Archives of Acoustics Vol. 48, No. 2, pp. 235–247 (2023), doi: 10.24425/aoa.2023.145235

Research Paper

Research on Scattering Feature Extraction of Underwater Moving Cluster Targets Based on the Highlight Model

Yang YANG, Jun FAN*, Bin WANG

Key Laboratory of Marine Intelligent Equipment and System of the Ministry of Education
Shanghai Jiao Tong University
Shanghai, China

*Corresponding Author e-mail: fanjun@sjtu.edu.cn

(*received October 5, 2022; accepted November 28, 2022*)

In detecting cluster targets in ports or near-shore waters, the echo amplitude is seriously disturbed by interface reverberation, which leads to the distortion of the traditional target intensity characteristics, and the appearance of multiple targets in the same or adjacent beam leads to fuzzy feature recognition. Studying and extracting spatial distribution scale and motion features that reflect the information on cluster targets physics can improve the representation accuracy of cluster target characteristics. Based on the highlight model of target acoustic scattering, the target azimuth tendency is accurately estimated by the splitting beam method to fit the spatial geometric scale formed by multiple highlights. The instantaneous frequencies of highlights are extracted from the time-frequency domain, the Doppler shift of the highlights is calculated, and the motion state of the highlights is estimated. Based on the above processing method, target highlights' orientation, spatial scale and motion characteristics are fused, and the multiple moving highlights of typical formation distribution in the same beam are accurately identified. The features are applied to processing acoustic scattering data of multiple moving unmanned underwater vehicles (UUVs) on a lake. The results show that multiple small moving underwater targets can be effectively recognized according to the highlight scattering characteristics.

Keywords: motion small cluster targets; feature fusion; azimuth trend scale; azimuth trend Doppler.

1. Introduction

The early warning and defence of ports, islands, and large floating platforms is mainly conducted by small targets such as frogman teams and underwater robot clusters (Jiang *et al.*, 2009; Huang *et al.*, 2020). In recent years, many countries have invested great efforts in studying the detection technology for small underwater targets. Diver detection sonar (DDS) is an important equipment for small underwater target detection (Sarangapani *et al.*, 2005; Lo, Ferguson, 2004). The Sentinel intruder detection sonar system by the American Sonardyne Company has a working frequency of 70 kHz, a bandwidth of 20 kHz, and an operating range of 600–900 m. The AquaShield detection sonar by the Israeli DSIT Company has a receiving array length of 1250 mm, a working frequency of 60 kHz, and an operating range of 700–1000 m. The X-type underwater monitoring system of the British Oceanscan Company can automatically track multiple targets at the same time; the operating frequency is 100 kHz, the operating range is 500–1000 m, and the direction-finding accuracy is 1°. The centre frequency of most diver detection sonar is mainly concentrated in the range of 60~100 kHz, the bandwidth is 20 kHz, and the detection range is hundreds of metres to kilometres. Diver detection sonar is a high-resolution image sonar that takes small targets as detection objects. The higher the working frequency is, the more detailed the description of the targets; however, the higher the frequency is, the closer the operating distance and its design are restricted by the resolution and operating distance. In the measurement test of small targets at sea, at 100 kHz, the target intensity of diver is −25~−20 dB, and the average target intensity of the frogman's exhalations (bubbles) is −14 dB (Hollett *et al.*, 2006). In addition, the average travelling speed of the frogman under load is approximately 1.8 km/h,

and the driving speed of the frogman carrier under load can reach 4.5–5.4 km/h. "Tuna 9" is a hand-held unmanned vehicle that is used to perform coastal and port monitoring and anti-mine warfare tasks in shallow water environments, and the maximum speed of the "Tuna 21" – a submersible craft of the same series – is 4.5 knots (Zhang, Tong, 2008). It can be seen that the target intensity of small underwater targets is low, the echo is weak, and the shallow sea interface reverberation makes the signal mixing ratio of the received signal low. A small target is mobile and flexible, but its moving speed is slow, the echo Doppler shift is low, and the classical frequency domain filtering makes it difficult to complete the high-resolution frequency shift estimation. The size and shape of different kinds of small targets lead to different spatial distributions of the target intensity. For an underwater cluster composed of multiple small targets, the multiple acoustic scattering interaction between the targets increases the complexity of the physical characteristics of the cluster targets. Therefore, it is important to study objective acoustic features that are easy to extract and apply to describe underwater moving cluster targets. The characteristics of the target acoustic scattering echo are related to the natural vibration mode of the target and the properties of the incident sound wave. The approximate calculation method of the target acoustic scattering field is commonly used in engineering, and the generation point of the scattering echo is regarded as a highlight (Tang, 1994; Chen *et al.*, 2013). The acoustic scattering highlights of small cluster targets with a certain distribution law will cover a large scale in space and will be equivalent to large-scale volume targets. However, due to the different spatial distribution, motion speed, and motion state of individual small moving targets constituting the cluster, the target echo is reflected in the modulation of Doppler; that is, the Doppler shift of the highlight of individual small targets relative to the receiving array is different. Although the clear echo structure of the target cannot be obtained using active sonar to detect small cluster targets, when the interface interference is serious or multiple targets are in the same or adjacent beam, the scattered waves of the strong highlights of the target still contain rich target feature information. Aiming to identify large-scale targets such as underwater crafts and small cluster targets in port monitoring, this paper develops an extraction method of the spatial scale and motion parameter features of the highlights of the target. Based on the highlight model theory, the acoustic scattering echo model of the relative motion between multiple highlights and the receiving hydrophone array is established. The azimuth tendency of the target is estimated by the high-resolution split beam method, and the spatial geometric scale covered by multiple highlights is calculated. The target multi-highlight structure is obtained in the Wigner-Weill distribution (WVD) time-frequency domain (Latif *et al.*, 2003; Rodríguez *et al.*, 2004; Sabra, Anderson, 2014), the reverberation and cross-term interference in the time-frequency domain are removed by morphological filtering (Ronse *et al.*, 2005; Bouaynaya *et al.*, 2008), and the instantaneous frequency of the underwater target highlight is extracted. According to the derived instantaneous frequency spectral line slope and frequency shift relationship, the broadband spectral Doppler shift of the highlight is calculated, and the motion state of the highlight is analysed. Combined with the target orientation, scale and motion information, a multiparameter feature space is established to describe the essential attributes of the target. The effectiveness of the features for identifying small cluster targets is verified by simulations and lake experiments.

2. Multitargets resolution method and theoretical derivation

For active sonar target detection, the spatial spectrum estimation method (Li *et al.*, 2014) is often used in engineering to suppress background interference and determine target orientation. When different targets are located in different scanning beams, they can be easily distinguished. However, when the azimuth of multiple targets is less than the beam resolution, multi targets appear in the same beam, and only a bright spot of multi-target energy superposition is displayed in the angle-distance spectrum, it is difficult to distinguish different target highlights. On the basis of spatial spectrum estimation results, this paper studies and extracts the acoustic scattering characteristics of the beam domain echo, and identifies the target through the characteristics of the strong scattering highlights superimposed in the echo. For an underwater vehicle with a certain scale, the reflected echo generated by it is an extended body with a certain distribution in the range and orientation. This extended body is approximately equivalent to the sum of several "highlight" signals with a certain distribution in space (Yang, Li, 2016). Small cluster targets work together underwater in a planned formation, and the acoustic scattering highlights of multiple small targets cover large-volume targets with a certain scale in space. For large-scale targets such as warships and submarines, the motion parameters of each strong scattering highlight on the same target are the same. However, for the volume target echo composed of multiple small targets, due to the different spatial distribution and motion state of each target, the phase and Doppler shift of each highlight echo are also different. So the spatial orientation distribution, dispersion scale structure features and motion features of highlights are extracted and fused to identify multiple underwater moving targets in this paper. The method flow chart is shown in Fig. 1.

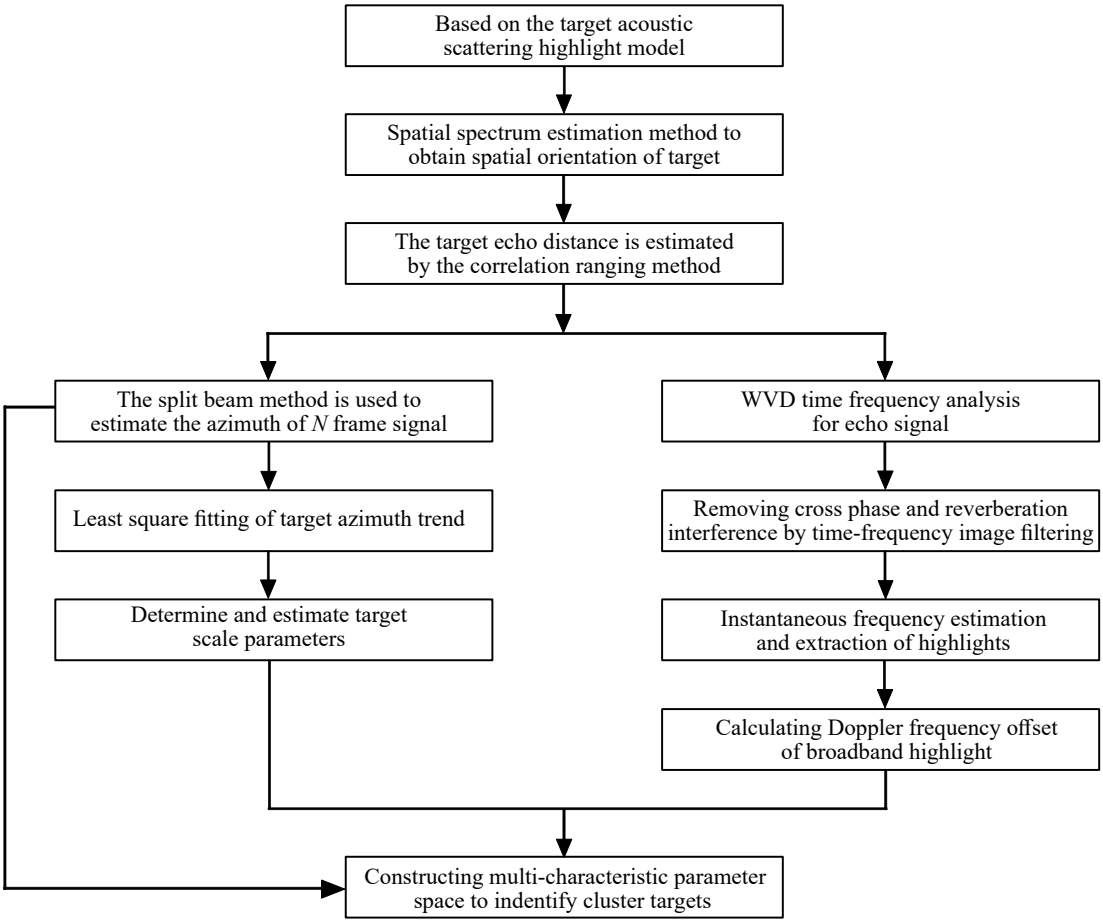

Fig. 1. Flow chart of multi-target resolution method.

First, a spatial spectrum estimation method is used to obtain the target azimuth. Next, the scale of the target is obtained by estimating the target azimuth trend, and then Doppler frequency shift of the highlights is estimated by extracting the instantaneous frequency of the target in the time-frequency domain. Finally, the multiparameter joint feature space is established to identify different targets in the same beam. The relevant theories and derivation of azimuth trend target scale estimation and Doppler parameter estimation are given in Subsecs. 2.1 and 2.2, respectively.

2.1. Azimuth trend target scale estimation

Target scale estimation based on the target azimuth trend is an effective scale recognition method (Ma *et al.*, 2004; Zhou, Yi, 2004; Jiang, Xiang, 2019) that has been applied in practical engineering. The basic idea of azimuth trend scale recognition is as follows: the target echo signal is divided into N subframes. According to the signals received by multiple array elements, the split beam method is used to obtain the azimuth estimation θ_i $(i = 1, 2, ..., N)$ of each subframe signal with N azimuth estimation values in total, and then the least squares linear fitting is used to obtain the azimuth trend value, that is, the linear estimation of the maximum angle range of the echo distribution over the whole target coverage space within the echo duration. Then, the target azimuth trend T_r is estimated.

It is assumed that the signals received by the two array elements with d space are $x(t)$ and $y(t) = x(t+\tau)$, and the relative delay is $\tau = d\sin\theta/c$. The two signals are Fourier transformed, and the cross-power spectrum of the sequence is calculated as follows:

$$P_{xy} = X(\omega)Y^H(\omega) = X(\omega)\left[X(\omega)e^{j\omega\tau}\right]^H = |X(\omega)|^2 e^{-j\omega\tau}. \quad (1)$$

It can be seen from the above equation that the phase angle of the cross-power spectrum of the output sequence of the two array elements is the phase difference generated by the incident signal on the two array elements, and the azimuth θ of the target can be calculated according to the phase angle of the cross-power spectrum. If the azimuth θ_i $(i = 1, 2, ..., N)$ of each subframe is obtained by splitting beam accurate azimuth estimation, it is assumed that the fitting equation is:

$$\theta = a + b \cdot i. \quad (2)$$

The least square minimizes the value in Eq. (3):

$$\Theta = \sum_{i=1}^{N}(\theta_i - \theta)^2 = \sum_{i=1}^{N}(\theta_i - (a + b \cdot i))^2. \tag{3}$$

We can obtain two equations that determine a and b:

$$\begin{aligned} a &= \frac{6}{N(N+1)}\left(\frac{2N+1}{3}A - B\right), \\ b &= \frac{12}{N(N^2-1)}\left(B - A\frac{N+1}{2}\right). \end{aligned} \tag{4}$$

In Eq. (4), assume $\sum_{i=1}^{N}\theta_i = A$, $\sum_{i=1}^{N}\theta_i i = B$.

The azimuth trend scale value T_r is obtained as follows:

$$\begin{aligned} T_r &= \theta(N) - \theta(1) = b(N-1) \\ &= 12\left(\frac{B}{N(N+1)} - A\frac{1}{2N}\right). \end{aligned} \tag{5}$$

Although azimuth trend scale identification has been applied in practical engineering, it is worth studying how the general law of the azimuth trend changes when cluster targets are incident in different spaces. In this paper, the azimuth trend values under different incident azimuth angles are obtained by simulating the superimposed multi-highlight echo. The general variation law is given, and the identification angle range with good effect can be obtained by using the azimuth trend for scale estimation.

2.2. Azimuth trend Doppler parameter estimation

In active detection, when there is relative motion between the sonar system and the target, the carrier frequency of the secondary transmitted wave received by the sonar system will produce a Doppler effect, which is manifested in the echo signal expanding in the time domain and shifting in the frequency domain (Boashash, 1992). Broadband transmitting signals are often used for detection. The echo signal contains more target attribute information and can improve the time-frequency resolution. Analysing and deriving the expression of the Doppler shift of wideband echo is the basis of studying the feature extraction method of the Doppler shift of target acoustic scattering highlights.

As shown in Fig. 2, the combined transceiver mode is adopted. The acoustic source emits a signal $s(t)$ in the far field. The point source target moves at a uniform speed from starting position A at speed μ, the angle relative to the combined transceiver system is θ, and v is the radial velocity component of the target. Assuming that the target moves towards the sonar, the pulse width of the transmitted signal is T. Derived from the geometric relationship, the received signal pulse width (Sheng *et al.*, 2020) is:

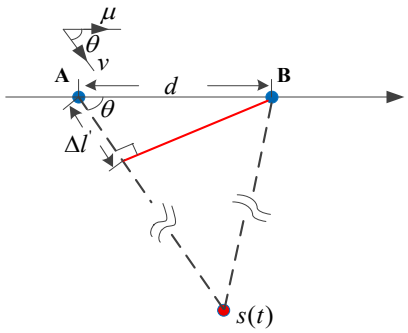

Fig. 2. Target motion model.

$$T_1 = \frac{1 - \frac{v}{c}}{1 + \frac{v}{c}}T = s \cdot T. \tag{6}$$

The speed of the underwater vehicle is markedly smaller than the speed of sound $v \ll c$, assuming

$$s = \frac{1 - \frac{v}{c}}{1 + \frac{v}{c}} = \frac{c - v}{c + v} \approx 1\frac{2v}{c} = 1 - \Delta,$$

$\Delta = \frac{2v}{c}$ will be called the Doppler factor, $f_d = \frac{2v}{c}f_0$ represents Doppler frequency shift in the frequency domain, and f_0 is the transmission signal frequency.

For broadband signals, different frequency points correspond to different Doppler shifts. The starting frequency of the linear frequency modulation (LFM) signal is set as f_L, the cut-off frequency is f_H, and the frequency modulation slope is $k = \frac{f_H - f_L}{T}$. Then, after the Doppler effect occurs, the initial frequency of the echo signal is set to $f_{L\text{echo}} = f_L + \Delta f_L$, and the cut-off frequency is set to $f_{H\text{echo}} = f_H + \Delta f_H$. According to Eq. (6), the frequency modulation slope of the frequency shift is:

$$\begin{aligned} k_{\text{echo}} &= \frac{f_{H\,\text{echo}} - f_{L\,\text{echo}}}{T_1} \\ &= \frac{f_H + \Delta f_H - f_L - \Delta f_L}{sT} = \frac{1 + \Delta}{s} \cdot k. \end{aligned} \tag{7}$$

After reorganization:

$$\Delta = \frac{2v}{c} = \frac{k_{\text{echo}} - k}{k_{\text{echo}} + k}. \tag{8}$$

According to Eq. (8), the Doppler factor of the broadband signal is related to the frequency modulation slope of the transmitted signal and the received highlight. As long as the frequency modulation slope of the received highlight is detected, the Doppler shift of the target can be estimated, and then the radial motion velocity of the target can be obtained.

3. Numerical simulation and analysis

The performance of azimuth trend estimation, Doppler parameter estimation of moving multiple

highlight targets, and recognition of typical formation motion highlights in feature space are verified through three simulation experiments.

The target model composed of three small balls is established. The coordinate system of the transceiver array and the target model is shown in Fig. 3. The target model and the L-shaped array are located in the xoz plane, in which the distance l_1 between ball 1 and ball 2 is 23 m, and the distance l_2 between ball 2 and ball 3 is 53 m. The 28-element receiving array is placed perpendicular to the z-axis, its first array element (set as the reference array element) is located along the z-axis, the no. 2 small ball is located at the coordinate (0, 0, 0), that is, the no. 0 beam is facing the no. 2 small ball, and the transmitting array of the L-shaped array is located near the reference array element, so the azimuth angle of the no. 2 small ball relative to the reference array element is 0°. The distance between the first array element of the array and the centre of the model is set as R, the target is located in the far field, and the distance taken in the simulation is $R = 600$ m. The angle between the longitudinal line of the model and the z-axis direction is the incident angle φ, and ϕ_1 and ϕ_2 are the azimuth angles of the reference array element corresponding to ball 1 and ball 3, respectively. The propagation time of the target model echo reaching the reference array element can be calculated according to the spatial position coordinates of the transmitting point and the spatial position coordinates of the receiving array element, and the echo signal received by the reference array element can be obtained. The echo signal of other array elements can be calculated through the array popularity matrix. The frequency range of LFM signal transmission is from f_L = 100 kHz to f_H = 200 kHz, and the signal pulse width is 20 ms.

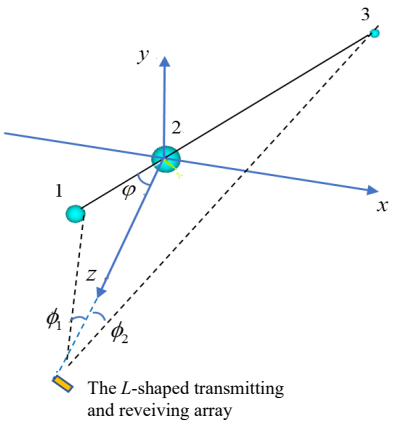

Fig. 3. Three-ball composition model.

Simulation analysis I: Azimuth trend estimation of single and multiple highlight targets

The azimuth trend estimation of single highlight and three highlight models at a certain angle and the omnidirectional incidence is studied. The physical properties of azimuth trend estimation are explained with this trend estimation. The validity of azimuth trend estimation affects the accuracy of target scale estimation. The three highlight echo is composed of the superposition of highlight components with different time delays, amplitudes and phases, and the amplitude corresponding to the three components is set to 0.8:1.2:0.6. The results are shown in Fig. 4.

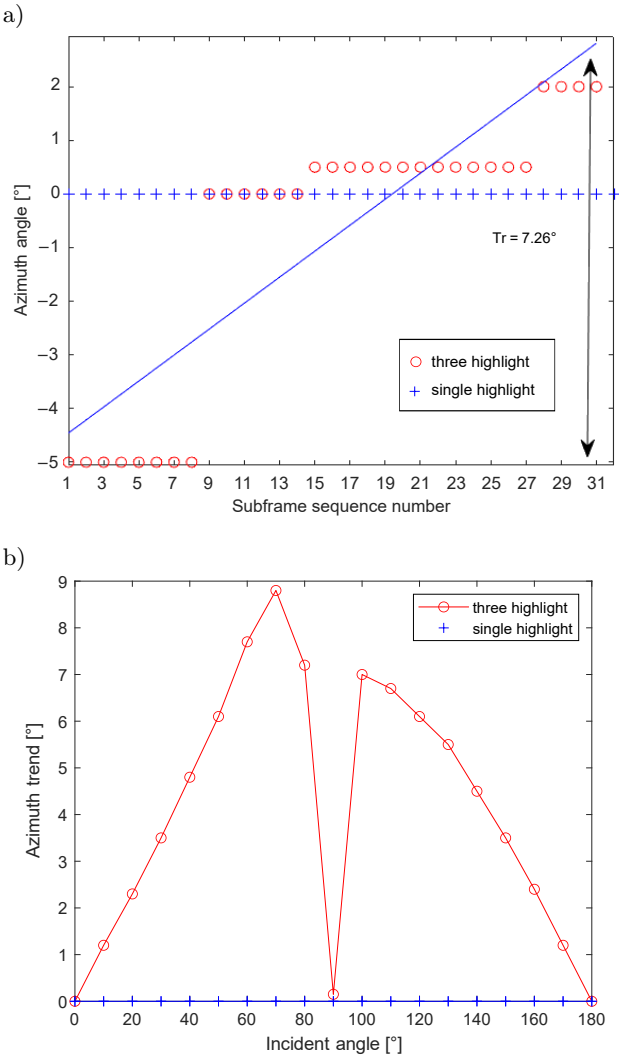

Fig. 4. Echo azimuth trend estimation: a) azimuth trend of 60° incident angle; b) azimuth trend of different incident angles.

Taking the incident angle φ = 60°, Fig. 4a shows that for single highlight targets, the azimuth trend value is very small, which is basically close to 0. Multiple highlights have a certain angle and distance distribution in space. The subframe segmentation in the azimuth trend estimation process distinguishes each highlight from the echo time or distance, accurately estimates the formation angle of each highlight by splitting the beam, and then obtains the maximum angle distribution range of the target highlight by the least

squares linear fitting. The estimated target azimuth trend Tr = 7.6°, $R1$ = 593.1 m, and $R2$ = 604.4 m. The calculated coverage scale of the three highlights is 76.6 m, which is consistent with the simulation condition of 76 m.

Figure 4b shows that the azimuth trend of a single highlight changes little with the incident angles and is also close to 0. The azimuth trend of multiple highlights with different incident angles is different. This is mainly because the relative positions of multiple highlights change when the incident angles are different, and the angular distribution of their echoes in space is different. However, the azimuth trend near the positive cross tends to 0 because the time delay difference of each ball echo is very small, most of the echoes overlap, and it is difficult to distinguish the highlights through subframe segmentation.

From the above analysis, we can see that the essence of the so-called azimuth trend is to distinguish the highlights through distance segmentation within the echo duration and accurately estimate the azimuth angle of the highlights by splitting the beam to obtain a linear estimation of the maximum angular range value of the echo distribution on the spatial distance covered by multiple highlights. The azimuth trend is an important feature that distinguishes multiple highlight targets from point targets and can be used for target scale recognition.

Simulation analysis II: Doppler parameter estimation of moving multiple highlight targets

To analyse the effect of the Doppler shift estimation method of the moving highlights, the instantaneous frequency distribution of the highlight is extracted in the time-frequency domain, and the Doppler shift is estimated by the relationship between the received highlight time-frequency spectrum slope and the transmitted signal frequency modulation slope. The initial state of the three highlights is shown in Fig. 3, and those are set to move at constant speeds of 1.2, 0, and −0.8 m/s. According to the derivation in Subsec. 2.2, the expression of the echo signal obtained from the moving highlight model is:

$$x(t) = \sum_{i=1}^{3} x_i(t) = \sum_{i=1}^{3} A_i \cos(2\pi(f_l + f_{di})t - \pi k_{di} t^2),$$

$$k_{di} = \frac{B}{\left(1 - \frac{f_{di}}{f_j}\right)T}, \tag{9}$$

where B is the bandwidth, T is the pulse width of the transmission signal, f_{di} is the Doppler frequency shift of the i-th highlight, and $f_j \in [f_L f_H]$.

To calculate the Doppler frequency shift of the moving target, it is necessary to obtain the instantaneous frequency of the highlight. In this paper, WVD is used to map the echo data to the time-frequency domain. The WVD of the LFM signal has the best energy aggregation in the time-frequency distribution method, and the signal energy is concentrated on a straight line that can reflect the instantaneous frequency change of the signal. However, there is cross-term interference in the multicomponent WVD. Morphological filtering is used to filter out cross-term interference by selecting appropriate structural elements, and the instantaneous frequency distribution of echo highlights is extracted without changing the time-frequency resolution and signal energy.

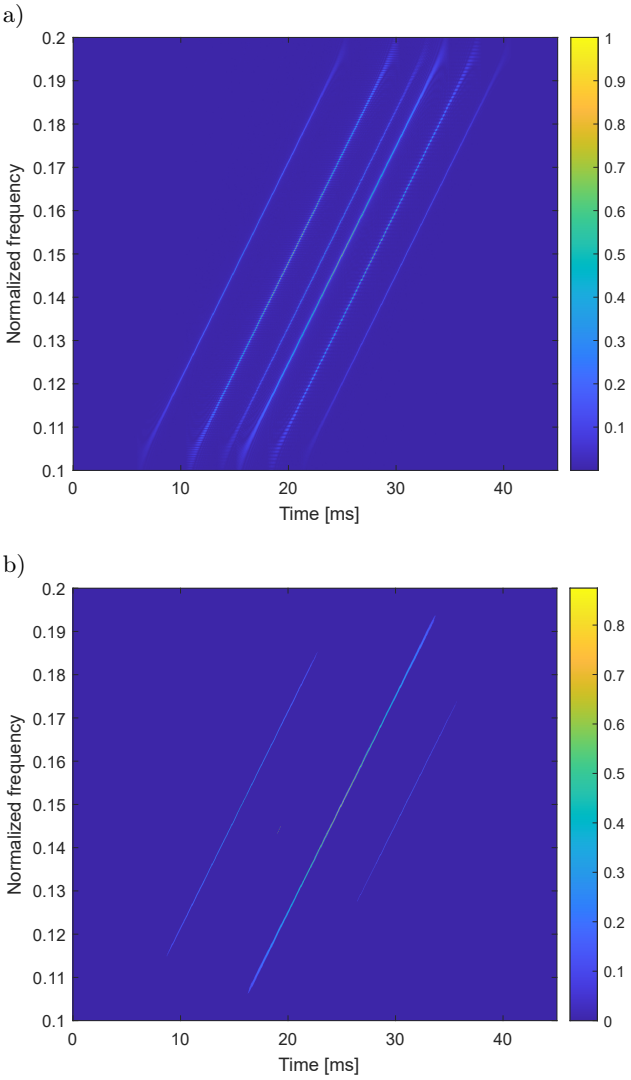

Fig. 5. Multiple highlight time-frequency distribution: a) time-frequency distribution of multiple highlights; b) highlight distribution after removing cross-term interference.

Figure 5a shows the time-frequency distribution of three highlights. The horizontal axis represents time, and the vertical axis represents normalized frequency. The normalized frequency is the ratio of signal frequency to the sampling frequency. In the experiment, the sampling frequency is 1 MHz. Theoretically, the time-frequency distribution should be three continu-

ous line spectra in the time-frequency domain, but the number of line spectra in the figure is greater than three, which is due to cross-term interference generated by the WVD algorithm. According to the morphological difference between highlight echo and cross-item interference, here, an inclined line with the same slope as the transmitted signal, is selected as the structural element, its length is half of the pulse width of the transmitted signal, and image objects of corrosion and expansion operations are used. The processing result is shown in Fig. 5b. The cross-term interference in the time-frequency plane is removed by morphological filtering, but the weak energy at the beginning and end of the highlight line spectrum is corroded, which does not affect the extraction and calculation of the line spectrum slope.

According to the time delay energy distribution of highlights, the line spectrum components of three highlights are searched and extracted, and the instantaneous frequency line spectrum of each highlight is obtained, as shown in Fig. 6. The slope of each highlight line spectrum is calculated as 5.0145e6, 4.9972e6, and 4.9917e6 by using the method of overlapping and averaging, and the slope of the transmitted signal is 5.0e6. The spectral slope of the no. 1 sphere increases relative to the transmitted signal, indicating that the highlight moves towards the transceiver transducer. The spectral slope of the no. 3 sphere decreases relative to the transmitted signal, indicating that the highlight moves to the transceiver transducer in a reverse manner. The no. 2 sphere is located on the centreline of the transceiver transducer, and its radial velocity in the direction of the transceiver transducer is 0, so the line spectrum slope is closer to the spectrum slope of the transmitted signal. According to the instantaneous frequency slope and Eq. (9), the radial velocity corresponding to the line spectrum is calculated as shown

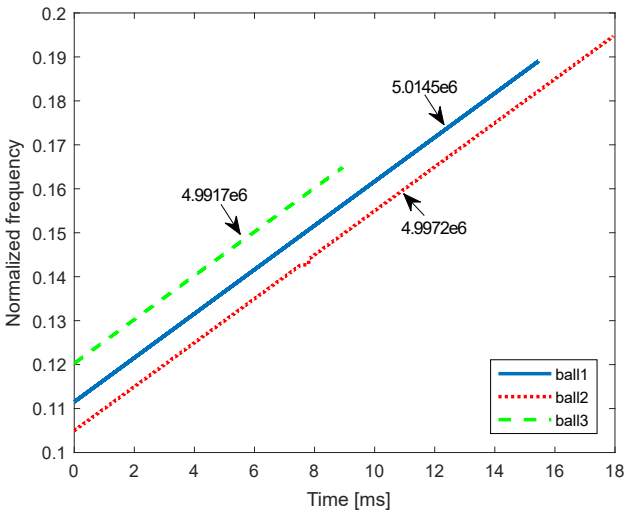

Fig. 6. Multiple highlight instantaneous frequency extraction.

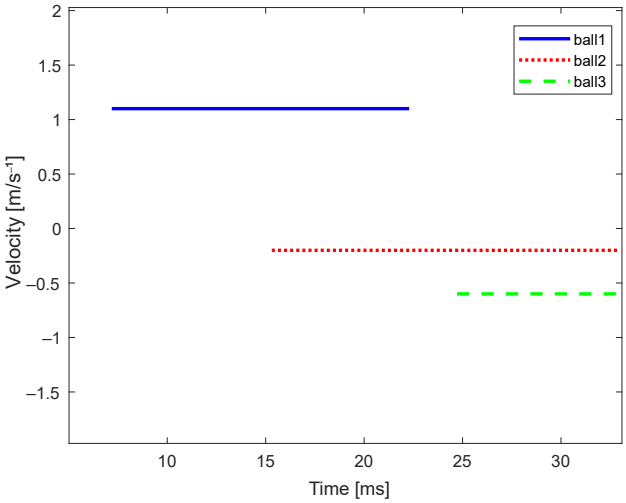

Fig. 7. Speed estimation.

in Fig. 7, which is 1.1, 0.2, and −0.63 m/s, respectively, which is basically consistent with the simulation conditions. The speed estimation accuracy is related to the calculation accuracy of the frequency modulation slope of the highlight, and it depends on the resolution of the instantaneous frequency of the highlight. The frequency resolution is inversely proportional to the effective length of the analysis signal. The simulation analysis signal length is 20 ms, and the corresponding frequency resolution is 50 Hz. According to the Eq. (8), the radial speed error is approximately 0.2 m/s.

Simulation analysis III: Recognition of typical formation motion highlights in feature fusion space

Verify the effectiveness of highlight identification in feature fusion space. The three-dimensional feature space of the target highlights is constructed by fusing the spatial orientation, velocity, and azimuth trend scale information, and the feature space distribution under the typical formation of multiple highlights is studied to identify the cluster targets. The spatial resolution of the receiving array is approximately 1°, and the three highlights move parallel to the x-axis with uniform speeds of 0.8, 1.6, and 2.2 m/s.

Assuming that the three highlights are arranged in a horizontal line, the vertical distance from the transceiver system is 600 m, and the spatial azimuth angles of the three highlights are −28, −30, and −31.5°, as shown in Fig. 8a. The orientation, azimuth trend scale, and Doppler parameters of the three highlights are extracted, the motion speed and coverage space scale of the echo are calculated and obtained, and the feature space is constructed, as shown in Fig. 8b. The spatial orientation of the three highlights is the same as the simulation setting, with speeds of 0.68, 1.53, and 2.1, respectively, and the coverage space scale is 48.2 m.

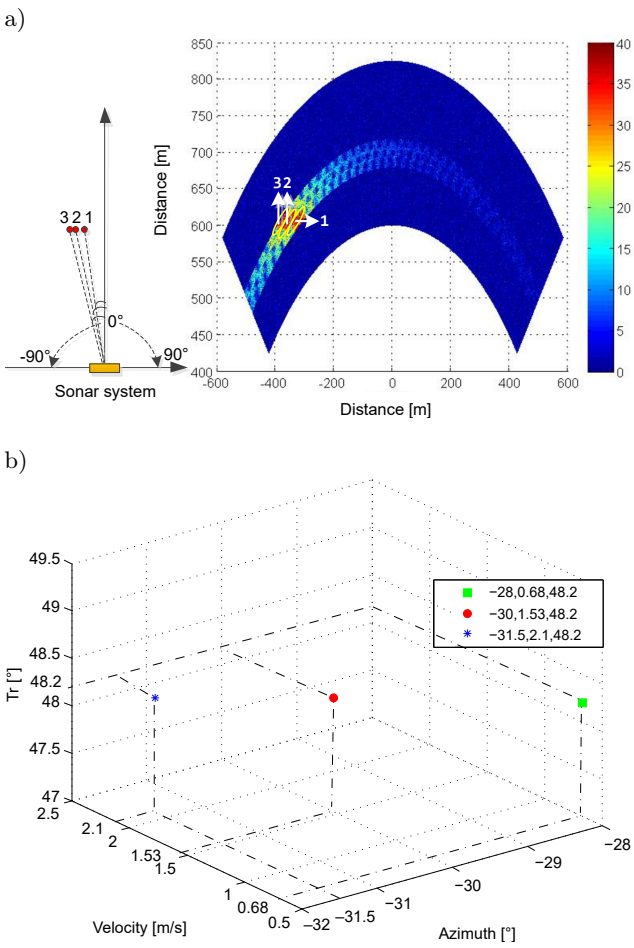

Fig. 8. Spatial characteristic distribution of the highlights in a line: a) spatial distribution of the highlights in a line; b) characteristic space of three highlights in a line.

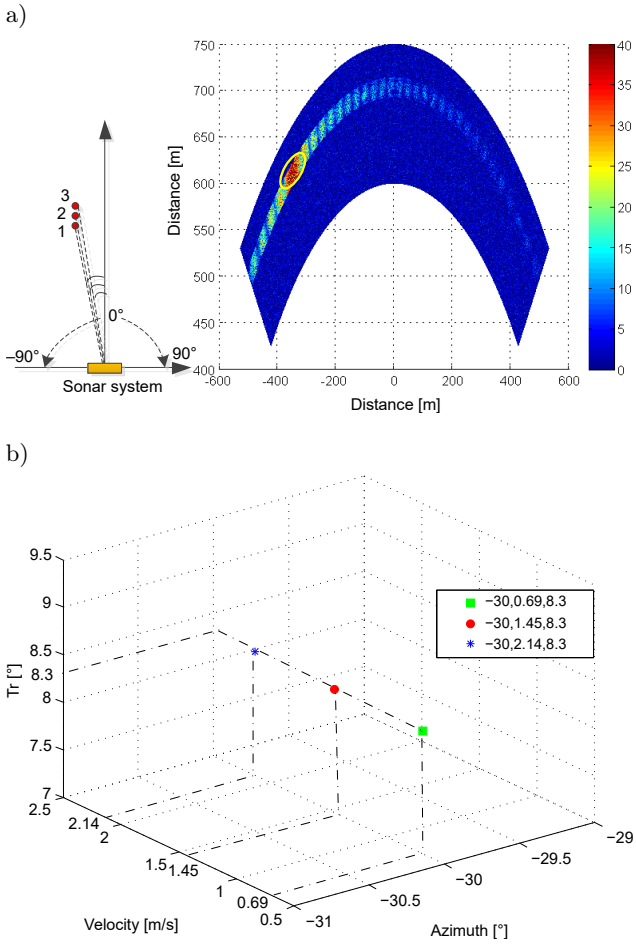

Fig. 9. Spatial characteristic distribution of three highlights in a column: a) spatial distribution of the highlights in a column; b) characteristic space of three highlights in a column.

Assuming that the three highlights are arranged in a straight line, the vertical distances from the transceiver system are 600, 603, and 608 m, and the vertical distances from the three highlights to the y-axis are all 10 m, as shown in Fig. 9a. The feature quantity of the three highlights is extracted and calculated, and the feature space is constructed, as shown in Fig. 9b. The spatial orientation of the three highlights is −30° because the spatial orientation distribution difference of the highlights is less than the resolution of the receiving array, resulting in the inability to accurately distinguish. The estimated velocities of the three highlights are 0.69, 1.45, and 2.14 m/s. The orientation estimation error is within the range of 1°, which has little impact on the velocity accuracy and can be ignored. The spatial scale covered by the three highlights is 8.3 m.

Assuming that the three highlights are arranged in a triangular formation, the vertical distances from the transceiver system are 600, 610, and 600 m, and the vertical distances from the three highlights to the y-axis are 5, 10, and 15 m, as shown in Fig. 10a. The construction feature space is shown in Fig. 10b. The spatial orientation of the three highlights is approximately −28.5°, the speeds are 0.68, 1.43, and 2.06 m/s, and the spatial scale covered is 10.4 m.

In summary, multiple highlights can be covered in large-scale targets in space. For clusters, due to the different motion states of small targets constituting clusters, the motion characteristics of multiple highlights in space are also different. Therefore, large-scale targets such as underwater crafts and cluster targets can be distinguished.

The simulation assumes that the target moves at a horizontal uniform speed; however, the target attitude and motion state of the actual cluster are usually unknown, so the actual motion speed of the target cannot be obtained, but the radial velocity of highlights still can be estimated according to the frequency modulation slope, and it can be determined that they belong to different targets according to the difference of the radial velocity in the same wave beam. In addition, during long-term underwater monitoring, the movement speed of the cluster also changes, and we

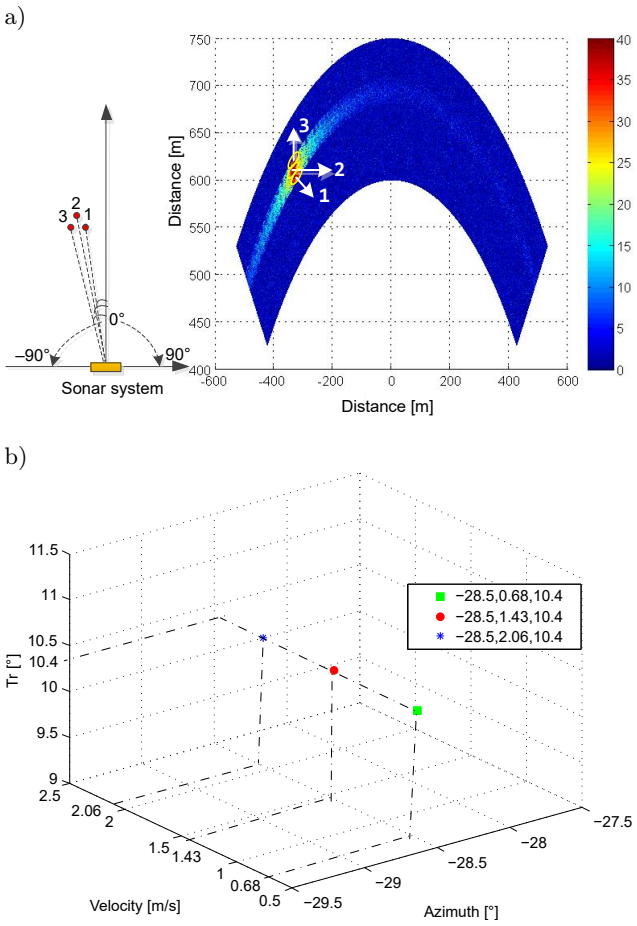

Fig. 10. Spatial characteristic distribution of three highlights in a triangle formation: a) spatial distribution of three highlights in a triangle formation; b) characteristic space of three highlights in a triangle formation.

can divide and analyse the length of the received data according to the resolution requirements, calculate the characteristic parameters of different data segments, and obtain real-time dynamic cluster characteristics for cluster target recognition.

4. Experimental analysis and verification

To verify the effectiveness of the multi-feature fusion method in identifying cluster targets, the experiment uses two small-scale UUVs moving underwater, with a length of 2.3 metres. The layout of the experiment is shown in Fig. 11. Using the combined transceiver detection mode, the pulse width of the transmitted linear frequency modulated signal is 2 ms, the frequency band is 100~200 kHz, the sampling rate is 1 MHz, the reception is a 28-element linear array, and the angular resolution of the conventional beamforming is 3.6°. The test conditions are as follows: one UUV and two UUVs sail autonomously on the side of the measuring ship at different speeds. The echo data of one UUV moving in a straight line at the side

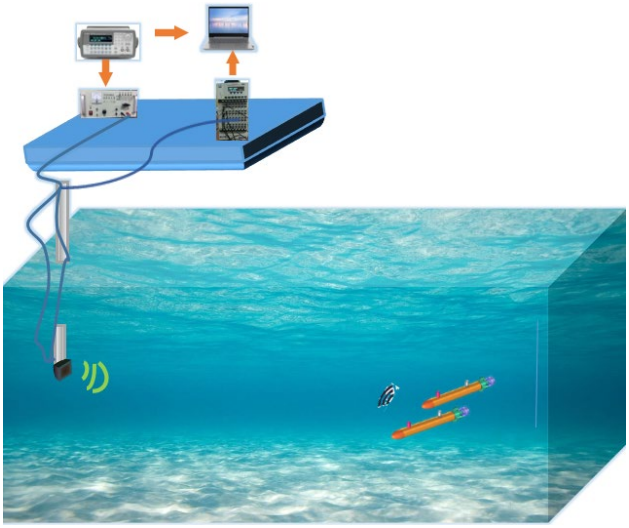

Fig. 11. Experimental layout.

of the measuring platform at a speed of 1.2 m/s and the echo data of two UUVs moving in circles at the side of the platform at speeds of 1.2 and 0.8 m/s are selected for data analysis and processing. The treatment results of typical working conditions are as follows.

5. Experimental results and analysis

(1) A single UUV sails in a straight line with a speed of 1.2 m/s parallel to the ship's side underwater. The time distance diagram of the target echo is shown in Fig. 12. Taking the 220th echo data for analysis, the spatial orientation of the target is shown in Fig. 13, and the target appears in the −12.4° orientation of the transceiver. Then, WVD time-frequency analysis is performed on the target echo, and the time-frequency distribution results are shown in Fig. 14a. Because there is interference in WVD, image morphological filtering is used to remove the cross term and the

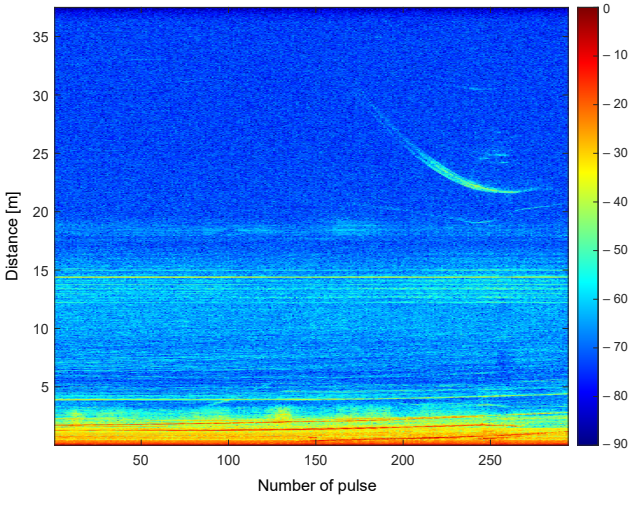

Fig. 12. Target time-distance diagram.

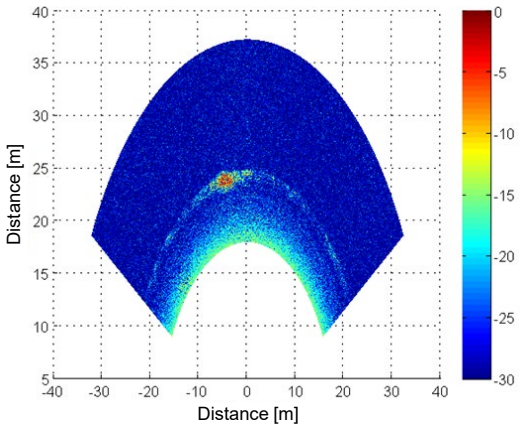

Fig. 13. Target space azimuth diagram.

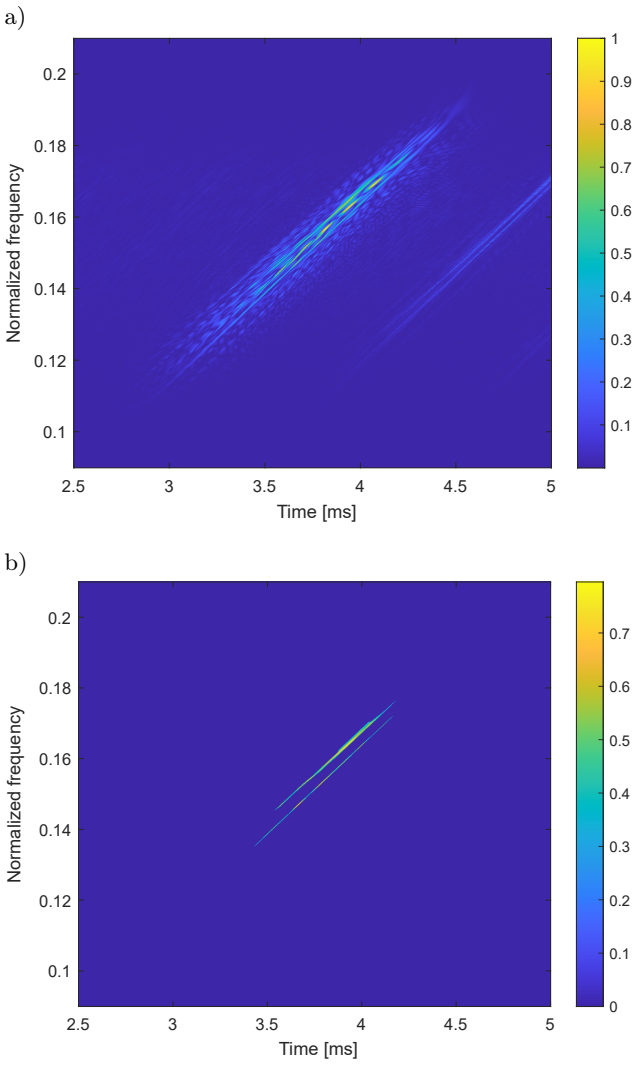

Fig. 14. Time-frequency distribution of target highlights: a) time-frequency distribution of the target echo; b) time-frequency distribution of the echo with cross-term interference removed.

reverberation. The results are shown in Fig. 14b. It can be seen that the target echo contains two scattering highlight components, which are the echoes scattered by the UUV cylinder. The instantaneous frequency distribution of highlights obtained by energy search in the time-frequency domain is shown in Fig. 15. It is estimated that the slopes of the instantaneous frequency spectral lines of the two highlights are 5.0036e7 and 5.0033e7, which are slightly larger than the slope of the transmitted signal of 5e7. The target moves towards the transceiver, and the calculation speeds are 1.23 and 1.13 m/s, respectively, which are approximately equal within the error range. It is estimated that the azimuth trend of the target Tr = 4.8°, as shown in Fig. 16, that is, the opening angle of the reference hydrophone array element relative to the scatterer is 4.8°. Using the maximum detection method, the distance between the first segment echo and the 32nd segment echo relative to the reference array element is 24.4 and 24.1 m, respectively. According to the triangular relationship, the target scale covered by the two highlights can be estimated to be approximately 2.1 m. Because the velocity

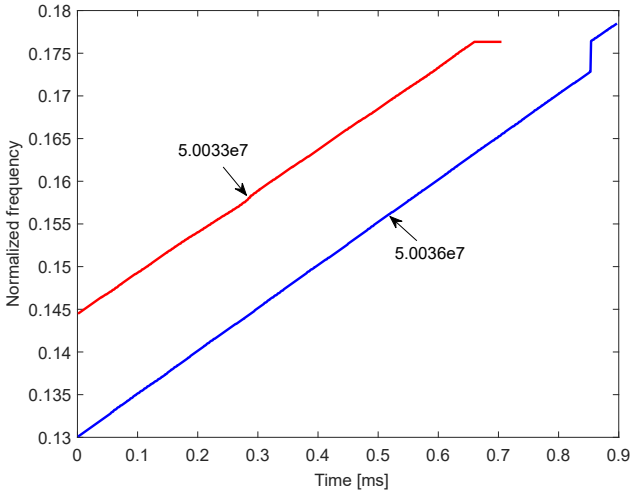

Fig. 15. Instantaneous frequency extraction of target highlights.

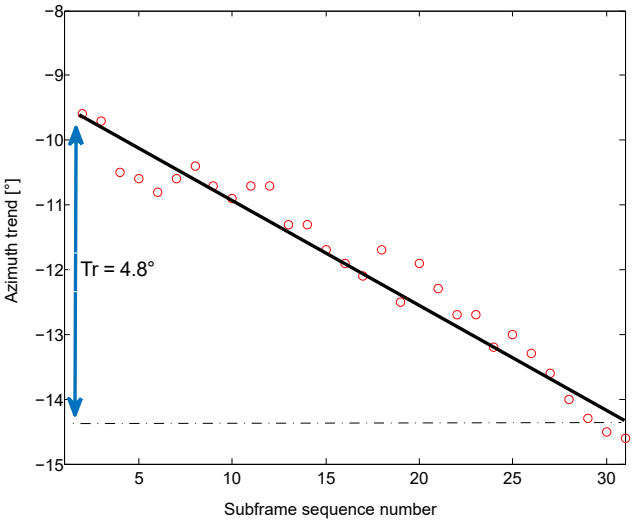

Fig. 16. Target azimuth trend diagram.

of the two highlights is approximately the same, different highlight components scattered by a target can be analysed, which is consistent with the test situation.

(2) Two UUVs move in circles on one side of the platform at speeds of 1.2 and 0.8 m/s. The time distance diagram of the target echo is shown in Fig. 17. We take the 903rd echo data for analysis. At this time, the no. 1 UUV passes the positive horizontal position and moves away from the transceiver along the circumference. The no. 2 UUV moves towards the positive horizontal position of the transceiver along the circumference. The spatial orientation of the target relative to the transceiver is shown in Fig. 18. Both UUVs appear in the 3.5° orientation of the transceiver. Then, the WVD time-frequency analysis is performed on the target echo, and the time-frequency distribution results are shown in Fig. 19a. Image morphological filtering is used to remove the interference, and the results are shown in Fig. 19b. The target echo contains three scattering highlights, and the instantaneous frequency distribution of the highlight echo is found in the time-frequency plane, as shown in Fig. 20. It is estimated that the instantaneous frequency line spectral slopes of the three highlights are 4.992e7, 5.004e7,

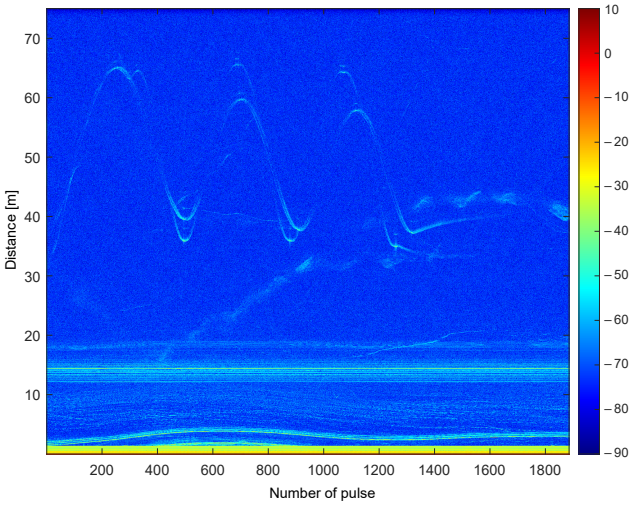

Fig. 17. Target-time distance diagram.

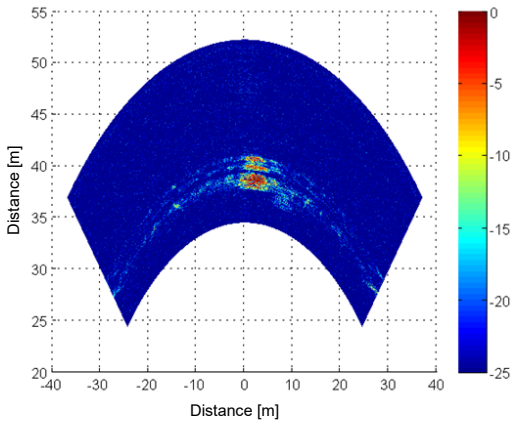

Fig. 18. Target space azimuth diagram.

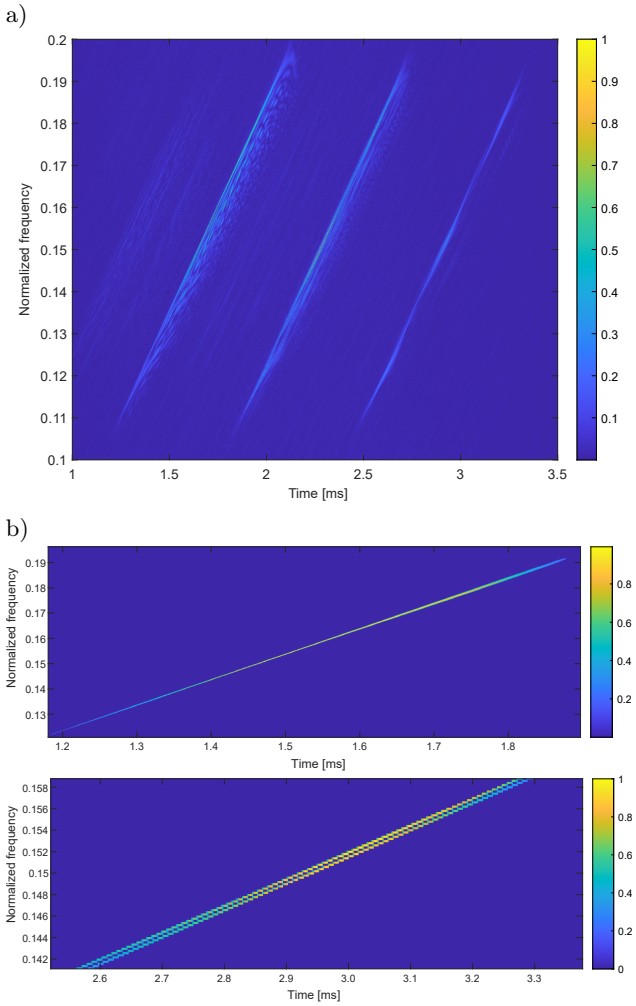

Fig. 19. Time-frequency distribution of target highlights: a) time-frequency distribution of the target echo; b) time-frequency distribution of the echo with cross-term interference removed.

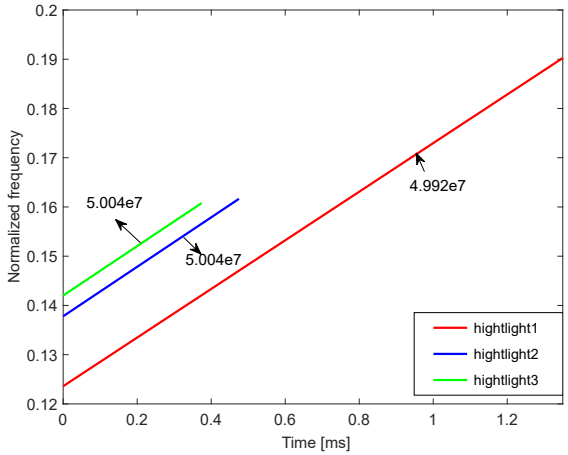

Fig. 20. Instantaneous frequency extraction of target highlights.

and 5.004e7, and the corresponding radial velocities are −0.6, 0.3, and 0.3 m/s. Because the motion attitude of the target is not accurately known, that is,

the angle between the target velocity and the radial velocity is unknown, the target motion velocity cannot be estimated from the radial velocity, but from the radial velocity, highlights 2 and 3 are the scattering components of the same target, moving towards the transceiver, and highlight 1 is far away from the transceiver. The estimation results of the target azimuth trend are shown in Fig. 21. Highlight 1 is the echo component scattered by the tail of the no. 1 UUV, and the azimuth trend of the single highlight is 0. Highlights and highlight 3 are the echo components scattered by the cylinder structure of the no. 2 UUV, and their azimuth trend is 2.1°. According to the geometric relationship, the distance between the echo and the reference array element is 39.8 and 41.1 m, respectively, and the target scale covered by the two highlights can be estimated to be approximately 1.9 metres. According to the analysis, highlight 1 is the echo of the no. 1 UUV, and highlights 2 and 3 are the echoes of the no. 2 UUV. Integrating the target orientation, azimuth trend and Doppler characteristics, it is recognized that the echo in a beam is composed of two UUV scattering components, which is consistent with the test situation.

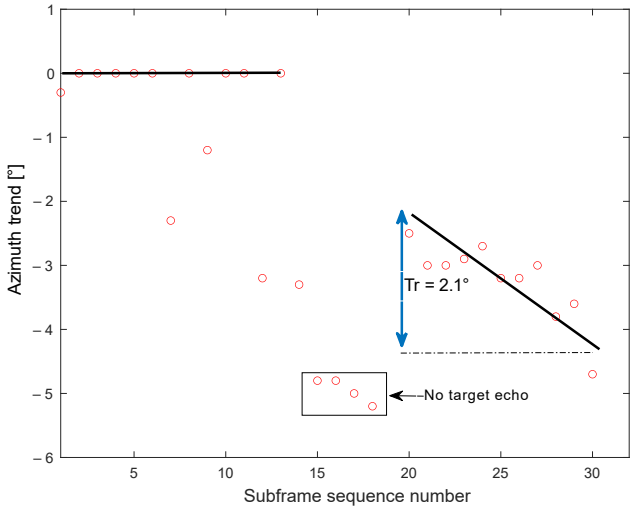

Fig. 21. Target azimuth trend diagram.

6. Conclusions

From the application perspective of underwater cluster target recognition, this paper studies and establishes the multi-feature space of echo highlights under typical formation and motion statuses. Through the analysis, the following conclusions can be drawn:

(1) The combined transceiver scattering characteristics of underwater cluster small targets are extracted, and the spatial orientation, geometric distribution and motion state differences of cluster targets are converted into physical features that are easy to use for sonar recognition;

(2) The objective of the azimuth trend is to distinguish each highlight through distance segmentation within the echo duration and obtain a linear estimation of the maximum angle range of the echo distribution on the spatial distance covered by multiple highlights. The azimuth trend is an important feature that distinguishes multi-highlight scale targets from point targets and can be used for target scale recognition.

(3) The corresponding relationship between the slope of the instantaneous frequency spectral line of highlights and the Doppler frequency shift is derived. The instantaneous frequency of multiple highlights of the target is extracted from the WVD time-frequency domain, and the motion information is calculated and extracted to describe the motion state of the highlights.

(4) The three-dimensional feature space of target highlight orientation, spatial scale and motion information is constructed, which realizes the effective analysis of multiple targets and solves the problem of fuzzy identification of multiple targets in the same beam. It provides effective and stable cluster echo recognition features for the early warning and monitoring of underwater ports and shallow waters.

Acknowledgments

The authors would like to acknowledge the China Postdoctoral Foundation Science Foundation (Grant no. 2020M681303) and the National Natural Science Foundation of China (Grant no. 61901079).

References

1. Boashash B. (1992), Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals, *Proceedings of the IEEE*, **80**(4): 520–538, doi: 10.1109/5.135376.

2. Bouaynaya N., Charif-Chefchaouni M., Schonfeld D. (2008), Theoretical foundations of spatially variant mathematical morphology, Part I: Binary images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **30**(5): 823–836, doi: 10.1109/TPAMI.2007.70754.

3. Chen Y.-F., Li G.-J., Wang Z.-S., Zhang M.-W., Jia B. (2013), Statistical feature of underwater target echo highlight [in Chinese], *Acta Physica Sinica*, **62**(8): 084302, doi: 10.7498/aps.62.084302.

4. Hollett R.D., Kessel R.T., Pinto M. (2006), At-sea measurements of diver target strengths at 100 kHz: Measurement technique and first results, [in:] *UDT-Europe 2006, Undersea Defence Technology Conference and Exhibition*, Germany.

5. Huang Y.S. *et al.* (2020), Review on the development of diver detection sonar system [in Chinese], *Journal of Unmanned Undersea Systems*, **28**(1): 1–8.

6. Jiang C.O., Xiang X.M. (2019), The analysis of the azimuth trend sequence distortion in the target scale identification [in Chinese], *Acta Acustica*, **38**(2): 275–277.
7. Jiang L.J., Yang J., Xu F. (2009), Technological progress of diver detection sonar [in Chinese], *Science in China Press*, **54**(3): 269–272.
8. Latif R., Aassif E., Moudden A., Faiz B. (2003), High-resolution time–frequency analysis of an acoustic signal backscattered by a cylindrical shell using a modified Wigner–Ville representation, *Measurement Science and Technology*, **14**(7): 1063–1065, doi: 10.1088/0957-0233/14/7/322.
9. Li X.K., Li T.T., Gu X.Y. (2014), Array gain of fourth-order cumulants beamforming under typical probability density background [in Chinese], *Acta Acustica*, **39**(5): 557–564, doi: 10.15949/j.cnki.0371-0025.2014.05.010.
10. Lo K.W., Ferguson B.G. (2004), Automatic detection and tracking of a small surface watercraft in shallow water using a high-frequency active sonar, *IEEE Transactions on Aerospace and Electronic Systems*, **40**(4): 1377–1388, doi: 10.1109/TAES.2004.1386890.
11. Ma G.Q., Xu D.M., Wang X.X. (2004), A study of simulation for underwater target scale identification with the technique of target azimuth tendency [in Chinese], *Ship Science and Technology*, **26**(3): 39–42.
12. Rodríguez M.A., San Emeterio J.L., Lázaro J.C., Ramos A. (2004), Ultrasonic flaw detection in NDE of highly scattering materials using wavelet and Wigner–Ville transform processing, *Ultrasonics*, **42**(1–9): 847–851, doi: 10.1016/j.ultras.2004.01.063.
13. Ronse C., Najman L., Decencière E. [Eds.] (2005), *Mathematical Morphology: 40 Years On. Proceedings of the 7th International Symposium on Mathematical Morphology, April 18–20, 2005*, Springer Dordrecht.
14. Sabra K.G., Anderson S.D. (2014), Subspace array processing using spatial time-frequency distributions: Applications for denoising structural echoes of elastic targets, *The Journal of the Acoustical Society of America*, **135**(5): 2821–2835, doi: 10.1121/1.4871183.
15. Sarangapani S., Miller J.H., Potty G.R., Reeder D.B., Stanton T.K., Chu D. (2005), Measurements and modeling of the target strength of diver, [in:] *Europe Oceans 2005*, **2**: 952–956, doi: 10.1109/OCEANSE.2005.1513185.
16. Sheng X.L., Mu M.F., Yin J.W., Yang C., Liu T. (2020), Sparse decompositon-based estimation method of Doppler frequency shift of underwater moving target, *Journal of Harbin Engineering University*, **41**(10): 1429–1435, doi: 10.11990/jheu.202007055.
17. Tang W.L. (1994), Highlight model of echoes from sonar targets, *Acta Acustica*, **19**(2): 92–100.
18. Yang Y., Li X.-K. (2016), Blind source extraction based on time-frequency characteristics for underwater object acoustic scattering [in Chinese], *Acta Physica Sinica*, **65**(16): 164301, doi: 10.7498/aps.65.164301.
19. Zhang Y.-M., Tong S. (2008), Tactics appliance of anti-diver technique in harbor defenses [in Chinese], *Ship Science and Technology*, **30**(6): 168–171.
20. Zhou D.B., Yi H. (2004), A study of submarine target bearing trend measurement and target discrimination [in Chinese], *Torpedo Technology*, **12**(4): 24–28.