitable option for global optimization of the SRP-PHAT.

Another problem of SRP-PHAT is that the calculated time delay is not equal to the actual time delay with the employment of coarser spatial grids, which will result in losing the energy information and reducing the location accuracy. A local accumulation method is proposed in [25], [36], [37] to solve this problem well, but it will reduce the spatial resolution of SRP-PHAT.

Since the direct method need not extract time information in advance and utilizes all the information of collected signals, the direct method is relatively more suitable than the indirect method for locating the shallow underground acoustic source. However, the achievement of a high-quality velocity model is costly and not feasible in subsurface sound source applications sometimes, since both direct and indirect methods require accurate velocity to achieve high-precision localization.

In this paper, an improved SRP-PHAT with unknown velocity (SRP-PHAT-UNVEL) is proposed for the shallow underground acoustic source localization. In fact, SRP-PHAT-UNVEL is proposed by combining velocity scanning with SRP-PHAT, which is equivalent to optimizing both space and velocity simultaneously. However, the computational complexity of SRP-PHAT-UNVEL is very high. The SRP-PHAT-UNVEL is transformed into a two-stage optimization to reduce its complexity. The experimental results show that the proposed method is feasible for locating acoustic sources with an unknown velocity and a high accuracy can be achieved by the joint estimation of acoustic source location and velocity.

II. SRP-PHAT WITH UNKNOWN VELOCITY

SRP-PHAT is a method that combines phase transform (PHAT) and steered response power (SRP). The SRP can be expressed as the output power of the filter-and-sum beamformer [23], which is given by

$$\mathrm{P}(\boldsymbol{x}) = \int_{-\infty}^{+\infty} Y(\omega, \boldsymbol{x})\, Y^*(\omega, \boldsymbol{x}) d\omega \tag{1}$$

where $Y^*(\omega, \boldsymbol{x})$ is the complex conjugate of $Y(\omega, \boldsymbol{x})$, and $Y(\omega, \boldsymbol{x})$ is the output of the filter-and-sum beamformer, defined as follows:

$$Y(\omega, \boldsymbol{x}) = \sum_{l=1}^{M} G_l(\omega) F_l(\omega) e^{-i\omega\tau(\boldsymbol{x}, l)} \tag{2}$$

where $\boldsymbol{x} = [x, y, z]^T \in \mathbb{R}^3$ denotes a candidate for the acoustic source position; $F_l(\omega)$ is the Fourier transform of the collected signal from the l-th sensor; $G_l(\omega)$ is the Fourier transform of the temporal filter, and $\tau(\boldsymbol{x}, l)$ is the direct time of wave travel from location $\boldsymbol{x}$ to the l-th sensor, which are selected to focus or steer the array to the source spatial location.

Substituting (2) into (1), the SRP becomes

$$\mathrm{P}(\boldsymbol{x}) = \sum_{l=1}^{M}\sum_{q=1}^{M} \int_{-\infty}^{+\infty} G_l(\omega) G_q^*(\omega) F_l(\omega) \times F_q^*(\omega) e^{-i\omega(\tau(\boldsymbol{x}, l) - \tau(\boldsymbol{x}, q))} d\omega \tag{3}$$

In (3), $G_l(\omega) G_q^*(\omega)$ can be considered as a weighting function and represented by $\Psi_{lq}(\omega)$. Then, (3) can be expressed as follows:

$$\mathrm{P}(\boldsymbol{x}) = \sum_{l=1}^{M}\sum_{q=1}^{M} \int_{-\infty}^{+\infty} \Psi_{lq}(\omega) F_l(\omega) F_q^*(\omega) e^{-i\omega(\tau(\boldsymbol{x}, l) - \tau(\boldsymbol{x}, q))} d\omega \tag{4}$$

when $\Psi_{lq}(\omega)$ is the PHAT function, (4) is the SRP-PHAT function. By scanning the spatial position $\boldsymbol{x}$, the determined maximum value $\mathrm{P}(\boldsymbol{x})$ is the estimated position of the sound source $\tilde{x}$; that is,

$$\tilde{x} = \mathrm{argmax}_{\boldsymbol{x}}\, \mathrm{P}(\boldsymbol{x}). \tag{5}$$

PHAT is a popular frequency-dependent weighting function of GCC [11], which is defined as follows:

$$\Psi_{lq}(\omega) = \frac{1}{\left|F_l(\omega) F_q^*(\omega)\right|}. \tag{6}$$

where $F_l(\omega)$ and $F_q(\omega)$ are the Fourier transforms of a pair of collected signals $f_l(t)$ and $f_q(t)$, and $F_q^*(\omega)$ is the complex conjugate of $F_q(\omega)$. The PHAT function broadens the signal spectrum and makes the SRP-PHAT can obtain sharper cross-spectrum/cross-correlation peaks, which improves the spatial resolution of the SRP.

Equation (4) can be rewritten by GCC, as follows:

$$\mathrm{P}(\boldsymbol{x}) = \sum_{l=1}^{M}\sum_{q=1}^{M} \mathrm{R}_{lq}\left(\tau_{lq}(\boldsymbol{x})\right) \tag{7}$$

$$\mathrm{R}_{lq}\left(\tau_{lq}(\boldsymbol{x})\right) = \int_{-\infty}^{+\infty} \frac{F_l(\omega) F_q^*(\omega) e^{-i\omega(\tau_{lq}(\boldsymbol{x}))}}{\left|F_l(\omega) F_q^*(\omega)\right|} d\omega \tag{8}$$

$$\tau_{lq}(\boldsymbol{x}) = \tau(\boldsymbol{x}, l) - \tau(\boldsymbol{x}, q) = \frac{|\boldsymbol{x} - \boldsymbol{x}_l| - |\boldsymbol{x} - \boldsymbol{x}_q|}{v} \tag{9}$$

where $\tau_{lq}(\boldsymbol{x})$ is the calculated time delay of arrival of the sensor pair (l,q) corresponding to an acoustic source located at $\boldsymbol{x}$. v is the ray velocity. $\mathrm{R}_{lq}(\cdot)$ is the generalized cross-correlation with phase transform (GCC-PHAT).

In SRP-PHAT, the maximization of GCC-PHAT output is achieved for acoustic source localization by scanning the spatial grids. Essentially, SRP-PHAT is the steering samples accumulation of GCC-PHAT, which can be expressed by Dirac delta function as follows:

$$\mathrm{P}(\boldsymbol{x}) = \sum_{l=1}^{M}\sum_{q=1}^{M} \int_{-\infty}^{+\infty} \mathrm{R}_{lq}(\tau)\, \delta\left(\tau - \tau_{lq}(\boldsymbol{x})\right) d\tau \tag{10}$$

From (9), the shape of the calculated time delay curve is adjusted to match the real delay curve by scanning the candidate spatial position in SRP-PHAT. As shown in Fig 1(a), different time delay curve shapes correspond to different spatial positions. When the spatial position is scanned to the acoustic source position, the calculated time delay curve is consistent with the real one. At this time, the largest energy of the GCC-PHAT samples accumulation is reached along the calculated time delay curve, and the acoustic source location is the spatial parameter corresponding to the maximum energy, which is the basic principle of SRP-PHAT localization When the velocity is unknown, the calculated time delay is a function of both space and velocity, as follows:

$$\tau_{lq}(\boldsymbol{x}, v) = \tau(\boldsymbol{x}, l) - \tau(x, q) = \frac{|x - x_l| - |x - x_q|}{v} \tag{11}$$

The (10) can be expressed as follows:

$$\mathrm{P}(\boldsymbol{x}, v) = \sum_{l=1}^{M}\sum_{q=1}^{M} \int_{-\infty}^{+\infty} \mathrm{R}_{lq}(\tau)\, \delta\left(\tau - \tau_{lq}(x, v)\right) d\tau \tag{12}$$

This is the SRP-PHAT with unknown velocity (SRP-PHAT-UNVEL), whose basic idea is still the process shown in Fig. 1a. However, the key difference is the consideration of the velocity influence on the time delay curve. In other words, we need to determine that the function $\mathrm{P}(x, v)$ has a unique maximum, and its corresponding spatial position and velocity should correspond to the actual sound source location and ray velocity. Given a fixed spatial position, the curvatures of the delay curves vary from different ray velocities. And given a certain velocity, the delay curve shape is different at different positions. Only when the spatial position and velocity match the actual situation, the shape

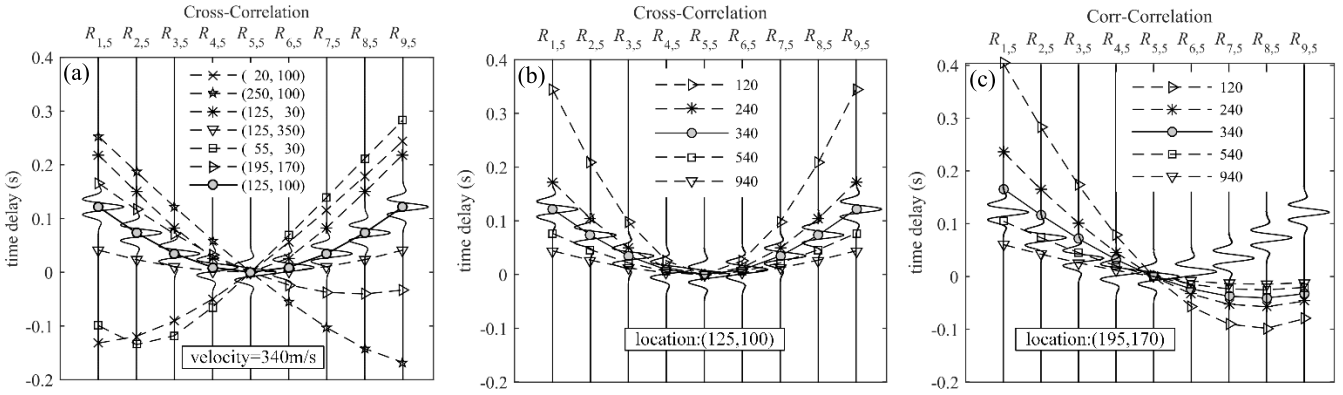


FIGURE 1. The changing patterns of the shape and curvature of the calculated delay curves with the spatial position and velocity. (a) For fixed velocity, the changing patterns of the shape with the spatial position. (b)-(c) For fixed position, the changing patterns of the curvature with the velocity from 120 m/s to 940 m/s. In (b), the fixed position is the actual acoustic source position (125 m, 100 m). $R_{i,5}$ represents the GCC of the signal collected by i-th sensor with the reference signal, wherein the reference signal is collected by fifth sensor. The circle-marked curve represents both the actual delay curve and the calculated delay curve in (a) and (b). But in (c), it only represents the calculated delay curve.

and curvature of calculated delay curves agree with real delay curves, as shown in Fig.1 (b,c). The result verifies that function $\mathrm{P}(x, v)$ has a unique maximum. A more detailed analysis of the SRP-PHAT-UNVEL principle is shown in Appendix 1. In fact, Appendix 1 essentially reveals the same principles as Fig.1. The SRP-PHAT-UNVEL achieves the simultaneous optimization of space and velocity, but it has low calculation efficiency. Since the candidate position only affects the delay curve shape, and the ray velocity only affects the curvature, the optimization of space and velocity are two relatively independent processes. Therefore, the simultaneous optimization can be modified into two-stage optimization of spatial position and velocity to improve computational efficiency.

The First-Stage Optimization:

With a given velocity v_i, $\mathrm{P}(x, v)$ is the function of x, and it can be denoted by $\mathrm{P}_{v_i}(x)$.

$$\tilde{x}_{v_i} = \mathrm{argmax}_x \, \mathrm{P}_{v_i}(x) \tag{13}$$

where $\tilde{x}_{v_i}$ is the optimal spatial position for velocity v_i. The output energy is $\mathrm{P}_{v_i}\left(\tilde{x}_{v_i}\right)$, and it can be denoted by $\mathrm{S}(v_i) = \mathrm{P}_{v_i}\left(\tilde{x}_{v_i}\right)$.

The Second-Stage Optimization:

Calculate the first-stage optimization for different velocity, then the energy function $\mathrm{S}(v)$ of velocity is obtained. From Fig.1 (a,b,c), the calculated delay curve only matches a real delay curve when v is the real velocity. At this point, the GCC-PHAT accumulation along the delay curve is the largest, that is $\mathrm{S}(v)$ reaches its maximum. An estimate of the velocity v_{opt} can be obtained by optimizing the $\mathrm{S}(v)$ function.

$$v_{opt} = \underset{v}{\mathrm{argmax}} \, \mathrm{S}(v) \tag{14}$$

The optimal spatial position $\tilde{x}_{v_{opt}}$ of the first-stage optimization can be determined according to v_{opt}. Therefore, the joint estimation of the velocity and acoustic source location is achieved.

However, the numerical calculation delay of (10) and (12) have slight inconformity with their actual delay,

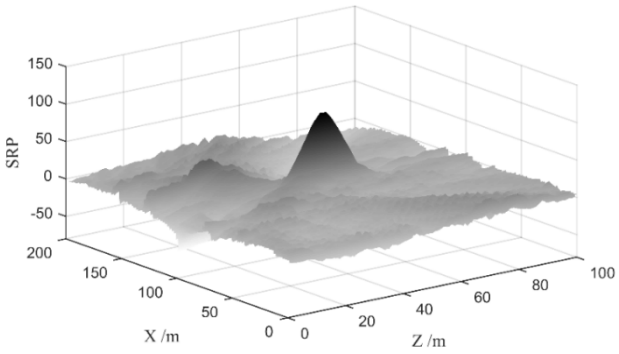


FIGURE 2. Energy distribution map of full space scanning in SRP space for modeling data. The process of generating simulation data is described in Section 4.1. The velocity used in the SRP-PHAT calculation is 750 m/s, which is the average velocity of the established velocity model. The arrows point out some local extremums, and S represents the maximum energy in the SRP space. The positioning result is the space coordinate corresponding to S.

which can be expressed as $\tau \neq \tau_{lq}(x, v)$ and can be well solved in [25], [36], [37]. The samples $\mathrm{R}_{lq}\left(\tau_{lq}(x, v)\right)$ for $\tau_{lq}(x, v) \neq \tau$ can be addressed with simple linear interpolation, which can largely reserve the original energy relationship with less influence on the spatial resolution of SRP.

III. TWO-STAGE OPTIMIZATION OF SRP-PHAT-UNVEL

A. FIRST-STAGE OPTIMIZATION ALGORITHM

According to (13), the first-stage optimization is actually to optimize SRP-PHAT. As the SRP space contains many local extrema (see Fig. 2), conventional optimization methods, such as gradient descent, are not applicable to SRP-PHAT. Various global optimization methods have been proposed to address this problem, such as SRC and CFRC, etc. The convergence rate of most of these optimization methods is relatively low for a large search range [2]. Furthermore, these optimization methods are implemented at the expense of accuracy and robustness. To this end, the bat algorithm [32]–[35] is used to accomplish

the first-stage optimization, which is mainly based on two aspects. First, the bat algorithm can be regarded as providing a balance between standard particle swarm algorithm and enhanced local search to achieve improved optimization effects and robustness [32], [34]. Second, the bat algorithm offers a faster convergence rate than existing optimization methods [32], as verified in section 5.2.

After idealizing some of the echolocation characteristics, the update process of the bat algorithm is described as follows. Assume that bats fly randomly with a velocity v_{b_i} at position x_i and emit pulses with a frequency f_i in the search space. Bats vary their frequency f_i and loudness A to search for prey, and automatically adjust their pulse frequency and pulse emission rate $r \in [0, 1]$. The new solutions, namely position $\boldsymbol{x}_i^t$ and velocity $v_{b_i}^t$ at time step t, are given by

$$f_i = f_{min} + \alpha(f_{max} - f_{min}) \tag{15}$$

$$v_{b_i}^t = v_{b_i}^{t-1} + (\boldsymbol{x}_i^{t-1} - \boldsymbol{x}_0)f_i \tag{16}$$

$$\boldsymbol{x}_i^t = \boldsymbol{x}_i^{t-1} + v_{b_i}^t. \tag{17}$$

where α is a random number with a uniform distribution within [0, 1] and $\boldsymbol{x}_0$ is the current optimal global location. Equations (15) to (17) are used for global optimization. For the local search part, once a solution is selected among the current optimal solutions, a new solution for each bat is generated locally, using a random walk:

$$\boldsymbol{x}_{new} = \boldsymbol{x}_{old} + \mu A^t. \tag{18}$$

where μ is a random number within $[-1, 1]$ and A^t is the average loudness of all of the bats. The loudness A_i and rate r_i of the pulse emission can be updated as follows:

$$A_i^{t+1} = \beta A_i^t,\ r_i^{t+1} = r_i^0\left[1 - \exp(-\gamma t)\right] \tag{19}$$

The specific process of the first-stage optimization based on the bat algorithm is as follows.

Step 1: Objective function $\mathrm{P}_{v_i}(\boldsymbol{x})$. Define the range $[\boldsymbol{x_{min}}, \boldsymbol{x_{max}}]$ of the search space. Initialize the velocity $\boldsymbol{v_{b_i}}$ of bats flying. Define the maximum frequency f_{max} and minimum frequency f_{min} of the emitted pulse. Initialize the pulse rate r^0 and define its enhancement factor $\gamma > 0$. Initialize the loudness A^0 and define its attenuation coefficient $\beta \in (0, 1)$. Define the maximum number *Niter* of iterations.

Step 2: Randomly initialize the bat location $\boldsymbol{x}_i$, and the current optimal location/solution $\boldsymbol{x}_0$ is searched according to the fitness value $\mathrm{P}_{v_i}(\boldsymbol{x}_i)$.

Step 3: Generate new solutions by adjusting the frequency, and updating the velocities and locations using (15) to (17).

Step 4: Generate a uniformly distributed random number *rand*: if (rand > r), select a solution among the best solutions. Generate a local solution around the best selected solution using (18).

Step 5: Generate a uniformly distributed random number*rand*: if (rand < A & $\mathrm{P}_{v_i}(\boldsymbol{x}_i) > \mathrm{P}_{v_i}(\boldsymbol{x}_0)$), accept the new solution $\boldsymbol{x_i}$. Increase r and reduce A according to (19).

Step 6: Rank the fitness values of all bats and determine the current optimal location $\boldsymbol{x_0}$ and fitness value $S_0 = \mathrm{P}_{v_i}(\boldsymbol{x}_0)$.

Step 7: While (iteration < niter), return to step 2. Otherwise, the algorithm ends with the total best solution $\tilde{x}_{v_i} = \boldsymbol{x_0}$ and fitness value $S(v_i) = S_0$.

Although the bat algorithm is highly robust, during the optimization process, it involves the pulse rate r, comparison of the loudness A with random numbers, and randomness of global and local optimization, which will result in small changes in $\tilde{x}_{v_i}$ and $S(v_i)$. To improve the solution stability further, the final optimal locations $\tilde{x}_{v_i}$ and fitness values $S(v_i)$ are the average of the last five iterations; that is,

$$\begin{cases} \tilde{x}_{v_i} = \dfrac{1}{5}\displaystyle\sum_{i=\mathrm{N_{iter}}-5}^{\mathrm{N_{iter}}} \boldsymbol{x}_0(i), \\ S(v_i) = \dfrac{1}{5}\displaystyle\sum_{i=\mathrm{N_{iter}}-5}^{\mathrm{N_{iter}}} \mathrm{P}_{v_i}(\boldsymbol{x}_0(i)) \end{cases} \tag{20}$$

B. SENCOND-STAGE OPTIMIZATION ALGORITHM

The object of second-stage optimization is the energy function $S(v)$. The specific executing steps are as follows:

Step 1 Possible value range of velocity V is decided based on prior knowledge. Then calculate $S(v_i)$ using (20) after selecting the sample v_i, $(i = 1, 2, \ldots N)$ from the V randomly. The N velocities v_i are distributed uniformly in the V space

Step 2 Choose M (<N) larger $S(v_j) \in S(v_i), j = 1, 2, \ldots, M$, and find the maximum value v_{max} and the minimum value v_{min} of velocity in v_j.

Step 3 Reduce the velocity range to V=$[v_{min}, v_{max}]$ if $\Delta v = v_{max} - v_{min}$ does not meet the velocity accuracy requirement, and return to step1.

Step 4 If Δv satisfies the velocity accuracy requirement, the velocity $v_{S_{max}}$ corresponding to the maximum value in $S(v_j)$ is taken as the final estimated value, namely, $v_{opt} = v_{S_{max}}$.

The pseudo-code for velocity optimization is as follows.

IV. SHALLOW UNDERGROUND ACOUSTIC SOURCE LOCATION

A. SIMPLIFIED UNDERGROUND VELOCITY MODEL

The velocity of wave propagation in a shallow underground medium is unknown. For inhomogeneous media, the velocity is a function of the spatial location. This means that the velocities of different spatial points differ, which results in great difficulties or even makes it impossible to locate the underground acoustic source. In general, the horizontal anisotropy of shallow underground media is not evident in a small range, and the transversal velocity change can be ignored. Therefore, the velocity model can be simplified to a layered or continuous model, as illustrated in Fig. 3. The velocity change between layers is generally small, as we are only concerned with a shallow underground medium, which means that $\sin\gamma \approx \sin\theta_i$ $(\gamma \approx \theta_i)$; that is, $\theta_1 \approx \theta_2 \approx \ldots \approx \theta_i = \theta$. Then, (21) can be obtained from the geometric relationship in Fig. 3.

$$\mathrm{d}_i = \frac{\mathrm{h}_i}{\cos\theta} \tag{21}$$

Algorithm 1 Pseudo-Code for Velocity Optimization

Objective function $\boldsymbol{S}(\boldsymbol{v})$, $\boldsymbol{v} \in \boldsymbol{V}$
Initialize the velocity scanning range V
Define the convergence precision $\boldsymbol{\varepsilon}$ or maximum number iter_num of iterations
while (iteration < iter_num) do
  Randomly select N velocities $\boldsymbol{v_i}$, $(\boldsymbol{i} = \mathbf{1}, \mathbf{2}, \ldots \boldsymbol{N})$ in V
  for $i = 1{:}\,N$ do
    Calculate $\boldsymbol{S}(\boldsymbol{v_i})$ by (20)
  end for
  Rank the $\boldsymbol{S}(\boldsymbol{v_i})$ and take the first M $(M < N)$ velocities $\boldsymbol{v_j}, \boldsymbol{j} = \mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{M}$.
Find $\boldsymbol{v_{max}} = \mathbf{max}(\boldsymbol{v_j})$, and $\boldsymbol{v_{min}} = \mathbf{min}(\boldsymbol{v_j})$
Compute $\boldsymbol{\Delta v} = \boldsymbol{v_{max}} - \boldsymbol{v_{min}}$
if $(\boldsymbol{\Delta v} > \boldsymbol{\varepsilon})$ then
  Reduce the velocity range to $\mathbf{V} = [\boldsymbol{v_{min}}, \boldsymbol{v_{max}}]$
else
  Drop out of the while loop
  end if
end while
Rank the $\boldsymbol{S}(\boldsymbol{v_i})$ and find the best solution $\tilde{\boldsymbol{v}}$
Post-process results

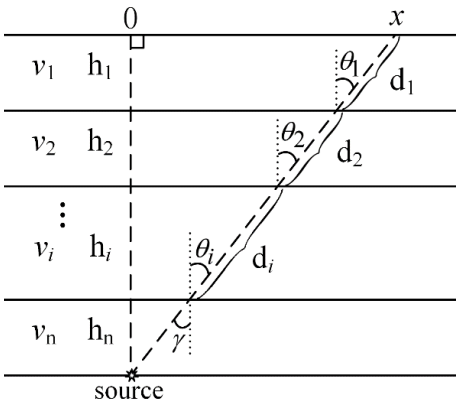


FIGURE 3. Simplified subsurface velocity model.

The propagating velocity v_{ray} along the ray path is

$$v_{\text{ray}} = \frac{\sum_{i=1}^{n} \mathrm{d}_i}{\sum_{i=1}^{n} \frac{\mathrm{d}_i}{v_i}} = \frac{\sum_{i=1}^{n} \frac{\mathrm{h}_i}{\cos\theta}}{\sum_{i=1}^{n} \frac{\mathrm{h}_i}{v_i \cos\theta}} = \frac{\sum_{i=1}^{n} \mathrm{h}_i}{\sum_{i=1}^{n} \frac{\mathrm{h}_i}{v_i}} = \frac{h_{\text{total}}}{t_{\text{total}}} = v_{\text{ave}} \tag{22}$$

where v_{ave} is the average velocity of the wave propagating vertically in the medium, which is an unknown constant for an identified medium. This means that the ray velocity is approximately equal to the average velocity for shallow underground media.

For localization, we mainly use direct time, which is related to the ray velocity. Therefore, the shallow underground acoustic source localization can be implemented by SRP-PHAT-UNVEL.

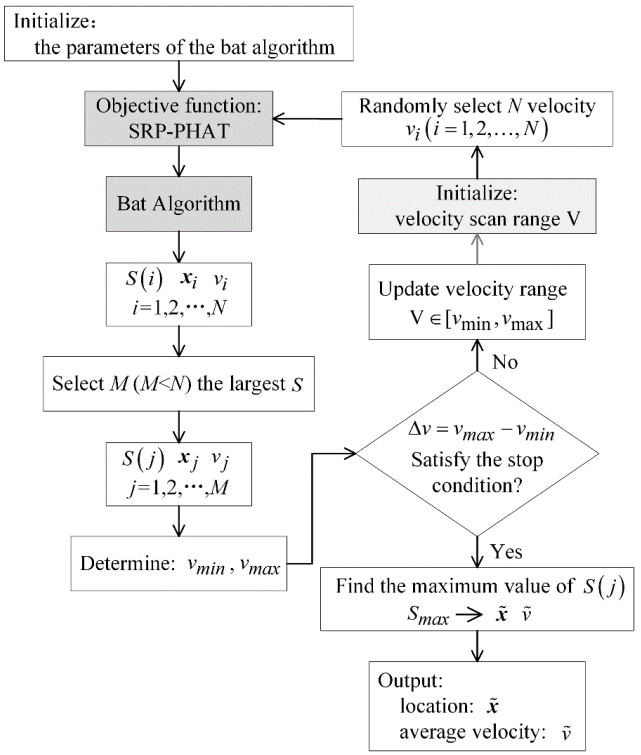


FIGURE 4. General flowchart of the shallow subsurface source localization.

B. LOCALIZATION FLOW

SRP-PHAT-UNVEL can be implemented through a two-stage interrelated optimizing process. The overall framework of two-stage optimization can be summarized by the flowchart displayed in Fig. 4, which can achieve both source localization and average velocity estimation.

V. EXPERIMENTS AND ANALYSIS

A. VELOCITY MODEL AND SIMULATION DATA

Owing to the non-repeatability and non-verifiability of real events in a shallow underground medium, it is difficult to measure the error between the calculated and real source. However, simulation can be flexibly set according to different needs, which is convenient for an algorithm performance analysis. Based on the above two points, the performance of the proposed algorithm is first analyzed using simulation data.

The velocity model for shallow underground media has been simplified to a depth-dependent model, of which the continuum model is typical. A continuum model with a velocity gradient of 2 m/s is established, as illustrated in Fig. 5(a), which also indicates the sensor layouts and source position. The average velocity of the medium above the source is 750 m/s. The established model has a horizontal distance of 200 m and a depth of 100 m. The sensors are arranged in the range of (z = 0, x = 50∼150 m), and the distance between sensors is 2 m. The source is located at (x = 100 m, z = 50 m).

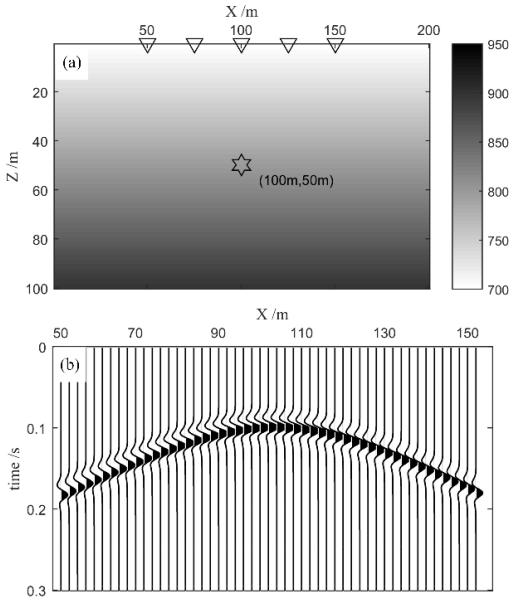


FIGURE 5. (a) Velocity model and location of source and sensors. (b) Synthetic data by finite difference forward modeling.

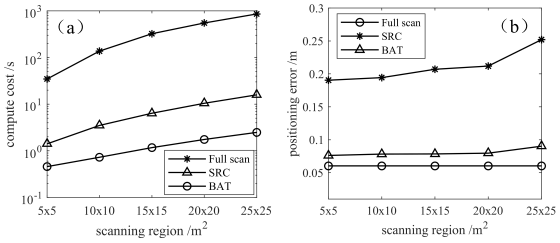


FIGURE 6. Calculation time and positioning errors of full scan, SRC and BAT methods for different spatial ranges. (a) is for calculation time and (b) is for positioning error. The error is defined as the distance between the actual and estimated positions.

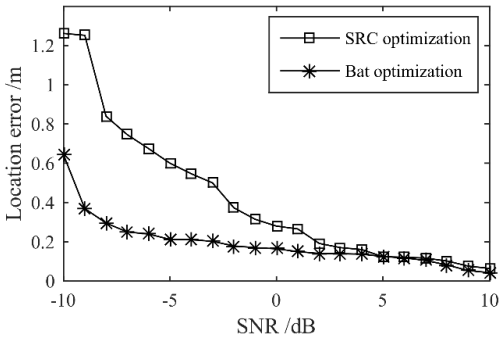


FIGURE 7. The localization results of simulation data with different SNR levels under known velocity.

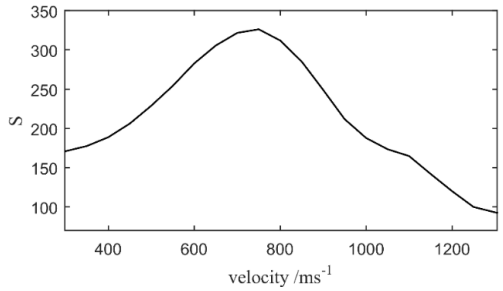


FIGURE 8. The $S - v$ curve for modeling data in Subsection 4.1.

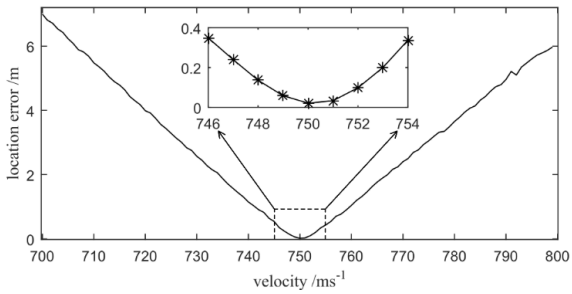


FIGURE 9. The relationship between the velocity variation and the location error.

The simulation data are obtained by finite difference forward modeling [47], as shown in Fig. 5(b).

B. FIRST-STAGE OPTIMIZATION ANALYSIS

Compared with CFRC, SRC has higher computational efficiency in the case of low noise [28]. To analyze the performance of the bat algorithm for SRP-PHAT optimization, the calculation efficiency and positioning accuracy of the bat and SRC algorithms are compared by the simulation data at an average velocity of 750 m/s. The below parameters are used in this paper. SRC uses the suggested parameters [27]: the number of random points $J = 3000$, the number of points used to define the new source space $N = 100$, final grid resolution 0.01 m. Bat algorithm parameters: population size: 40, the number of generations increases with the enlarge of search space, loudness $A = 2$, pulse rate $r = 1$.

The calculation time of full space scan, bat optimization and SRC optimization for different spatial ranges were compared, as shown in Fig. 6. The final grid resolution of full-space scanning is 0.01 m. Fig. 6a shows that the computational efficiency of full-space scanning is the lowest. The calculation efficiency of bat optimization is much higher than that of SRC optimization. For different scanning spatial ranges, the average calculation time of SRC and bat optimization is 7.54 s and 1.32 s respectively. The computational efficiency of the bat algorithm is better than that of SRC 5 times, while the previous existing optimization methods have only twice as fast as SRC optimization [35]. Meanwhile, the positioning error of Bat optimization is smaller than that of SRC optimization, as shown in Fig. 6b. The positioning error of the full-space scanning is the smallest and remains unchanged. For different spatial ranges, the average positioning error of bat optimization is 0.081 m and that of SRC optimization is 0.211 m. Therefore, the bat algorithm provides better computational efficiency and localization accuracy. Moreover, the experiment results demonstrate that the bat algorithm optimization is more robust than the SRC optimization.

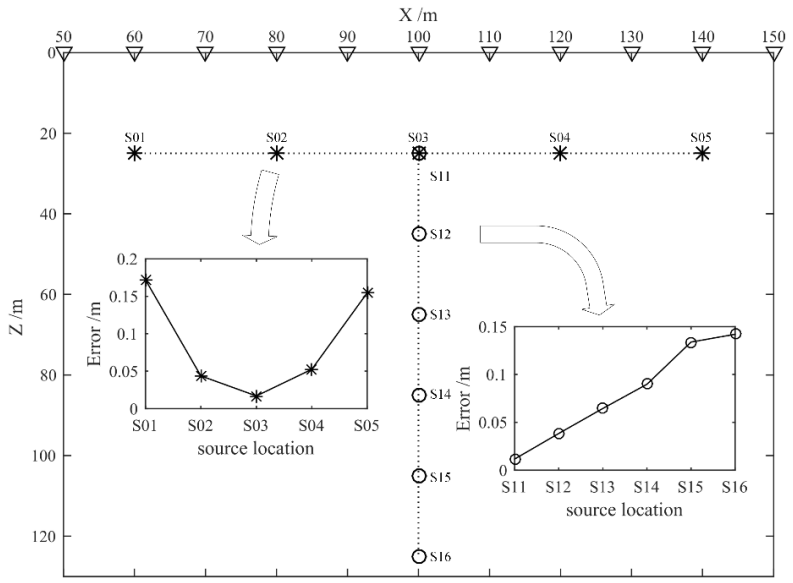


FIGURE 10. Positioning errors of sound sources at different locations.

To analyze the performance of the bat algorithm for SRP-PHAT optimization further, different levels of Gaussian white noise are introduced into the modeling data to generate various simulation data with different signal-to-noise ratios (SNR $= -10$ to 10 dB). The relationship between positioning error and SNR is obtained by Monte Carlo simulation as shown in Fig. 7. With the increasing of SNR, the positioning error of both bat algorithm and SRC decrease greatly. When SNR < 1, bat algorithm has higher accuracy than SRC, which indicates that the bat algorithm is insensitive to noise. When SNR > 1, the accuracy of bat algorithm is almost the same as SRC. But the accuracy of bat algorithm is better overall. For bat algorithm, the maximum error is 0.64 m with 0.31% error percentage[1] and the minimum error is 0.04 m with a 0.02% error percentage. While for SRC, the maximum error is 1.26 m with 0.61% error percentage and the minimum error is 0.06 m with a 0.03% error percentage. The average errors of bat algorithm and SRC are 0.19 m and 0.42 m respectively.

It demonstrates that the first-stage optimization of bat algorithm possesses a faster convergence rate and a higher positioning accuracy than the common methods.

C. SECOND-STAGE OPTIMIZATION ANALYSIS

When velocity is unknown, the verification of the unique maximum of energy function $\boldsymbol{S}(\boldsymbol{v})$ is the key issue for positioning by velocity scanning. In section II, the unique maximum of $\boldsymbol{S}(\boldsymbol{v})$ was verified theoretically. Next, simulation experiments are conducted to further verify it. The $\boldsymbol{S}-\boldsymbol{V}$ curve can be obtained by full velocity scanning with a scanning range from 300 to 1300 m/s and an interval of 10 m/s, as illustrated in Fig. 8. It can be observed that S first increases and then decreases with an increase in velocity. When the velocity $\boldsymbol{v} = 750$ m/s (average velocity of the model in subsection 5.1), S reaches its maximum value. These changes indicate that S reaches its maximum value only when the velocity is scanned to the actual average velocity.

As illustrated in Fig. 8, the SRP energy reaches its maximum when the average velocity is close to the true average velocity. To investigate the influence of the velocity change on the localization accuracy, repeated experiments are conducted at various velocity with noise-free modeling data. The velocity varies from 700 to 800 m/s with a 1 m/s interval, and the location error is illustrated in Fig. 9. With further deviation of the velocity from the true average velocity, the localization error increases, and the error is approximately linear with the deviation degree. The velocity change within a small region near the true average velocity has little influence on the positioning accuracy. As indicated in the subfigure of Fig. 9, the positioning error is less than 0.4 m at a velocity deviation of below 4 m/s, which still represents high positioning accuracy. Therefore, the velocity scanning interval of 1 to 2 m/s can achieve relatively high localization accuracy in practical applications, and the velocity scanning interval can be appropriately increased according to actual requirements.

As indicated in Fig. 8 and Fig. 9, the $\boldsymbol{S}(\boldsymbol{v})$ is a convex function with a unique extremum. And the spatial position at the $\boldsymbol{S}(\boldsymbol{v})$ extremum point corresponds to the real source position. Therefore, the source localization can be achieved by the SRP-PHAT-UNVEL.

D. LOCATION ANALYSIS WITH SIMULATION DATA

The simulation results of the two-stage optimization process demonstrate that the sound source localization under unknown velocity can be achieved by SRP-PHAT-UNVEL theoretically. In this section, an overall analysis of the

[1]The error percentage is defined as the percentage of the positioning error from the actual sound source position to the coordinate origin

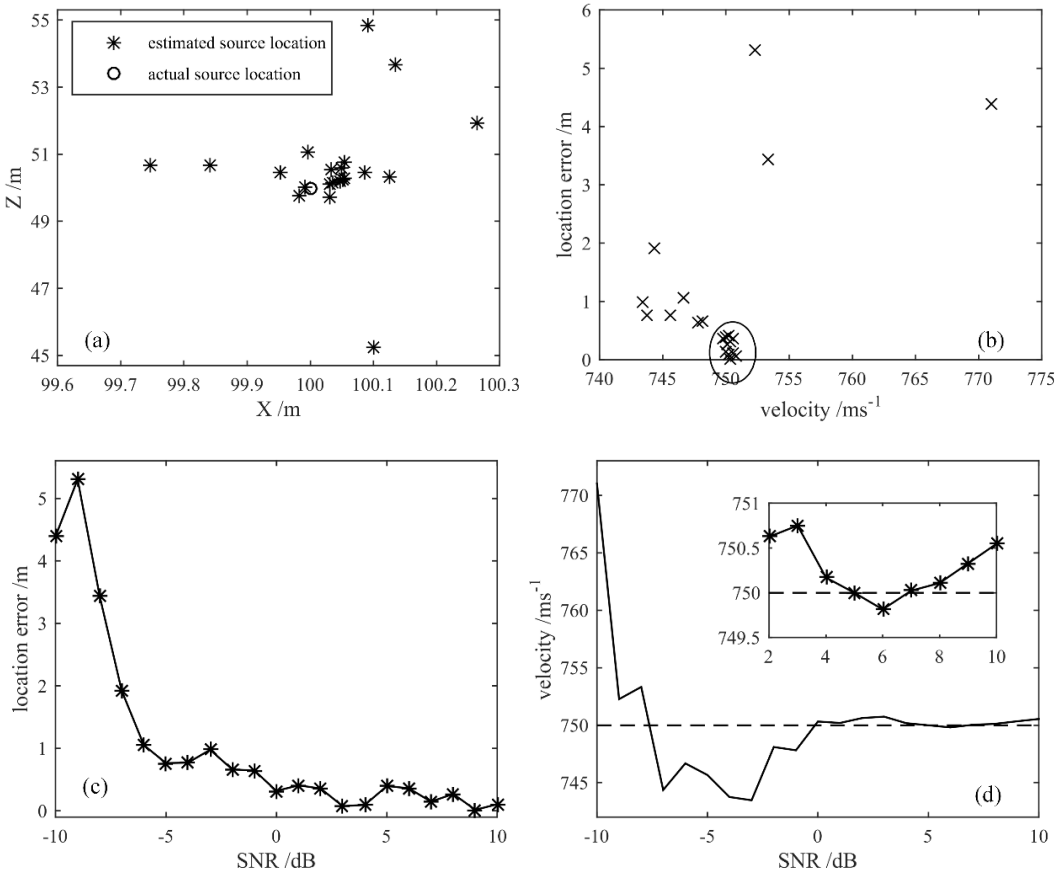


FIGURE 11. The location and velocity estimation for noisy simulation data with different SNR. (a) Spatial distribution of the estimated location. (b) Relationship between velocity and localization error. (c) Relationship between localization error and SNR. (d) Relationship between estimated velocity and SNR.

SRP-PHAT-UNVEL positioning performance is given. The joint estimation performance of SRP-PHAT-UNVEL for the position and velocity of noise-free simulation data is firstly studied. The localization result is (100.087 m, 50.262 m) with a positioning error of 0.184 m, while the estimated average velocity is 749.642 m/s with an error of 0.358 m/s, which is very close to the actual average velocity of 750 m/s.

The above analysis is only applicable for a fixed sound source and it is not universal. Therefore, the localization performance of the acoustic source at different locations should be further verified. The positioning errors of different acoustic source locations were analyzed, where the locations of acoustic source change in the horizontal direction or in the vertical direction, as shown in Fig. 10. The positioning error of the sound source at S_{03} is the minimum. The positioning error is larger with the acoustic source farther away from S_{03} along the horizontal direction, and it is symmetrically distributed around S_{03}. This may be the ambiguous estimation of X coordinate since the energy of the SRP spectrum in the X direction is not concentrated when the acoustic source changes in the horizontal direction. While, the positioning error increases with the acoustic source is far away from the center of the sensors array along the vertical direction. Compared with the array aperture, the location of the acoustic source is far from the center, which will inevitably reduce the positioning accuracy. However, as long as the acoustic source is guaranteed to be a near-field source, its positioning error is sufficiently small. No matter the location of the acoustic source changes horizontally or vertically, the estimated velocity varies slightly, and its average value is 750.85 m/s.

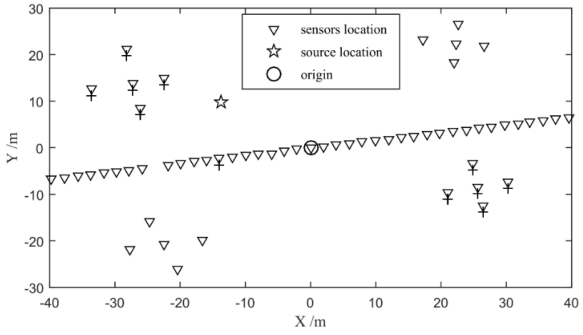


FIGURE 12. The surface locations of sensors and source. The sensors with mark '+' belong to one group.

The results demonstrate that the method can achieve source positioning in a shallow underground unknown medium with high localization accuracy and a small estimated velocity error.

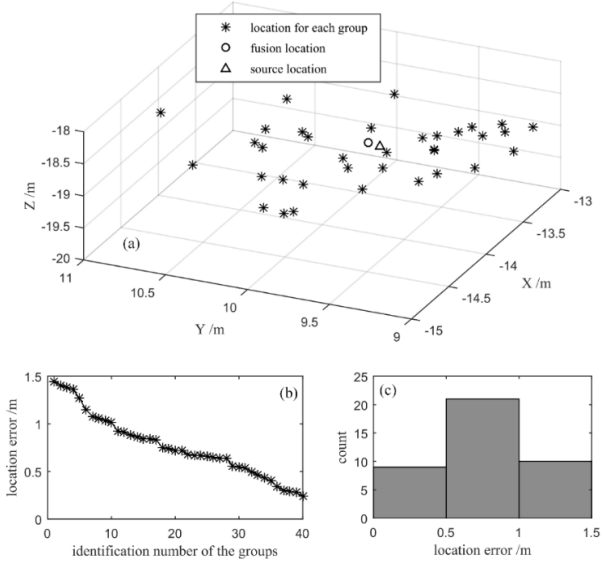


FIGURE 13. Grouping localization results and errors. (a) Spatial distribution of location for each group, (b) the location error of each group are sorted from largest to smallest and (c) histogram of the number of groups with different error ranges.

TABLE 1. Location results of the actual data.

	Source location (m)	**Estimated location** (m)	Location error of the grouping (m)			**location error** (m)
			Max	Min	Ave	
X	-13.80	**-13.83**	1.45	0.24	0.77	**0.1**
Y	9.74	**9.79**				
Z	-18.86	**-18.79**				
Estimated velocity	873 m/s					

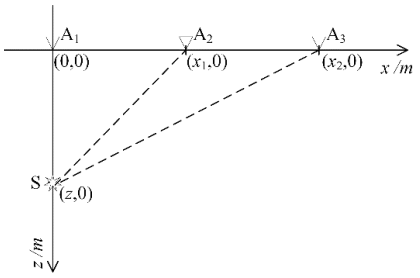


FIGURE 14. Geometric relationship between acoustic source and sensors array.

For noisy data with different SNRs (SNR $= -10$ to 10 dB), the estimated location and velocity results obtained by the proposed method are illustrated in Fig. 11. As can be observed from Fig. 11(a), the horizontal location error is small, while the vertical (depth) error is large. However, the number of position estimations with large error is not significant (only 4), which indicates that the algorithm is not sensitive to noise. From Fig. 11(b), it can be observed that localization results with a location error of less than 1 m account for the largest proportion, and the corresponding velocity is approximately 750 m/s. The estimated velocity is close to the real velocity for different SNRs, which indicates that the velocity estimation is not sensitive to noise. Fig. 11(c) illustrates that the localization error decreases with an increase in the SNR. The localization error is relatively large when the SNR is less than 0 dB, while it is relatively small when the SNR is greater than 0 dB. Fig. 11(d) illustrates that the estimated velocity is closer to the real velocity with an increase in the SNR; when the SNR is greater than 0 dB, the velocity gradually changes.

Therefore, the proposed method can locate the acoustic source and obtain the estimated velocity at the same time with a small error when the SNR > 0.

E. LOCATION ANALYSIS WITH ACTUAL DATA

The collected signal from an underground acoustic source with a known position by the sensor array on the ground surface is referred to as actual data. Strictly speaking, actual data belong to semi-simulated data, as the source location is known. The position of the underground acoustic source is set to (X $= -13.80$ m, Y $= 9.74$ m, Z $= -18.86$ m). A total of 40 sensors are arranged in a straight line, and four five-element cross-arrays are arranged on both sides of the line, as illustrated in Fig. 12. The actual data are collected with a 1 ms sampling interval for a 1 s sampling time. The collected actual data are pre-processed, including data validity analysis and SNR enhancement. The pre-processed data are divided into 40 groups. Each group consists of the diagonal of two five-element crosses and an odd or even number of sensors in the linear array, which contains 11 sensors in total, such as the sensors marked by '+' in Fig. 12. Since the subsurface medium is unknown and often complex, the use of all collected data for positioning may not satisfy the simplified velocity model in section IV. Grouping increases the likelihood of meeting the simplified velocity model. Even if some groups do not satisfy the simplified velocity model, its impact on positioning can be weakened by data fusion. So the main purpose of the grouping is to improve the robustness of the proposed method further and cause the final localization result to be more stable, reliable, and accurate. Moreover, grouping can improve computational efficiency to a certain extent.

Each group data is used for positioning by the proposed method, and the localization results are illustrated in Fig. 13. Fig. 13(a) indicates that the localization result for certain groups contains a large location error, while the fused localization result is very close to the actual source. This demonstrates that the location accuracy can be improved by grouping and fusion. The location errors of each group are presented in Fig. 13(b), from which it can be observed that different groupings have different localization errors, and the error of most groupings is between 0.5 m and 1 m, as indicated in Fig. 13(c). Thereafter, the 40 localization results are fused by principal component analysis [39], and the final localization result and location error are displayed in Table 1. As indicated in Table 1, the maximum positioning error is approximately 1.45 m, the minimum positioning error is approximately 0.24 m, and the error average is 0.77 m.

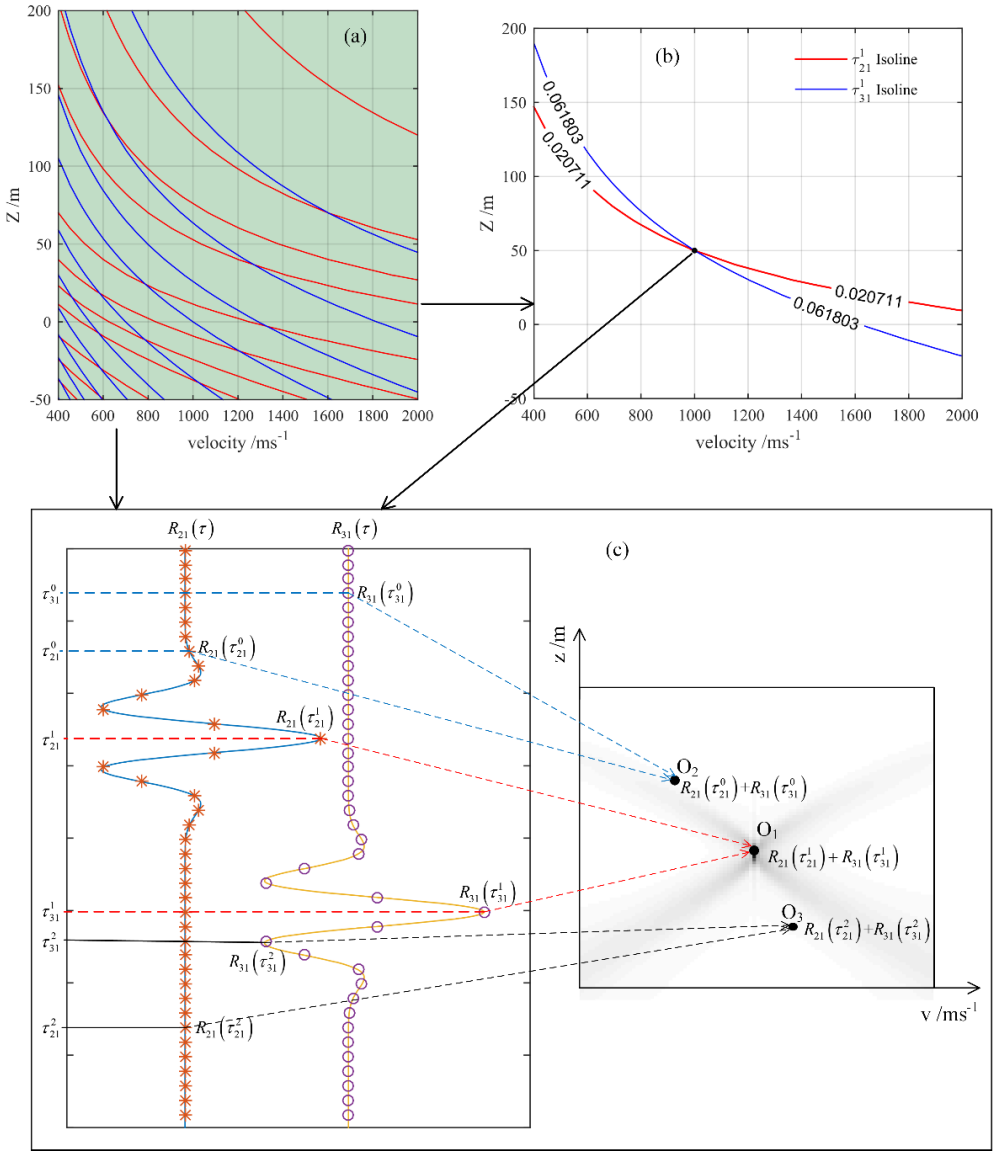


FIGURE 15. (a) Isotime line of different τ_{21} and τ_{31}, (b) isotime line of given τ_{21} and τ_{31}, and (c) illustration of the principle of position and velocity separation. $R_{21}(\tau)$ represents the GCC-PHAT of the sensor A2 and the sensor A1, $R_{31}(\tau)$ represents the GCC-PHAT of the sensor A3 and the sensor A1.

However, the fusion location error is only 0.1 m, which is less than the minimum value of the grouping localization error. The localization error percentage following fusion is only 0.39%. So, the grouping and fusion can improve the localization accuracy.

The actual data localization results demonstrate that the proposed method can achieve positioning of a shallow underground acoustic source with unknown velocity, while grouping and fusion can improve the final localization accuracy.

VI. CONCLUSION

As one of the most robust and accurate acoustic source localization methods, SRP-PHAT only applies to known-velocity cases. However, the accurate velocity models are hard to be achieved in the shallow underground acoustic source localization. In this paper, a special SRP-PHAT with unknown velocity is proposed to make the shallow acoustic source localization possible. The principles of SRP-PHAT-UNVEL are studied from the positioning essence and are solved by a specific two-stage optimization with efficient computational cost. The feasibility of the specific optimization method is discussed from the aspects of theory and experiment. The results demonstrate that the first-stage optimization based on bat algorithm provides better computing efficiency and higher estimation accuracy than traditional methods. Besides, the second-stage optimization can be achieved by the common linear search technique since the function of velocity and energy is an ideal convex function. The simulation results demonstrate that the proposed method can achieve the highly accurate joint estimation of

acoustic source position and velocity. Meanwhile, the proposed method is also applicable in the case with low SNR. Furthermore, the robustness and accuracy of the positioning for actual data can be further improved through the grouping data and fusion positioning results of all groups. Although the SRP-PHAT-UNVEL is proposed for shallow underground acoustic source localization, it can also be extended to similar cases in other fields.

APPENDIX

In order to clearly explain the idea and facilitate the visualizing, the geometric relationship between the sound source and the sensors array is shown in Fig. 14. This is a special case. But the same analysis can also be produced in the general situation, only the non-linear equation is more complex.

τ_{21} represents the time difference of arrive (TDOA) from S to sensor A2 and the sensor A1. τ_{31} represents the TDOA from S to the sensor A3 and the sensor A1. From the geometric relations in Fig. 14, we can get:

$$\tau_{21} = \frac{\sqrt{z^2 + x_1^2} - z}{v} \tag{A.1}$$

$$\tau_{31} = \frac{\sqrt{z^2 + x_2^2} - z}{v} \tag{A.2}$$

(A-1) and (A-2) can be further simplified as follow:

$$\begin{cases} \tau_{21} v + 2\tau_{21} vz = x_1^2 \\ \tau_{31} v + 2\tau_{31} vz = x_2^2 \end{cases} \tag{A.3}$$

For variables v and z, the equations (A-3) is nonlinear and nonhomogeneous. If τ_{21} and τ_{31} are unknown, the number of variables is always more than the number of equations, resulting in (A-3) underdetermined with infinite solutions. If both τ_{21} and τ_{31} are known, we can get v and z by solving (A-3) to achieve separation of position and velocity. This conclusion can be extended to the general geometric relations of arrays, but the nonlinear nonhomogeneous equations are more complex and more difficult to solve.

For a given τ_{21}^1, equation (A-1) has infinite sets of solutions to (z, v). Similarly, for a given τ_{31}^1, equation (A-2) has also infinite sets of solutions to (z, v). However, according to the equations (S3), it can be known that there must be a unique solution (z1, v1) for (A-1) and (A-2) to be true at the same time. This means that there is only one intersection point between the isotime lines determined by the function $\tau_{21}(z, v) = \tau_{21}^1$ and $\tau_{31}(z, v) = \tau_{31}^1$ (see Fig. 15(b)). Different τ_{21} and τ_{31} will determine different intersection points (see Fig. 15(a)).

In summary, when TDOA is known, the simultaneous estimation of velocity and position can be achieved by solving the equations of TDOA-distance-velocity.

The SRP-PHAT-UNVEL is not intended to solve the TDOA-distance-velocity equations to achieve a joint estimation of position and velocity. The joint scanning space and velocity are used to obtain $\tau_{21}^0, \tau_{21}^1, \tau_{21}^2, \cdots; \tau_{31}^0, \tau_{31}^1, \tau_{31}^2, \cdots$. And then their corresponding generalized cross-correlation values are added as output. Only when the scanned spatial position and velocity match the actual ones, the output energy is the largest, which is just the sum of the maximum value of the generalized cross correlation, as shown in O1 in Fig. 15c. But beyond that, any other solution (z, v) does not make the generalized cross-correlation maximums add up at the same place, as shown by O2 and O3 in Fig. 15c.

REFERENCES

[1] J. C. Chen, K. Yao, and R. E. Hudson, "Source localization and beamforming," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 30–39, Mar. 2002.

[2] S. Asgari, J. Z. Stafsudd, R. E. Hudson, K. Yao, and E. Taciroglu, "Moving source localization using seismic signal processing," *J. Sound Vib.*, vol. 335, pp. 384–396, Jan. 2015.

[3] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech Signal Process.*, vol. ASSP-24, no. 4, pp. 320–327, Aug. 1976.

[4] J. Benesty, J. Chen, and Y. Huang, "Time-delay estimation via linear interpolation and cross correlation," *IEEE Trans. Speech Audio Process.*, vol. 12, no. 5, pp. 509–519, Sep. 2004.

[5] Y. A. Huang and J. Benesty, "Adaptive multi-channel least mean square and Newton algorithms for blind channel identification," *Signal Process.*, vol. 82, no. 8, pp. 1127–1138, 2002.

[6] J. K. Tugnait, "Time delay estimation with unknown spatially correlated Gaussian noise," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 549–558, Feb. 1993.

[7] J. Benesty, Y. Huang, and J. Chen, "Time delay estimation via minimum entropy," *IEEE Signal Process. Lett.*, vol. 14, no. 3, pp. 157–160, Mar. 2007.

[8] Y. Huang, J. Benesty, G. W. Elko, and R. M. Mersereati, "Real-time passive source localization: A practical linear-correction least-squares approach," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 8, pp. 943–956, Nov. 2001.

[9] J. Benesty, M. M. Sondhi, and Y. Huang, *Springer Handbook of Speech Processing*. Berlin, Germany: Springer-Verlag, 2008.

[10] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 8, pp. 1223–1225, Aug. 1987.

[11] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 12, pp. 1661–1669, Dec. 1987.

[12] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994.

[13] L. Geiger, "Probability method for the determination of earthquake epicenters from the arrival time only," *Bull. St. Louis. Univ.*, vol. 8, no. 1, pp. 56–71, 1912.

[14] A. Douglas, "Joint Epicentre determination," *Nature*, vol. 215, pp. 47–48, Jul. 1967.

[15] R. S. Crosson, "Crustal structure modeling of earthquake data: 1. Simultaneous least squares estimation of hypocenter and velocity parameters," *J. Geophys. Res.*, vol. 81, no. 17, pp. 3036–3046, 1976.

[16] W. Spence, "Relative epicenter determination using P-wave arrival-time differences," *Bull. Seismol. Soc. Amer.*, vol. 70, no. 1, pp. 171–183, 1980.

[17] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

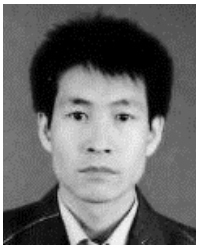
[18] B. D. Rao and K. V. S. Hari, "Performance analysis of root-music," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.

[19] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[20] G. Liu and X. Sun, "Efficient method of passive localization for mixed far-field and near-field sources," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, pp. 902–905, 2013.

[21] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[22] M. Omologo and P. Svaizer, "Use of the crosspower-spectrum phase in acoustic event location," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 3, pp. 288–292, May 1997.
[23] J. H. DiBiase, "A high-accuracy, low-latency technique for talker localization in reverberant environments using microphone arrays," Ph.D. dissertation, Brown Univ., Providence, RI, USA, 2000.
[24] J. P. Dmochowski, J. Benesty, and S. Affes, "A generalized steered response power method for computationally viable source localization," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 8, pp. 2510–2526, Nov. 2007.
[25] M. Cobos, A. Marti, and J. J. Lopez, "A modified SRP-PHAT functional for robust real-time sound source localization with scalable spatial sampling," *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 71–74, Jan. 2011.
[26] D. Salvati, C. Drioli, and G. L. Foresti, "Sensitivity-based region selection in the steered response power algorithm," *Signal Process.*, vol. 153, pp. 1–10, Dec. 2018.
[27] H. Do, H. F. Silverman, and Y. Yu, "A real-time SRP-PHAT source location implementation using stochastic region contraction(SRC) on a large-aperture microphone array," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Apr. 2007, pp. I-121–I-124.
[28] H. Do and H. F. Silverman, "A fast microphone array SRP-PHAT source location implementation using coarse-to-fine region contraction(CFRC)," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, Oct. 2007, pp. 295–298.
[29] H. Do and H. F. Silverman, "Stochastic particle filtering: A fast SRP-PHAT single source localization algorithm," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, Oct. 2009, pp. 213–216.
[30] L. O. Nunes, W. A. Martins, M. V. S. Lima, L. W. P. Biscainho, M. V. M. Costa, F. M. Gonçalves, A. Said, and B. Lee, "A steered-response power algorithm employing hierarchical search for acoustic source localization using microphone arrays," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 5171–5183, Oct. 2014.
[31] M. B. Coteli, O. Olgun, and H. Hacihabiboglu, "Multiple sound source localization with steered response power density and hierarchical grid refinement," (in English), *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 11, pp. 2215–2229, Nov. 2018.
[32] X. S. Yang, "A new metaheuristic bat -inspired algorithm," in *Nature Inspired Cooperative Strategies for Optimization (NISCO)* (Studies in Computational Intelligence). Berlin, Germany: Springer, 2010, pp. 65–74.
[33] X.-S. Yang and A. H. Gandomi, "Bat algorithm: A novel approach for global engineering optimization," *Eng. Comput.*, vol. 29, no. 5, pp. 464–483, 2012.
[34] K. Khan and A. Sahai, "A comparison of BA, GA, PSO, BP and LM for training feed forward neural networks in e-learning context," *Int. J. Intell. Syst. Appl.*, vol. 4, no. 7, pp. 23–29, 2012.
[35] I. Fister, S. Rauter, X.-S. Yang, and K. Ljubič, and I. Fister Jr., "Planning the sports training sessions with the bat algorithm," *Neurocomputing*, vol. 149, pp. 993–1002, Feb. 2015.
[36] A. Marti, M. Cobos, J. J. Lopez, and J. Escolano, "A steered response power iterative method for high-accuracy acoustic source localization," *J. Acoust. Soc. Amer.*, vol. 134, no. 4, pp. 2627–2630, 2013.
[37] D. Salvati, C. Drioli, and G. L. Foresti, "Exploiting a geometrically sampled grid in the steered response power algorithm for localization improvement," *J. Acoust. Soc. Amer.*, vol. 141, no. 1, pp. 586–601, 2017.
[38] A. R. Levander, "Finite-difference forward modeling in seismology," in *Geophysics Encyclopedia of Earth Science*. Boston, MA, USA: Springer, 1989.
[39] M. T. Sadeghi, M. Samiei, and J. Kittler, "Fusion of PCA-based and LDA-based similarity measures for face verification," *EURASIP J. Adv. Signal Process.*, vol. 2010, p. 23, Feb. 2010.



PENGFEI NIE received the B.Sc. and M.S. degrees from the College of Geo-Exploration Science and Technology, Jilin University, in 2005 and 2008, respectively, and the Ph.D. degree from the College of Communication Engineering, Jilin University, in 2012.

He was with the Qingdao Institute of Marine Geology for marine seismic data processing and imaging. He is currently a Lecturer with the North University of China. His research interests include seismic signal processing and imaging, array signal processing, and pattern recognition.

BIN LIU received the Ph.D. degree from the School of Information and Communication Engineering, North University of China, in 2014. He is currently an Associate Professor with the North University of China, where he is involved in the field of light field imaging, image processing, and signal processing and recognition.

PING CHEN received the B.Sc., M.S., and Ph.D. degrees from the School of Information and Communication Engineering, North University of China.

He is a Communications Commentator of the Ministry of Informatics, National Natural Science Foundation of China, a member of the Taiyuan Branch, Chinese Computer Society, and the Director of the Information Branch, Shanxi Society of Experts and Scholars. He is currently a Professor with the North University of China, where he is mainly involved in signal and information processing, image processing and reconstruction, and photoelectric detection.

KUN LI received the B.E. degree from the School of Information and Communication Engineering, North University of China (NUC), Taiyuan, China, in 2014, where he is currently pursuing the Ph.D. degree. He was a Visiting Researcher with the School of Chemical and Process Engineering, University of Leeds (UoL), Leeds, U.K., in 2017. His research interests include signal processing and electrical tomography.

YAN HAN received the B.S. degree in electrical engineering from the East China Engineering College (now Nanjing University of Technology), Nanjing, China, in 1982, the M.S. degree in testing technology and instrument from the North China Institute of Technology, Taiyuan, China, in 1987, and the Ph.D. degree in information and communication engineering from the Beijing Institute of Technology, Beijing, China, in 1998.

He is currently a Professor and the Director of the Shanxi Key Laboratory of Signal Capturing and Processing and the Shanxi Engineering Technology Research Center of Modern Nondestructive Testing. His research interests include communications technology, signal processing and recognition, automatic detection technology, digital image processing, and information reconstruction.