Ocean Engineering 322 (2025) 120442

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Research paper

Improvements of bearing time records target extraction method based on culstering

Hao Yin [a,b], Chao Li [a,b,*], Haibin Wang [a,b], Jun Wang [a,b], Zili Qin [a,b], Yannick Benezeth [c], Fan Yang [d]

[a] State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China
[b] University of Chinese Academy of Sciences, Beijing 101408, China
[c] IMVIA, UR 7535, Université de Bourgogne, Dijon 21078, France
[d] CNRS UMR5022, Université de Bourgogne, Dijon 21078, France

ARTICLE INFO

Keywords:
Underwater signal detection
Computer vision
K-means
Bearing time records

ABSTRACT

This study aims to improve automated target detection in Bearing Time Records (BTRs) images, addressing key challenges such as low signal-to-noise ratios and false alarms. Motivated by the need for more reliable detection methods in marine environments, we propose three key techniques: (1) Global–Local Peak initialization to optimize cluster center setup, (2) an adaptive method using SSE derivatives for precise determination of cluster count, and (3) a fuzzy rule based on correlation coefficient histograms to reduce false alarms. The proposed approach demonstrates 100% accuracy in simulations and has proven highly effective in sea trials, significantly enhancing the reliability of target detection in marine settings.

1. Introduction

Underwater acoustic signals are more extensively utilized and more advanced than electromagnetic and light waves for information transmission, owing to their properties of low attenuation and long-range propagation. Passive underwater acoustic target detection analyzes noise from targets captured by passive acoustic sensor systems. Bearing Time Records (BTRs) are essential visualization tools in sonar and radar systems for underwater target detection and tracking, as shown in Fig. 1. BTRs are created by plotting the power of beamforming signals using algorithms like Conventional Beamforming (CBF) or Minimum Variance Distortionless Response (MVDR) over a two-dimensional plane of time and bearing angle. In a BTR, the horizontal axis represents the bearing angle, while the vertical axis represents time. High-intensity regions correspond to strong signal returns from specific directions at particular times, indicating the presence and movement of targets. Traditionally, sonar operators with extensive experience have manually analyzed and detected targets in the post-processing stage of underwater acoustic signals. However, with the rapid development of unmanned devices, there is a pressing need to develop a reliable, intelligent target detection system for unmanned platforms to replace traditional manual judgments. To address this need, existing target detection algorithms based on BTRs have been developed, categorized into fixed threshold methods, enhancement algorithms, motion analysis, and unsupervised learning techniques. Fixed threshold methods struggle with adaptive settings in complex marine environments, while enhancement algorithms, like direction estimation and background equalization, add computational burden. Motion analysis methods, such as Hidden Markov Models and Kalman filtering, require manual setup, limiting automation. Deep learning approaches, like segmentation networks, show promise but face real-time performance challenges. These existing methods, though valuable, have clear limitations that hinder their full potential in autonomous underwater target detection.

Given these challenges, our research focuses on advancing computer vision-based passive acoustic processing technologies to overcome the limitations of existing methods. With advancements in intelligent equipment and unmanned devices, computer vision-based technologies for underwater acoustic information processing have become increasingly significant. Employing image processing, pattern recognition, and deep learning, these technologies automate the analysis of acoustic information and target recognition, thus replacing traditional manual methods. The evolution of this technology not only boosts the automation and efficiency of processing acoustic information but also can be integrated into intelligent underwater devices to enhance their autonomy and perceptual abilities. Our research is dedicated to advancing computer vision-based passive acoustic information processing technologies to address the complexities of the marine environment and meet the increasing demands of maritime tasks. It aims to provide more

* Corresponding author at: State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China.
E-mail address: chao.li@mail.ioa.ac.cn (C. Li).

https://doi.org/10.1016/j.oceaneng.2025.120442
Received 17 October 2024; Received in revised form 13 January 2025; Accepted 18 January 2025
Available online 31 January 2025

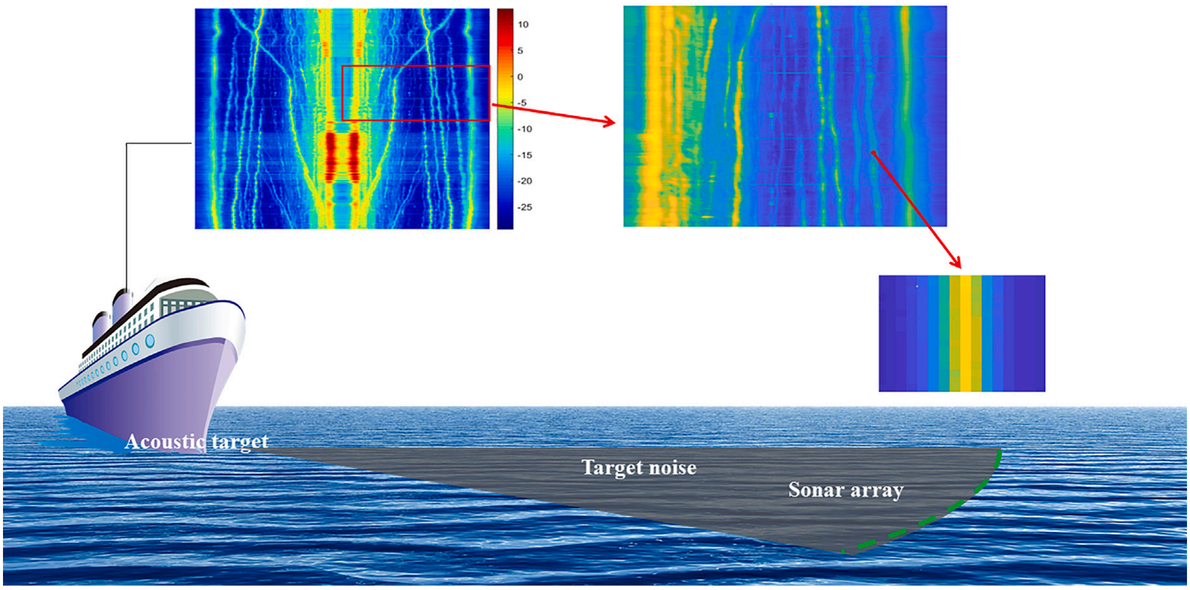

Fig. 1. Underwater acoustic target detection based on BTRs.

Table 1
Summary of methods used in underwater target detection.

Method category	Main concept	Limitations	Main references
Threshold methods	Utilizes threshold to extract peaks for potential targets.	Setting thresholds for different marine environments remains challenging.	Li et al. (2012)
Enhancement algorithms	Focuses on post-processing methods.	May enhance side lobes or lose weak target information.	Zhang et al. (2020), Carbone and Kay (2012), Zhu et al. (2021), Yang (2017) and Yin et al. (2023)
Motion analysis algorithms	Establishes motion models for dynamic state analysis of targets; Employs methods like HMM and Kalman filtering to track targets before detection.	Requires manual intervention for the choice of initial point.	Yin et al. (2019), Zhang et al. (2024), Li et al. (2019), Kaba and Temeltas (2022), Kim et al. (2017), Xin et al. (2015, 2017) and Northardt and Nardone (2018)
Supervised learning methods	Utilizes deep learning networks for extracting and tracking targets.	Requires large datasets for training; lacks a clear physical mechanism.	Shin et al. (2023)
Unsupervised learning methods	Constructs bases using unsupervised learning techniques.	Validated to achieve high accuracy but may require more accurate tuning.	Yin et al. (2024)

efficient and reliable technical support for maritime operations and drive their progression towards greater intelligence and automation.

The remainder of this paper is organized as follows: Section 2 provides a comprehensive review of related work; Section 3 presents the proposed false alarm suppressing method in detail; Sections 4 and 5 analyze the simulation and experiment results, respectively; finally, Section 6 concludes this work.

2. Related work

The existing target detection algorithms based on BTRs can be divided into five categories, as summarized in Table 1. The first category includes fixed threshold methods, which struggle with adaptive threshold settings due to the complexity of marine environments. The second category features enhancement algorithms, including advanced direction estimation and background equalization techniques. The former aims to increase target detection accuracy, though it may add computational burden and have limited applicability in practical sea experiments. The latter improves detection performance through post-processing but may also enhance side lobes or obscure weak target information. The third category includes motion analysis algorithms that develop motion models for analysis or employ Hidden Markov Models (HMM) and Kalman filtering for pre-detection tracking. These methods often require manual setup, which hinders full automation. The fourth category utilizes unsupervised learning, specifically segmentation networks, for detection. However, these methods lack a clear physical mechanism and struggle with real-time performance in complex marine environments. Lastly, we previously introduced an unsupervised learning approach for beam pattern extraction that effectively reduces false alarms.

Conventional passive sonar target detection methods primarily use energy detection techniques, which set a fixed threshold based on peaks extraction in BTRs to identify potential targets. However, due to the complexity of underwater environments, this method often causes false alarms and false negatives. In response, Li et al. (2012) developed an algorithm that employs adaptive thresholds. Despite this progress, accurately distinguishing between noise and true targets to set adaptive thresholds for various marine conditions remains a significant challenge.

In underwater target detection, image enhancement technology is crucial for improving detection performance. This technology is primarily divided into two categories: improved azimuth estimation algorithms and background equalization techniques. Improved azimuth estimation algorithms aim to enhance target localization accuracy by improving algorithm precision (Zhang et al., 2020). Although these methods theoretically improve accuracy, they also increase computational demands and may affect system robustness. While improved azimuth estimation algorithms offer higher precision, their increased computational load makes them less ideal for time-sensitive applications. In contrast, background equalization techniques, as a post-processing method, are better suited for practical applications. Historically, these techniques were mainly used in frequency domain analysis, such as improving the image quality of low-frequency analysis and detecting noise spectrum envelope modulation. Early algorithms such as the Two-Pass Mean (TPM), Split Three-Pass Mean (S3PM), Order Truncated Average (OTA), and Split Averaging Exclusion Average (SAXA) have demonstrated distinct advantages in various applications (Struzinski and Lowe, 1984). The spectral background of narrowband passive acoustic detection systems was normalized using a windowed three-pass mean noise spectrum estimator to reduce bias under colored Gaussian noise conditions (Shapiro and Green, 2000). The 3D Minimum Variance Spectrum Estimation Normalizer further improved these effects (Carbone and Kay, 2012). Recently, background equalization techniques have continuously evolved. The introduction of L1 norm regularization and total variation regularization methods enhanced point-based and area-based features (Lei et al., 2016). The effectiveness of these algorithms is influenced by variables such as window size and thresholds. To address this issue, the Inverse Beam Characteristic Scanning (IBCS) algorithm was developed, optimizing parameter settings to enhance detection by processing intervals between local peaks and valleys (Zhu et al., 2021). Additionally, deconvolution beamforming technology further improved beamforming resolution (Yang, 2017), while the sub-band peak energy detection (SPED) algorithm applied to BTR denoising achieved significant results (Yin et al., 2023). A novel approach employing α-stable distribution modeling has been developed to supplant the traditional Gaussian model for processing Discrete Fourier Transform (DFT) coefficients of ship radiation noise, replacing the variance parameter with the scale parameter γ, thereby effectively enhancing BTR performance and increasing Peak Signal-to-Noise Ratio (PSNR) by 3.1 dB (Yu et al., 2023). Despite these advancements, challenges persist. Some normalization algorithms function as low-pass filters and may neglect weak energy peaks, potentially resulting in the loss of high-frequency information. Moreover, while some enhancement algorithms suppress noise and preserve high-frequency details, they may inadvertently amplify side lobes, leading to new false alarms.

Target motion analysis, a third method for detecting targets by determining their dynamic state information, can be broadly classified into two primary categories: establishing motion models and track-before-detect strategies. Both categories present complex challenges of nonlinear parameter estimation, which are critical to effectively understanding and predicting target movements (Nardone et al., 1984). Traditional systems, known as Bearing-Only Target Motion Analysis (BOTMA), have often struggled with consistency and effectiveness due to their inherent unobservability issues (Ince et al., 2009). Early BOTMA algorithms relied heavily on prior information to formulate deterministic solutions, limiting their broader applicability (Pham, 1993). Classical methods, such as Linear Least Squares (LLS) and Pseudo Linear Estimation (PLE), faced difficulties due to system nonlinearity (Ho and Chan, 2006). While Nguyen's method is computationally efficient and relatively straightforward to implement, its reliance on multiple iterations can hinder convergence and introduce estimation biases (Nguyen and Doğançay, 2017). To address the inherent biases of PLE, researchers have developed more refined algorithms that incorporate corrected auxiliary variables (Zhang and Xu, 2010). With the evolving demands of real-time processing, recursive methods have emerged as alternatives to traditional batch processing in target motion analysis. These methods aim to address the limitations of iterative processes, which are sensitive to initial conditions and search steps. However, in practical applications, especially in scenarios involving quasi-stationary platforms and single vector hydrophones, these methods still face challenges such as bias and divergence. These scenarios are underexplored, and obtaining accurate motion analysis results from low-quality BTRs remains difficult. Recent efforts have focused on refining these techniques. For example, Fan (Yin et al., 2019) guided BTR positioning information processing based on target motion analysis results, and Zhang et al. (2024) proposed a motion analysis method for low-quality BTRs that enhances the pseudo-linear estimation algorithm by extracting peak and dispersion data, setting custom constraints, and establishing adaptive parameters (Zhang et al., 2024). This method has been validated for its effectiveness in improving motion analysis. In conclusion, despite progress in target motion analysis, significant challenges persist in practical applications, particularly when dealing with low-quality data and in quasi-stationary scenarios. Future research is needed to optimize these algorithms, enhancing their adaptability and accuracy under different conditions.

In underwater target detection, track-before-detect strategies represent a key approach in decision analysis methods. These strategies extensively employ HMM and Kalman filtering to effectively monitor targets. Aidala was a pioneer in utilizing Kalman filtering for target state estimation (Aidala, 1979). Following his work, the adoption of nonlinear filtering methods, such as the Extended Kalman Filter (EKF), has increased. These methods are prized for their recursive algorithms, which significantly reduce computational demands (Lin et al., 2002; Li et al., 2019; Kaba and Temeltas, 2022; Kim et al., 2017), making them highly effective in complex underwater environments. HMM are well-suited for describing trajectory state transitions and have found widespread application in underwater signal processing. Xin et al. (2015) introduced a track-before-detect algorithm based on HMM designed for detecting and tracking weak BTR trajectories, achieving a claimed 3 dB gain in detection performance. Another study by Xin et al. (2017) developed an HMM algorithm optimized for scenarios involving multiple targets and indeterminate target quantities. Although these motion analysis-based target detection methods perform well theoretically, there are some issues in practical applications. The main problem is that these methods fail to adequately handle the association of measurement tracks in complex situations, leading to frequent operator interventions to find auxiliary points. Northardt and Nardone (2018) proposed a track-before-detect algorithm that does not require manual intervention and can continuously track target crossings, mergers, and splits, avoiding track loss. However, the likelihood function designed by this algorithm is not suitable for complex sea trials because it sacrifices the tracking capability for weak signals to enhance the robustness of tracking strong but intermittent signals.

As the fourth method, Shin et al. (2023) developed a deep learning-based segmentation network (DLV3+MSC) specifically for extracting targets from noisy and discontinuous BTR images. This network employs supervised learning, training on noisy BTR images alongside their

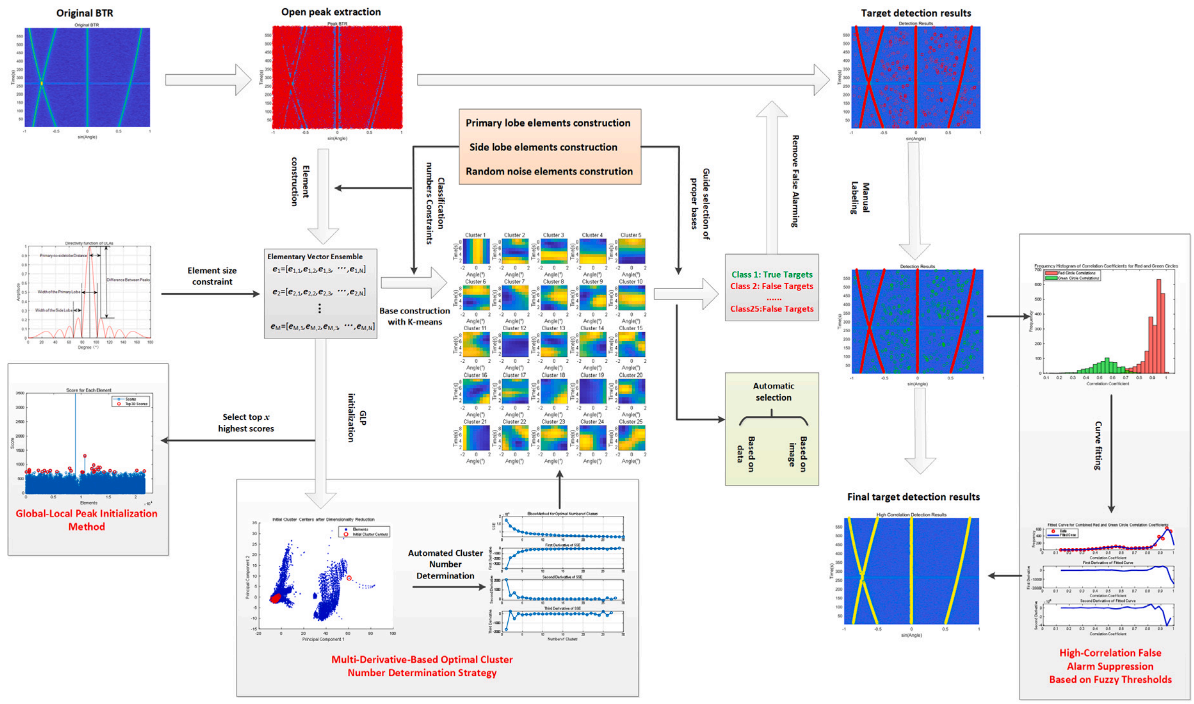

Fig. 2. Technical method.

labeled counterparts to predict continuous target trajectories. It also introduces innovative loss functions, which have significantly enhanced recall rates and F3-score to 92.28% and 73.28%, respectively. Despite these advancements, this machine learning approach faces challenges, including the absence of a clear physical mechanism, dependency on extensive datasets, and limited adaptability in complex underwater environments. Moreover, it lacks the capability to operate in real-time.

In the evolution of the k-means clustering algorithm, various initialization methods have been proposed to improve the performance and stability of the results. MacQueen initially suggested random selection of initial cluster centers (MacQueen et al., 1967), an approach that often resulted in unstable outcomes due to its inherent randomness. Subsequently, Gonzalez introduced the MaxMin Distance method (Gonzalez, 1985), which strategically optimizes the initial point selection by maximizing the minimum distance between cluster centers, enhancing both clustering quality and algorithm robustness. In the early 21st century, Ding et al. combined Principal Component Analysis (PCA) with k-means clustering (Ding and He, 2004), introducing an initialization strategy that not only preserved major data features but also reduced dimensionality, thus refining the clustering process. Further developments led Arthur and Vassilvitskii to propose the k-means++ method (Arthur and Vassilvitskii, 2006), which significantly improved clustering quality by employing a probabilistic method to select initial centers. Kaufman and Rousseeuw later developed the KKZ method (Kaufman and Rousseeuw, 2009), which selected initial centers by identifying points that maximize distance from each other, intending to cover the data space more effectively. However, these traditional methods exhibited limitations when applied to our specific tasks.

The method we previously proposed is an unsupervised learning approach for beam pattern extraction, aimed at suppressing false alarms. This method constructs bases (these refer to the visual patterns of the main lobe, side lobes, or random noise, extracted through unsupervised learning, which hold clear physical significance.) for the main lobe, side lobes, and noise using unsupervised learning to suppress false alarms, providing a clear physical mechanism. This article is based on the above content and proposes an improved method. Based on our previous unsupervised learning approach, future research could benefit from the latest advancements in clustering algorithms. For instance, Singh and Huang (2024a) proposed the Ambiguous Kernel Distance Clustering (AKDC) algorithm, which, based on ambiguous set theory, effectively handles ambiguous pixel information and has shown promising performance in complex data analysis. This provides useful insights for data processing in underwater acoustic target detection based on BTRs. Additionally, Singh and Huang (2024b) introduced the Ambiguous Edge Detection Method (AEDM), which, also based on ambiguous set theory, effectively identifies target boundaries in noisy images. Furthermore, Singh and Bose (2021) developed a method combining K-means clustering with the Fast Forward Quantum Optimization Algorithm (FFQOA), which has achieved significant results in high-dimensional data processing. These clustering-based approaches offer new directions for optimizing future underwater acoustic target detection systems based on BTRs.

3. Method

In our last study (Yin et al., 2024), we employed unsupervised learning methods for base construction and used data-driven and algorithm-driven methods for base automatic classification. Although theoretically, automatic classification of bases can achieve 100% accuracy, we observed an approximate 3% false alarm rate in practical applications, compared to human eye recognition. We found that this 3% false alarm rate mainly originates from the unsupervised learning process in the base construction phase. During this phase, unsupervised learning identifies meaningful structures from a large amount of unlabeled data, which aids in the automatic classification of bases. However, since this method does not utilize prior label information, it occasionally struggles to distinguish between very similar categories in certain scenarios, resulting in false alarms.

(1) The first issue is the uncertainty in the number of clusters k. In theory, the optimal number of clusters should include the main lobe, left sidelobe, right sidelobe, and various categories of random noise, which are typically numerous and mutually orthogonal. As shown in the section on base construction using K-means in Fig. 2, Cluster 1 corresponds to the visual pattern of the main lobe, while Clusters 19 and 21 correspond to the visual patterns of the right and left sidelobes, respectively. All other clusters represent the visual patterns

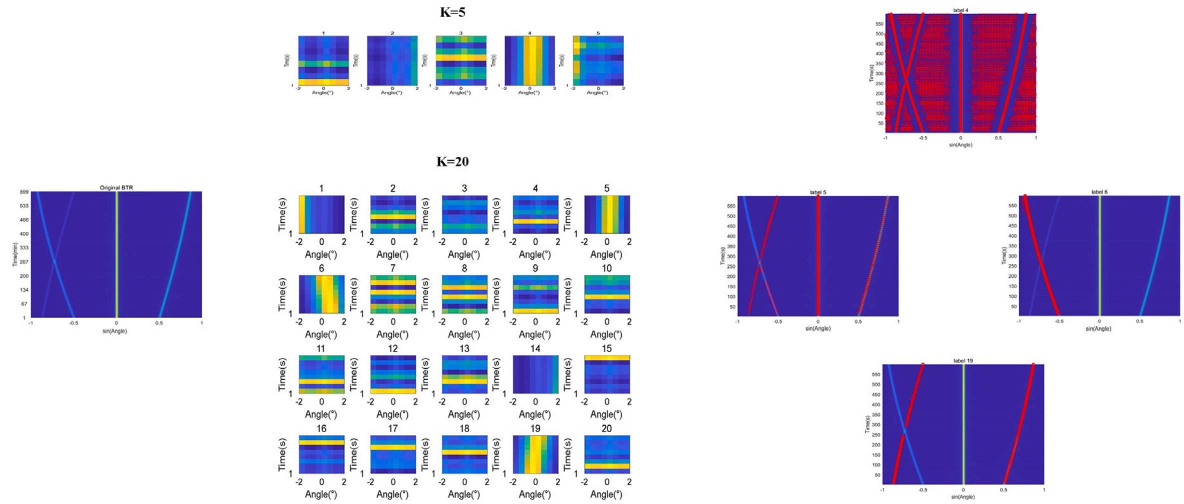

Fig. 3. Fewer k numbers, more false alarms; more k numbers, higher risk of false negative.

of random noise. The number of distinguishable noise categories varies across different scenarios. As shown in Fig. 3, when $k = 5$, Cluster 4 corresponds to the visual pattern of the main lobe, and the target detection result contains a large number of false alarms. When $k = 20$, Clusters 5, 6 and 19 all correspond to the visual pattern of the main lobe. Increasing the number of clusters(k) increases the risk of false negatives. Therefore, setting a fixed number of clusters does not meet the task requirements, as it may result in random noise being incorrectly classified into the main lobe cluster. It is important to clarify that increasing k does not inherently guarantee improved results. A larger k can lead to overfitting, where the main lobe pattern may be divided into multiple sub-clusters. The objective is to identify the optimal value of k, as different values of k generate distinct clustering outcomes, and the goal is to select the most appropriate k for the given task. There is a pressing need to develop an algorithm that can adaptively determine the number of clusters to enhance the accuracy and efficiency of underwater acoustic target detection. Such an algorithm would enable more precise differentiation and classification of complex patterns, thereby optimizing clustering outcomes and reducing false alarm rates.

(2) The second issue is the misclassification of random noise as the main lobe pattern. In practical applications, we observe that even with a sufficiently large number of clusters, random noise is still occasionally misclassified as part of the main lobe pattern, which significantly increases the false alarm rate. This issue arises because of the inherent uncertainty and complexity of random noise, which often makes it indistinguishable from the main lobe pattern. Current algorithms may lack sufficient discriminatory power when dealing with these random noises, leading to their incorrect classification into the main lobe pattern. To address this, it is crucial to explore an effective method that accurately distinguishes these random noises from the main lobe pattern, thereby reducing the false alarm rate and improving classification accuracy.

(3) The third issue is ensuring alignment between the azimuthal coordinates used in BTRs and those in Conventional Beamforming (CBF) directional function. When employing uniform linear arrays for CBF, the beamforming directivity functions are typically expressed in terms of sinusoidal angles. However, in practical applications such as the BTRs, the horizontal axis is commonly represented in angle values (in degrees). This difference leads to inconsistencies in the representation and response characteristics of azimuthal data. A critical aspect of addressing this issue is the development of a coordinate transformation method that reliably aligns the azimuthal information from beamforming directivity functions with that presented in BTRs.

By addressing the issues proposed above, we expect to enhance the performance of underwater acoustic target detection, reduce false alarms, and improve classification accuracy. The overall framework of our proposed method is shown in Fig. 2. First, we extract peaks from the original BTRs. Then, bases are constructed, followed by the creation of bases using the Global–Local Peak (GLP) initialization method and an improved clustering approach. Next, we apply automated base visual pattern selection, leading to target detection results after suppressing false alarms. Finally, we employ curve fitting of the high-correlation histogram to suppress false alarms based on fuzzy thresholds.

3.1. Global–local peak initialization method

For potential targets extracted from the BTRs, define appropriate sizes to construct elements $U(\theta_j, t)$. Use unsupervised methods to build a base, the unsupervised learning method used here is clustering. Initial cluster centers need to be set and optimized first, followed by iterative clustering and dynamic adjustment of the number of clusters based on a quantification strategy. The performance of clustering algorithms depends on the selection of initial cluster centers. If the initial cluster centers are poorly chosen, the clustering results may become trapped in local optima, failing to reflect the global characteristics of the data. To overcome this issue, we proposed the GLP initialization method to improve the clustering algorithms by optimizing the initialization process of cluster centers. This approach not only enhances the stability of the algorithms but also improves their ability to capture the global structure of the data, leading to more accurate clustering results.

The proposed method first calculates the global similarity G_i for each data point, followed by determining the local optimal similarity L_i. Based on these two metrics, we compute a comprehensive score S_i for each data point and select the top k points with the highest scores as the initial cluster centers. The specific steps are as follows:

(1) Initialization of Cluster Number, k:

Begin by selecting the initial number of clusters, k. This number can be determined through prior data analysis or heuristic methods.

(2) Calculation of Global Similarity, G_i:

For each e_i in the dataset, calculate the global similarity to the dataset. This is defined as the average correlation coefficient with all other elements, calculated as follows:

$$G_i = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^{n} R_{ij} \tag{1}$$

where R_{ij} is the correlation coefficient between elements e_i and e_j, and n is the total number of elements. The coefficient R_{ij} is given by:

$$R_{ij} = \frac{\sum_{p=1}^{P} e_{ip} e_{jp}}{\sqrt{\left(\sum_{p=1}^{P} e_{ip}^2\right) \cdot \left(\sum_{p=1}^{P} e_{jp}^2\right)}} \tag{2}$$

where e_{ip} and e_{jp} are the values of elements e_i and e_j at peak p, and P is the number of peaks.

(3) Calculation of Local Optimal Similarity, L_i:

Identify the local optimal similarity for each element e_i by finding the maximum correlation coefficient among elements that have a global similarity greater than G_i. The local optimal similarity is defined as:

$$L_i = \max_{j: G_j > G_i} R_{ij} \tag{3}$$

(4) Selection of Initial Cluster Centers:

Calculate a comprehensive score S_i for each element by multiplying the global similarity G_i and the local optimal similarity L_i. The formula for the score is:

$$S_i = G_i \times L_i \tag{4}$$

Rank the elements in descending order based on their scores S_i. Select the top k elements with the highest scores as the initial cluster centers:

$$C^{(0)} = \{c_1^{(0)}, c_2^{(0)}, \dots, c_k^{(0)}\} \tag{5}$$

3.2. Multi-derivative-based optimal cluster number determination strategy

Theoretically, the number of clusters in beam pattern analysis should correspond to the main lobe V_{main}, left sidelobe V_{left}, right sidelobe V_{right}, and various modes of random noise V_{noise}. Random noise can typically be divided into multiple categories, which are orthogonal to each other. The exact number of these noise categories may depend on the scenario. The main lobe visual pattern V_{main} is defined as the true target base visual pattern V_{true}, while the other visual patterns are categorized as false target base V_{false} (false alarms).

We propose a strategy called the Multi-Derivative-Based Optimal Cluster Number Determination Strategy for the automatic determination of the optimal number of clusters, which involves a combined analysis of first, second, and third-order derivatives. By thoroughly evaluating these derivatives, the strategy precisely identifies critical points for adjusting the number of clusters, thereby optimizing the clustering process to ensure high precision and reliability.

As detailed in Algorithm 1, the clustering iteration process involves several key steps: element assignment based on correlation coefficients, updating cluster centers, checking for convergence, and finally, outputting the results.

The proposed method dynamically adjusts the number of clusters based on the trend of the SSE as a function of the number of clusters k. The steps are as follows:

(1) Calculation of SSE and its Derivatives:
Each time the cluster centers are updated, calculate the SSE for the current number of clusters k and its derivative. For each number of clusters k, run the clustering algorithm and calculate the SSE for that k value:

$$\text{SSE}(k), \text{SSE}'(k), \text{SSE}''(k), \text{SSE}'''(k) \tag{6}$$

These derivatives help judge the trend of SSE with changes in the number of clusters. Among them:

- $\text{SSE}'(k)$ shows the rate of change of SSE with respect to k.
- $\text{SSE}''(k)$ shows the acceleration of the change in SSE.
- $\text{SSE}'''(k)$ shows the rate of change of the acceleration of SSE.

Algorithm 1 Detailed Clustering Iteration Process

Input: Set of elements $\{e_1, e_2, \dots, e_n\}$, number of clusters k
Output: Cluster labels for each element, cluster centers $C^{(t)}$
procedure ELEMENT ASSIGNMENT
  for $i = 1$ to n **do**
    for $j = 1$ to k **do**
      Compute $r(e_i, c_j^{(t)})$:

$$r(e_i, c_j^{(t)}) = \frac{\sum_{p=1}^{P} e_i^{(p)} c_j^{(t,p)}}{\sqrt{\sum_{p=1}^{P} (e_i^{(p)})^2 \cdot \sum_{p=1}^{P} (c_j^{(t,p)})^2}}$$

    end for
    Assign e_i to the cluster with the highest correlation:

$$y_i^{(t)} = \arg\max_j r(e_i, c_j^{(t)})$$

  end for
end procedure
procedure UPDATE CLUSTER CENTERS
  for $j = 1$ to k **do**
    Calculate new center for cluster j:

$$c_j^{(t+1)} = \frac{1}{|S_j|} \sum_{i \in S_j} e_i$$

    where $S_j = \{i \mid y_i^{(t)} = j\}$ and $|S_j|$ is the count of elements in S_j
  end for
end procedure
procedure CONVERGENCE CHECK
  if $\max_j \|c_j^{(t+1)} - c_j^{(t)}\| < \varepsilon$ **then**
    Convergence achieved
  else
    Return to ELEMENT ASSIGNMENT
  end if
end procedure
Output: Cluster labels $y_i^{(t)}$, final cluster centers $C^{(t)}$

(2) Determining the Optimal Number of Clusters k:
Select the k that satisfies the following conditions as the number of clusters:

- $\text{SSE}'(k) < 0$
- $\text{SSE}''(k) < 0$
- $\text{SSE}'''(k) > 0$ and is the largest among these k values.

The theoretical base for this strategy is that a negative first derivative ($\text{SSE}'(k) < 0$) indicates that as the number of clusters k increases, the SSE decreases. This implies that increasing the number of clusters can better fit the data. A negative second derivative ($\text{SSE}''(k) < 0$) indicates that the rate of decrease of SSE is accelerating, meaning that the improvement in clustering performance is increasing as k increases. A positive third derivative that is at its maximum ($\text{SSE}'''(k) > 0$) indicates that at this value of k, the rate of decrease of SSE starts to slow down, and this slowing down reaches a relative maximum point. This suggests that although increasing the number of clusters still reduces SSE, the effectiveness of this reduction begins to diminish significantly. Beyond this point, increasing the number of clusters may not provide significant benefits and may even degrade the model's performance due to over-clustering.

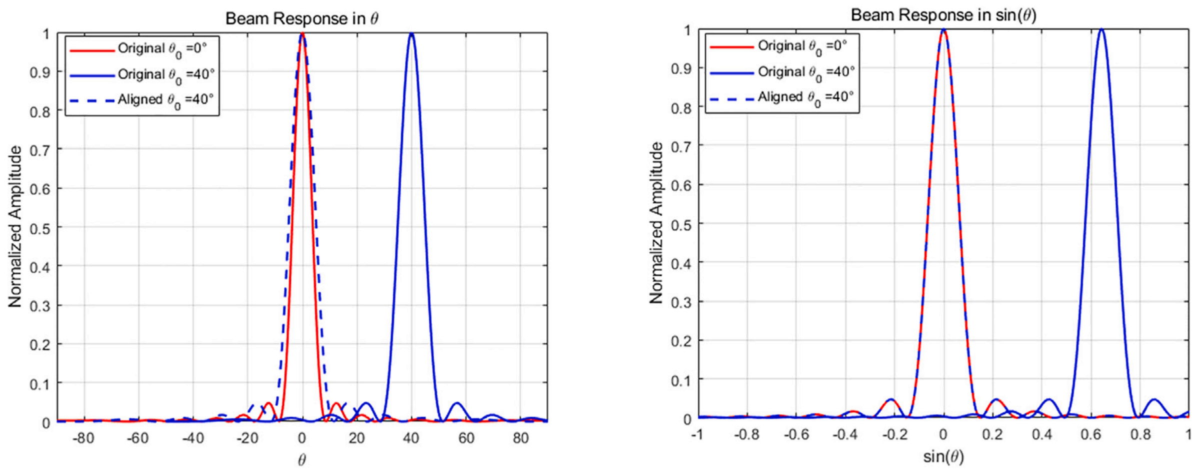

Fig. 4. Shift Invariance of CBF Directional Function.

The GLP initialization method and Multi-derivative-based optimal cluster number determination strategy are essential for improving clustering performance in underwater acoustic target detection. Poor initialization can lead to local optima, hindering target-noise differentiation, while an incorrect choice of k can cause false alarms or false negative. Therefore, using robust initialization techniques like GLP and adaptive methods to determine k is crucial for enhancing clustering accuracy and optimizing detection performance.

3.3. High-correlation false alarm suppression based on Fuzzy thresholds

Due to the inherent uncertainty and complexity of random noise, current algorithms may lack sufficient distinction when dealing with these noises, leading to their incorrect classification as the main lobe pattern, thereby generating false alarms. Since the core function of the algorithm is a correlation function, we propose a method called High-Correlation False Alarm Suppression Based on Fuzzy Thresholds. Algorithm steps for suppressing false alarms based on the correlation coefficient and fuzzy thresholds are as follows:

(1) Frequency histogram construction and curve fitting.

Plot all calculated correlation coefficients into a frequency histogram to observe their distribution, which is essential for subsequent curve fitting, if B is the class interval, then b_k represents the center value of the kth histogram bin. The range of each bin usually spans from $b_k - B/2$ to $b_k + B/2$. The frequency $H(b_k)$ of each bin can be represented as:

$$H(b_k) = \#\{r \mid b_k - B/2 \le r < b_k + B/2\} \tag{7}$$

where B is the class interval and b_k is the center value of the kth histogram bin.

Using the center point c_i and frequency n_i of each histogram bin, calculate the spline curve $S(c)$ such that it approximates the frequencies n_i at each bin center c_i. This involves applying a spline curve fitting method to match the spline curve $S(c)$ to the histogram.

(2) Derivative Calculation and fuzzy threshold.

In this task, there is often a significant difference between the correlation coefficients of true targets and false alarms. On the fitted frequency histogram, there is a clear demarcation point in the distribution of correlation coefficients for targets and false alarms, often corresponding to the maximum value of the second derivative. By identifying this demarcation point, false alarms can be effectively distinguished from true targets. Compute the first and second derivatives of the curve $S(c)$ to identify the key point in the distribution:

$$S'(c) = \frac{dS}{dc} \tag{8}$$

$$S''(c) = \frac{d^2S}{dc^2} \tag{9}$$

Find the maximum value of $S''(c)$, denoted as $\max(S''(c))$, and set the threshold τ to $\max(S''(c))$. Retain all detection results where $r_i > \tau$ and suppress the others as false alarms.

This method is crucial for distinguishing random noise from the main lobe pattern. By applying fuzzy thresholds, it reduces false alarms, improving classification accuracy and the reliability of the detection system.

3.4. Shift invariance in CBF using Sine of angle coordinates

To ensure shift invariance with the directional function of CBF, convert the horizontal axis of the BTRs from linear angular values to corresponding sine values. The essence of this transformation lies in preserving the shift invariance of the beam's directional response. Alongside this coordinate transformation, we have performed mathematical proofs to validate the efficacy of this approach. Shift invariance is a fundamental property in signal processing, indicating that a shift in the input signal results in a proportional shift in the output signal. This section provides a formal proof of shift invariance for the CBF directional function associated with a uniform linear array. The proof demonstrates that shifting the time variable t by τ results in a corresponding shift in the output $F_{CBF}(\vartheta, \tau)$.

Consider the CBF directional function for a uniform linear array defined as:

$$D(\theta) = A_0 \left| \frac{\sin\left(N\pi\frac{d}{\lambda}(\sin\theta - \sin\theta_0)\right)}{N\sin\left(\pi\frac{d}{\lambda}(\sin\theta - \sin\theta_0)\right)} \right| \tag{10}$$

where A_0 is a scaling factor, N is the number of array units, d is the units spacing, λ is the wavelength, θ is the angle of incidence, and θ_0 is the steering angle.

The BTR pixel value, denoted as $F_{CBF}(\vartheta, t)$, at a given direction ϑ and time t is expressed by:

$$\begin{aligned} F_{CBF}(\vartheta, \tau) &= |D_{CBF}(\vartheta, \tau)|^2 \\ &= \left| A_0 \frac{\sin[N\pi\frac{d}{\lambda}(\sin(\theta-\vartheta) - \sin\theta_0(t-\tau))]}{N\sin[\pi\frac{d}{\lambda}(\sin(\theta-\vartheta) - \sin\theta_0(t-\tau))]} \right|^2. \end{aligned} \tag{11}$$

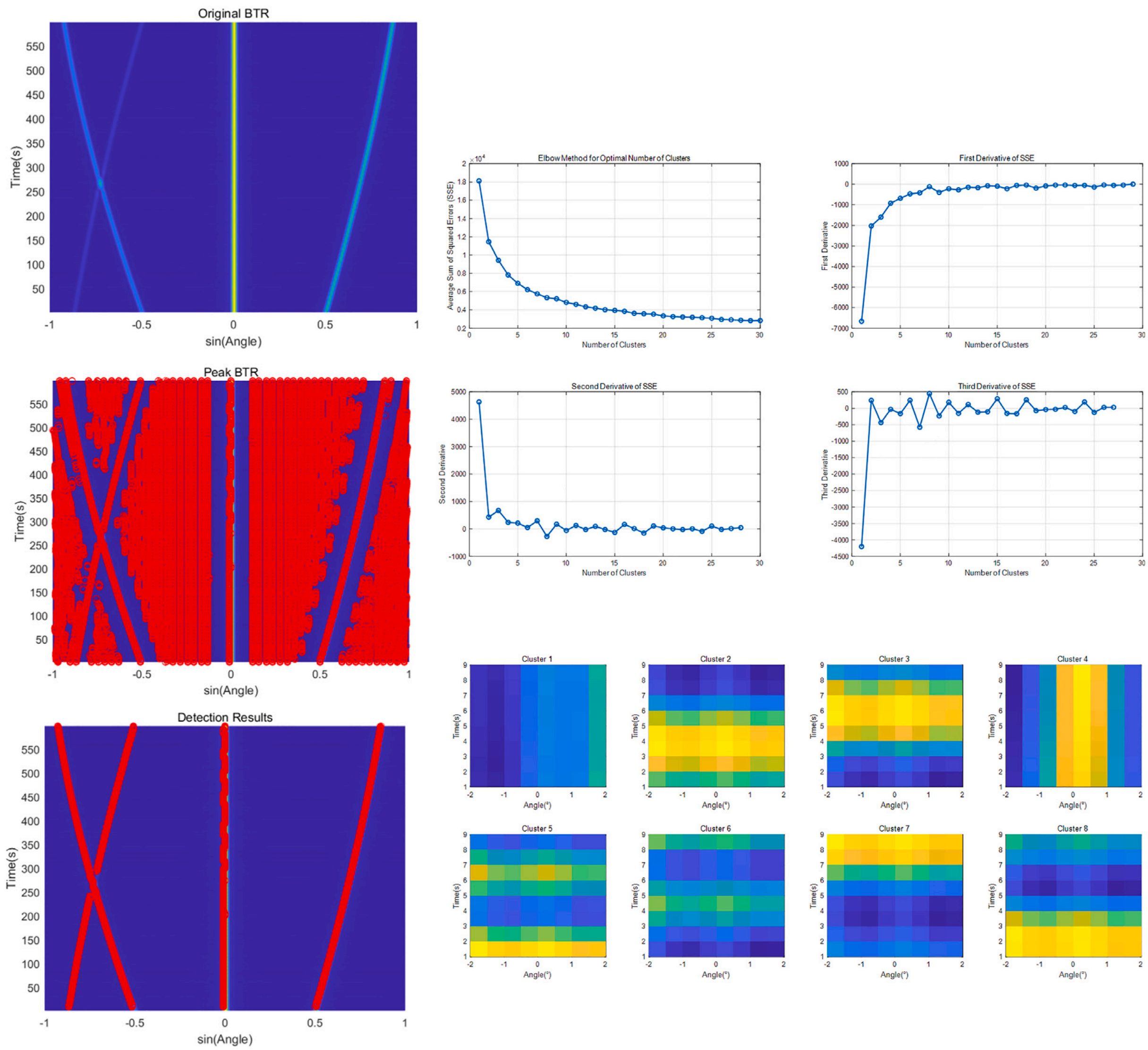

Fig. 5. BTRs at an SNR of 0 dB.

To prove shift invariance, assume the time variable t is shifted by $\Delta\tau$. The corresponding output becomes:

$$F_{CBF}(\vartheta, \tau + \Delta\tau) = \left|D_{CBF}(\vartheta, \tau + \Delta\tau)\right|^2$$
$$= \left|A_0 \frac{\sin\left(N\pi\frac{d}{\lambda}(\sin(\theta - \vartheta) - \sin\theta_0(t - (\tau + \Delta\tau)))\right)}{N\sin\left(\pi\frac{d}{\lambda}(\sin(\theta - \vartheta) - \sin\theta_0(t - (\tau + \Delta\tau)))\right)}\right|^2 \tag{12}$$

Through the properties of the sine function and the linearity of the shift operation, it can be shown that:

$$\sin\left(N\pi\frac{d}{\lambda}(\sin(\theta - \vartheta) - \sin\theta_0(t - (\tau + \Delta\tau)))\right)$$
$$= \sin\left(N\pi\frac{d}{\lambda}(\sin(\theta - \vartheta) - \sin\theta_0(t - \tau - \Delta\tau))\right) \tag{13}$$

Confirming the shift invariance:

$$F_{CBF}(\vartheta, \tau + \Delta\tau)$$
$$= \left|A_0 \frac{\sin\left(N\pi\frac{d}{\lambda}(\sin(\theta - \vartheta) - \sin\theta_0(t - \tau - \Delta\tau))\right)}{N\sin\left(\pi\frac{d}{\lambda}(\sin(\theta - \vartheta) - \sin\theta_0(t - \tau - \Delta\tau))\right)}\right|^2 \tag{14}$$

As shown in Fig. 4, when the horizontal axis is represented by angles, the main lobes of the beamforming patterns at 0 degrees and 40 degrees cannot align. However, when using the sine of the angle as the horizontal axis, the main lobes can align due to shift invariance. Therefore, in the sine coordinate system, the beam response exhibits ideal shift-invariant properties, meaning that the shape and position of the beam remain consistent relative to changes in the input angle.

Based on the shift invariance property, we propose converting the horizontal axis of the BTRs from linear angular values to their corresponding sine values, which is crucial to ensure the system maintains shift invariance, as utilizing $\sin(\theta)$ rather than θ aligns with the directional characteristics of the CBF function. By adopting this approach, the shift-invariant property of the beam response is preserved, thereby enhancing the accuracy and consistency of directional estimates.

4. Simulations

Ship radiated noise is comprised of three principal components: mechanical noise, propeller noise, and hydrodynamic noise. Hydrodynamic noise remains constant over time and exhibits a continuous

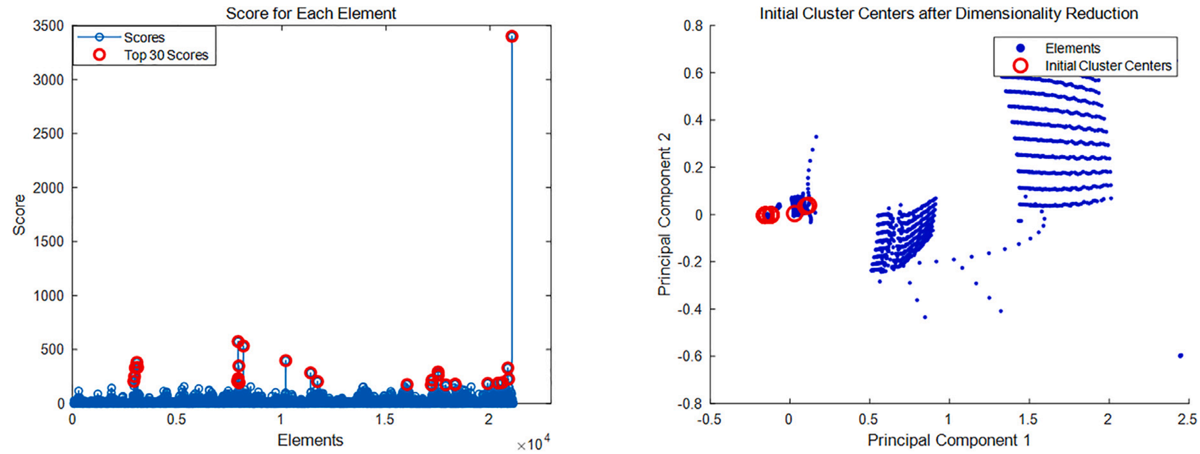

Fig. 6. Visualization of the GLP initialization method, SNR = 0 dB.

spectrum. Mechanical noise, generated by the operation of fans, air conditioners, and pumps, is typically characterized by line spectra. Owing to the periodic mechanical movements and gear operations, ship engine mechanical noise features both a broadband continuous spectrum and line spectra. Propeller noise, a major source of ship radiated noise, not only includes a low-frequency line spectrum but also predominantly displays a continuous spectrum, spanning frequencies from 5 Hz to 100 kHz. Cheng et al. (2018) proposed that the simulation signal for representing ship radiated noise is modeled using the Eq. (15):

$$x(t) = \left(1 + \sum_{k=1}^{K} m_k \cos(k\Omega t)\right) \cdot \sum_{n=N_L}^{N_H} A_n \cos(2\pi n f_0 t + \varphi_n) \tag{15}$$

where:

- t is the time variable,
- K represents the number of harmonics included in the modulation, with $K = 4$,
- m_k are the modulation indices for each harmonic, satisfying $0 < m_k < 1$, and are set equally for $k = 2, 3, 4$,
- Ω is the fundamental modulation frequency, set at 5 Hz,
- A_n denotes the amplitude of the nth carrier frequency,
- φ_n represents the initial phase of the nth carrier,
- f_0, the base frequency of the carriers, is defined as $f_0 = \frac{1}{T}$,
- N_L and N_H define the lower and upper bounds of the carrier frequency index, calculated as $N_L = \lceil 100/f_0 \rceil$ and $N_H = \lfloor 200/f_0 \rfloor$ respectively,
- The sampling frequency is set to 2 kHz,
- The total duration T of the signal is 599 s.

In this study, we systematically verified the performance of newly proposed methods under various Signal-to-Noise Ratio (SNR) conditions (ranging from 0 dB to −10 dB). The experimental design included a signal frequency range of 100–200 Hz and a sampling rate of 2000 Hz. The receiving equipment consisted of a uniform linear array with 256 units, with an inter-unit spacing of 1.5 m. The target signal's angular measurement range was from 0 to 180 degrees, with the accuracy of 0.5 degrees. In the simulation, four targets initiated movement from 30°, 60°, 120°, and 90°, moving to 60°, 22°, 150°, and 90° respectively, over a duration of 599 s. We employed CBF to process the received signals, initially performing open peak extraction on the broadband signals, selecting 200 Hz as the representative frequency for calculation. According to beamforming theory, the minimum peak distance at this time (corresponding to the main lobe width) should be 3 elements, approximately 1.5°; and the elemental width (corresponding to twice the main-side lobe distance) should be 9 elements, or 4.5°. Due to the property of the sine function that multiple input values correspond to one output value, we observed that converting the original 0 to 180° angular range using the sine transform could result in multiple x-values corresponding to the same y-value. To ensure the conversion's uniqueness, we adjusted the angular range to −90 to 90°. This adjustment ensured the unique correspondence property of the angular conversion, benefiting the accuracy and efficiency of subsequent data processing.

Experiments were conducted in an environment with a SNR of 0 dB, and the results are shown in the first subplot of Fig. 5. Here, the horizontal axis represents the sine of angles, covering a total of 180 angles corresponding to 360 pixel values, and the vertical axis indicates time, measured in snapshots or seconds. Utilizing the proposed GLP method to initialize the cluster centers, with the visualization of the top 30 highest scoring cluster centers shown in Fig. 6. These experiments indicate that the maximum value of the third derivative of SSE is 8. At this point, $\mathrm{SSE}'(k) < 0$, $\mathrm{SSE}''(k) < 0$, and $\mathrm{SSE}'''(k) > 0$. Therefore, the adaptively determined number of clusters is set to 8, the fourth class visual pattern accurately represents the beam's main lobe pattern. After selecting this fourth class, the results indicate a false alarm rate of zero.

The overall clustering algorithm process described in this paper involves initializing cluster centers using various methods and then performing clustering to automatically determine the number of clusters. The resulting patterns are representative of the main lobe, left and right side lobe, and random noise. In Section 3 of this paper, we introduce a novel GLP initialization algorithm, which selects initial clustering centers based on both global similarity and local peak similarity of data points. This method aims to identify the "initial centers" of data distributions, thereby enhancing the stability and accuracy of the clustering results. As shown in Fig. 7, our comparative analysis on simulation BTR are as follows:

(1) Random initialization: Cluster 6 represents the main lobe pattern. However, after selecting cluster 6, false alarms still exist in the target detection results.

(2) Max-Min Distance initialization method: Clusters 6 and 7 represent the main lobe patterns. However, dividing the main lobe pattern into two classes leads to false negative.

(3) KKZ initialization method: Cluster 6 represents the main lobe pattern. After selecting cluster 6, false alarms still exist in the target detection results.

(4) K-means++ initialization method: Cluster 1 represents the main lobe pattern. However, after selecting cluster 1, false alarms still exist in the target detection results.

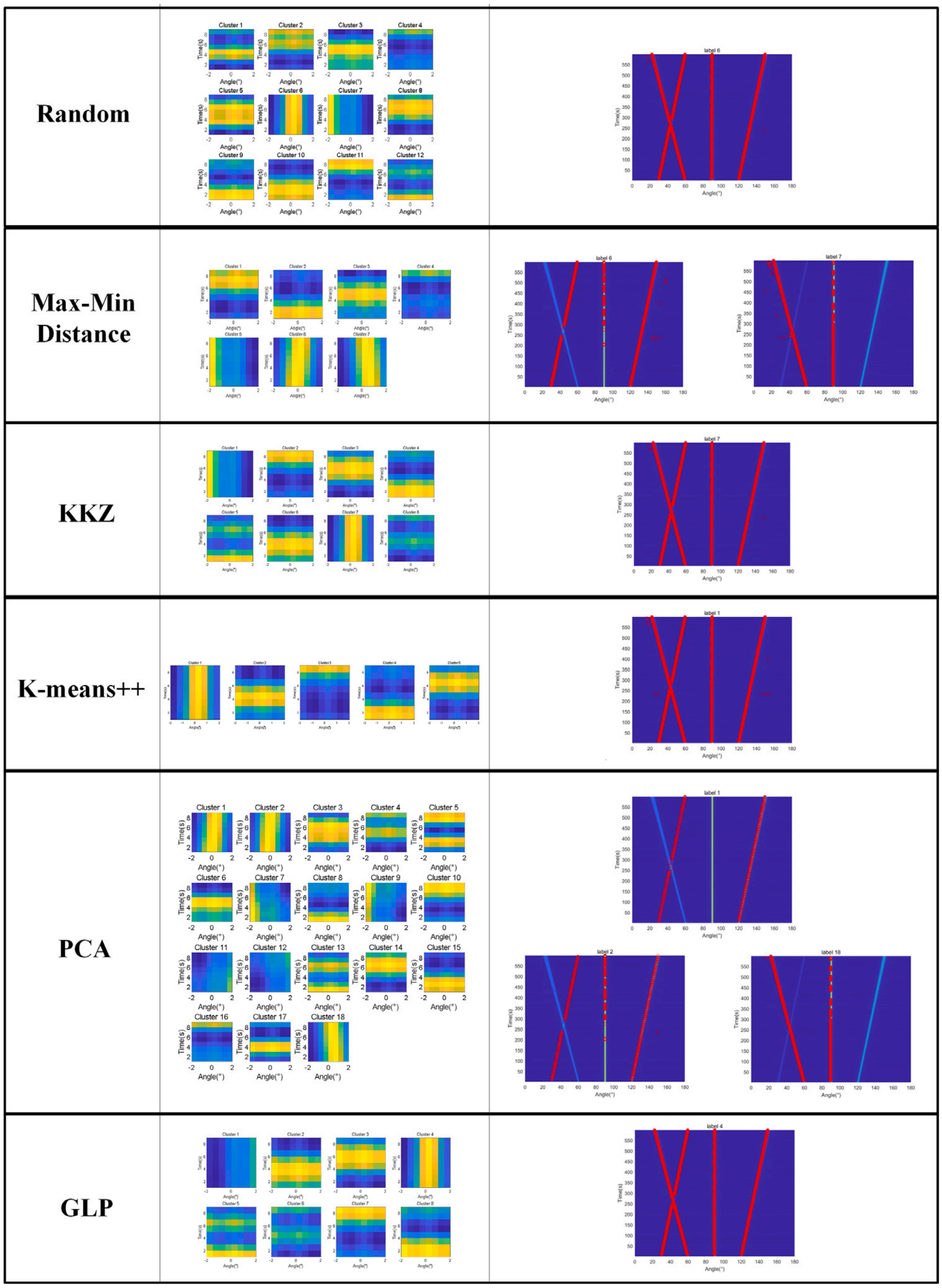

Fig. 7. Comparison of six initialization methods.

(5) PCA initialization method: Clusters 1, 2, and 18 represent the main lobe patterns. Dividing the main lobe pattern into three classes leads to false negative.

(6) GLP initialization method: Cluster 4 represents the main blade pattern. After selecting cluster 4, neither false alarms nor false negative occur.

Using the same BTRs and keeping all parameters consistent except for the initialization methods, we compared the results of different initialization techniques. Our proposed GLP initialization method does not produce false alarms and also avoids dividing the main lobe pattern into multiple classes, which leads to false negative.

Experiments were also conducted at an SNR of −10 dB, using the proposed GLP method to initialize the cluster centers, with the visualization of the top 30 highest scoring cluster centers shown in Fig. 8. According to the adaptive threshold method for determining the number of clusters, the maximum value of the third derivative of the SSE is 25, where $\mathrm{SSE}'(k) < 0$, $\mathrm{SSE}''(k) < 0$, and $\mathrm{SSE}'''(k) > 0$, as shown in Fig. 9, where the first class of clustering represented the beam main lobe pattern. However, false alarms still remain, as shown in the third subfigure of Fig. 10, where red circles represent 2288 correct detections and green circles represent 544 false alarms, resulting in a false alarm rate of 19.2% and detection rate of 80.8%. This indicates that under low SNR conditions, some random noise might be incorrectly classified as

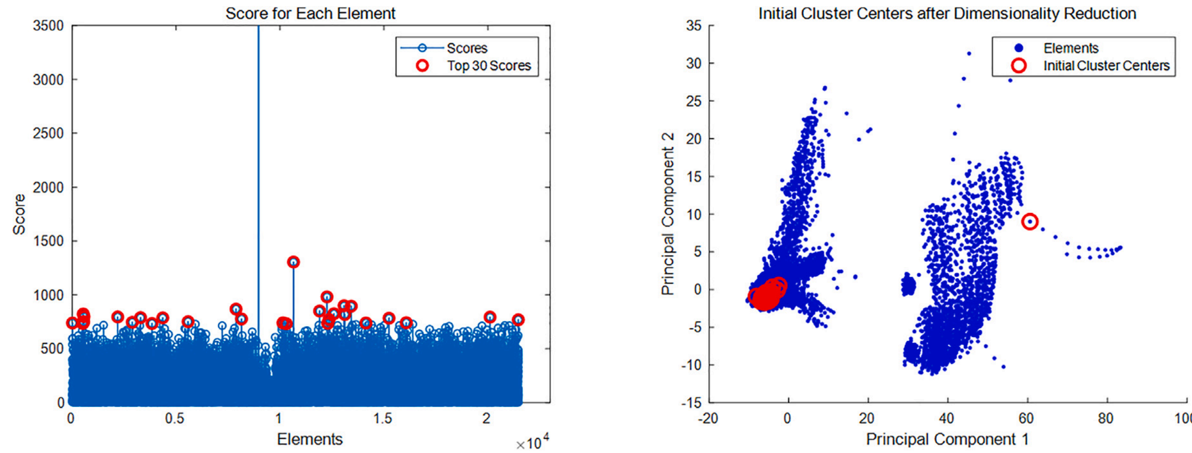

Fig. 8. Visualization of the GLP initialization method, SNR = −10 db.

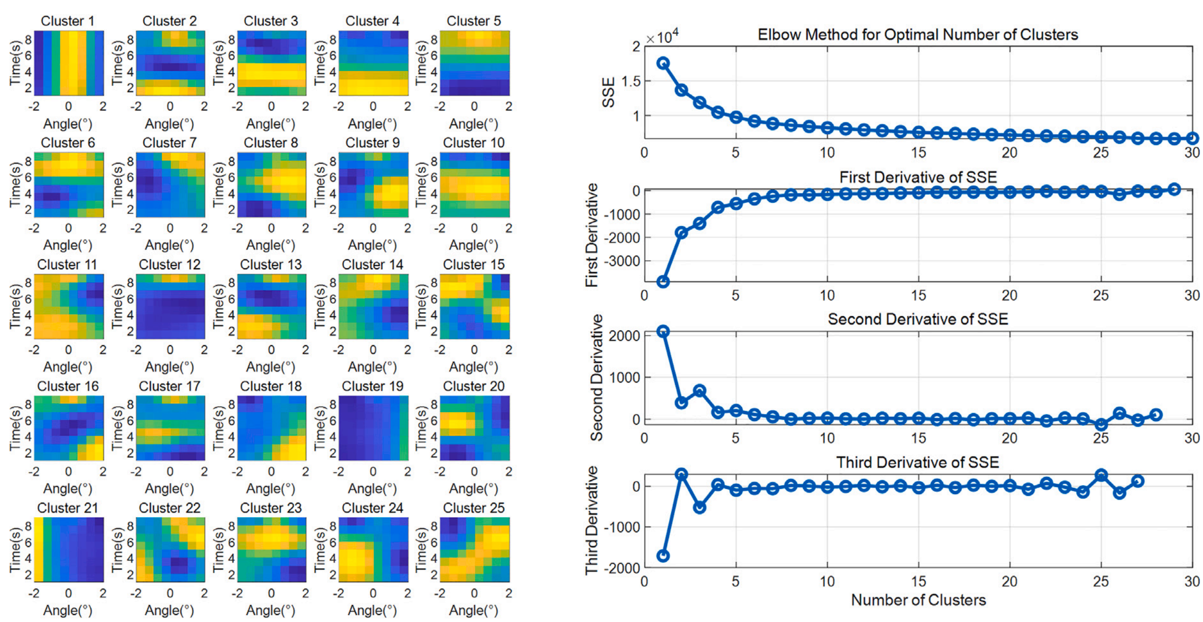

Fig. 9. The number of k at an SNR of −10 dB.

part of the main lobe pattern. To more accurately distinguish these random noises from the main lobe pattern, we employed a high-correlation false alarm suppression method based on fuzzy thresholds. This method involved calculating the correlation coefficients of all elements on the main lobe pattern with the main lobe pattern itself and fitting a probability distribution histogram to these correlation coefficients. As shown in Fig. 11, the correlation coefficients ranged from 0.14964 to 0.99231, with the correlation coefficients of actual targets typically above 0.8, while those of false alarms were below 0.7. Based on high-correlation based on fuzzy thresholds, where $\tau = 0.8250$, elements with correlation coefficients above this fuzzy threshold were retained, while those below the threshold were discarded. After applying the High-Correlation False Alarm Suppression method, as shown in the fourth subfigure of 10, with yellow circles indicating 2288 correct detections, the accuracy was improved to 100%, and the false alarm rate was reduced to 0%. To calculate the false negative rate, we used 599 snapshots of data. Since the last element is not detected, the theoretical total number of peaks should be $590 \times 4 - 1 = 2359$. Of these, 2288 were true positives, and 71 were false negatives, resulting in a recall rate of 97%.

5. Sea-trial experiments

5.1. Dataset description

Experiments of this section evaluate the method by using sea-trial data collected in the South China Sea during the summer of 2021. The sonar array utilized in the experiment is a towed horizontal line array with 256 elements, all data were processed using the CBF method

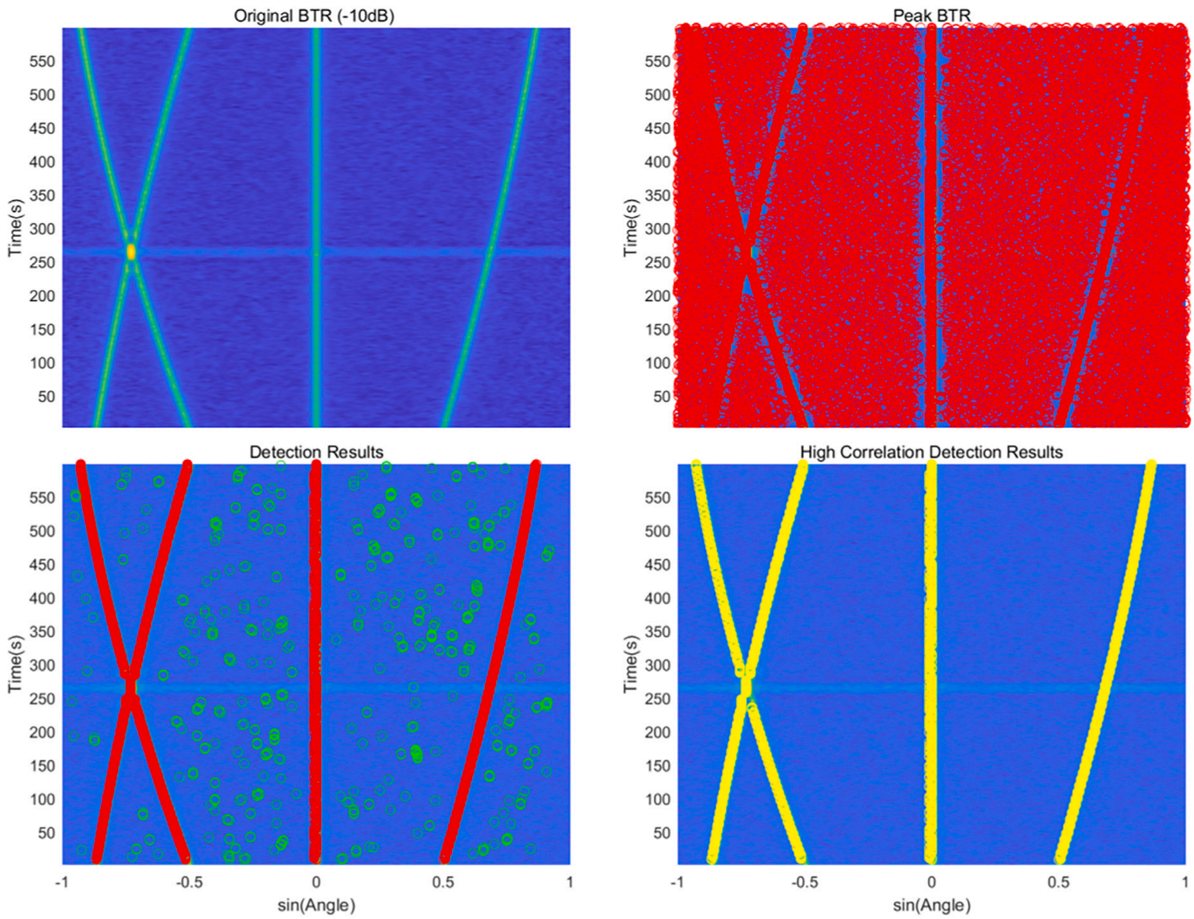

Fig. 10. BTRs at an SNR of −10 dB (Red circles represent true targets, green circles represent false alarms, and yellow circles represent true targets after false alarm suppression.).

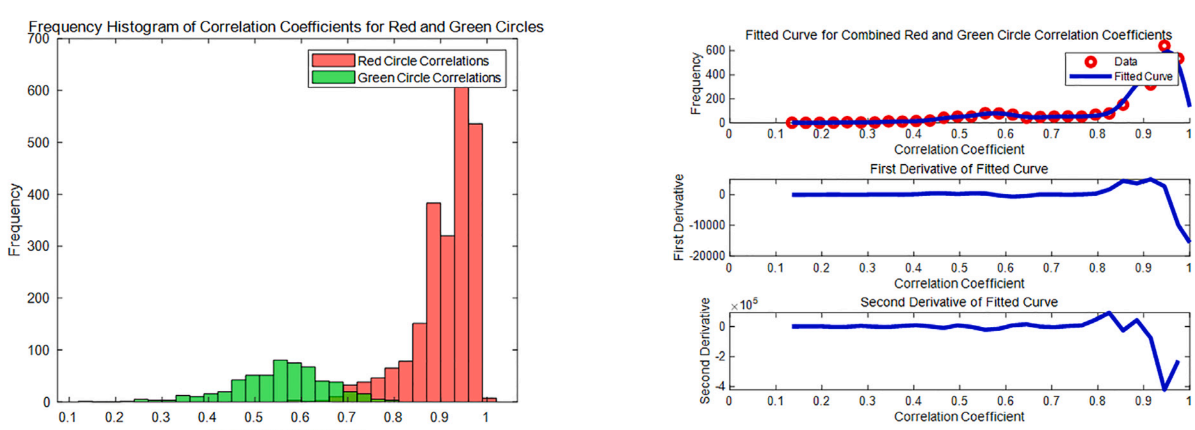

Fig. 11. Frequency histogram of correlation coefficients and fitting curve at an SNR of −10 dB.

for array signal processing. Experiments were conducted separately in shallow and deep sea areas, a detailed description of the original dataset are presented in Table 2. For the "Shallow Marine" dataset, depths range from a minimum of 71.15 m to a maximum of 91.74 m, while the "Deep Marine" dataset, depths vary from 3750.48 m to 3827.72 m.

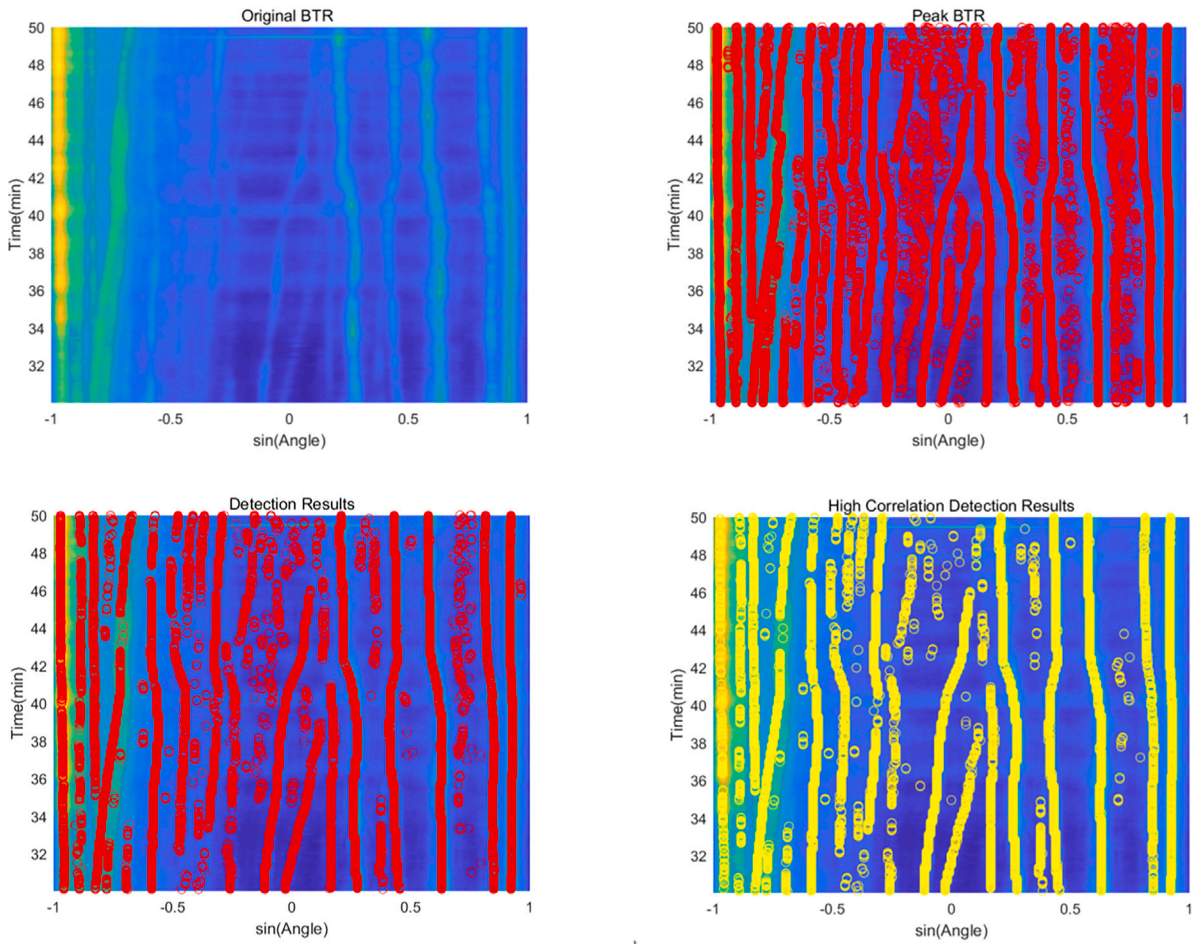

Fig. 12. BTRs in shallow marine.

Table 2
Parameters of datasets.

	Shallow marine	Deep marine
Duration	9728	9984
Snapshot interval	1 s	1 s
Snapshot period	0 s	0 s
Number of array units	256	256
Units spacing	1.5 m	1.5 m
Frequency	100–250 Hz	100–250 Hz
Beamforming algorithm	CBF	CBF
Water depth	71.15–91.74 m	3750.48–3827.72 m

5.2. Performance evaluations

For the study of shallow marine environments, the original BTR is visualized in the first subplot of Fig. 12, where the horizontal axis represents the sine of angles, covering a total of 180 angles corresponding to 360 pixel values, and the vertical axis indicates time, measured in seconds. According to the adaptive threshold method for determining the number of clusters, the maximum value of the third derivative of SSE is 3. At this point, $\mathrm{SSE}'(k) < 0$, $\mathrm{SSE}''(k) < 0$, and $\mathrm{SSE}'''(k) > 0$, as shown in Fig. 13. The first class of clustering represented the beam's main lobe pattern. Despite only displaying the visual pattern of the first cluster, a significant number of false alarms continued to be observed. To address this, we employ a high-correlation false alarm suppression method based on fuzzy thresholds. This method involves calculating the correlation coefficients between all elements on the main lobe pattern and itself, and fitting these coefficients to a probability distribution histogram, as shown in Fig. 14. The correlation coefficients ranged from 0.01363 to 0.99872. Typically, coefficients for true targets were consistently above 0.8, suggesting strong correlation, while those identified for false alarms generally fell below 0.7, indicating weaker or erroneous associations. Employing a high-correlation method based on fuzzy thresholds, the maximum value of the second derivative was calculated, yielding $\tau = 0.8550$. Elements corresponding to correlation coefficients above this threshold were retained, while those below were discarded, leading to a significant reduction in the rate of false alarms. The final target detection results demonstrated a substantial decrease in the rate of false alarms, as shown in the fourth subplot of Fig. 12.

For the study of deep marine environments, these samples are visualized in the first subplot of Fig. 15. According to the adaptive threshold method for determining the number of clusters, the maximum value of the third derivative of the Sum of SSE is 4. At this point, $\mathrm{SSE}'(k) < 0$, $\mathrm{SSE}''(k) < 0$, and $\mathrm{SSE}'''(k) > 0$, as shown in Fig. 16. The last cluster represents the main lobe beam pattern. Despite only displaying the visual pattern of the last cluster, a significant number of false alarms continued to be observed. To address this issue, a high-correlation false alarm suppression method based on fuzzy thresholds was employed. This method involves calculating the correlation coefficients between all elements on the main lobe pattern and itself and fitting these

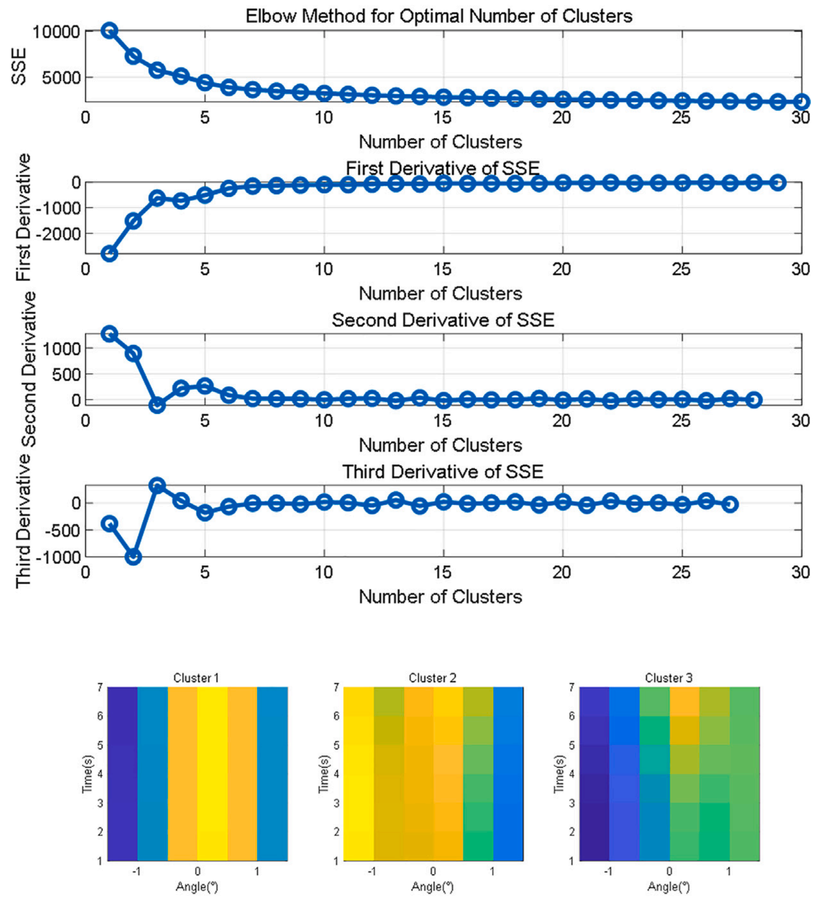

Fig. 13. The number of k in shallow marine.

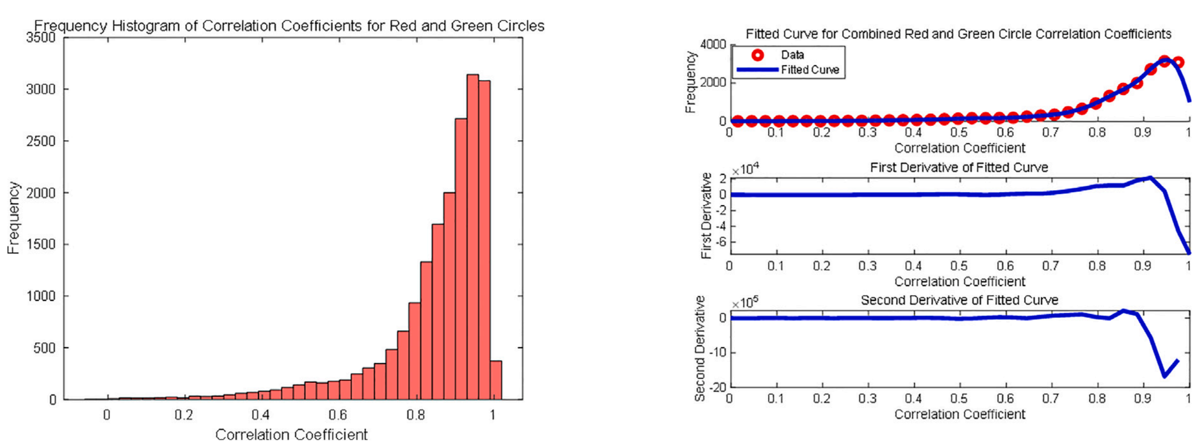

Fig. 14. Frequency histogram of correlation coefficients and its fitting curve in shallow marine.

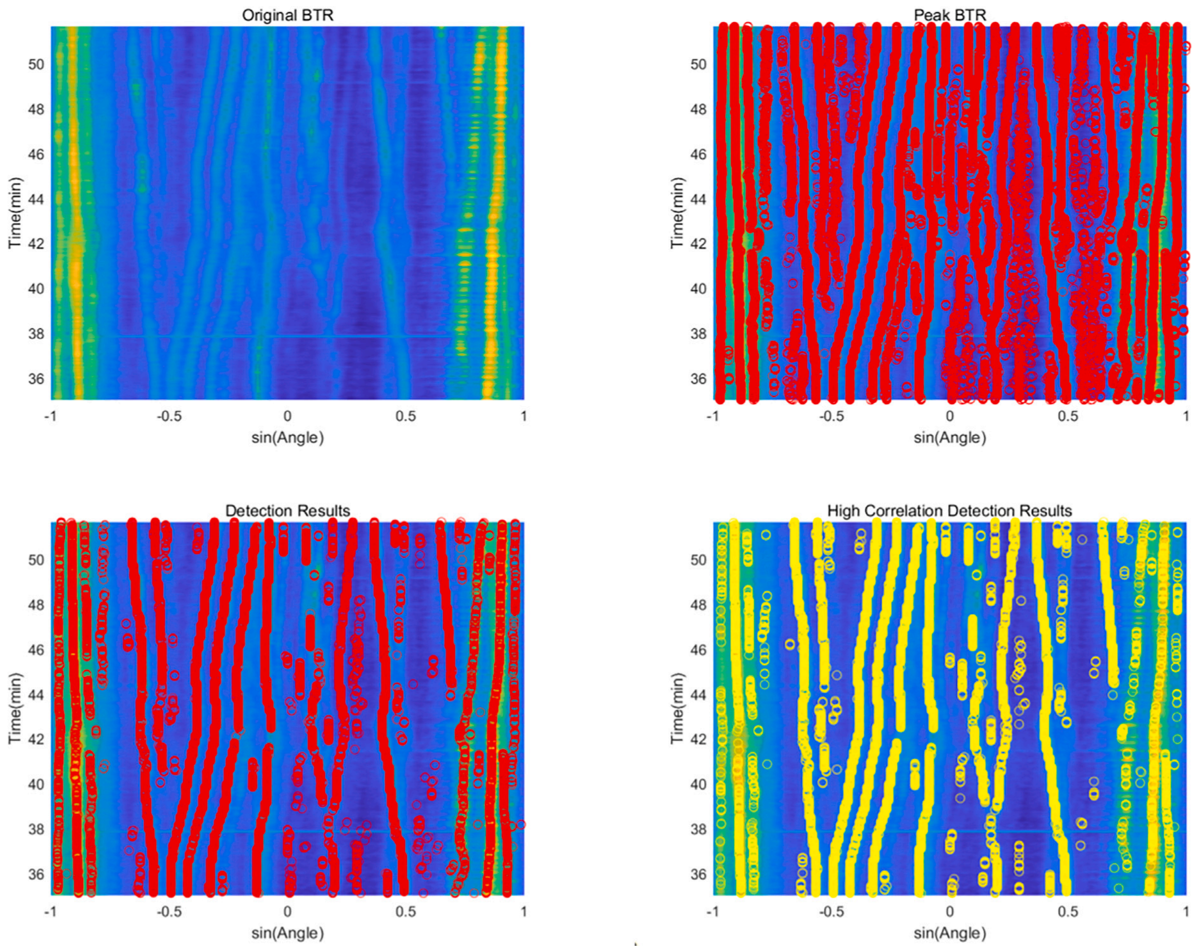

Fig. 15. BTRs in deep marine.

coefficients to a probability distribution histogram, as shown in Fig. 17. The correlation coefficients ranged from 0.05729 to 0.99798. Typically, coefficients for true targets were above 0.75, whereas those for false alarms were below 0.65. Employing a high-correlation method based on fuzzy thresholds, the maximum value of the second derivative was calculated, yielding $\tau = 0.7950$. Elements corresponding to correlation coefficients above this threshold were retained, while those below were discarded, leading to a significant reduction in the rate of false alarms. The final target detection results demonstrated a substantial decrease in the rate of false alarms, as shown in the fourth subplot of Fig. 17.

6. Conclusion

This paper improves upon previous research on automated target detection methods for BTR images. We introduce a clustering center initialization method suitable for this task, the GLP initialization method. This method adaptively determines the number of clusters through quantitative analysis of the first, second, and third derivatives of the SSE. To address the issue of minor false alarms being misclassified as the main lobe pattern under low SNR conditions, we propose a determination rule based on the derivatives of the fitting curve for the frequency histogram of correlation coefficients using fuzzy thresholds. This approach achieved a 100% accuracy rate on simulated datasets and also demonstrated excellent results with sea trial datasets. The research not only optimizes the initialization process of clustering centers, ensuring high accuracy and robustness in complex marine environments but also effectively reduces the rate of false alarms in low SNR conditions by introducing a decision-making mechanism based on fuzzy thresholds. These innovations enhance the overall performance of the system, making it more potent for practical marine detection tasks. The proposed method is primarily suitable for targets such as ships that continuously radiate noise. However, it is not applicable to targets like acoustic bombs that emit noise transiently. In weak target detection under low SNR conditions, this method can effectively suppress sidelobes and random noise, enabling earlier detection of targets and longer-term tracking of them. Regarding future work, we plan to further identify different targets based on our existing foundation. In actual sea trial data, one major challenge we anticipate is the discontinuity of targets displayed in BTR images. Typically, due to environmental noise, target movement, or sensor limitations, acoustic signals from underwater objects may not be continuously recorded. This discontinuity in signals can obscure the true shape and trajectory of targets, thus increasing the complexity of detection and tracking processes. To address this challenge, we plan to develop more advanced algorithms to more effectively handle data discontinuities and more accurately infer the positions and movements of discontinuous signals. This will involve enhancements to our existing models, incorporating advanced temporal and spatial analysis techniques to predict positions in scenarios where observational data are incomplete or intermittent. This strategy not only targets the existing challenges in data processing

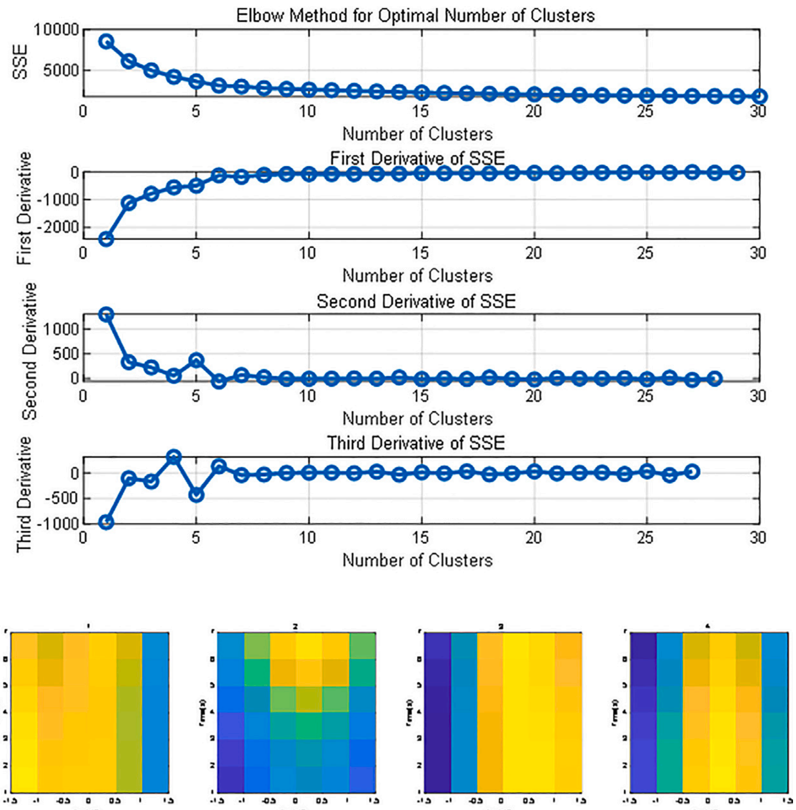

Fig. 16. The number of k in deep marine.

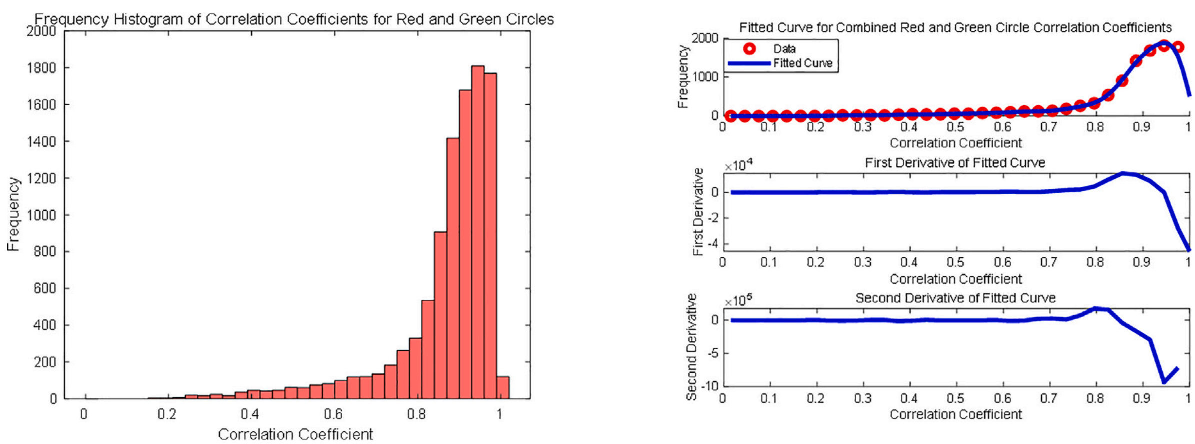

Fig. 17. Frequency histogram of correlation coefficients and its fitting curve in deep marine.

but also provides a methodological foundation for handling target detection in more complex marine environments.

CRediT authorship contribution statement

Hao Yin: Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Chao Li:** Writing – review & editing, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Haibin Wang:** Writing – review & editing, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization. **Jun Wang:** Writing – review & editing, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization. **Zili Qin:** Investigation, Conceptualization. **Yannick Benezeth:** Writing – review & editing, Investigation, Conceptualization. **Fan Yang:** Writing – review & editing, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This article is funded by the General Program of National Natural Science Foundation of China (ID: 62171440), the Chinese Academy of Sciences Project (ID: XDB0700403), and supported by the China Scholarship Council, China. We would like to express my gratitude to Cormac Mooney, Associate Professor at Capital University of Economics and Business(China), for reviewing this article and providing valuable feedback on grammar, style and syntax.

Data availability

The data that has been used is confidential.

References

Aidala, V.J., 1979. Kalman filter behavior in bearings-only tracking applications. IEEE Trans. Aerosp. Electron. Syst. (1), 29–39.

Arthur, D., Vassilvitskii, S., 2006. k-Means++: The Advantages of Careful Seeding. Technical Report, Stanford.

Carbone, C.P., Kay, S.M., 2012. A novel normalization algorithm based on the three-dimensional minimum variance spectral estimator. IEEE Trans. Aerosp. Electron. Syst. 48 (1), 430–448.

Cheng, Y.-s., Li, Z.-z., Qiu, J.-x., 2018. Underwater acoustic target identification. Underwater Acoustic Target Identification.

Ding, C., He, X., 2004. K-means clustering via principal component analysis. In: Proceedings of the Twenty-First International Conference on Machine Learning. p. 29.

Gonzalez, T.F., 1985. Clustering to minimize the maximum intercluster distance. Theoret. Comput. Sci. 38, 293–306.

Ho, K., Chan, Y.T., 2006. An asymptotically unbiased estimator for bearings-only and Doppler-bearing target motion analysis. IEEE Trans. Signal Process. 54 (3), 809–822.

Ince, L., Sezen, B., Saridogan, E., Ince, H., 2009. An evolutionary computing approach for the target motion analysis (TMA) problem for underwater tracks. Expert Syst. Appl. 36 (2), 3866–3879.

Kaba, U., Temeltas, H., 2022. Generalized bias compensated pseudolinear Kalman filter for colored noisy bearings-only measurements. Signal Process. 190, 108331.

Kaufman, L., Rousseeuw, P.J., 2009. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley & Sons.

Kim, J., Suh, T., Ryu, J., 2017. Bearings-only target motion analysis of a highly manoeuvring target. IET Radar Sonar Navig. 11 (6), 1011–1019.

Lei, Z., Yang, K., Zhang, Q., Xia, H., 2016. Two dimensional TV-L1 regularization for underwater acoustic source tracking. In: OCEANS 2016-Shanghai. IEEE, pp. 1–4.

Li, Z., Li, Y., Huang, H., 2012. Study on automatic continuous trackingand location algorithm for underwater target. Chin. J. Sci. Instrum. 33 (3), 520–528.

Li, X., Zhao, C., Yu, J., Wei, W., 2019. Underwater bearing-only and bearing-Doppler target tracking based on square root unscented Kalman filter. Entropy 21 (8), 740.

Lin, X., Kirubarajan, T., Bar-Shalom, Y., Maskell, S., 2002. Comparison of EKF, pseudomeasurement, and particle filters for a bearing-only target tracking problem. In: Signal and Data Processing of Small Targets 2002. 4728, SPIE, pp. 240–250.

MacQueen, J., et al., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. 1, (14), Oakland, CA, USA, pp. 281–297.

Nardone, S., Lindgren, A., Gong, K., 1984. Fundamental properties and performance of conventional bearings-only target motion analysis. IEEE Trans. Autom. Control 29 (9), 775–787.

Nguyen, N.H., Doğançay, K., 2017. Improved pseudolinear Kalman filter algorithms for bearings-only target tracking. IEEE Trans. Signal Process. 65 (23), 6119–6134.

Northardt, T., Nardone, S.C., 2018. Track-before-detect bearings-only localization performance in complex passive sonar scenarios: A case study. IEEE J. Ocean. Eng. 44 (2), 482–491.

Pham, D.-T., 1993. Some quick and efficient methods for bearing-only target motion analysis. IEEE Trans. Signal Process. 41 (9), 2737–2751.

Shapiro, J.H., Green, T., 2000. Performance of split-window multipass-mean noise spectral estimators. IEEE Trans. Aerosp. Electron. Syst. 36 (4), 1360–1370.

Shin, W., Kim, D.-S., Ko, H., 2023. Target tracking from weak acoustic signals in an underwater environment using a deep segmentation network. J. Mar. Sci. Eng. 11 (8), 1584.

Singh, P., Bose, S.S., 2021. A quantum-clustering optimization method for COVID-19 CT scan image segmentation. Expert Syst. Appl. 185, 115637. http://dx.doi.org/10.1016/j.eswa.2021.115637, URL https://www.sciencedirect.com/science/article/pii/S0957417421010319.

Singh, P., Huang, Y.-P., 2024a. AKDC: Ambiguous kernel distance clustering algorithm for COVID-19 CT scans analysis. IEEE Trans. Syst. Man Cybern.: Syst. 54 (10), 6218–6229. http://dx.doi.org/10.1109/TSMC.2024.3418411.

Singh, P., Huang, Y.-P., 2024b. An ambiguous edge detection method for computed tomography scans of coronavirus disease 2019 cases. IEEE Trans. Syst. Man Cybern.: Syst. 54 (1), 352–364. http://dx.doi.org/10.1109/TSMC.2023.3307393.

Struzinski, W.A., Lowe, E.D., 1984. Performance comparison of four noise background normalization schemes proposed for signal detection systems. J. Acoust. Soc. Am. 76 (6), 1738–1742.

Xin, J., Le, B., Bo, L., Luo, L., 2015. Track before detect of weak trajectory using Hidden Markov Model. In: 2015 4th International Conference on Computer Science and Network Technology. ICCSNT, Vol. 1, IEEE, pp. 1473–1477.

Xin, J.-r., Zhao, Y., Luo, L., 2017. Detection of bearing-only trajectory of multi-target using Hidden Markov model. In: 2017 9th International Conference on Modelling, Identification and Control. ICMIC, IEEE, pp. 76–80.

Yang, T., 2017. Deconvolved conventional beamforming for a horizontal line array. IEEE J. Ocean. Eng. 43 (1), 160–172.

Yin, F., Li, C., Wang, H., Nie, L., Zhang, Y., Liu, C., Yang, F., 2023. Weak underwater acoustic target detection and enhancement with BM-SEED algorithm. J. Mar. Sci. Eng. 11 (2), 357.

Yin, F., Li, C., Wang, H., Yang, F., 2019. Automatic acoustic target detecting and tracking on the azimuth recording diagram with image processing methods. Sensors 19 (24), 5391.

Yin, H., Li, C., Wang, H., Yin, F., Yang, F., 2024. False alarm suppressing for passive underwater acoustic target detecting with computer visual techniques. Ocean Eng. 305, 117969.

Yu, G., Tang, B., Piao, S., 2023. A novel bearing-time record estimation method based on α-stabledistribution modeling. J. Acoust. Soc. Am. 154 (4_supplement), A211.

Zhang, Y., Wang, C., Zhang, Q., Da, L., Jiang, Z., 2024. Bearing-only motion analysis of target based on low-quality bearing-time recordings map. IET Radar Sonar Navig. 18 (5), 765–781.

Zhang, Y.J., Xu, G.Z., 2010. Bearings-only target motion analysis via instrumental variable estimation. IEEE Trans. Signal Process. 58 (11), 5523–5533.

Zhang, J., Xu, X., Chen, Z., Bao, M., Zhang, X.-P., Yang, J., 2020. High-resolution DOA estimation algorithm for a single acoustic vector sensor at low SNR. IEEE Trans. Signal Process. 68, 6142–6158.

Zhu, J., Peng, C., Zhang, B., Jia, W., Xu, G., Wu, Y., Hu, Z., Zhu, M., 2021. An improved background normalization algorithm for noise resilience in low frequency. J. Mar. Sci. Eng. 9 (8), 803.