



Article

High-Precision Sub-Wavelength Motion Compensation Technique for 3D Down-Looking Imaging Sonar Based on an Acoustic Calibration System

Jun Wang ^{1,2,3,†} , Peihui Liang ^{1,2,†} , Junqiang Song ^{1,2}, Pan Xu ^{1,2}, Yongming Hu ^{1,2}, Peng Zhang ^{1,2}, Kang Lou ^{1,2}, Rongyao Ren ³ and Wusheng Tang ^{1,2,*}

¹ College of Meteorology and Oceanography, National University of Defense Technology, Changsha 410073, China; wwjj103@nudt.edu.cn (J.W.); peinudt@nudt.edu.cn (P.L.); junqiang@nudt.edu.cn (J.S.); xupan09@nudt.edu.cn (P.X.); sdss@21cn.com (Y.H.); zhangpeng23a@nudt.edu.cn (P.Z.); loukang17@nudt.edu.cn (K.L.)

² Hunan Key Laboratory for Marine Detection Technology, Changsha 410073, China

³ Tianjin Institute of Advanced Technology, Tianjin 300450, China; renry9710@163.com

* Correspondence: tangwusheng@nudt.edu.cn; Tel.: +86-152-0087-0306

† These authors contributed equally to this work.

Abstract: Three-dimensional hydro-acoustic imaging is a research hot spot in the underwater acoustic signal processing field, which has a wide range of application prospects in marine environmental resource surveying, seabed topography and geomorphological mapping, and underwater early warning and monitoring. To solve the problem that the resolution of the current imaging sonar reduces rapidly with increase in distance and a scanning gap exists in side-scan sonar, we designed a down-looking 3D-imaging sonar with a linear array structure. The imaging scheme adopts a time-domain spatial beam-forming method with the Back Projection (BP) algorithm as the core, and the formation of a virtual plane array can effectively improve the along-track resolution. To cope with the interference of the carrier motion error on the imaging, we proposed a high-precision sub-wavelength motion compensation method based on a real-time acoustic calibration system. Simulation and real data experiments show that the motion compensation method can effectively eliminate the influence of motion error and make the imaging energy more focused, leading to higher-quality acoustic images. Under equal average energy, the maximum superimposed sound intensity values in the imaging results increased by 20.75 dB and 6.57 dB, respectively, for simulation and real data. After motion compensation, the resolution of this imaging system reached $3\text{ cm} \times 3\text{ cm} \times 2.5\text{ cm}$ @ Depth = 17 m, TBP = 30 s · Hz.

Keywords: 3D hydro-acoustic imaging; down-looking linear array; motion compensation; acoustic calibration



Academic Editors: Andrzej Stateczny and Jaroslaw Tegowski

Received: 14 November 2024

Revised: 23 December 2024

Accepted: 24 December 2024

Published: 27 December 2024

Citation: Wang, J.; Liang, P.; Song, J.; Xu, P.; Hu, Y.; Zhang, P.; Lou, K.; Ren, R.; Tang, W. High-Precision Sub-Wavelength Motion Compensation Technique for 3D Down-Looking Imaging Sonar Based on an Acoustic Calibration System. *Remote Sens.* **2025**, *17*, 58. <https://doi.org/10.3390/rs17010058>

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, with the increasing maturity of underwater acoustic transducer technology and the accelerating development of hydro-acoustic imaging technology, high-resolution imaging and high-confidence recognition of small underwater targets have become a research hot spot [1–3]. Three-dimensional (3D)-imaging sonars employ beam imaging technology popularly to reconstruct underwater 3D acoustic images [4,5]. However, the spatial resolution of conventional beam imaging is limited, which is usually determined by the array aperture [6–8]. Adaptive beam imaging is a technique which is capable of improving the spatial resolution. Stoica et al. [9] applied a natural extension of

the Capon beamformer to the case of uncertain steering vectors, addressing the issue of performance degradation in the Capon beamformer when the steering vector knowledge is imprecise. Liao et al. [10] proposed a robust adaptive beam imaging technique that controls the peak sidelobe level in adaptive arrays, which is important for maintaining the integrity of the imaging resolution by managing interference. Along with the adaptive beam imaging technique, Wang et al. [6] used the deconvolution algorithm DAMAS2 to enhance 3D sonar resolution.

High-resolution imaging must provide a sufficient processing gain to combat the background noise. One way to combat noise is through the design of image reconstruction [11]. Yazici et al. [12] designed a Wiener filter in the affine group Fourier transform domain to minimize the wideband noise range-Doppler reflectivity, which can enable an operator to precondition transmitted waveforms to reject noise. Limin et al. [13] explored a family of rational orthogonal wavelets in active sonar detection to achieve Doppler robustness and noise mitigation.

Furthermore, motion errors, waveform design schemes and the sensor array structure can affect the resolution of 3D underwater acoustic imaging, with motion errors particularly reducing the imaging resolution significantly [14–16]. Due to various random factors, such as wind, waves, and sonar carrier manipulation, the carrier will deviate from the ideal trajectory. The image quality will be compromised if motion errors are not accounted for during the imaging progress [17,18]. Motion errors exceeding $1/8$ of a wavelength will significantly affect the quality of 3D-imaging. General motion error compensation methods include error coarse compensation based on motion sensors [19,20] and motion error compensation based on sonar echo data [21,22]. The former is limited by the performance of the sensors and usually cannot achieve sub-wavelength accuracy. The latter has high compensation accuracy but has strict requirements for the structure and movement of the array.

In our study, the imaging scheme adopts the time-domain spatial beam-forming method with the Back Projection (BP) algorithm as the core. The formation of a virtual plane array can effectively improve the along-track resolution. To cope with the interference of the carrier motion errors on the imaging, we proposed a high-precision motion compensation method based on a real-time acoustic calibration system. The compensation accuracy can reach the sub-wavelength level, which can meet the needs of high-resolution imaging.

2. Moving Linear Array Down-Looking 3D-Imaging Model

2.1. Down-Looking Imaging Sonar System

Given the limited along-track resolution of the down-looking linear array sonar, we create a large virtual aperture along the trajectory by moving the linear array. The signal transceiver mode of the downward view can avoid the gaps problem of the side-scan imaging sonar and has unique advantages in 3D imaging, which is similar but different to the mode of the multi-beam sounder. In addition, adopting the down-looking imaging system, the spatial coordinate relationship is relatively simple, which is convenient for spatiotemporal registration and conducive to the fusion imaging of sonar data from multiple voyages.

In this paper, we studied a 3D image reconstruction method for the moving down-looking linear array imaging sonar. Figure 1 shows the working diagram of the down-looking linear array 3D-imaging system.

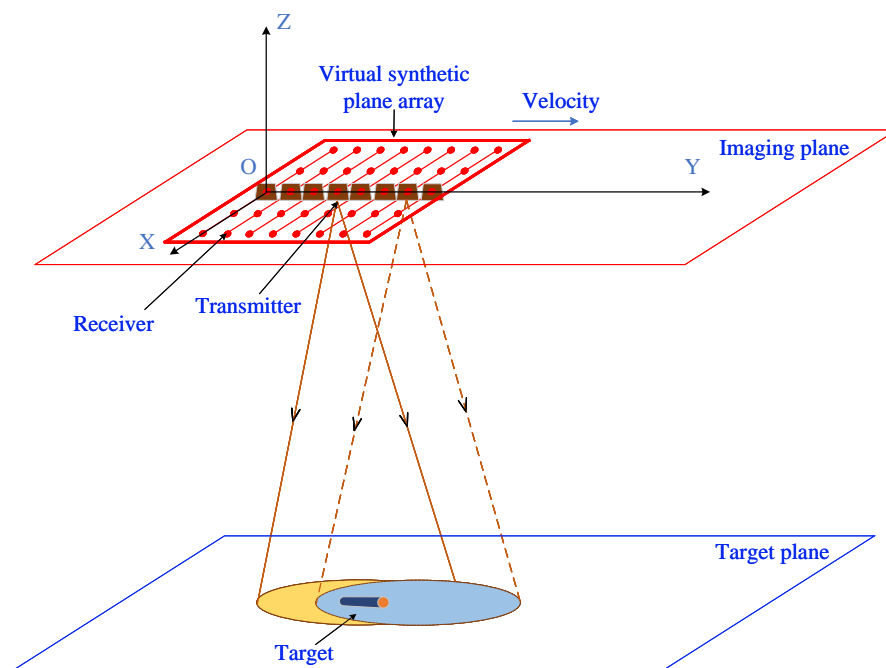


Figure 1. Schematic diagram of 3D-imaging sonar operation.

As shown in Figure 1, the carrier coordinate system is established with the center point of the linear array as the coordinate origin, also called the relative coordinate system. The extension direction of the linear array is the X-direction, the direction perpendicular to the linear array is the Y-direction, and the vertical upward direction is the Z-direction. The establishment of the coordinate system follows the right-hand rule. To better interpret the motion imaging process, the Y-direction in the diagram is redefined as the along-track direction. The X-direction and Z-direction in the diagram are redefined as the across-track direction and the depth direction, respectively.

Different from the general linear array, the specially designed receiving array has two hydrophones with a certain distance on both sides of the along-track direction, which is convenient for receiving three-dimensional underwater acoustic signals for the positioning of the sonar system [23]. The transmitter is a cylindrical acoustic source with a diameter of 3 cm located below the center of the horizontal array. The structure of the down-looking linear array imaging sonar is shown in Figure 2.

In motion imaging of the sonar system, the sound source of the down-looking linear array imaging sonar periodically actively emits a pulsed linear frequency modulation signal, also called a chirp signal. The chirp signal has good anti-interference ability due to its good autocorrelation characteristics, which makes it resistant to noise interference and improves the imaging quality in hydro-acoustic imaging. At the same time, the wide beam of the sound source can continuously irradiate the same target, and the imaging of the target can be completed by forming a virtual synthesis plane array by moving the linear array, which can effectively improve the along-track resolution of the target image [24].

The echo signals received by the linear receiving array, after pulse compression, can be recorded as a two-dimensional matrix $S_{N,M}(i, j; t)$, representing that this set of processed signals comes from M frames of echo signals received by N hydrophone elements. i is the index number of the hydrophone, and j is the frame number of the received signal.

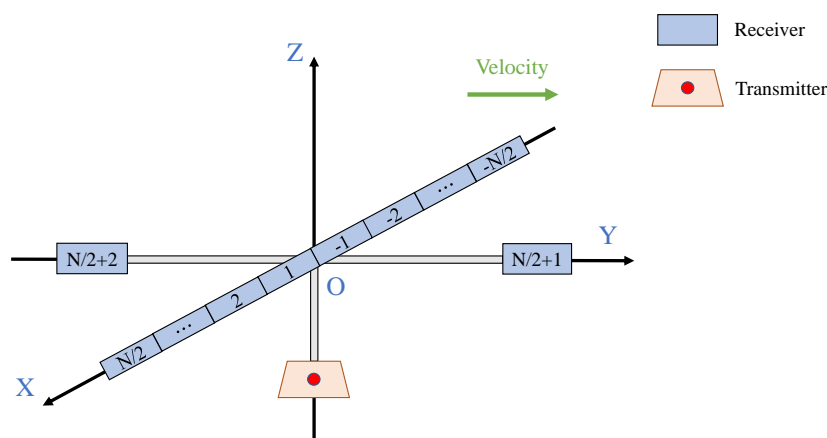


Figure 2. Structure of the transceiver array in the carrier coordinate system.

2.2. 3D BP Algorithm for the Down-Looking Linear Array Sonar

Despite the lower imaging efficiency of time-domain algorithms compared to frequency-domain and beam-domain algorithms, which results in high computational loads when dealing with long echo data and large-scale imaging scenes, both frequency-domain range-Doppler algorithms and beam-domain $w - k$ algorithms require the carrier to move in an almost strictly uniform linear motion. In the context of complex underwater environmental interference and difficult carrier attitude control, the advantages of time-domain imaging algorithms, which do not demand a specific trajectory for the carrier and offer good imaging stability and high precision, become prominent. Moreover, time-domain imaging algorithms are characterized by their simple imaging principles and ease of engineering implementation. Research into time-domain 3D-imaging algorithms has a more direct significance for array imaging on a larger scale [25]. The application of GPU parallel computing technology will greatly enhance the efficiency of time-domain imaging algorithms. And we have applied this technology to data processing.

Therefore, the 3D-imaging scheme adopts the time-domain spatial beam-forming method with the Back Projection (BP) algorithm as the core. The BP algorithm is an exact point-by-point search imaging method within the beam domain [26,27]. The basic principle for achieving 3D-imaging is based on the forward sound propagation model and pulse-compressed sonar echo data. It calculates the round-trip time delay of each grid point in the 3D target space to the sound source and hydrophone. Based on the round-trip time delay and the spatial geometric correspondence, the algorithm coherently superimposes the amplitudes of the processed signals that come from each hydrophone element and back-projects this sum onto the 3D grid point in the target space as the scattering intensity at that point. The described protocol will address the signal of each frame and amalgamate the outcomes of the 3D grid point intensity calculations. Thereby, the 3D image of the target space is reconstructed. It should be noted that when the platform is moving at a high speed, the resolution along the trajectory will be affected by the time delay. If this is not taken into account, it will affect the quality of the imaging.

The fundamental steps for 3D-imaging via the back-projection algorithm are establishing the forward model of sound propagation (i.e., ascertaining the positional information of the sound source, hydrophone elements, and target grid points), calculating the round-trip time delays, and performing back-projection. In the 3D BP imaging algorithm for the

down-looking linear array, the imaging space is defined within the absolute coordinate system. And then we need to obtain the absolute position coordinates of the hydrophones and the sound source in each frame. Based on these coordinates, the target's round-trip time delay is computable, and finally, back-projection is conducted to achieve 3D-imaging. The specific process is as follows:

Step I: Set the imaging space

Establish an absolute coordinate system, setting the plane center of the bottom surface of the target imaging space as the origin of coordinate O_{abs} . The imaging space is designated as $A\{x \mid X_{min} \leq x \leq X_{max}, y \mid Y_{min} \leq y \leq Y_{max}, z \mid Z_{min} \leq z \leq Z_{max}\}$, and let the spacing of the imaging grid points along each axis of the coordinate system be Dx, Dy, Dz . Consequently, the coordinates of the imaging grid points in the absolute coordinate system can be determined as $P(Gx, Gy, Gz)$. The number of the imaging grid points can be calculated as the following equation:

$$P_{\text{mum}} = \left[\frac{X_{\text{max}} - X_{\text{min}}}{Dx} + 1 \right] \times \left[\frac{Y_{\text{max}} - Y_{\text{min}}}{Dy} + 1 \right] \times \left[\frac{Z_{\text{max}} - Z_{\text{min}}}{Dz} + 1 \right] \quad (1)$$

Step II: Obtain the absolute position coordinates of the hydrophones and the sound source in each frame

Let the absolute coordinates of the i th hydrophone element in the j th frame be denoted as $H(x_H^{(i,j)}, y_H^{(i,j)}, z_H^{(i,j)})$, and the absolute coordinates of the sound source in the j th frame be denoted as $S(x_S^{(j)}, y_S^{(j)}, z_S^{(j)})$. Given the assumptions in the down-looking linear array 3D-imaging scheme, the positional relationship between each hydrophone element and the sound source is fixed. Therefore, by calculating the absolute position coordinates of a specific hydrophone element or the sound source in a frame, one can determine the absolute position coordinates of all hydrophone elements and the sound source in this frame.

To facilitate the calculation of the absolute position coordinates for each hydrophone element and the sound source, the carrier coordinate system (relative coordinate system) has already been established in the text, as shown in Figures 1 and 2. Here, let the relative coordinates of the i th hydrophone element be denoted as $H_0(x_0^{(i)}, y_0^{(i)}, 0)$, and the relative coordinates of the sound source be denoted as $S_0(0, 0, z_0)$. Let Ds represent the spacing between each hydrophone element along the x-axis, and Ls represent the distance from two hydrophone elements along the y-axis to the linear array.

The relative horizontal coordinate of the hydrophone element $x_0^{(i)}$ can be calculated using the following formula:

$$x_0^{(i)} = \begin{cases} (Ds) \times i - \frac{Ds}{2}, 1 \leq i \leq N/2 \\ (Ds) \times i + \frac{Ds}{2}, -N/2 \leq i \leq -1 \\ 0, i = N/2 + 1, N/2 + 2 \end{cases} \quad (2)$$

The relative vertical coordinate of the hydrophone element can be calculated using the following formula:

$$y_0^{(i)} = \begin{cases} 0, -N/2 \leq i \leq N/2 \\ Ls, i = N/2 + 1 \\ -Ls, i = N/2 + 2 \end{cases} \quad (3)$$

Step III: Calculate the target's round-trip time delay

Given the aforementioned conditions, the round-trip time delays from the target grid coordinate points to each hydrophone element and the sound source can be calculated in the j th frame.

The time delay from the target grid point to the sound source can be calculated using the following formula:

$$\tau_{TS} = \frac{1}{c} \sqrt{(Gx - x_S^{(j)})^2 + (Gy - y_S^{(j)})^2 + (Gz - z_S^{(j)})^2} \quad (4)$$

The time delay from the target grid point to the hydrophone elements can be calculated using the following formula:

$$\tau_{TH} = \frac{1}{c} \sqrt{(Gx - x_H^{(i,j)})^2 + (Gy - y_H^{(i,j)})^2 + (Gz - z_H^{(i,j)})^2} \quad (5)$$

Therefore, the round-trip time delays from the target grid coordinate points to the i th hydrophone element and the sound source in the j th frame can be written as:

$$\tau_{i,j} = \tau_{TS} + \tau_{TH} \quad (6)$$

Step IV: Delay and sum to achieve 3D-imaging

Let the signal sampling rate be denoted as f_s . After pulse compression, the time-domain signal $S_{N,M}(i, j; t)$ is sampled to obtain the discrete-time echo signal $S_{N,M}(i, j; n)$, and the round-trip time delay $\tau_{i,j}$ is sampled to obtain the round-trip delay points number $\Gamma_{i,j}$. For a target point P in space, the imaging result at this point is the accumulation of the amplitudes of the echo signals received by each hydrophone element in each frame at the time delay of that point. The imaging result for this point can be represented by the following formula:

$$\text{Im } P(Gx, Gy, Gz) = \left| \sum_{i=-N/2}^{N/2} \sum_{j=1}^M S_{N,M}(i, j; n - \Gamma_{i,j}) \right| \quad (7)$$

Applying the aforementioned processing to each grid point in the target imaging space, the 3D acoustic map of the target imaging space can be obtained.

2.3. Simulation of the 3D BP Algorithm

In order to validate the effectiveness of the time-domain BP 3D-imaging algorithm used in our paper, the point target simulation experiment under typical parameters is designed and carried out.

In the simulation, the pulse length of the linear frequency modulation signal $T_r = 1$ ms, the carrier center frequency $f_c = 85$ kHz, the signal bandwidth $B_r = 30$ kHz, and the frequency range is from 70 kHz to 100 kHz. The signal amplitude is 0.8, and the initial phase is 0. A total of frame number $M = 50$ chirp signals are emitted, and the sampling rate of the echo signal $f_s = 1$ MHz. The sound speed in water $c = 1449.93$ m/s.

In the simulation, the receiving array consists of a total of $(N + 2) = 64$ hydrophone elements. The spacing between the N hydrophone elements along the x-axis $D_s = 9.5$ cm, and the distance from the latest two hydrophone elements along the y-axis to the linear array $L_s = 0.5$ m. The velocity of the carrier $V_0 = 0.05$ m/s.

The relative coordinate of the sound source is $S_0(0,0,0.04)$. The imaging space is designed as $A\{x \mid -0.5 \text{ m} \leq x \leq 0.5 \text{ m}; y \mid -0.5 \text{ m} \leq y \leq 0.5 \text{ m}; z \mid -0.1 \text{ m} \leq z \leq 0.1 \text{ m}\}$, and the spacing of the imaging grid points along each axis of the coordinate system $Dx = Dy = Dz = 0.01 \text{ m}$. The single target grid point is set at the origin of the absolute coordinate system; its coordinate is $P(0,0,0)$. The distance from the target grid point to the imaging plane is $D_T = 10 \text{ m}$.

The simulated echoes received by the hydrophone array of the down-looking linear array sonar are processed using the time-domain 3D back-projection imaging algorithm. The imaging results are shown in Figure 3.

As the 2D-imaging results show in Figure 3a, apart from the slight side lobes in the across-track and along-track directions where the target point is located, the point target focusing imaging effect is good, truly reflecting the position of the underwater target. By setting a reasonable threshold, the imaging results of the point target can be highlighted from the target imaging space. Figure 3b presents the 3D-imaging result of the point target after adjusting the threshold.

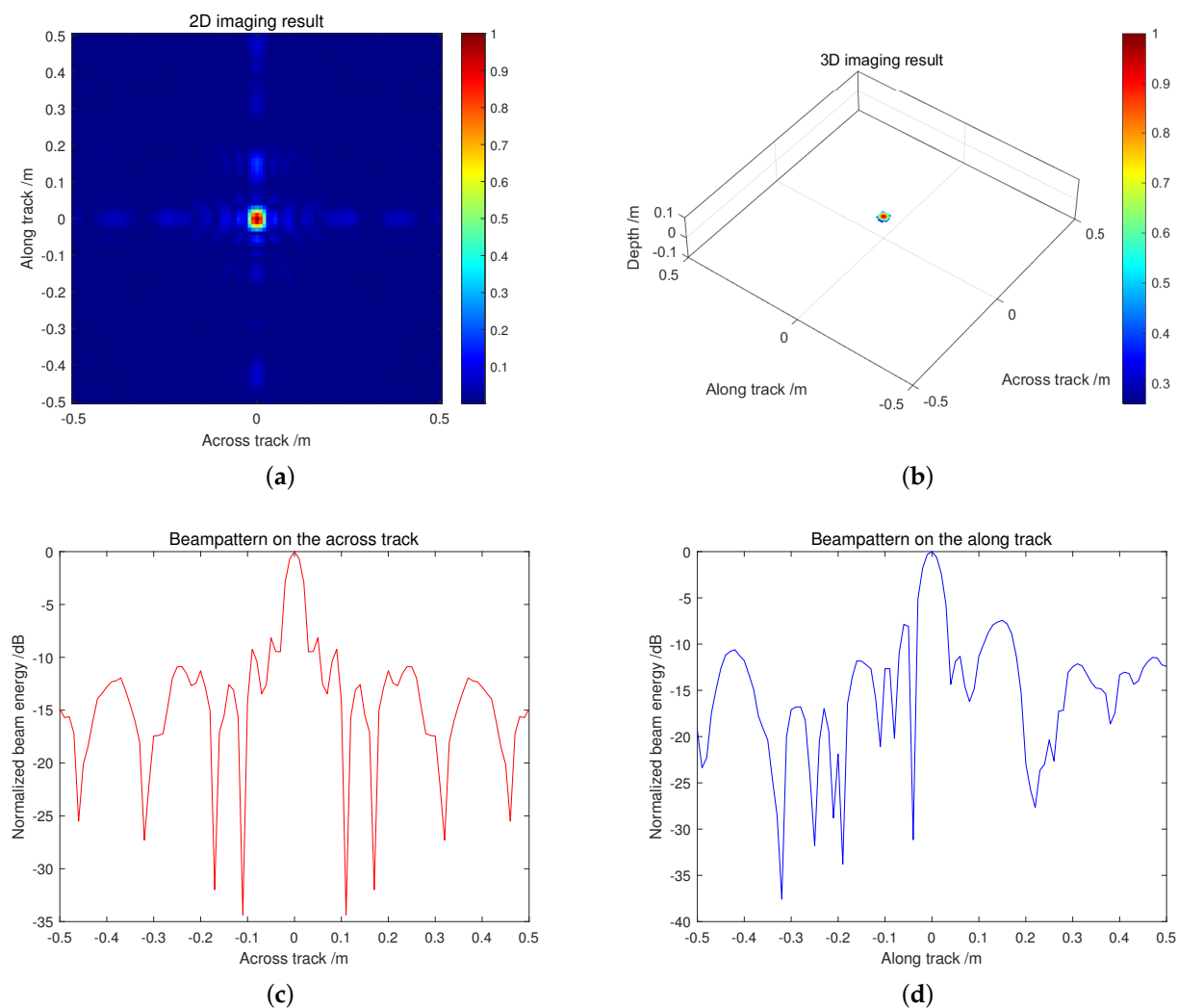


Figure 3. Cont.

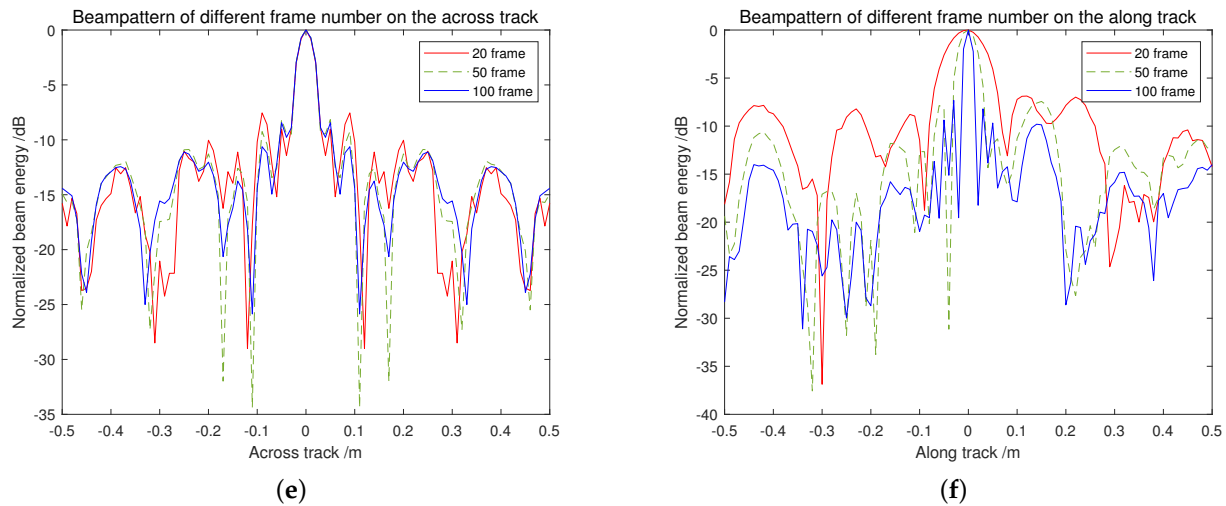


Figure 3. Simulation of single point target: (a) 2D-imaging result of the point target; (b) 3D-imaging result of the point target; (c) beampattern on the across track; (d) beampattern on the along track; (e) beampattern of different frame number on the across track; (f) beampattern of different frame number on the along track.

In Figure 3c, the beam pattern of the point target in the across-track direction has a 3 dB bandwidth of 0.04 m. The imaging result of the down-looking linear array sonar in the across-track direction is obtained by processing the echo signals received by N hydrophone elements in the across-track direction after pulse compression. The resolution capability cannot surpass the limitation of the physical aperture. The resolution capability in the across-track direction is determined by the across-track receiving beam width, which is related to the real aperture of the N -element linear array, the wavelength of the central wavelength of the sound wave λ_c , and the target depth D_T . Therefore, the resolution capability of the sonar system in the across-track direction is a constant value at the same depth.

In Figure 3d, the beam pattern of the point target in the along-track direction also has a 3 dB bandwidth of 0.04 m. However, the along-track resolution of the imaging system surpasses the limitation imposed by the physical aperture. It is determined by the along-track receiving beam width, which is related to the virtual synthetic aperture $L = V_0 t$ formed by the down-looking linear array during the motion imaging process, the central wavelength of the sound wave λ_c , and the target depth D_T .

For the down-looking linear array imaging system, the across-track imaging resolution is almost fixed, but the along-track imaging resolution can be improved by shortening the along-track receiving beam width, which is caused by increasing the length of the synthetic virtual aperture along the track. Different virtual aperture lengths can be chosen and the echo signals of the varied frames processed to obtain the beampattern in the across-track and along-track directions at the target depth, which is shown in Figure 3e,f, respectively. More imaging results can be found in Figure A1.

From Figure 3e,f, it can be observed that as the length of the synthetic virtual aperture increases, the resolution of the sonar system's acoustic image in the across-track direction (x-axis) remains constant, while the resolution in the along-track direction (y-axis) gradually improves. The results of the simulation experiment are consistent with the theoretical conclusions. However, due to the limitation of the transmitted beam width, the length of the virtual synthetic aperture will not increase indefinitely.

3. Motion Compensation Based on Real-Time Acoustic Calibration System

3.1. Effect of Motion Errors

Down-looking array imaging sonar is typically mounted on underwater Remote Operated Vehicles (ROVs), Autonomous Underwater Vehicles (AUVs), or fixed to the bottom of a ship. Due to the influence of dynamic factors such as wind and waves, as well as the limited motion control capabilities of the carrier, the carrier's movement will inevitably deviate from the set trajectory, making it challenging to obtain the absolute position coordinates of each hydrophone element and the sound source during motion. According to theoretical analysis, motion errors exceeding $1/8$ of a wavelength will significantly affect the quality of 3D-imaging. Using only GPS positioning or inertial navigation often cannot meet the accuracy requirement.

The carrier motion error refers to the discrepancy between the attitude and trajectory of the carrier during motion and the established mathematical model. The main sources of motion error for the sonar carrier can be summarized as follows:

1. Carrier trajectory measurement error: When performing sonar imaging, it is necessary to accurately measure parameters such as the position, direction, and velocity of the sonar carrier. However, due to the precision limitations of various sensing devices or environmental interference, there is often some error in the measurement results;
2. Insufficient carrier attitude stability: The sonar carrier is easily disturbed by natural factors such as waves, wind, and ocean currents during travel, causing changes in its attitude, which, in turn, affects the reception direction and propagation path of the sonar signals;
3. Human operation error: Mistakes made by operators when adjusting the attitude of the sonar carrier or failing to adjust the attitude in time at specified locations can also lead to motion errors.

The motion error of the carrier manifests as a time delay difference when the sonar receives the echo signals, which will affect the phase information of the received echoes and thus affect the quality of the acoustic image. In practice, however, hydrophone elements at different locations will experience a time delay error τ_e due to the carrier's motion error. Therefore, the actual expression of the echo signal received by a single hydrophone element is:

$$s_{re}(t) = \text{rect}\left(\frac{t - \tau - \tau_e}{T_r}\right) \exp\left\{j(2\pi f_c(t - \tau - \tau_e) + \pi K_r(t - \tau - \tau_e)^2)\right\} \quad (8)$$

From the above formula, it can be seen that motion errors will simultaneously affect the envelope and phase shift of the echo signal. Motion compensation actions are required to process the echoes to obtain a relatively ideal acoustic image.

When the sonar carrier is in motion, it typically has six degrees of freedom: three translational degrees of freedom and three rotational degrees of freedom. The translational degrees of freedom include movements along the x-, y-, and z-axes, while the rotational degrees of freedom encompass rotations around the x-, y-, and z-axes. For ease of description, the carrier motion errors introduced by these six degrees of freedom are as follows: swag, surge, heave, pitch, roll, and yaw [28]. A schematic illustration of these six types of motion errors is shown in Figure 4.

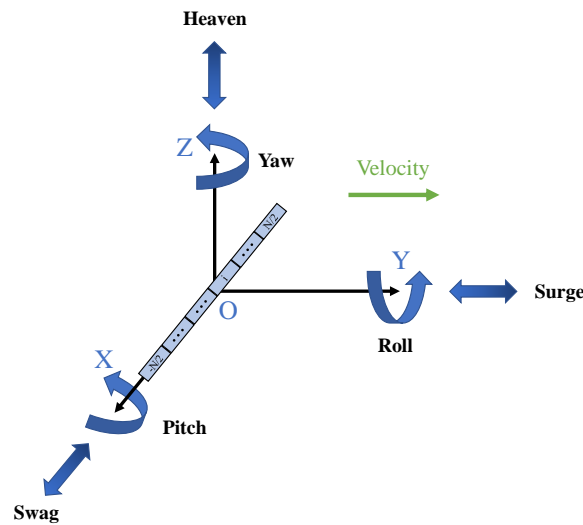


Figure 4. Six types of motion errors.

For the down-looking linear array 3D-imaging sonar, the first three translational motion error components will primarily affect the calculation of the distance from the sound source and the hydrophone elements to the target, thereby affecting the round-trip time delay and phase shift of the echo signals received by the hydrophone, and consequently having a significant impact on the imaging results. The last three rotational motion error components will mainly affect the range of the observation area for 3D sonar, and on the acoustic signal propagation, they will affect the Doppler center frequency of the signal and the phase between the various elements of the receiving array, which, in turn, affects the beamforming results. However, for our downward-looking linear array 3D-imaging sonar system mounted on an Autonomous Underwater Vehicle (AUV), it is required that there should be no significant attitude changes while it is operating. The presence of the three rotational motion error components may result in geometric distortion in the sonar imagery, but the impact on image quality is rather smaller. Therefore, our research focuses on compensating for the first three translational motion error components.

Given that the acoustic array is stationary, the three translational motion errors uniformly influence each hydrophone element. Therefore, in the analysis of translational motion errors, the motion error of the array center-point can be used to represent the entire hydrophone array. Let the position of the array center point at any moment be $(x(t), y(t), z(t))$. When the linear array moves from the coordinate origin along the y-axis at a speed of V_0 , the ideal position of the array center-point can be written as $(0, V_0t, 0)$. For the imaging target point $P(Gx, Gy, Gz)$ in the target space, the ideal distance from the array center to this point is:

$$R_{ideal}(t) = \sqrt{Gx^2 + (V_0t - Gy)^2 + Gz^2} \quad (9)$$

However, in reality, there is translational motion error of the carrier, denoted as $\Delta x = x(t) - 0$, $\Delta y = y(t) - V_0t$, $\Delta z = z(t) - 0$. For the target point P , the actual distance from the array center to this point is:

$$R_{actual}(t) = \sqrt{[x(t) - Gx]^2 + [y(t) - Gy]^2 + [z(t) - Gz]^2} \quad (10)$$

Let $R_0 = \sqrt{(Gx^2 + Gz^2)}$ represent the ideal distance from the target point to the moving trajectory, and let the ideal angle between the line connecting the target point

and the array center and the YOZ plane be fixed as θ . Then, we have $Gx = R_0 \sin \theta$ and $Gz = -R_0 \cos \theta$, and (9) can be written as:

$$R_{ideal}(t) = \sqrt{R_0^2 + (V_0 t - Gy)^2} \approx R_0 + \frac{(V_0 t - Gy)^2}{2R_0} \quad (11)$$

Equation (10) can be written as:

$$R_{actual}(t) = \sqrt{[y(t) - Gy]^2 + R_0^2 - 2x(t)R_0 \sin \theta + 2z(t)R_0 \cos \theta} \quad (12)$$

After derivation, the difference between the actual distance and the ideal distance from the array center to the target point can be expressed as:

$$\Delta R(t) = R_{actual}(t) - R_{ideal}(t) \approx \frac{\Delta y(V_0 t - Gy - \Delta y/2)}{R_0} - \Delta x \sin \theta + \Delta z \cos \theta \quad (13)$$

Analyzing the impact of the three error components of swag, surge, and heave separately: the swag and heave, after projection, will directly affect the distance measurement accuracy and have a significant impact on the imaging results, whereas the surge, compared to the distance from the array center to the target point, is a small account, and its impact on the imaging results is relatively minor. Therefore, the compensation accuracy for the carrier motion errors of swag and heave needs to reach the millimeter level to ensure that the carrier's motion error is controlled at the sub-wavelength level, thereby obtaining high-quality acoustic images.

3.2. Scheme of Motion Compensation

3.2.1. Motion Compensation System

In this section, the acoustic calibration system is first introduced. Figure 5 shows a schematic diagram including the acoustic calibration system. While the hydrophone elements of the down-looking linear array imaging sonar receive the target scattering echo signals, they also receive the calibration signals emitted by the underwater preset calibration sound source.

By reasonably setting the sound emission period and delay of the calibration sound source to match the emission period of the sound source in the sonar array and the target depth, the calibration sound source signal and the target echo signal can fall within the same sound signal acquisition window. This enables the calibration sound source signal to be used for positioning each hydrophone element. It should be noted that the calibration sound source has a limited range of effectiveness. When the calibration sound source is far from the sonar array, the calibration signal received by the hydrophone elements will be submerged in the background noise and cannot be effectively extracted.

There are two different processing methods for the received calibration sound source signals. One is the sonar positioning measurement algorithm based on three-point positioning, which has high computational efficiency but is relatively weak in adaptability to the entire hydrophone system. It can only reflect the position coordinates of the selected three hydrophone elements well, and the method will fail when the data of the selected hydrophone elements for positioning are abnormal. The other is the sonar positioning measurement algorithm based on multi-point positioning, which fully utilizes the calibration signals received by each hydrophone element and positions the sonar system based on the idea of least squares. It has good adaptability to the system but has lower computational efficiency. In engineering applications, it is essential to choose the appropriate method based on the specific circumstances. Both kinds of sonar positioning measurement algorithms will be introduced and compared with each other.

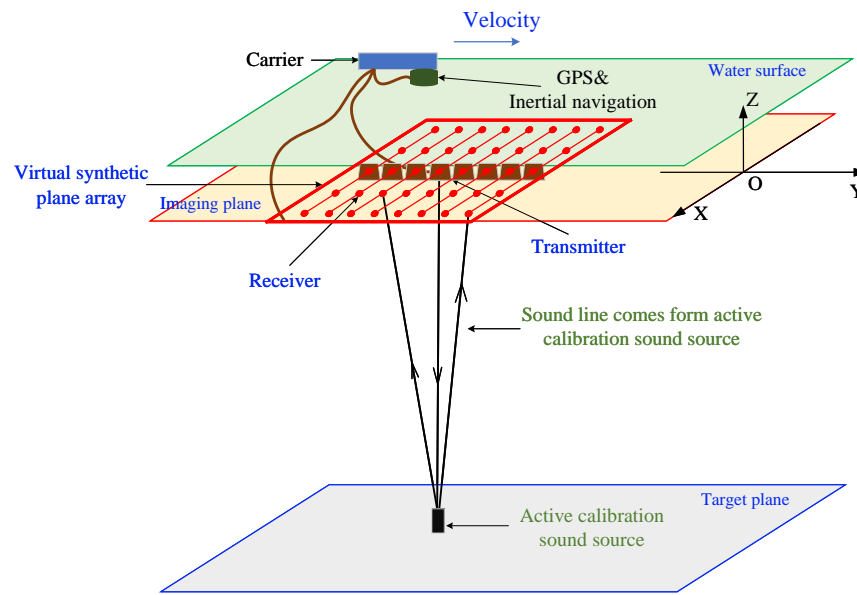


Figure 5. Schematic diagram of motion compensation system.

3.2.2. Motion Compensation Algorithm Based on Three-Point Positioning

The principle of the motion compensation algorithm based on three-point positioning involves selecting three hydrophone elements that are not collinear and using the calibration signals they receive to calculate the distances from these hydrophone elements to the calibration sound source. Through geometric relationships, the relative position coordinates of the calibration sound source can be calculated. Since the absolute position coordinates of the calibration sound source are known, the absolute position coordinates of each hydrophone element and the carrier sound source can be obtained through coordinate transformation, thereby compensating for the sonar echo signals.

Step I: Select the hydrophone elements for positioning

To attain optimal outcomes in acoustic positioning predicated on geometric relationships, it is crucial to select hydrophone elements with favorable spatial characteristics. Generally, the three hydrophone elements used for positioning will be chosen from the hydrophone elements at the ends of the linear array and the two hydrophone elements perpendicular to the linear array.

Step II: Calculate the distance from the selected hydrophone elements to the calibration sound source

During the down-looking linear array sonar imaging process that includes a calibration sound source, the echo signals received by the hydrophone elements consist of two parts: one is the imaging signal containing target information, and the other is the calibration signal containing the position information of the calibration sound source. The echo signals received by the sonar can be expressed in the following form:

$$s_r(n) = s_{imag}(n) + s_{carli}(n) \quad (14)$$

where $s_r(n)$ represents the sampled echo signal, while $s_{imag}(n)$ and $s_{carli}(n)$ denote the sampled imaging signal and calibration signal, respectively.

Perform pulse compression on the calibration signal and select the sampling point corresponding to the maximum amplitude of the calibration signal N_{max} ; the following condition should be met:

$$s_{carli}(N_{max}) = \max\{|s_{carli}(n)|\} \quad (15)$$

The distance from the hydrophone element receiving the calibration signal to the calibration sound source can be calculated using the following formula:

$$R_{HC} = (N_{max} + N_{start} - N_{delay}) \frac{c}{f_s} \quad (16)$$

where N_{start} is the delay point of the channel transmission, and N_{delay} is the delay point of the calibrated sound source.

For the three hydrophone elements used for positioning, the distances to the calibration sound source are denoted as $R_{HC}^{(j,n)}$, $n = 1, 2, 3$.

Step III: Calculate the relative position coordinates of the calibration sound source

Let the relative coordinates of the calibration sound source in the j th frame be denoted as $C(x_C^{(j)}, y_C^{(j)}, z_C^{(j)})$. Based on the distances from the three selected hydrophone elements to the calibration sound source R_{HC} and the relative position coordinates of the selected hydrophones $H_0(x_0^{(i)}, y_0^{(i)}, z_0^{(i)})$, $i = 1, 2, 3$, the following system of equations can be established:

$$\begin{cases} \sqrt{(x_C^{(j)} - x_0^{(1)})^2 + (y_C^{(j)} - y_0^{(1)})^2 + (z_C^{(j)} - z_0^{(1)})^2} = R_{HC}^{(j,1)} \\ \sqrt{(x_C^{(j)} - x_0^{(2)})^2 + (y_C^{(j)} - y_0^{(2)})^2 + (z_C^{(j)} - z_0^{(2)})^2} = R_{HC}^{(j,2)} \\ \sqrt{(x_C^{(j)} - x_0^{(3)})^2 + (y_C^{(j)} - y_0^{(3)})^2 + (z_C^{(j)} - z_0^{(3)})^2} = R_{HC}^{(j,3)} \end{cases} \quad (17)$$

Using mathematical tools, the numerical solution of the aforementioned equations can be calculated. This numerical solution includes two solutions that are symmetrical with respect to the plane. By analyzing the geometric relationship, the solution that corresponds to the desired criteria is selected as the relative position coordinates of the calibration sound source.

Step IV: Solve the absolute coordinates of each hydrophone element and the carrier sound source by coordinate transformation, and calculate the motion error compensation values

Since the operational steps after the calculation of the relative position coordinates of the calibration sound source are the same in both the three-point positioning and multi-point positioning algorithms, these two calculation methods will be described with a specific frame calculation in the motion compensation algorithm based on multi-point positioning.

3.2.3. Motion Compensation Algorithm Based on Multi-Point Positioning

The basic principle of the motion compensation algorithm based on multi-point positioning is similar to that of the ultra-short baseline (USBL) acoustic positioning method. In the USBL acoustic positioning method, the shipborne transceiver emits a pulse interrogation signal to the underwater transponder. After receiving the interrogation signal, the transponder emits a response pulse signal that is different from the interrogation signal. By receiving the response pulse signal through several hydrophone elements spaced a few centimeters apart, the distance and bearing of the transponder are calculated based on the round-trip time and phase difference of the sound wave propagation in water,

thereby calculating the coordinates of the transponder. With the absolute coordinates of the transponder known, the precise coordinates of the carrier can be determined through coordinate transformation.

The down-looking linear array sonar's $(N + 2)$ hydrophone elements are utilized to simultaneously receive imaging signals and calibration signals. Since the calibration sound source emits signals slightly later than the carrier sound source, the calibration signals can be separated from the signals used for imaging during the data processing. At the same time, the echo ranging is performed using the calibration signals received by all of the hydrophone elements, and the relative position coordinates of the underwater calibration sound source are calculated using the least squares method. Similarly to the three-point positioning method, since the absolute coordinates of the calibration sound source have been determined during deployment, the absolute coordinates of each hydrophone element and the carrier sound source can be obtained through coordinate transformation. Then, the motion errors embodied in the echo signal can be calculated.

Step I: Calculate the distance from each hydrophone element to the calibration sound source

Using the same method as in the three-point positioning to calculate the distance from the hydrophone element to the calibration sound source, extract the calibration signal received by each hydrophone element, and select the maximum amplitude value among them. The distance from each hydrophone element to the calibration sound source can be calculated using Equation (16).

Step II: Calculate the relative coordinates of the calibration sound source using the Newton–Raphson iterative method

During the motion imaging process of the down-looking linear array sonar, it is necessary to calculate the relative position coordinates of the calibration sound source for each frame. Let the distance from the $(N + 2)$ hydrophones to the calibration sound source be $R_{HC}^{(j,n)}$, $n = 1, 2, \dots, N + 2$. The relative coordinates of the calibration sound source $C(x_C^{(j)}, y_C^{(j)}, z_C^{(j)})$ in the j th frame can be obtained by solving a nonlinear system of $(N + 2)$ equations. Based on the principle of least squares, the Newton–Raphson iterative method is used to calculate the relative coordinates of the calibration sound source. The algorithm flow is as follows:

(1) Set the initial parameters

Let the initial relative coordinates of the calibration sound source be denoted as $C^{(k)}(x_C^{(j,k)}, y_C^{(j,k)}, z_C^{(j,k)})$, $k = 0$, the maximum number of iterations be N_{break} , and the maximum allowable error be σ_{max} . In order to minimize the number of iterations, the relative coordinates of the calibration sound source should be initialized to approximate the true values as closely as feasible.

(2) Calculate the Jacobian matrix of the system of equations

The i th equation in the system can be represented by the following formula:

$$\sqrt{(x_C^{(j)} - x_H^{(i,j)})^2 + (y_C^{(j)} - y_H^{(i,j)})^2 + (z_C^{(j)} - z_H^{(i,j)})^2} = R_{HC}^{(j,n=i)} \quad (18)$$

This system of nonlinear equations can be written in the following simplified form:

$$\begin{cases} f_1(x_C^{(j)}, y_C^{(j)}, z_C^{(j)}) = 0 \\ f_2(x_C^{(j)}, y_C^{(j)}, z_C^{(j)}) = 0 \\ \vdots \\ f_{N+2}(x_C^{(j)}, y_C^{(j)}, z_C^{(j)}) = 0 \end{cases} \quad (19)$$

Then, the Jacobian matrix of the system of nonlinear equations can be denoted as:

$$F'(C) = \begin{bmatrix} \frac{\partial f_1(X)}{x_c^{(j)}} & \frac{\partial f_1(Y)}{y_c^{(j)}} & \frac{\partial f_1(Z)}{z_c^{(j)}} \\ \frac{\partial f_2(X)}{x_c^{(j)}} & \frac{\partial f_2(Y)}{y_c^{(j)}} & \frac{\partial f_2(Z)}{z_c^{(j)}} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_{N+2}(X)}{x_c^{(j)}} & \frac{\partial f_{N+2}(Y)}{y_c^{(j)}} & \frac{\partial f_{N+2}(Z)}{z_c^{(j)}} \end{bmatrix} \quad (20)$$

- (3) Iteratively solve the relative coordinates of the calibrated sound source

The iterative format of the Newton–Raphson iterative method is:

$$C^{(k+1)} = C^{(k)} - F'(C^{(k)})^{-1}F(C^{(k)}) \quad (21)$$

The termination condition for the iteration is $\|C^{(k+1)} - C^{(k)}\|_{\infty} < \sigma_{\max}$, or the iteration reaches the maximum number N_{break} . At the end of the iteration, the relative coordinates of the calibrated sound source in the j th frame $C(x_C^{(j)}, y_C^{(j)}, z_C^{(j)})$ can be obtained.

Step III: Solve for the absolute coordinates of each hydrophone element and the carrier sound source by coordinate transformation

Once the system deployment is completed, the absolute position coordinates of the calibration sound source are thereby determined. The coordinates in the relative coordinate system can be transformed into the coordinates in the absolute coordinate system through a linear transformation. Let the absolute position coordinates of the calibration sound source be C_{abs} . In the j th frame, the absolute coordinates of each hydrophone element and the carrier sound source can be calculated using the following formula:

$$\begin{cases} H = (C_{abs} - C) + H_0 \\ S = (C_{abs} - C) + S_0 \end{cases} \quad (22)$$

Step IV: Calculate the compensation values for the echo motion error

The absolute coordinates of the aforementioned hydrophone elements and the carrier sound source are obtained through the joint calculation of signals from $(N + 2)$ hydrophone elements, which can be considered as the most ideal results obtained under the condition of limited data. The difference between the calculated distance R^* from the hydrophone element to the calibration sound source and the distance R_{HC} from the hydrophones to the calibration sound source measured by the echo signal is the error value $\varepsilon_{i,j}$ generated in the sonar positioning process. The compensation value for the sonar motion error is the opposite number of the error value.

In the j th frame, the ideal distance R^* from the i th hydrophone element to the calibration sound source can be calculated using the following formula:

$$R^* = \sqrt{(x_H^{(i,j)} - x_C^{(j)})^2 + (y_H^{(i,j)} - y_C^{(j)})^2 + (z_H^{(i,j)} - z_C^{(j)})^2} \quad (23)$$

The error value can be calculated by the following formula:

$$\varepsilon_{i,j} = R^* - R_{HC} \quad (24)$$

The compensation value can be calculated by the following formula:

$$\rho_{i,j} = -\varepsilon_{i,j} \quad (25)$$

In addition, depending on the sound speed c and the sampling rate f_s , the number of echo delay points N_{comp} used to compensate can be obtained. After compensation, the echo signals used for imaging can be expressed as:

$$s_{used}(t) = \text{rect}\left(\frac{t - \tau - \tau_e + \rho/c}{T_r}\right) \exp\left\{j(2\pi f_c(t - \tau - \tau_e + \rho/c) + \pi K_r(t - \tau - \tau_e + \rho/c)^2)\right\} \quad (26)$$

In mathematical terms, the three types of translational motion errors can be separated through approximate assumptions, looking at Equation (13). However, in the actual signal processing process, these three types of motion errors are deeply coupled together, such as Equation (8). Our motion error compensation method does not focus on measuring individual motion error components but compensates for motion errors in the time delay of the echos, ingeniously resolving the issue of the spatial coupling of motion errors, which has been shown in Equation (26). Additionally, the compensation accuracy and the different effects between the motion compensation algorithm based on three-point positioning and multi-point positioning can be found in Figure A2.

4. Experiments and Results

4.1. Validation Test of Simulation Data

This section tests the feasibility of the motion error compensation method based on the acoustic calibration system that we propose, using simulated data. The setting of simulation parameters is all based on the actual experimental system and experimental environment. Therefore, the simulated data can more closely approximate the real imaging environment.

The parameters of the acoustic signal in the simulation are as follows: the pulse length of the linear frequency modulation signal $T_r = 1$ ms, the carrier center frequency $f_c = 85$ kHz, the signal bandwidth $B_r = 30$ kHz, and the frequency range from 70 kHz to 100 kHz. The signal amplitude is 0.8, and the initial phase is 0. A total of frame number $M = 100$ chirp signals are emitted, and the sampling rate of the echo signal $f_s = 1$ MHz. The sound speed in water $c = 1449.93$ m/s.

The parameters related to the transducer array in the simulation are as follows: the receiving array consists of a total of $(N + 2) = 64$ hydrophone elements. The spacing between the N hydrophone elements along the x-axis $D_s = 9.5$ cm, and the distance from the latest two hydrophone elements along the y-axis to the linear array $L_s = 0.5$ m. The velocity of the carrier $V_0 = 0.05$ m/s. The relative coordinate of the sound source is $S_0(0, 0, 0.04)$.

The parameters related to the imaging space are as follows: the imaging space is designed as $A\{x \mid -0.5 \text{ m} \leq x \leq 0.5 \text{ m}; y \mid -0.5 \text{ m} \leq y \leq 0.5 \text{ m}; z \mid -0.1 \text{ m} \leq z \leq 0.1 \text{ m}\}$, and the spacing of the imaging grid points along each axis of the coordinate system $Dx = Dy = Dz = 0.01$ m.

The parameters related to the target are as follows: the interval between the target points is 0.2 m, and the distance from the target plane to the imaging plane is 10 m. The distribution of the simulation targets can be found in Figure 6a. Under ideal conditions, the imaging results of these point targets are shown in Figure 6b.

To verify the correction effect of the motion compensation algorithm, we added the sinusoidal errors with an amplitude of 0.2 m to the ideal trajectory in the along-track, across-track, and depth directions. At the same time, we used the motion error compensation method proposed in our paper to estimate the trajectory and compensate for the motion error. Under real experimental conditions, such a large translational motion error would not occur, but to vividly demonstrate the effect of our motion compensation method, we deliberately magnified the motion error in the simulation. In Figure 7a, the blue dashed line represents the ideal straight trajectory, the red solid line represents the actual trajectory

with errors, and a series of green dots represent the predictions of our proposed motion error compensation method, which is highly consistent with the actual trajectory. Figure 7b shows the difference between the actual motion error and the estimated value. We can observe that our method achieves a compensation accuracy of $-4 \sim 4 \times 10^{-3}$ m for the sway motion error on the x-axis and the heave motion error on the z-axis. Based on the analysis in Section 3.1, the compensation accuracy for the sway and heave components of the motion error is relatively high, requiring precision to the millimeter level. As Figure 7b shows, the motion error compensation method we proposed meets the corresponding accuracy requirements for the compensation of these two motion error components, resulting in high-quality imaging results.

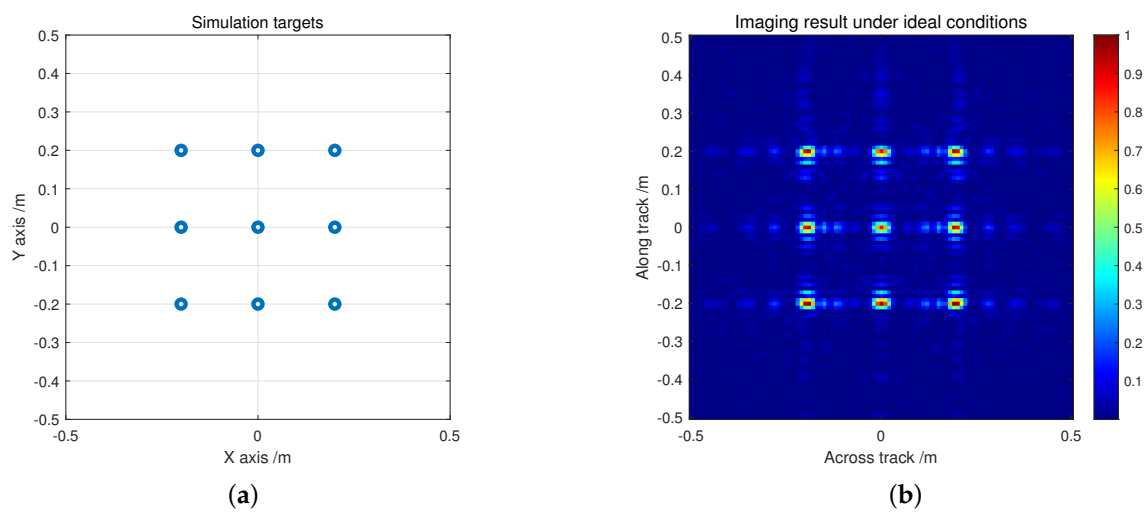


Figure 6. Target and ideal imaging result of simulation data: (a) distribution of the simulation targets; (b) ideal imaging result.

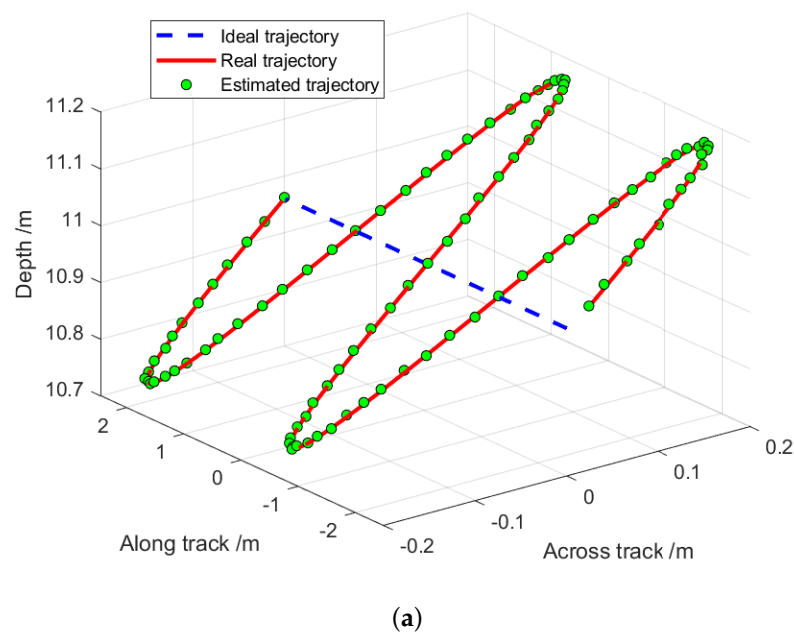


Figure 7. Cont.

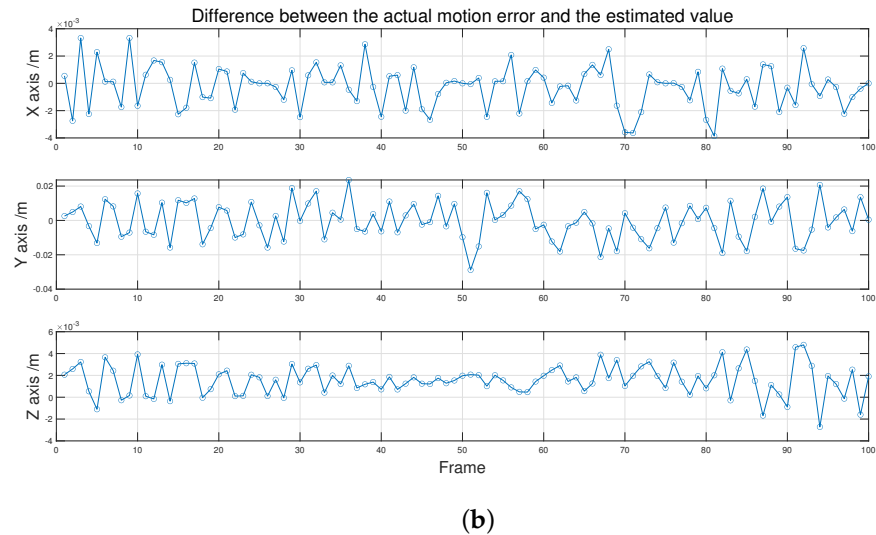


Figure 7. Effect of the motion compensation: (a) trajectory estimation result, (b) the difference between the actual motion error and the estimated value.

Ultimately, we can use the motion-compensated trajectory of the carrier for imaging processing. We compared the visual effects of the targets in the normalized imaging results and the changes in the maximum superimposed sound intensity value before and after motion compensation. Figure 8a shows the normalized imaging result of the point group by down-looking linear array sonar without motion error compensation. The imaging results exhibit severe defocus, making it impossible to discern the nine target points in the sonar image. And the maximum superimposed sound intensity value is 42.68 dB. Figure 8b displays the normalized imaging result after processing with the motion error compensation algorithm. The imaging results indicate that the motion error compensation algorithm effectively addressed the defocus issue, leading to a significant enhancement in the imaging quality of the sonar image from the downward-looking 3D-imaging sonar. At this time, the maximum superimposed sound intensity value in the sound field increases to 63.42 dB.

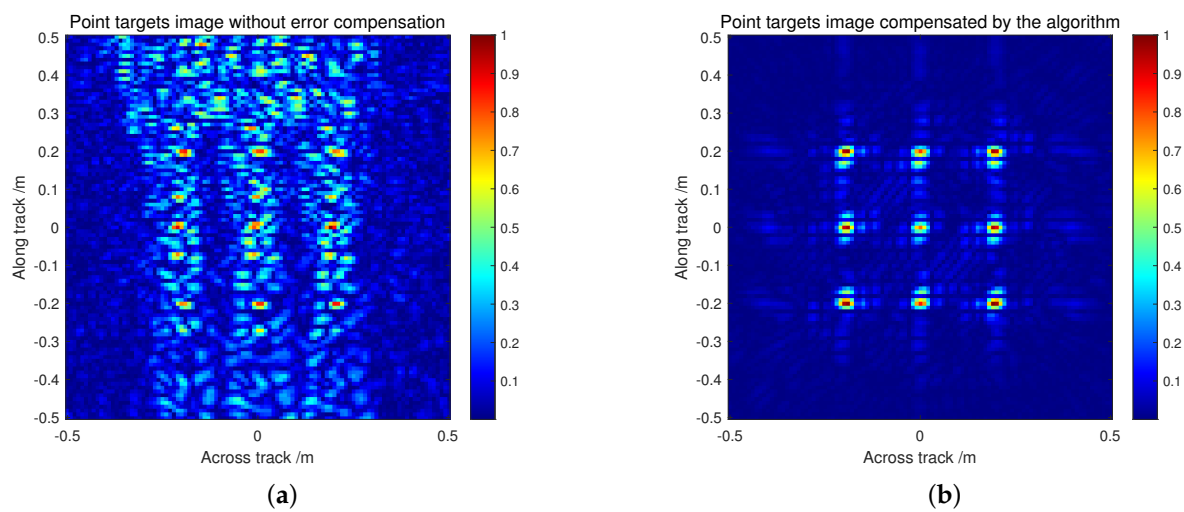


Figure 8. Impact of motion compensation on imaging results: (a) point targets image without error compensation; (b) point targets image compensated by the algorithm.

4.2. Validation Test of Real Data

In this section, we use the real data collected by the downward-looking linear array 3D imaging sonar during the lake trial to verify the motion error compensation method based on the real-time acoustic calibration system. The parameters related to the transducer array and the acoustic signal are the same as those in simulation. However, due to the variability of the underwater environment, in our experiments with real data, we have measured the sound velocity profile of the experimental water area before imaging to correct the sound velocity values used in the imaging process. And the total of frame number $M = 701$. The parameters related to the imaging space are changed: the imaging space is expanded to $A\{x \mid -1 \text{ m} \leq x \leq 1 \text{ m}; y \mid -1 \text{ m} \leq y \leq 1 \text{ m}; z \mid -1 \text{ m} \leq z \leq 1 \text{ m}\}$, and the spacing of the imaging grid points along each axis of the coordinate system $Dx = Dy = Dz = 0.01 \text{ m}$.

The target depth is 17 m, corresponding to the distance from the target plane to the imaging plane. The diameter of the circular point target on the target is 3 cm, and the interval between the point targets is 5 cm, and our emitted sound signal has a central wavelength of 2 cm. This target design is already very close to the limit of the resolution capability of sound waves to showcase the performance of imaging sonar with our motion compensation method. The shape of the target is shown in Figure 9.

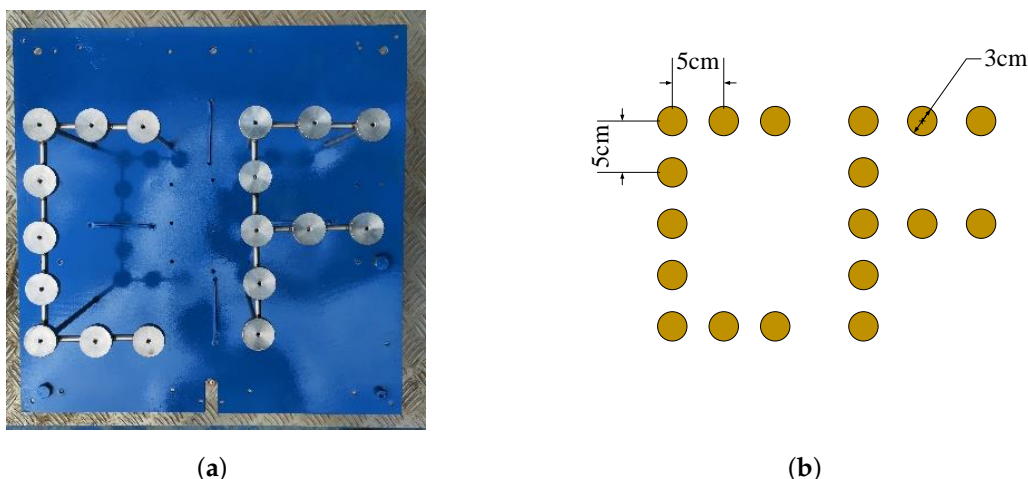


Figure 9. The shape of the target: (a) optical photograph of the target; (b) the dimensions of the target.

The imaging system during the lake trial consists of a transducer array, a carrier, and auxiliary equipment. The transducer array is rigidly connected to the carrier, which is also equipped with other auxiliary measurement devices, such as a GPS positioning system and an inertial navigation system. A motor can drive the carrier to move along with the transducer array. The central control platform can control the auxiliary measurement system to work in conjunction with the transducer array, completing the collection of the sonar echo signal data, GPS positioning data, and inertial navigation data during the movement. For the calibration sound source, as the imaging distance increases, the calibration accuracy tends to decrease. Raising the calibration sound source frequency can improve the positioning accuracy, but the effective range of the signal will also decrease. By increasing the beam opening angle of the calibration sound source, the effective range of positioning can be expanded, but the positioning distance will be reduced. Therefore, key factors such as the calibration sound source frequency, signal strength, and beam opening angle, which constrain the range and distance, need to be comprehensively designed according to specific application requirements. Within a bottom imaging height of 30 meters, the calibration sound source we use fully meets the precision requirements for sub-wavelength motion error compensation.

In the experiments with real data, we compared the visual effects of the targets in the normalized imaging results and the changes in the maximum superimposed sound intensity value between the error coarse compensation scheme based on motion sensors and our proposed motion error compensation method. The comparison of the imaging results is shown in Figure 10. Figure 10a presents the imaging result with error coarse compensation based on motion sensors. Figure 10b presents the imaging result with motion compensation we proposed. We also perform an internal comparative analysis in this imaging progress to demonstrate the BP algorithm's quality and clarity, which can be found in Figure A3. Figure 10c presents the 3D-imaging result after motion compensation.

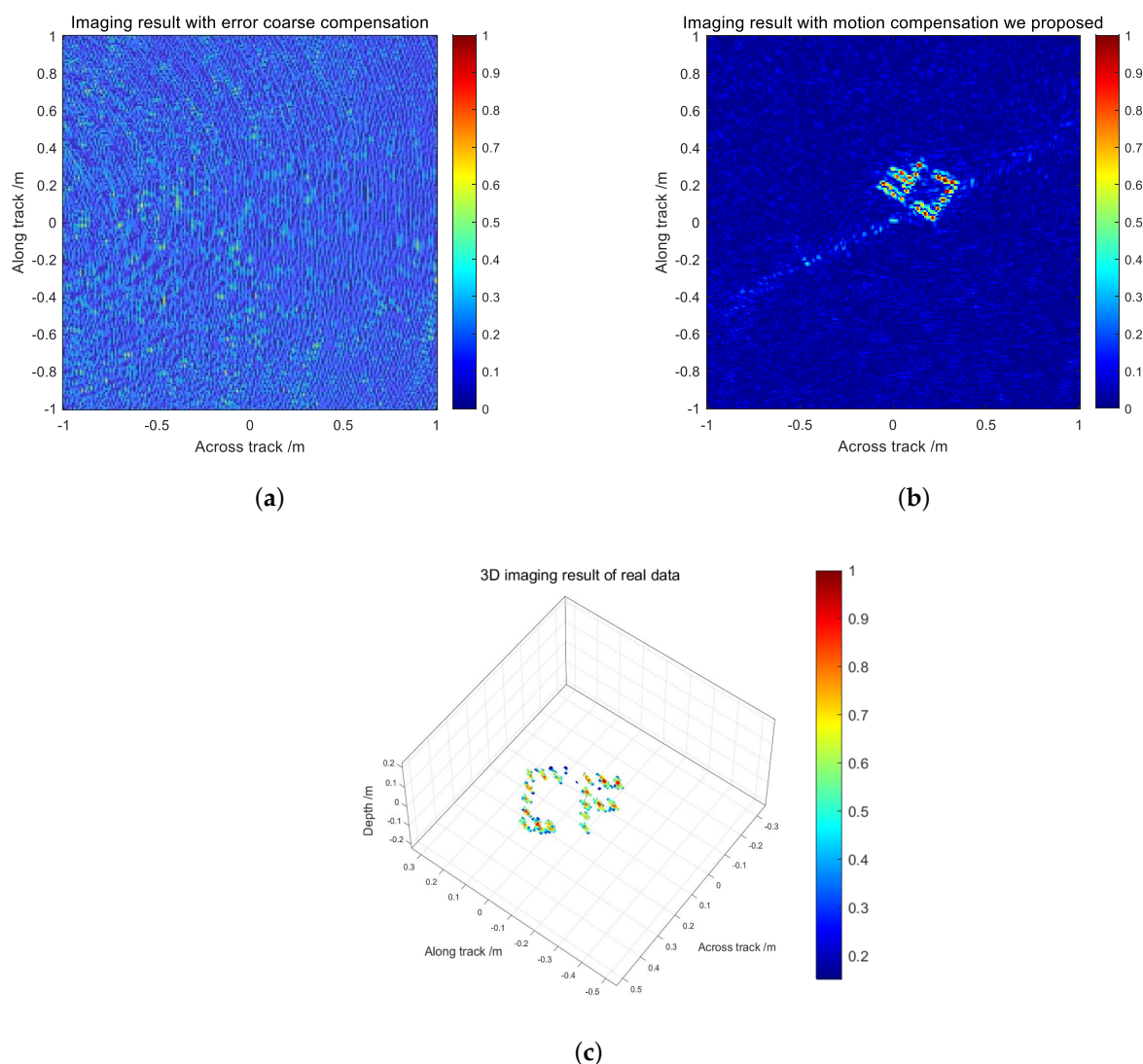


Figure 10. Imaging results of the lake trial: (a) imaging result with error coarse compensation based on motion sensors; (b) imaging result with motion compensation we proposed; (c) 3D-imaging result after motion compensation we proposed.

From Figure 10a, it can be seen that relying solely on the error coarse compensation scheme based on GPS positioning and inertial navigation systems, although the positioning error can be controlled within 2 cm, such positioning accuracy is far from sufficient for acoustic imaging. The acquired normalized imaging result is completely out of focus, and the target objects cannot be observed. Under this condition, the maximum superimposed sound intensity value is 31.43 dB. In contrast, Figure 10b presents the normalized imaging result with motion compensation we proposed. At the target depth of 17 m, the synthetic

aperture length along-track is 5.18 m, and the physical aperture length across-track is 5.78 m, theoretically achieving a resolution of 3 cm. Not only the shape of the target can be clearly seen in the figure, but also the points on the target can be distinguished. And the maximum superimposed sound intensity value increases to 38.02 dB. In the comparison, discernible targets could no longer be identified in Figure 10a, but clear resolution results are presented in Figure 10b. This demonstrates the effectiveness of our proposed motion compensation method compared to the error coarse compensation. For the motion error compensation scheme based on sonar echo data, it is not applicable to the sonar array structure used in this study according to references [21,22].

5. Discussion

According to our knowledge of public publications, there are no scholars or research institutions that have proposed motion error compensation methods based on acoustic calibration systems in underwater 3D acoustic imaging. In experiments, compared to the poor imaging result with error coarse compensation, the resolution of the imaging system after motion compensation we proposed has already reached $3\text{ cm} \times 3\text{ cm} \times 2.5\text{ cm}$ @ Depth = 17 m, TBP = $30\text{ s} \cdot \text{Hz}$. By employing the motion error compensation method proposed in this paper, the error compensation values for the signals received by most hydrophone elements are less than 1.5 mm, equivalent to $1/10\lambda$, demonstrating that the motion compensation accuracy has reached the sub-wavelength level. Furthermore, under the condition of equal average energy, the maximum superimposed sound intensity values in the imaging results increased by 20.75 dB and 6.57 dB, respectively, for simulated and actual data, resulting in better visual imaging results. These experimental data and results demonstrated the effectiveness of our proposed plan. However, it should be noted that although the acoustic calibration system can bring high motion error compensation accuracy, the useful range of a single calibration sound source is limited and cannot meet the needs of large-area imaging. In the future, distributed acoustic calibration systems will be one of the directions for research. Additionally, we must consider the implementation complexity and cost of the scheme based on the requirements of the task. If sub-wavelength level motion error compensation is required within a small range, our proposed motion compensation method not only costs less but also performs better compared to the error coarse compensation scheme based on motion sensors. However, for large-range but low-precision motion error compensation requirements, the error coarse compensation scheme based on motion sensors is sufficient, and at this time, the cost of large-range high-precision acoustic calibration is enormous.

6. Conclusions

In this paper, a moving linear array 3D down-looking imaging model is established, and the back-projection algorithm for point-by-point imaging in the time domain is adopted to achieve clear 3D-imaging of point targets. To address the challenge of high-precision motion error compensation, we proposed a sub-wavelength motion compensation method based on a real-time acoustic calibration system, establishing two motion compensation models of three-point positioning and multi-point positioning. To verify the key role of the motion compensation method in practical hydro-acoustic imaging, we added sinusoidal motion errors to the simulation data, proving that the proposed motion compensation method meets the theoretical accuracy requirements and has achieved good results in simulation experiments. In addition, we compared the imaging results before and after motion compensation using real data. The experimental results show that compared with the measurement method of GPS combined with inertial navigation, after using the motion error compensation method based on the real-time acoustic calibration system, the focusing

quality of the down-looking 3D-imaging was significantly improved. Ultimately, our system is capable of presenting high-resolution 3D target images with fine details.

Author Contributions: Conceptualization, J.W. and P.L.; methodology, J.W. and J.S.; software, W.T.; validation, J.W., P.L. and W.T.; formal analysis, P.L.; investigation, P.X.; resources, Y.H.; data curation, W.T.; writing—original draft preparation, P.L.; writing—review and editing, J.W.; visualization, K.L. and R.R.; supervision, J.S.; project administration, J.W. and W.T.; funding acquisition, W.T. and P.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant No. 62401601; the Science Foundation of National University of Defense Technology under Grant No. ZK21-30.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

As a supplement to Section 2.3, these are the imaging results for a single point target in the simulation with 20 frames and 100 frames.

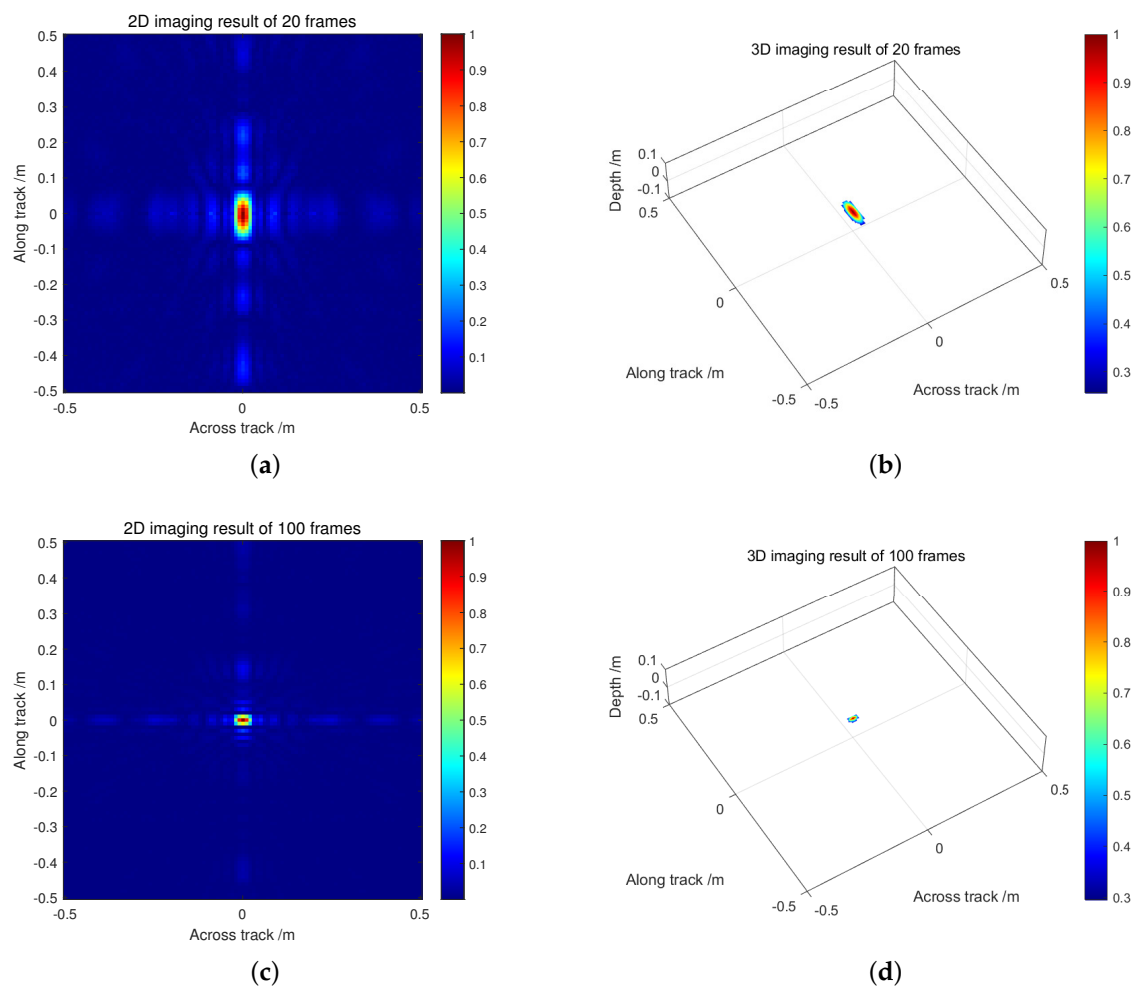


Figure A1. Imaging results of different frame number in simulation: (a) 2D-imaging result of 20 frames; (b) 3D-imaging result of 20 frames; (c) 2D-imaging result of 100 frames; (d) 3D-imaging result of 100 frames.

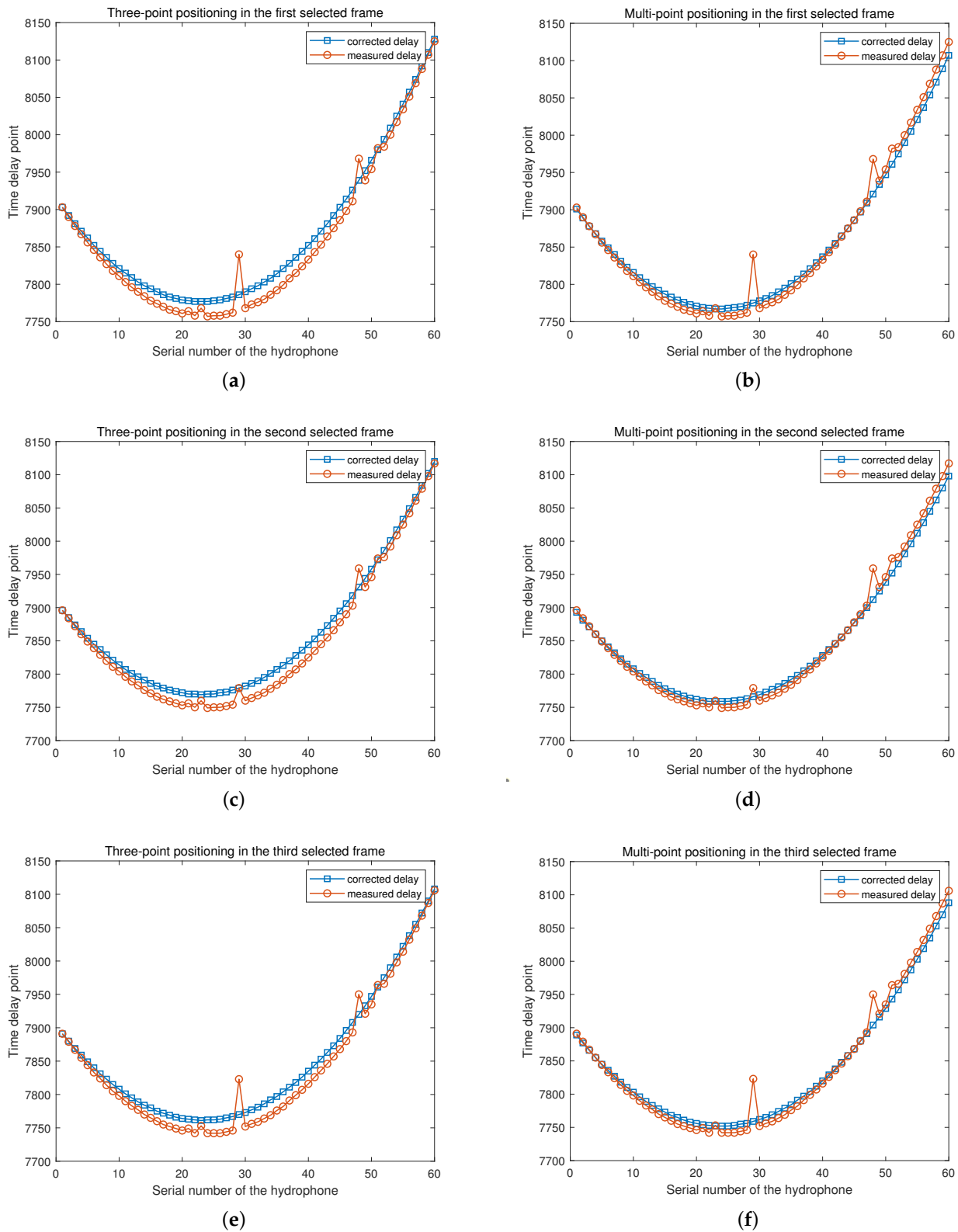


Figure A2. Comparison of the effects of three-point positioning and multi-point positioning: (a) three-point positioning in the first selected frame; (b) multi-point positioning in the first selected frame; (c) three-point positioning in the second selected frame; (d) multi-point positioning in the second selected frame; (e) three-point positioning in the third selected frame; (f) multi-point positioning in the third selected frame.

To illustrate the differences between three-point positioning and multi-point positioning, we randomly selected three consecutive frames from a segment of data and plotted their time delay point of each hydrophone in Figure A2. It can be seen from the figure that, since three-point positioning uses the hydrophones at the ends of the linear array for positioning, the time delay correction results fit well at both ends, with increased error at the center of the linear array. In contrast, multi-point positioning, which fully utilizes the echo signals received by all hydrophone elements, results in a corrected time delay that matches the measured time delay more closely. Additionally, the motion compensation accuracy of both positioning algorithms has reached the sub-wavelength level in experiments, which can be seen from the correction of the time delay points.

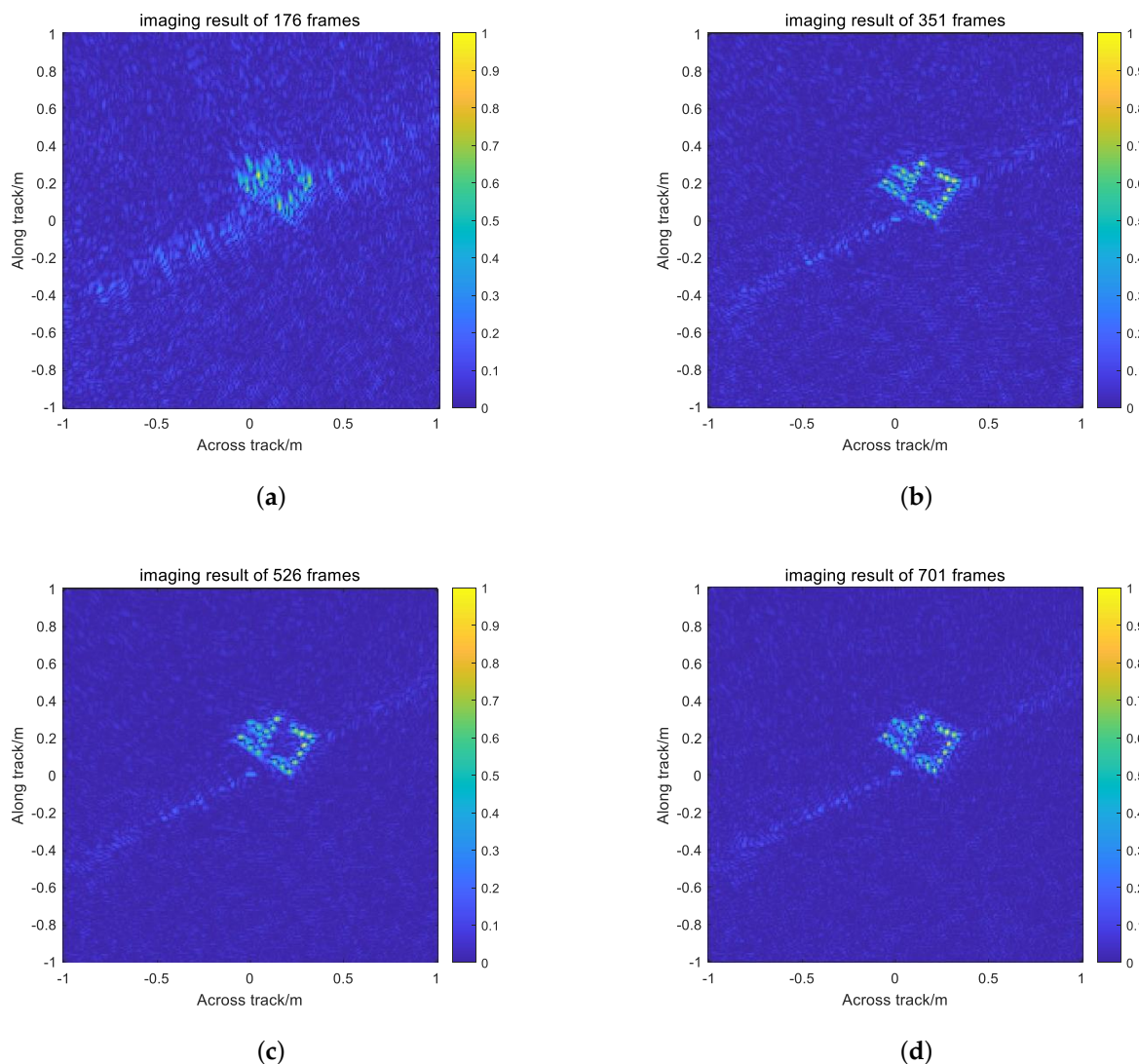


Figure A3. Reconstruction of different frame number: (a) imaging result of 176 frames; (b) imaging result of 351 frames; (c) imaging result of 526 frames; (d) imaging result of 701 frames.

In Figure A3, we performed an internal comparative analysis in the imaging progress of Figure 10b to demonstrate the BP algorithm's quality and clarity. We obtained a series of imaging results through iterative reconstruction of data from different frames, corresponding to different virtual synthetic apertures. In Figure A3a, the shape of the target cannot be discerned, and the maximum sound intensity value is 27.62 dB. As shown in Figure A3b, when the number of frames is increased to 351, the reconstructed result allows the shape of the target to be identified, and the noise points around the target are significantly reduced,

with the maximum sound intensity value being 32.63 dB. In Figure A3c, the target becomes even clearer, and the surrounding noise points are further reduced, with a maximum sound intensity value of 37.31 dB. Compared to Figure A3c, Figure A3d shows almost no visual difference, but the maximum sound intensity value has increased to 38.02 dB. With the increase in the number of frames used, the visual effects of the iterative reconstruction results are better, and the maximum sound intensity value is also increasing, highlighting the quality and clarity of the BP algorithm. However, as can be seen, continuing to increase the number of data frames has a limited effect on the enhancement of the imaging results. Therefore, we chose the imaging result of 701 frames as the final result presented in Figure 10b.

References

1. Li, H.; Wei, B.; Du, W. Technical Progress in Research of Multibeam Synthetic Aperture Sonar. *Acta Geod. Cartogr. Sin.* **2017**, *46*, 1760–1769. [\[CrossRef\]](#)
2. Chu, X.; Fu, Z.; Yu, S.; Tu, X.; Huang, Y.; Ding, X. Underwater Image Enhancement and Super-Resolution Using Implicit Neural Networks. In Proceedings of the 2023 IEEE International Conference on Image Processing (ICIP), Kuala Lumpur, Malaysia, 8–11 October 2023; pp. 1295–1299. [\[CrossRef\]](#)
3. Zhu, J.; Yin, T.; Guo, W.; Zhang, B.; Zhou, Z. An Underwater Target Azimuth Trajectory Enhancement Approach in BTR. *Appl. Acoust.* **2025**, *230*, 110373. [\[CrossRef\]](#)
4. Menakath, M.M.; Raveendranatha Panicker, M.; Hareesh, G. Exploring time domain beamforming algorithm for underwater 3D acoustic imaging. In Proceedings of the OCEANS 2022—Chennai, Chennai, India, 21–24 February 2022; pp. 1–6. [\[CrossRef\]](#)
5. Qin, L.; Huang, H.; Wang, P.; Pengfei, Z.; Liu, J. Broadband downward-looking multi-beam 3D imaging algorithm in frequency domain. *Acta Acust.* **2019**, *44*, 604–612. [\[CrossRef\]](#)
6. Wang, P.; Chi, C.; Liu, J.; Huang, H. Improving performance of three-dimensional imaging sonars through deconvolution. *Appl. Acoust.* **2021**, *175*, 107812. [\[CrossRef\]](#)
7. Chen, J.; Xiong, R.; Yu, H.; Xu, G.; Xing, M. Nonparametric Full-Aperture Autofocus Imaging for Microwave Photonic SAR. *IEEE Trans. Geosci. Remote. Sens.* **2024**, *62*, 5214815. [\[CrossRef\]](#)
8. Chen, J.; Li, M.; Yu, H.; Xing, M. Full-Aperture Processing of Airborne Microwave Photonic SAR Raw Data. *IEEE Trans. Geosci. Remote. Sens.* **2023**, *61*, 5218812. [\[CrossRef\]](#)
9. Stoica, P.; Wang, Z.; Li, J. Robust Capon beamforming. *IEEE Signal Process. Lett.* **2003**, *10*, 172–175. [\[CrossRef\]](#)
10. Liao, Z.; Duan, K.; He, J.; Qiu, Z.; Li, B. Robust Adaptive Beamforming Based on a Convolutional Neural Network. *Electronics* **2023**, *12*, 2751. [\[CrossRef\]](#)
11. Li, C.X.; Guo, M.F.; Zhao, H.F. An Iterative Deconvolution-Time Reversal Method with Noise Reduction, a High Resolution and Sidelobe Suppression for Active Sonar in Shallow Water Environments. *Sensors* **2020**, *20*, 2844. [\[CrossRef\]](#)
12. Yazici, B.; Xie, G. Wideband Extended Range-Doppler Imaging and Waveform Design in the Presence of Clutter and Noise. *IEEE Trans. Inf. Theory* **2006**, *52*, 4563–4580. [\[CrossRef\]](#)
13. Yu, L.; Ma, F.; Lim, E.; Cheng, E.; White, L.B. Rational-Orthogonal-Wavelet-Based Active Sonar Pulse and Detector Design. *IEEE J. Ocean. Eng.* **2019**, *44*, 167–178. [\[CrossRef\]](#)
14. Myers, V.; Quidu, I.; Zerr, B.; Sæbø, T.O.; Hansen, R.E. Synthetic Aperture Sonar Track Registration With Motion Compensation for Coherent Change Detection. *IEEE J. Ocean. Eng.* **2020**, *45*, 1045–1062. [\[CrossRef\]](#)
15. Zhu, J.; Xie, Z.; Jiang, N.; Song, Y.; Han, S.; Liu, W.; Huang, X. Delay-Doppler Map Shaping through Oversampled Complementary Sets for High-Speed Target Detection. *Remote Sens.* **2024**, *16*, 2898. [\[CrossRef\]](#)
16. Chi, C. Basic Theory for Underwater Real-Time 3-D Acoustical Imaging. In *Underwater Real-Time 3D Acoustical Imaging: Theory, Algorithm and System Design*; Springer: Singapore, 2019; pp. 11–19. [\[CrossRef\]](#)
17. Wu, M.; Yan, S. Motion Compensation for OFDM Inverse Synthetic Aperture Sonar Imaging Based on Compressed Sensing. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–13. [\[CrossRef\]](#)
18. Wu, H.; Zhou, F.; Xie, Z.; Tang, J.; Zhong, H.; Zhang, J. Two-Dimensional Space-Variant Motion Compensation Algorithm for Multi-Hydrophone Synthetic Aperture Sonar Based on Sub-Beam Compensation. *Remote Sens.* **2024**, *16*, 2144. [\[CrossRef\]](#)
19. Zhang, Y.; Li, G.; Zhang, P.; Wei, L.; Liu, J. Motion Compensation Algorithm of Synthetic Aperture Sonar Based on Multisensor Data Fusion. *J. Beijing Univ. Posts Telecommun.* **2017**, *40*, 82–86. [\[CrossRef\]](#)
20. Yang, F.; Lu, X.; Li, J.; Guo, J. Correction of Imperfect Alignment of MRU for Multibeam Bathymetry Data. *Geomat. Inf. Sci. Wuhan Univ.* **2010**, *35*, 816–820.
21. Bellettini, A.; Pinto, M. Theoretical accuracy of synthetic aperture sonar micronavigation using a displaced phase-center antenna. *IEEE J. Ocean. Eng.* **2002**, *27*, 780–789. [\[CrossRef\]](#)

22. Pinto, M.A. Long Term Accuracy of Synthetic Aperture Sonar Micronavigation using a Displaced Phase Centre Antenna. In Proceedings of the OCEANS 2018 MTS/IEEE Charleston, Charleston, SC, USA, 22–25 October 2018; pp. 1–6. [\[CrossRef\]](#)
23. Liu, X.; Zhou, F.; Zhou, H.; Tian, X.; Jiang, R.; Chen, Y. A Low-Complexity Real-Time 3-D Sonar Imaging System with a Cross Array. *IEEE J. Ocean. Eng.* **2016**, *41*, 262–273. [\[CrossRef\]](#)
24. Zhang, X.; Yang, P.; Cao, D. Synthetic aperture image enhancement with near-coinciding Nonuniform sampling case. *Comput. Electr. Eng.* **2024**, *120*, 109818. [\[CrossRef\]](#)
25. Chen, P.; Tian, X.; Chen, Y. Optimization of the Digital Near-Field Beamforming for Underwater 3-D Sonar Imaging System. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 415–424. [\[CrossRef\]](#)
26. Yegulalp, A. Fast backprojection algorithm for synthetic aperture radar. In Proceedings of the Proceedings of the 1999 IEEE Radar Conference. Radar into the Next Millennium (Cat. No. 99CH36249), Waltham, MA, USA, 22 April 1999; pp. 60–65. [\[CrossRef\]](#)
27. Guo, Y.; Suo, Z.; Jiang, P.; Li, H. A Fast Back-Projection SAR Imaging Algorithm Based on Wavenumber Spectrum Fusion for High Maneuvering Platforms. *Remote Sens.* **2021**, *13*, 1649. [\[CrossRef\]](#)
28. Zhang, J.; Cheng, G.; Tang, J.; Wu, H.; Tian, Z. A Subaperture Motion Compensation Algorithm for Wide-Beam, Multiple-Receiver SAS Systems. *J. Mar. Sci. Eng.* **2023**, *11*, 1627. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.