Automatic Overlapping Area Determination and Segmentation for Multiple Side Scan Sonar Images Mosaic

Xiaodong Shang, Jianhu Zhao, and Hongmei Zhang

***Abstract*—Seafloor image is important for marine scientific research and ocean engineering. With the wide use of side scan sonar (SSS), multiple measured SSS images need to be mosaicked to form a large-scale and high-resolution seafloor image. This article proposes an automatic mosaic method. First, the overlapping areas between adjacent strips are determined automatically by the track lines and swath width. In the overlapping areas, the image matching is conducted using the speeded-up robust features algorithm with the constraint of geographic coordinates of detected feature points. Then, the overlapping area is segmented using the k-means cluster method. Within each segmented region, the coordinate transformation model is established to make the positions of common features in the overlapping areas unique. With regards to multiple SSS images mosaic, the method to mosaic in sequence is also proposed according to the permutation parameter of each strip. Finally, the SSS images in the overlapping areas are fused and a large-scale seafloor image is formed. This proposed method was applied in one water area of Shenzhen, demonstrating good performance in terms of coordinate consistency of the same features in the mosaicked image. The mean coordinate deviations of the same feature points in overlapping areas were nearly zero and the standard deviation of them also decreased. This proposed method is easily transferable to other study areas and provides an objective, repeatable means for multiple SSS images mosaic.**

***Index Terms*—K-means, mosaic, side scan sonar (SSS) image, speeded-up robust features (SURF), swath width, track line.**

I. Introduction

SEAFLOOR images reveal submarine substrate distributions and underwater features, which are of significance to marine scientific research and ocean engineering, such as habitat mapping [1], design of offshore structures [2], underwater navigation [3], pipeline inspection [4], and subsea engineering [5]. Side scan sonar (SSS) towed behind a survey vessel emits a wide-angle beam and receives backscatter data at fixed time intervals to form a high-resolution seafloor image [6]. The surveying works along a track line each time with one strip image as output. Since the width of a single SSS image is relatively smaller than the length of the track line, multiple parallel surveying lines are designed for an acquisition of a large-scale seafloor map. To ensure the data quality and the full surveying of the seafloor, the proper acquisition of the seafloor area located below the sonar requires a redundant coverage of at least 40% for image matching [7]. Therefore, the generation of a large-scale seafloor image is through processing several parallel SSS images, respectively, and then mosaicking them together with the help of redundant coverage.

Most SSS image mosaic studies tend to apply either geocoding method [8], [9] or feature-based method [6]. The geocoding method constructs a large-scale seafloor map in the defined geographic frame using multiple strips based on image coordinates. In the overlapping areas, the average pixel values of adjacent strips are used [9]. This method is efficient, but the performance of the mosaicked image depends on the position accuracy of each strip. When ultrashort base line (USBL) or short base line (SBL) is not used in the surveying process because of the high cost, the towing operation, potential currents dragging and flat bottom assumption and other factors induce errors in the positions of the SSS image [10]–[12]. Even with the USBL or SBL, the complicated seafloor topography still makes the flat bottom assumption invalid and the position accuracy of the SSS image low [10]. Thus, if the SSS images are mosaicked according to position coordinates, obvious artifacts (e.g., target ghosting, dislocation) appear on the mosaicked image [6]. Compared with the geocoding method, feature-based methods apply image matching for neighboring strips and build a coordinate transformation model based on the detected couple feature points (CFPs) in the overlapping areas. The CFPs represent the detected FPs for the same target in the overlapping area of adjacent SSS images. In the transformation model, one SSS image works as the reference and the others are aligned to it. Consequently, the common targets are assigned with the same position coordinates and artifacts are likely to disappear.

Feature-based image mosaicking method has been a research hot point of remote sensing field for decades [13]–[17], including using speeded-up robust features (SURF) algorithm [18], [19]. Li *et al*. studied some import remote sensing image mosaicking works and discussed the current achievements and the future challenges [20]. These methods used in remote sensing image

Manuscript received November 22, 2020; revised January 28, 2021 and February 17, 2021; accepted February 18, 2021. Date of publication February 24, 2021; date of current version March 17, 2021. This work was supported in part by National Natural Science Foundation of China under Grant 41576107 and Grant 41376109, in part by National Key R&D Program of China under Grant 2016YFB0501703, in part by China Scholarship Council under Grant 201906270174, and in part by Research and application of subsea multi-source survey data processing and fusion technology for offshore wind farm under Grant KY2017-02-75. *(Corresponding author: Jianhu Zhao.)*

Xiaodong Shang and Jianhu Zhao are with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China (e-mail: xiaodongshang@whu.edu.cn; jhzhao@sgg.whu.edu.cn).

Hongmei Zhang is with the School of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, China (e-mail: hmzhang@whu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3061747

mosaicking filed can also be applied to SSS images mosaic according to the SSS data characteristics.

In application of the feature-based method to mosaic SSS images, the first step is to detect CFPs followed by image matching. Daniel *et al*. made use of the highlighted area and shadow zone in SSS images to detect FPs, which only found large features in the image [21]. Chailloux *et al*. were the first to combine the correlation ratio and mutual information (CR & MI) of adjacent image intensity to obtain the common features between two strips. However, their method is affected by the complex noise from the ocean environment and variational gains [7]. What's worse, the method is bound to fail for homogeneous areas with low entropy content [22]. SURF algorithm has proved the effectiveness to detect FPs for SSS images [23]–[26]. It locates key points of high variation and focuses on the spatial distribution of gradient information. Meanwhile, it has the invariant ability in rotation, scale, brightness and contrast [27]. After getting the FPs, some scholars suggested that setting prior position knowledge as a constraint for SSS image matching reduces the mismatching ratio [6], [25]. Since only the FPs in the overlapping areas is of interest, the determination of the overlapping areas between adjacent strips is desirable but scarcely. Using the CFPs, a coordinate transformation model is established. However, due to the water current, wave, vessel velocity, cable length variation and complicated topography, the coordinate deviations of CFPs in different positions are variable and the local image distortions may be different across the whole SSS image. Thus, establishing a single coordinate transformation model may be unpractical for the whole overlapping area. It is prefer to segment the overlapping area into several regions based on the target distribution and image textures [6], [23]. By establishing the coordinate transformation model for each segmented region, the local image distortions are taken into consideration. Then, the utility of the established model corrects the coordinates of sensed SSS images, which are fused with the reference one in the last step. Nonetheless, obvious drawbacks of most segmentation operations are subjective and empirical. Thus, an objective and automatic method of overlapping areas segmentation is in desperate need.

This article proposes a multiple SSS images mosaic method with higher degree of automation and objectivity. It is also an extension of our previous study [6]. In Section II, an automatic SSS image mosaic method is presented to obtain the large-scale seafloor image. In Section III, a detailed description of the experiments and results are presented. In Section IV, we discussed the importance of overlapping area determination and segmentation for SSS images mosaic. This proposed method is also compared with some state-of-the-art methods. Another experiment in Bohai sea is added to further verify our method. Meanwhile, the limitations of the proposed method and the future work are discussed. Finally, main conclusion is drawn in Section V. By proposing a frame for multiple SSS images mosaic, the main contributions of this article are summarized as follows.

1) Based on the track lines and swath width, a method to automatically determine the overlapping area is proposed, replacing manual operation.
2) The improved image matching method increases the correct ratio of initial matched CFPs by adding geographic coordinate constraints to the SURF algorithm.
3) To fix the subjective and empirical problem in manual labeling, k-means cluster method is utilized to divide the overlapping areas into several regions, according to the distribution of FPs and track line points.

II. Materials and Methods

A. Survey Area and Data

The development of this method is validated by the SSS data collected in a water area of Shenzhen as shown in Fig. 1(a), where the water depth ranges from 10 to 15 m. In this experiment, EdgeTech 4125 with the operating frequency of 400 kHz and towed by a cable of 5 m behind the vessel was adopted. Four surveying lines with the lengths of 2015, 2390, 2474, and 1936 m were designed. The swath width was set as 200 m and the overlapping rate was 50%. During the surveying, raw surveying data were recorded in ".xtf" files.

To process these obtained files, the original data were first decoded to form waterfall images ping by ping. Afterward, a comprehensive bottom tracking method which based on the threshold principle, the last peak method, the assumption of seabed continuous variation and the symmetry assumption was used to detect the first backscatter echo of each ping after emitting the acoustic wave [28]. Next, the radiometric correction was utilized to compensate for the propagation and spreading attenuation as well as the beam angle variation [12]. Then, slant-rang correction was employed to project temporal echoes on to the ground, converting slant range to across-track coordinates [29]. Finally, by combining vessel-mounted GPS positioning results, cable length and the towfish heading, the SSS image pixel coordinates were reckoned [29] and the geocoded SSS images were obtained, as shown in Fig. 1(b)–(e). Due to the towing operation, the flat bottom assumption, towing distance and bearing controlled by the towing speed and the currents, the accuracies of geographic coordinates of SSS image are usually low. Thus, using the geocoding method to mosaic SSS images together may cause artifacts, for example, one target may be presented in two different positions.

B. Method

1) Overlapping Areas Determination According to the Track Lines and Swath Width: Usually, the track lines during surveying are parallel and the swath width remains constant [7]. Fig. 2 indicates that the overlapping areas between adjacent strips *i*-1 and *i* are determined by two lines $L_{2(\text{i-1})\text{-1}}$ and $L_{2(\text{i-1})}$. The track lines associated with these two strips are $TL_{i\text{-}1}$ and TL_i, which are almost straight and can be described by a function $y_i = a_i x + b_i$ ($i = 2, 3, \ldots, n$). Using the track line points, the *slope* parameter a_i and *intercept* parameter b_{i} are easily computed by least square approximation. As the $L_{2(i\text{-}1)\text{-}1}$ and $L_{2(i\text{-}1)}$ are parallel to the track lines and the distance between them is the swath width D, the equation to describe the edge of overlapping

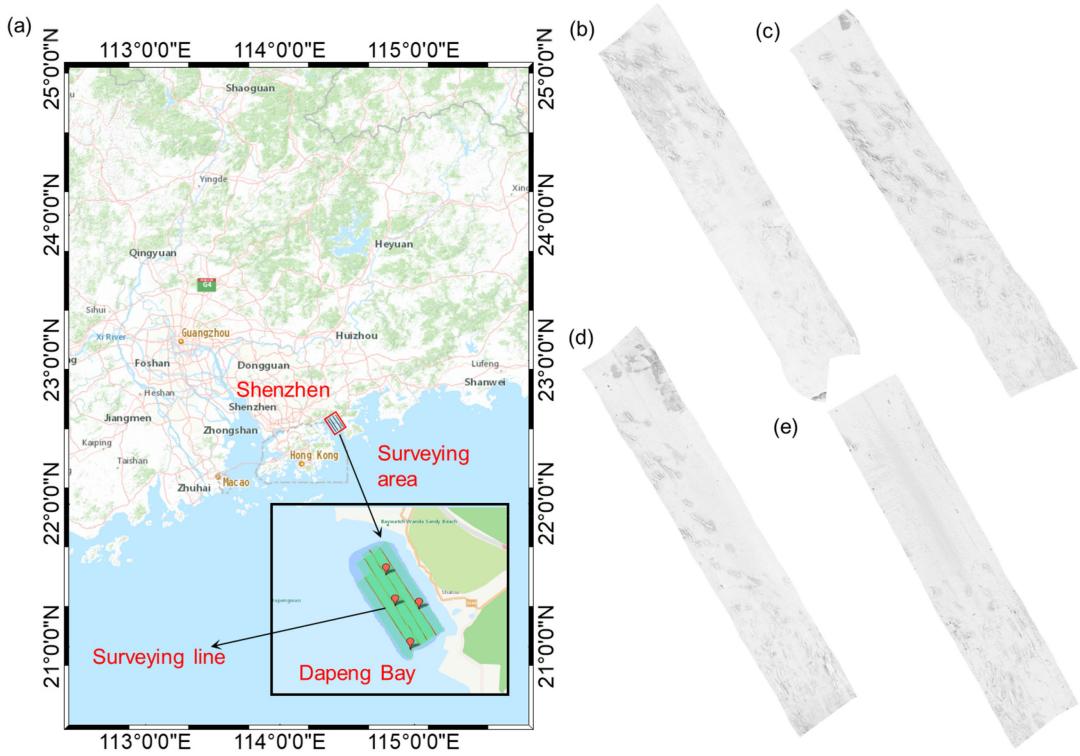

Fig. 1. (a) Surveying area and (b)–(e) Four geocoded SSS images.

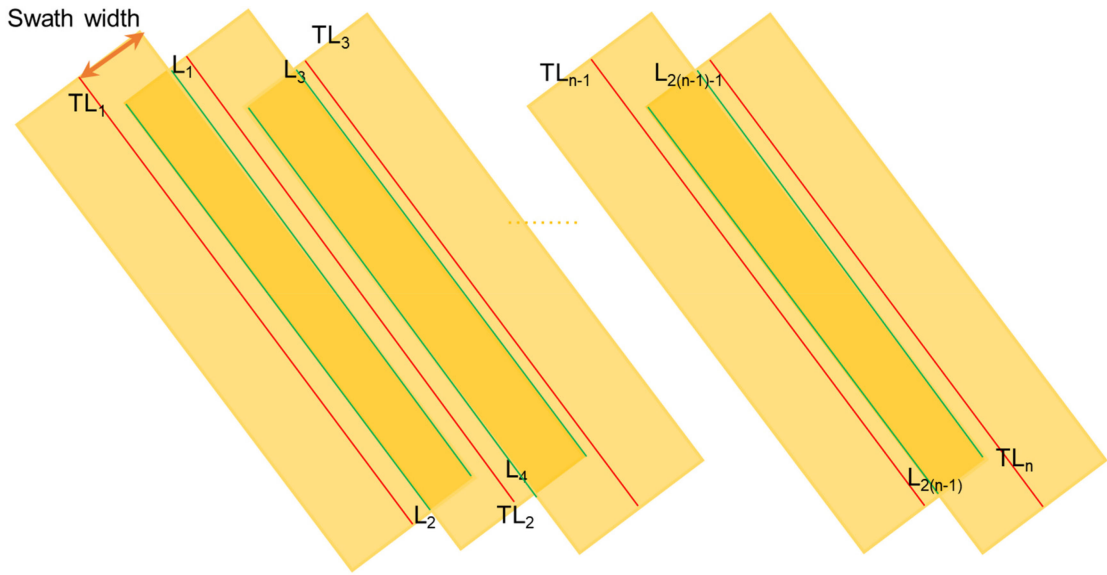

Fig. 2. Overlapping areas determination according to the track lines and swath width. TL is the abbreviation of track line and L_{i} indicates the edge of the i_{th} strip.

area can be set as

$$\begin{aligned} Y_{2(i-1)-1} &= a_{i-1}x + b_{i-1} + D \\ Y_{2(i-1)} &= a_i x + b_i - D. \end{aligned} \tag{1}$$

Based on (1) and Fig. 2, the points (x,y) located in the overlapping area between adjacent two strips i-1 and i should satisfy

$$(x, y) \in \left\{(x, y) | Y_{2(i-1)} < y < Y_{2(i-1)-1}\right\}. \tag{2}$$

2) FPs Detection and Image Matching Within the Overlapping Area Under the Constraint of Image Geographic Coordinates: The SURF algorithm has proved its effectiveness for SSS image feature detection. By using (2), we selected the detected FPs in the overlapping areas. When no FPs are retained, it

indicates that the overlapping areas are homogeneous and no obvious seafloor features exist. In this case, the pixel values of adjacent strips can be averaged directly in these positions. When lots of FPs is detected in the overlapping areas, image matching is usually conducted based on Euclidean distance between the SURF feature vectors of FPs [27]. However, only adopting SURF feature vectors for SSS image matching may cause lots of mismatches [6], [25]. Though there are image coordinate differences between FPs for the same feature target, this kind of difference is within a range. Thus, when the geographic coordinate deviations of the detected FPs in adjacent SSS images are too large, these FPs do not make the correct matches. Based on this, the geographic coordinates of FPs are adopted as a constraint to make the initial image matching more correct. The main steps are as follows.

1) Take strip *i*-1 as reference image and strip *i* as sensed image. For each FP in the reference image, *m* nearest FPs in the sensed image are selected as the candidate matches. This process can be easily achieved by employing *k-d* tree searching method [30] and *m* is set as 20 in this article.
2) Compute the Euclidean distances of feature vectors between the FPs in the reference image and the selected *m* ones in sensed image.
3) Choose the candidate matched point associated with the minimum Euclidean distance between SURF feature vectors as the matched one.
4) Repeat the above process for all FPs in the reference image, the initial image matching result for adjacent strips can be obtained.

Even with the constraint of geographic coordinates, mismatches are practically inevitable. It is necessary to eliminate the incorrect matches through random sample consensus (RANSAC) algorithm to obtain the correct matched CFPs [31].

3) Overlapping Area Segmentation Using K-Means Cluster Method: After getting the correct matches, the following step is to transform the coordinates of FPs in sensed SSS image by referring to these in reference strip. Since the local distortions in different positions of SSS image are variable, it is prefer to segment the overlapping areas into separate regions and establish the corresponding transformation models [6], [23]. To make the segmentation operation automatically and objectively, the k-means cluster algorithm is adopted. Meanwhile, track line points serve as a good constraint to ensure the global stability of the coordinate transformation of FPs [6]. And they are combined with the FPs to establish the coordinate transformation model. The use of k-means cluster algorithm to segment the overlapping areas according to the distribution of FPs and track line points is performed by following steps.

1) Conduct k-means cluster algorithm for the matched FPs in the overlapping area of sensed SSS image and the track line points of adjacent strips. The optimal number of clusters is determined by computing within group sum of squares (WGSS) and the change in gradient of WGSS against number of clusters [32].
2) Choose the matched FPs and corresponding track line points in each cluster to establish a coordinate transformation model, which is described in the following section.
3) Compute the distance between the pixel geographic coordinates of the sensed image and the cluster centers. The transformation model associated with the minimum distance is adopted to transform the pixel coordinates in the sensed SSS image.

4) Local Coordinate Transformation With Constraint of Track Line Points in Segmented Regions: Since the SSS image coordinates are affected by many factors (e.g., currents, flat seabed assumption), the distortions in different positions of the image are variable. The thin plate spline (TPS) function can model complex local distortions without undue effects being propagated throughout the remainder of the image [33]. Thus, it is adopted to establish a coordinate transformation model. Meanwhile, considering both the local distortions between FPs and the global stability in the overlapping area, evenly distributed points in SSS track lines are taken as a constraint when establishing the transformation model [6].

For each segmented region, take coordinates of matched CFPs as (*Xf*, *Yf*) and the selected track line point coordinates of adjacent strips as (*Xc*, *Yc*), the transformation model based on TPS function is shown as

$$\begin{pmatrix} K & Q \\ Q^T & 0_{3^*3} \end{pmatrix} \begin{pmatrix} W & V \\ A & B \end{pmatrix} = \begin{pmatrix} F & G \\ 0_{3^*1} & 0_{3^*1} \end{pmatrix}. \tag{3}$$

where *n* is the sum of number of matched CFPs and track line points.

$$K = \begin{pmatrix} r_{11}^2 \log r_{11}^2 & r_{12}^2 \log r_{12}^2 & \dots & r_{1n}^2 \log r_{1n}^2 \\ r_{21}^2 \log r_{21}^2 & r_{22}^2 \log r_{22}^2 & \dots & r_{2n}^2 \log r_{2n}^2 \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1}^2 \log r_{n1}^2 & r_{n2}^2 \log r_{n2}^2 & \dots & r_{nn}^2 \log r_{nn}^2 \end{pmatrix}$$

$$r_{ij}^2 = [r_i x_j', y_j']^2 = (x_j' - x_i')^2 + (y_j' - y_i')^2$$

$$Q = (\,\mathrm{ones}_{n^*1}\ X_{\mathrm{sen}}\ Y_{\mathrm{sen}}\,)'$$

$$X_{\mathrm{ref}} = (Xf_{\mathrm{ref}}', Xc_{\mathrm{ref}}')' = (x_1, x_2, \dots, x_n)'$$

$$Y_{\mathrm{ref}} = (Yf_{\mathrm{ref}}', Yc_{\mathrm{ref}}')' = (y_1, y_2, \dots, y_n)'$$

$$X_{\mathrm{sen}} = (Xf_{\mathrm{sen}}', Xc_{\mathrm{sen}}')' = (x_1', x_2', \dots, x_n')'$$

$$Y_{\mathrm{sen}} = (Yf_{\mathrm{sen}}', Yc_{\mathrm{sen}}')' = (y_1', y_2', \dots, y_n')'$$

$$W = (w_1, w_2, \cdots, w_n)'\, V = (v_1, v_2, \cdots, v_n)'$$

$$A = (a_1, a_2, a_3)'\, B = (b_1, b_2, \cdots, b_n)'$$

$$F = X_{\mathrm{ref}} - X_{\mathrm{sen}}\, G = Y_{\mathrm{ref}} - Y_{\mathrm{sen}}.$$

To solve (3), we calculate the parameters W, V, A, and B by

$$\begin{pmatrix} W \\ A \end{pmatrix} = \begin{pmatrix} K & Q \\ Q' & 0_{3*3} \end{pmatrix}^{-1} \begin{pmatrix} F \\ 0_{3*1} \end{pmatrix} \\ \begin{pmatrix} V \\ B \end{pmatrix} = \begin{pmatrix} K & Q \\ Q' & 0_{3*3} \end{pmatrix}^{-1} \begin{pmatrix} G \\ 0_{3*1} \end{pmatrix}. \tag{4}$$

Through (3) and the computed parameters, we can acquire the transformation of arbitrary point (x, y) in the overlapping area of sensed image. Establish the transformation model for each segmented region and transform the pixel coordinates of sensed image, the positions of the same targets in different SSS images become constant and the artifacts disappear. By averaging the gray values between reference SSS image and transformed sensed SSS image in overlapping areas and retaining the raw gray values in nonoverlapping areas, the mosaicked SSS image between adjacent strips is achieved. In addition to compute the average value in overlapping areas, many related works have resorted to weighted average, such as the cosine distance weighted blending method [14]. However, when using SSS for seafloor mapping, the emitting intensity and time varying gain parameters were set constant for one surveying task [8], [9], [29], [34]. As a result, for the same area on the seafloor, the recorded SSS data intensity levels remain constant after employing the processing steps including radiation correction and slant-range correction. Consequently, averaging the pixel values in overlapping areas is practical to ensure that the mosaicked image is seamless.

5) Multiple SSS Images Mosaic Strategy: For multiple SSS images, the image mosaic is conducted for adjacent two strips in order. Fig. 2 indicates that the SSS images are permuted according to track lines. The schematic diagram of this automatic and objective multiple SSS images mosaic method is shown in Fig. 3, which includes the following steps.

1) Compute *slope* and *intercept* parameters of each track line. Then, select the x coordinate of arbitrary one point in the track line and calculate the y coordinates using the obtained track line functions. By sorting the computed y coordinates from minimum to maximum as $y_1 < y_2 < \ldots < y_n$, the obtained y_i is named *permutation* parameter.
2) Choose the adjacent SSS images associated with two minimum *permutation* parameters.
3) Conduct FPs detection for the chosen SSS images using SURF algorithm.
4) Compute the overlapping area between the two strips. When no overlapping area exists, the raw data are retained; when there exist overlapping areas, select the detected FPs in the overlapping areas and conduct image matching using SURF algorithm with the constraint of geographic coordinates.
5) Segment the overlapping area into several regions using k-means algorithm. For each segmented region, when no matched FPs are detected, the geocoding method is adopted to get the mosaicked result; when enough FPs are detected, the coordinate model is established. By the established model, the sensed SSS image coordinates are corrected.
6) Conduct image fusing for the sensed SSS image with corrected coordinates and the reference one, the mosaicked result for adjacent SSS images is acquired. In the overlapping areas, the averaging pixel values are used; in nonoverlapping areas, raw pixel values are retained.
7) Mosaic this previous mosaicked one and the following strip with *permutation* parameter y_i according to steps (c)–(f). When no SSS image is left, the mosaicked SSS image is acquired.

C. Assessment

To assess the effectiveness of the proposed method, the direct and subjective method is to check whether there are artifacts in the finally mosaicked image. The objective criteria to validate our method is to conduct the interior checking and exterior checking. The interior checking involves choosing the CFPs from reference and sensed images to establish the coordinate transformation model and computing their coordinate deviations (in east and north directions) as well as statistical parameters (maximum / minimum / mean / standard deviation) before and after the mosaicking process. The exterior checking involves choosing the CFPs not used to establish the transformation model and also computing those statistical parameters.

III. Results

A. Image Matching With the Constraint of Geographic Coordinates

The image matching results using SURF algorithm with and without the constraint of geographic coordinates between two adjacent strips were shown in Fig. 4, which indicated that with the constraint of geographic coordinates, more correct initial matches were obtained [see Fig. 4(a) and (c)]. After conducting the RANSAC algorithm for initial matches, Fig. 4(d) still presented the obvious mismatches. This is because the features that existed on the seafloor surface were relatively simple and the detected edges and FPs could be represented by similar SURF descriptors, which may decrease the distinctiveness of SURF descriptors. If not using the geographic coordinate as a constraint, two distant FPs that were represented by similar SURF descriptors may be treated as matches. As a result, many mismatches appeared. Since the RANSC algorithm is sensitive to high exterior rate, it fails to eliminate the wrong matches when they accounted for a large proportion [35]. On the contrary, using the constraint of geographic coordinates can help avoid this problem because even when the SURF descriptors of FPs were similar, only the FPs that were not far away were treated as matches. Thus, initial image matching performed better and the correct matches were obtained after applying RANSAC [see Fig. 4(b)]. This result proved the importance of using the geographic coordinates as a constraint.

B. Distribution of Coordinate Deviations of Matched CFPs

Compute the coordinate deviations of matched CFPs, the distribution of them was shown in Fig. 5. The starting points of the arrows were located at the positions of the FPs and the length

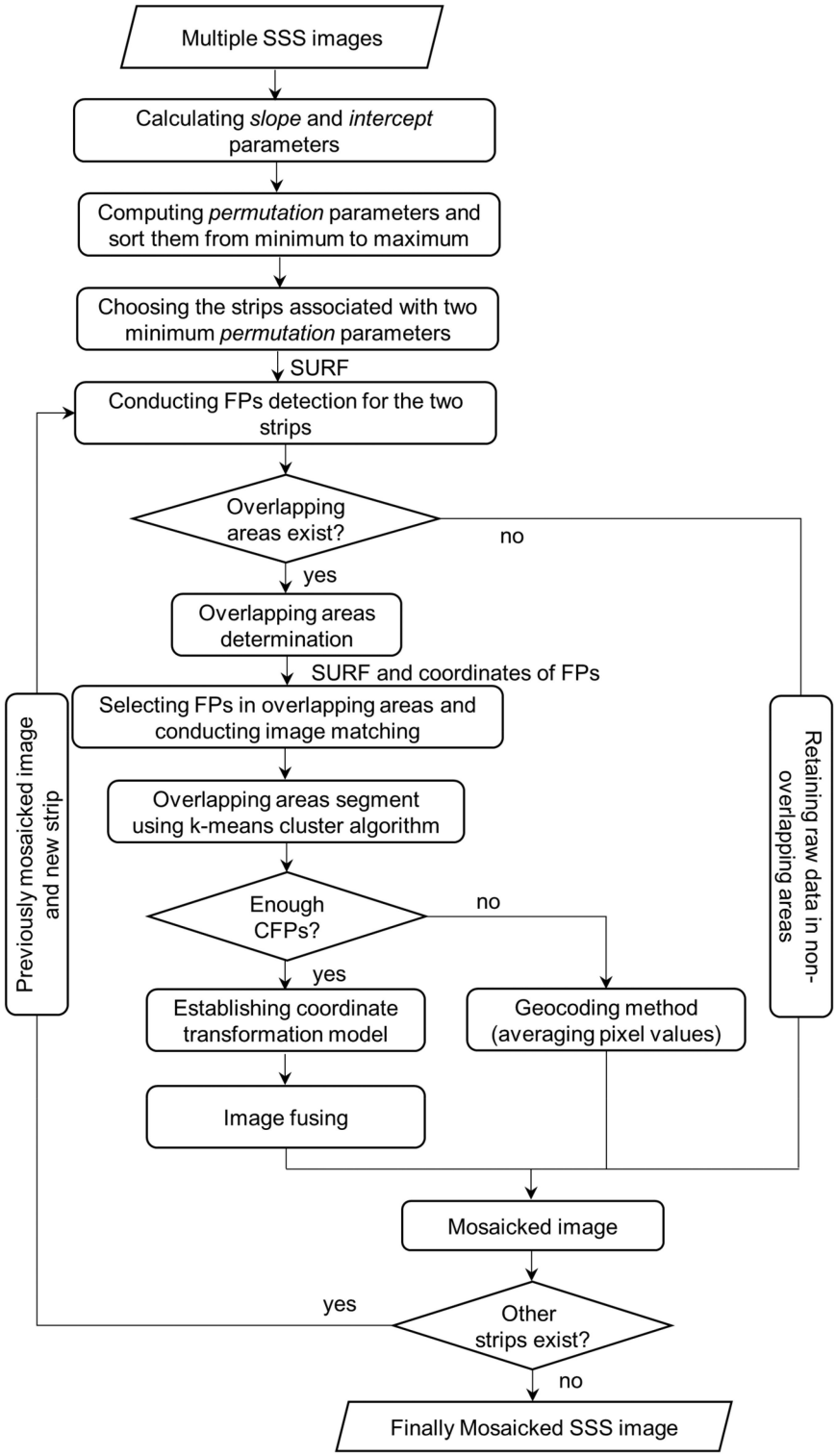

Fig. 3. Schematic diagram of multiple SSS images mosaic method.

of arrows indicated the magnitudes of coordinate deviations. Though all the arrows pointed the southeast direction, some arrows were pointing near the east direction while others pointed almost the south direction. Moreover, the lengths of arrows occupied a wide range. The variation of the directions and lengths of the arrows suggested that the coordinate deviations between matched CFP were variable. When establishing a coordinate transform model, these variations should be considered. Since TPS function could model the local coordinate deviations between matched CFPs without undue effects being propagated throughout the remainder of the image [33], [36], it was adopted to establish the coordinate transform model.

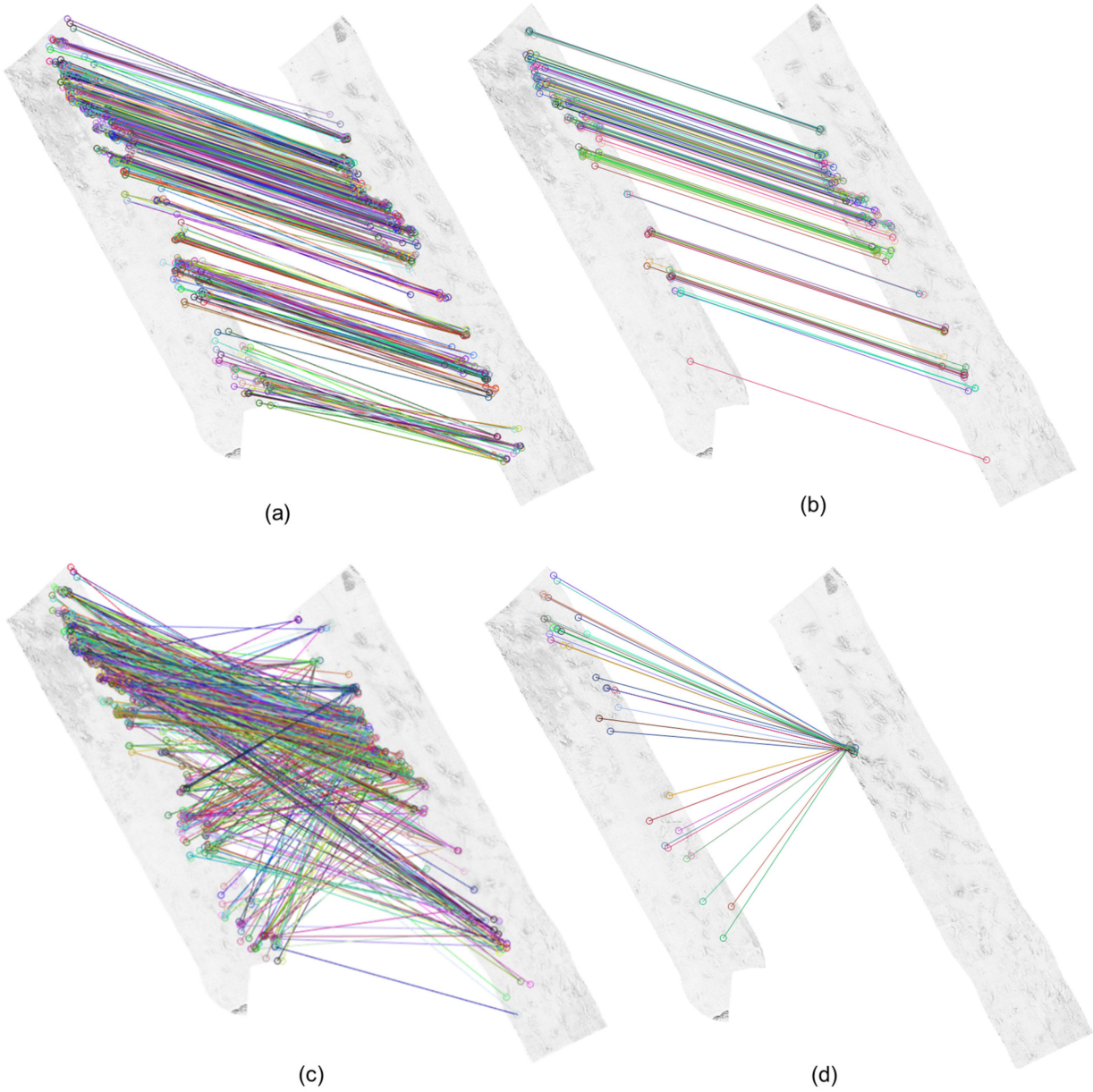

Fig. 4. Initial image matching results (a) with and (c) without constraint of geographic coordinates. (b) and (d) are the matched results after applying RANSAC algorithm for (a) and (c).

C. Overlapping Area Segmentation Using K-Means Algorithm and Adjacent SSS Image Mosaic

Segmentation operation was also conducted for SSS image mosaic when considering the variable coordinate deviations between matched CFPs. Using the k-means algorithm to cluster the FPs and track line points into different classes, the plots of WGSS and the change in gradient of WGSS against cluster numbers were shown in Fig. 6(b). Fig. 6(b) suggested that when the cluster number was larger than four, the plot of WGSS and gradient in WGSS became stable. Thus, the overlapping area was segmented into four regions. The FPs and track lines were assigned into individual regions [see Fig. 6(a)]. For yellow, blue and red regions in Fig. 6(a), abundant FPs existed. Thus, using the FPs and track line points for each region, the corresponding coordinate transformation model could be established according to (3). As no enough FPs existed in the green region, no transformation model was built.

After getting three transformation models and computing the geographical distance between the pixel coordinates of the sensed image in the overlapping area and the four cluster centers, the transformation model associated with the minimum geographical distance was adopted to transform the pixel coordinates of the sensed image in the overlapping areas. When the pixel coordinates of the sensed image in the overlapping area were nearest from the cluster center of green region, no coordinate transformation operation was conducted and the average gray values between adjacent strips were directly used. By

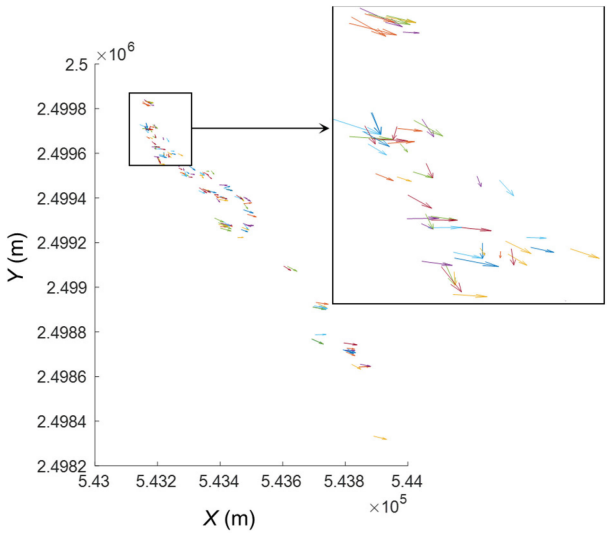

Fig. 5. Distribution of coordinate deviations of matched CFPs. The arrows and the length of them indicate the directions and the magnitude of the coordinate deviations. A zoomed area is shown in black rectangle.

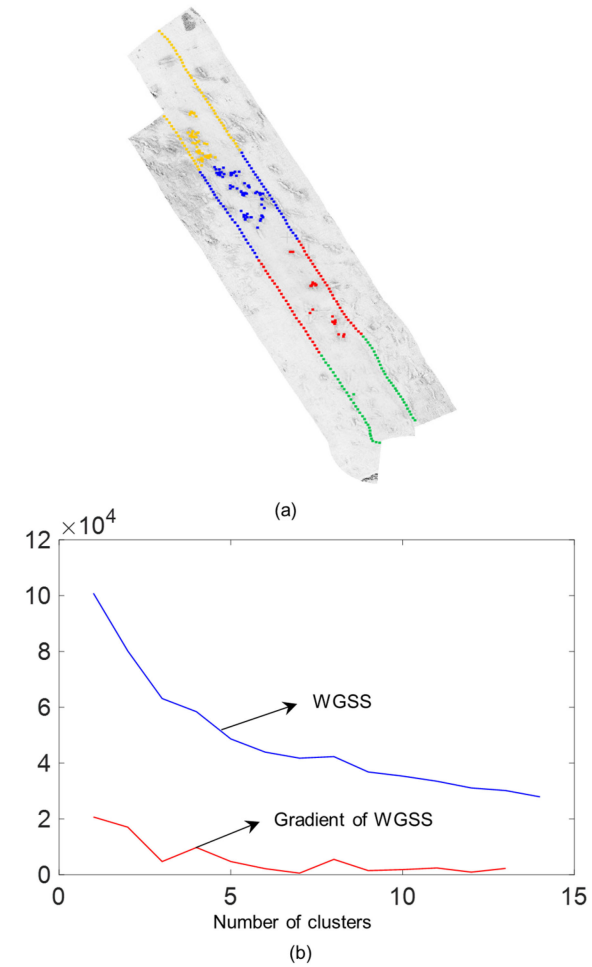

Fig. 6. (a) Segmented result (yellow, blue, red and green) for overlapping area between adjacent strips. (b) Plot of cluster numbers against WGSS and the gradient of WGSS.

computing the average gray values between reference SSS image and transformed sensed SSS image in overlapping areas and retaining the raw data in nonoverlapping areas, the mosaicked SSS image of adjacent strips was obtained [see Fig. 6(a)].

D. Comparison Between Using the Proposed Method and Geocoding Method

To get the large-scale seafloor image using the four strips in Fig. 1, the *permutation* parameters of each strip were calculated (y_1 = 2.4989e+6; y_2 = 2.4993e+6; y_3 = 2.4997e+6; y_4 = 2.5002e+6) and sorted from minimum to maximum. Next, two strips with the minimum two parameters were chosen to form a mosaicked image using the proposed method. Sequentially, the third and fourth strips with the parameters y_3 and y_4 were mosaicked with the former mosaicked one to form the final large-scale seafloor image [see Fig. 7(a)]. This mosaicked image using the proposed method was also compared with the one using the geocoding method [see Fig. 7(b)]. It indicated that using the geocoding method, the artifacts (the same targets were seen in different positions) were presented in the mosaicked image [see Fig. 7(b1), (b2), and (b3)]. While using our proposed method, the same targets were associated with unique locations in the mosaicked image and no artifacts appeared [see Fig. 7(a1), (a2), and (a3)]. This experiment was performed using the hardware with the i7, 3.40 GHz Intel Core and 8.00 GB RAM. To get the mosaicked image, it took 15.47 minutes.

E. Assessment

To further objectively evaluate the effectiveness of the proposed mosaic method, seventy percent of matched CFPs were used for interior checking and the remaining points were used for exterior checking. The histograms and statistical parameters of coordinate deviations of the CFPs were shown in Fig. 8 and Table I. In Table I, dE1 and dE2 were the coordinate deviations of CFPs for interior checking before and after conducting the coordinate transformation in east direction; dN1 and dN2 were those for interior checking in north direction; dE1' and dE2' were those for exterior checking in east direction; dN1' and dN2' were those for exterior checking in north direction. Before applying the transformation model to the overlapping area in sensed image, there were obvious coordinate deviations for matched CFPs in reference and sensed images (see Fig. 8 and Table I). The mean coordinate deviations in terms of east and north directions were 17.37 and -10.74 m for interior checking points and 13.98 and 8.91 m for exterior checking ones. That was why the same target appeared in different positions in the mosaicked image when using geocoding method [see Fig. 7(b)].

While after transforming the FP coordinates of the sensed image, the coordinate deviations of matched CFPs decreased. For the interior checking points, mean coordinate deviations in terms of east and north directions decreased to 0.01 and 0.00 m; standard deviation (STD) decreased from 10.30 and 6.62 to 4.22 m and 3.47 m. With regards to exterior checking points, mean coordinate deviations for east and north directions decreased to 0.81 and 1.73 m; STD decreased from 10.18 and 22.34 to

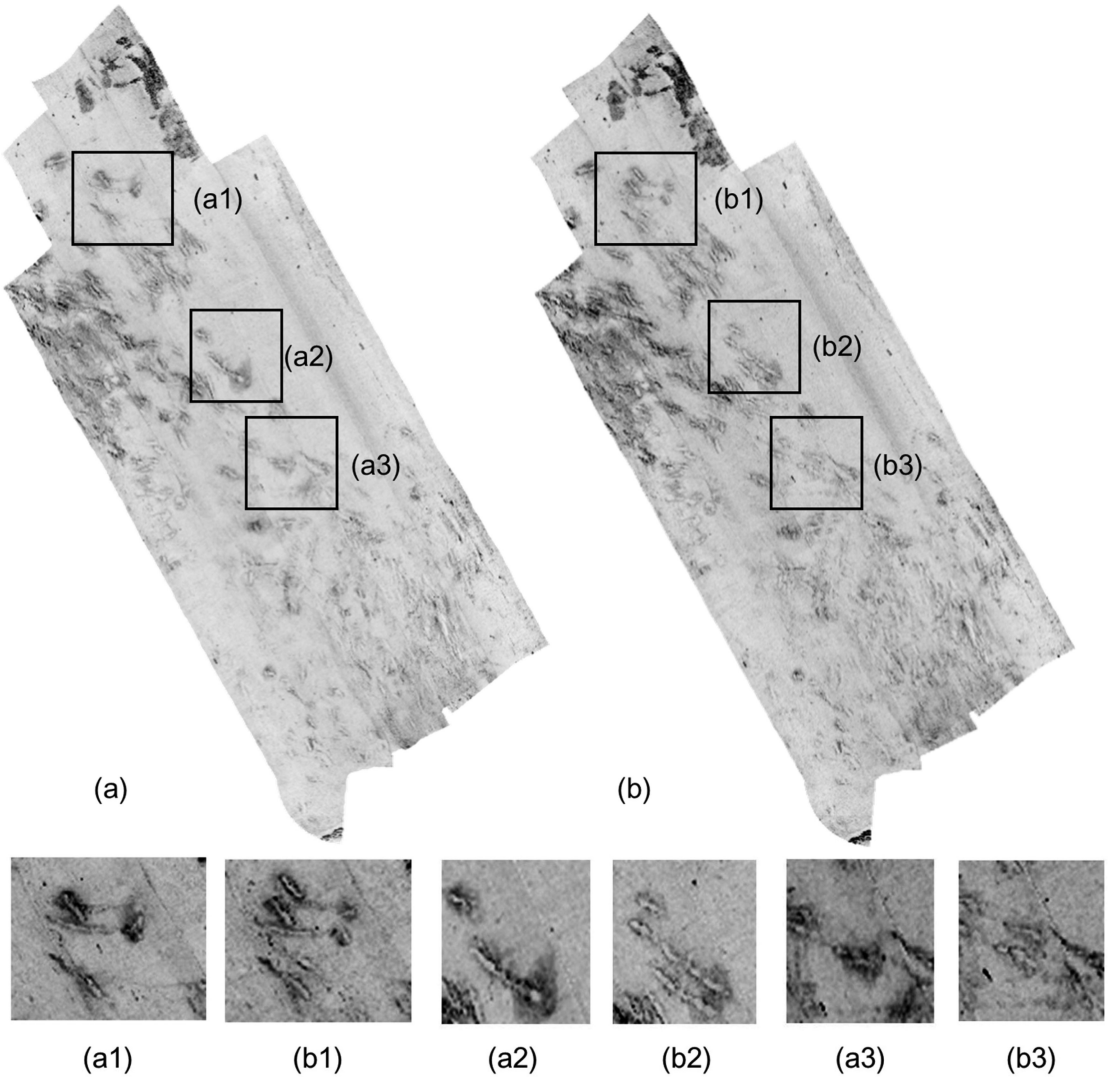

Fig. 7. SSS image mosaic results using (a) the proposed and (b) geocoding method. The contrast has been increased.

TABLE I
ACCURACY ASSESSMENT FOR WHOLE AREA

		Max. (m)	Min. (m)	Mean (m)	STD (m)
Interior checking	dE1	42.50	-1.94	17.37	10.30
	dE2	15.86	-14.09	**0.01**	**4.22**
	dN1	3.75	-25.25	-10.74	6.62
	dN2	11.94	-12.08	**0.00**	**3.47**
Exterior checking	dE1’	41.50	-4.12	13.98	10.18
	dE2’	23.59	-18.18	**0.81**	**8.93**
	dN1’	47.25	-23.25	8.91	22.34
	dN2’	27.20	-24.06	**1.73**	**11.36**

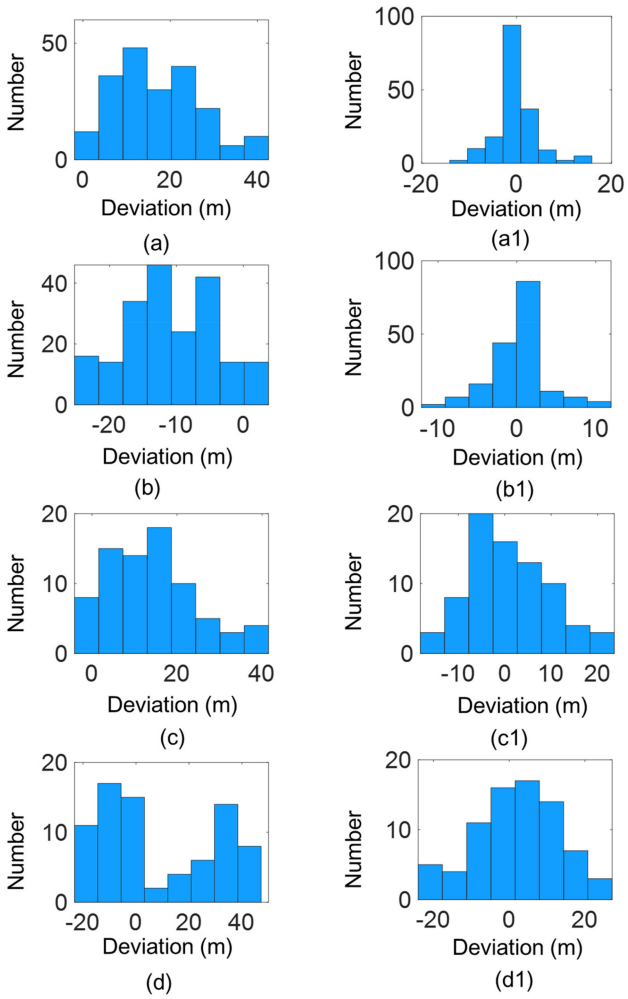

Fig. 8. (a) and (a1) Histograms of coordinate deviations of CFPs for interior checking before and after conducting the coordinate transformation in east direction. (b) and (b1) Those for interior checking in north direction. (c) and (c1) Those for exterior checking in east direction. (d) and (d1) Those for exterior checking in north direction.

8.93 m and 11.36 m. This indicated that the same targets were finally assigned with the almost same coordinates. For both interior checking and exterior checking, the consistency of CFP coordinates was improved. Meanwhile, the statistical parameters and histograms for interior checking seemed better than these for exterior checking. But they were both better than these for the original data before conducting the coordinate transformation. These histograms and statistical parameters proved that using the proposed multiple SSS images mosaic method, the coordinates of matched CFPs in the mosaicked image became unique.

IV. Discussion

A. Impact of Overlapping Area Determination on SSS Image Matching

To evaluate the importance of overlapping area determination for SSS image matching, the SSS image matching results before and after applying the RANSAC algorithm for two adjacent strips without overlapping area determination were shown in Fig. 9 and Table II. Fig. 9 and Table II indicated that the number of initial matched points (1301 matched points) was greatly more than that (570 matched points) when only conducting image matching in the overlapping area. But most of initial matches were wrong in Fig. 9(a). Compared with the image matching results shown in Fig. 4(c) and (d), which used the SURF algorithm for initial image matching and the RANSAC to eliminate the mismatches, Fig. 9(b) still presented the correct matches. This further proved the importance of using the geographic coordinates as a constrain for SSS image matching. However, the number of retained corrected matched points (116 matched points) was less than that (154 matched points) when conducting image matching in the overlapping area (see Table II). The ratio of correct matches was low (8%) and some correct matches were not retained [see black circles in Fig. 9(b)]. Moreover, compared to Fig. 9(b), the correct matches distributed more evenly when conducting image matching in overlapping areas [see Fig. 4(b)].

B. Importance of Segmentation Operation for SSS Image Coordinate Transformation

Considering that the coordinate deviations of matched CFPs were variable across the SSS image, the overlapping area was segmented into several regions when establishing the coordinate transformation model. When no segmentation operation was conducted, the SSS image mosaic result also seemed desirable [see Fig. 10(b)] because the using of TPS function to establish the transformation model could accommodate the local coordinate deviations between CFPs. However, when some areas were zoomed, the artifact would appear (the same feature was seen in different positions), as shown in Fig. 10(b1). On the contrary, no artifacts appeared in Fig. 10(a1) and the same target was associated with unique coordinates in the mosaicked image when using the segmentation operation.

The accuracy assessment for adjacent SSS image mosaic without and with segmentation operation was given in Table III. For the interior checking points when using segmentation operation, mean coordinate deviations in terms of east and north directions after coordinate transformation decreased from 3.09 and -1.14 to -0.01 m and 0.00 m; STD decreased from 9.40 and 7.26 to 3.75 m and 4.06 m. With regards to exterior checking points when using segmentation operation, mean coordinate deviations in terms of east and north directions after coordinate transformation decreased from 4.71 and -1.54 to 1.35 m and 0.34 m; STD decreased from 9.79 and 7.29 to 9.34 m and 7.02 m. These statistical parameters indicated that using the segmentation operation, the coordinate deviations of matched CFPs decreased more. Both Fig. 10 and Table III suggested that the use of segmentation operation could help to avoid the appearance of artifacts in the mosaicked image.

C. Comparison With Other Image Matching Methods

Image matching is the premise of image mosaic and provides accurate matched CFPs for building a coordinate transformation model. In Section III-A, the image matching results with and without constraint of geographic coordinates of FPs had been compared, which indicated that initial image matching

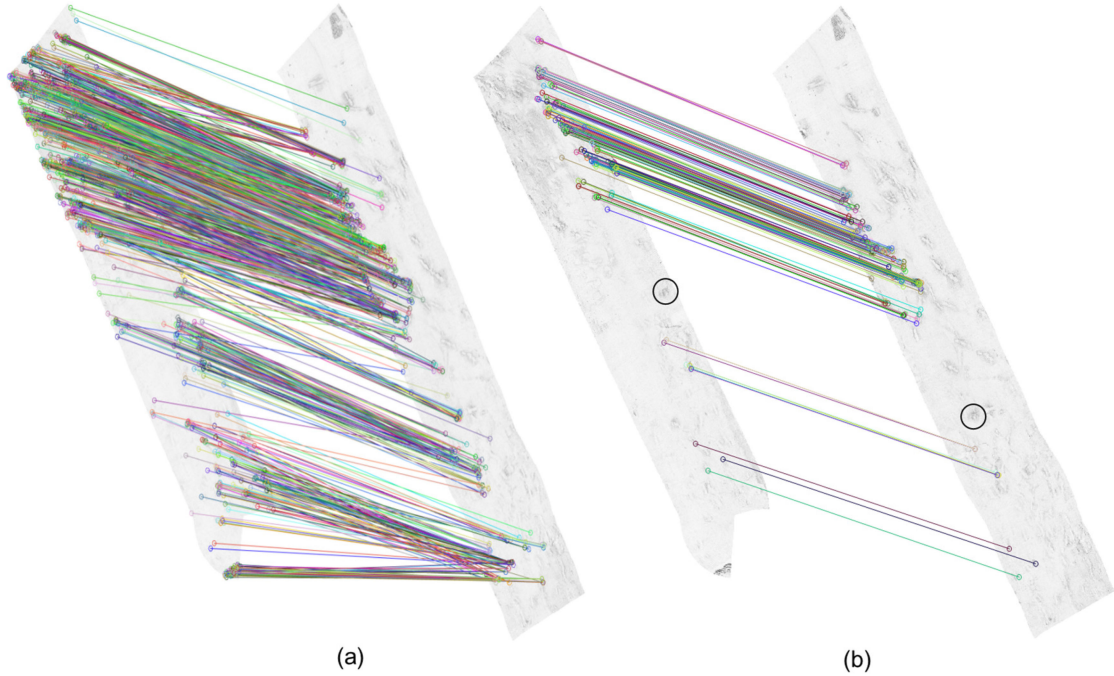

Fig. 9. (a) Initial image matching result without determining overlapping area. (b) Image matching result after applying RANSAC algorithm.

TABLE II
COMPARISON OF WHETHER OVERLAPPING AREA DETERMINATION WAS ADOPTED FOR IMAGE MATCHING

	Initial matches	Correct matches	Correct ratio
With overlapping areas determination	570	154	27%
Without overlapping areas determination	1301	116	8%

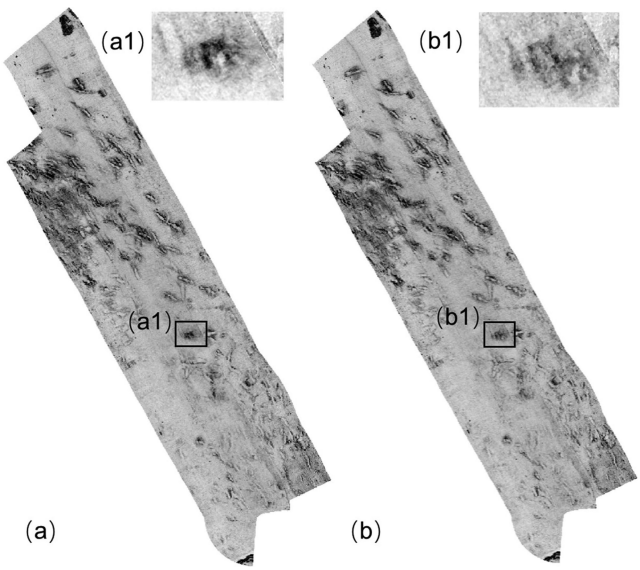

Fig. 10. Comparison of mosaicking two adjacent strips (a) with and (b) without segmentation operation. The contrast has been increased.

performed better under the constraint of geographic coordinates. After applying the RANSAC algorithm for the initial matching, the correct matches were obtained.

In addition to using the SURF algorithm for SSS image matching, Chailloux *et al*. also proposed an intensity-based block matching algorithm [7]. They compared different similarity measures of FPs and concluded that the combination of CR & MI would be a good descriptor for SSS image FPs [7]. Thus, CR & MI without and with constraint of geographic coordinates were used for initial matching. Meanwhile, many other feature-based methods exist in image matching field [37] and the representative oriented fast and rotated BRIEF (ORB) algorithm is an efficient alternative to SURF [38]. Thus, the ORB algorithm without and with constraint of geographic coordinates were also adopted for initial matching.

After applying the RANSAC algorithm for the initial matches, the finally matched results were shown in Fig. 11 (a)–(d). It can be seen that under the constraint of geographic coordinates, the matched results were correct. On the contrary, wrong matches still existed without the constraint of geographic coordinates.

This experiment used another two image matching methods to detect and describe CFPs of two SSS images. No matter what kind of algorithm was used, the constraint of geographic coordinates could always ensure better initial matches. After applying the RANSAC algorithm for the initial matches, the correct matched results could be obtained. The results further proved the importance of considering the geographic coordinates as a constraint.

TABLE III
ACCURACY ASSESSMENT FOR ADJACENT SSS IMAGE MOSAIC WITH AND WITHOUT SEGMENTATION OPERATION

			Max. (m)	Min. (m)	Mean (m)	STD (m)
Without Segmentation	Interior checking	dE1	42.50	-1.94	16.99	10.01
		dE2	35.62	-14.95	**3.09**	**9.40**
		dN1	7.25	-21.25	-7.03	7.27
		dN2	13.81	-17.34	**-1.14**	**7.26**
	Exterior checking	dE1′	41.50	-0.25	17.26	9.39
		dE2′	13.97	-15.95	**4.71**	**9.79**
		dN1′	6.00	-22.00	-5.38	7.53
		dN2′	13.97	-18.29	**-1.54**	**7.29**
With Segmentation	Interior checking	dE1	36.43	-7.12	11.77	8.90
		dE2	8.86	-16.12	**-0.01**	**3.75**
		dN1	5.75	-22.50	-7.23	6.94
		dN2	13.66	-15.96	**0.00**	**4.06**
	Exterior checking	dE1	41.62	-2.56	13.92	10.74
		dE2′	37.53	-15.57	**1.35**	**9.34**
		dN1′	5.75	-20.00	-7.72	7.37
		dN2′	12.38	-19.20	**0.34**	**7.02**

D. Method Repeatability: Applicability of the Proposed Method in Bohai Sea Waters

To further verify the performance of the proposed method, another experiment using the collected data in Bohai sea waters, where the depth ranges from 10 to 20 m, was conducted. In this experiment, two surveying lines with lengths of 701 and 720 m were designed. The swath width was set as 50 m and the overlapping rate between two adjacent strips was 50%. After processing the obtained SSS data, two strips were shown in Fig. 12 (a) and (b), which indicated there exist the same features and targets in the overlapping area. Applying the proposed mosaic method for the two strips, the mosaicked image was presented in Fig. 12 (c). This mosaicked image revealed the features of both strips I and II. Moreover, the same features in the overlapping area were clearly presented and no artifacts appeared. This experiment was performed using the same hardware with the i7, 3.40 GHz Intel Core and 8.00 GB RAM, and the computing time was 10.82 minutes.

Both the experiments using the collected data in Shenzhen and Bohai sea were conducted along the coast and the depth is not very deep. Since the SSS is mainly towed by a cable behind a vessel for scanning surveying, it is always near the seafloor surface and used for high-resolution mapping in different underwater scenarios including deep sea, shallow waters, lakes or rivers [12], [29]. This operating mechanism of SSS will not be affected by the water depth when the sea conditions are stable [29]. Therefore, the proposed method can be used in different water areas.

E. Limitations of the Proposed Method and Future Direction

The two experiments using the collected data in Shenzhen and Bohai sea both used the same-survey data for image mosaic. For the same survey, the parameters of SSS during the surveying time would be unchanged [8], [9], [29], [34]. Thus, the same features of the seafloor surface present the same intensity on different SSS images. After processing the SSS images, no tonal difference exists between different strips. Consequently, the SURF algorithm is sufficient to detect the CFPs.

However, some long-term survey tasks require several periods. In each surveying period, the parameters of SSS may be reset. Moreover, a different SSS with a different frequency may be used. As a result, the SSS images obtained from different times and sonars would present the tonal difference. This kind of intensity difference is sometimes nonlinear [39], [40]. In this case, using SURF algorithm may fail to detect the accurate CFPs as SURF is not robust to nonlinear intensity difference [41], [42]. To settle this problem, many new image matching methods and novel descriptors of FPs have appeared in remote sensing image matching field, including the RIFT method [41] and the DLSS descriptor [42]. Meanwhile, the tonal adjustment should be considered before mosaicking because of the SSS image intensity differences. A local moment matching algorithm has proved its effectiveness to solve this problem [14]. Based on the existed research and SSS data characteristics, the future work would focus on developing new methods to detect and describe the FPs on the SSS images obtained from different sonars and

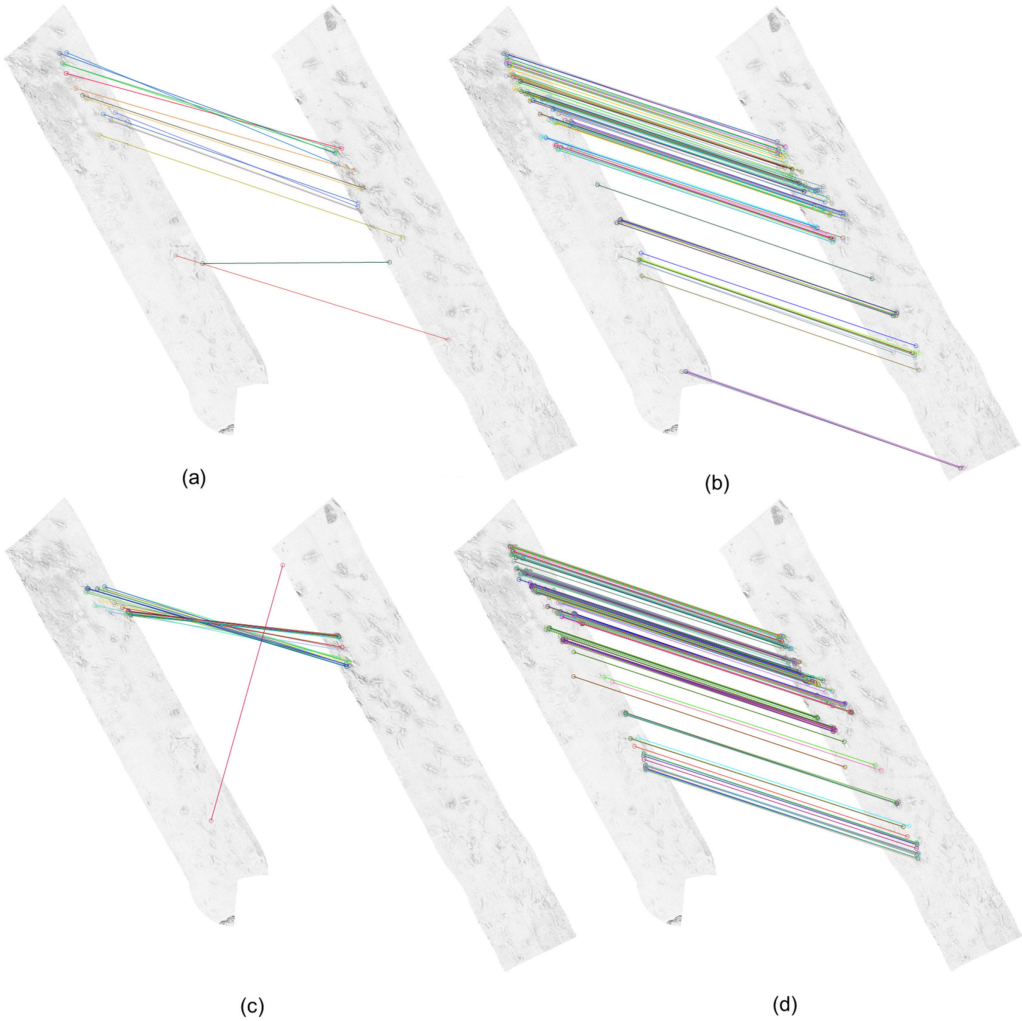

Fig 11. (a) and (b) SSS image matching results after applying RANSAC algorithm for the initial matching using CR & MI without and with constraint of geographic coordinates. (c) and (d) These using ORB algorithm without and with constraint of geographic coordinates.

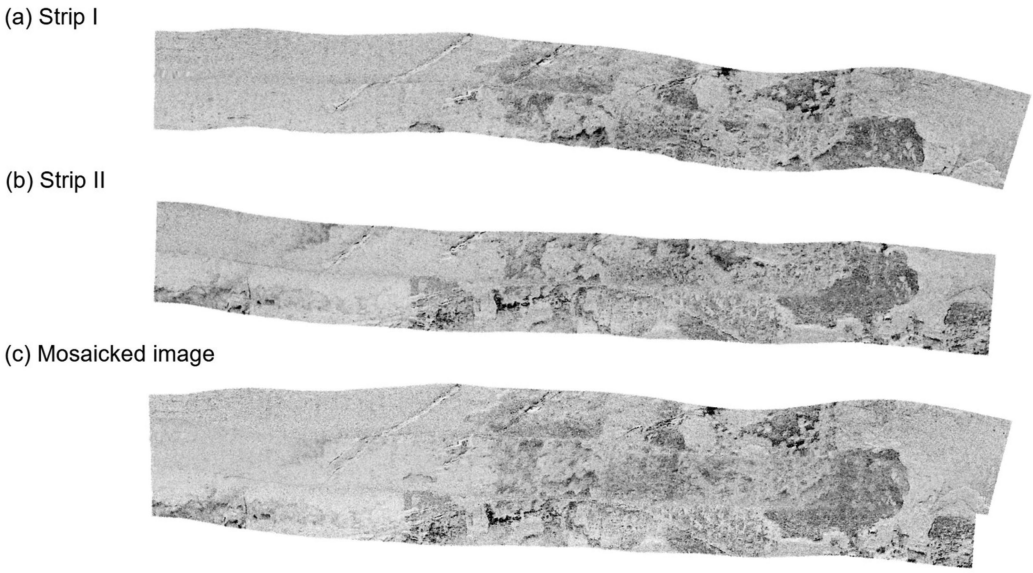

Fig 12. Two strips and the mosaicked image in Bohai sea waters.

periods. Meanwhile, the tonal adjustment is also a research point for these kinds of SSS images.

V. Conclusion

This article proposes a multiple SSS images mosaic method to form a large-scale seafloor image. When multiple strips are used, the mosaic operation is conducted between two adjacent strips in sequence according to the permutation parameters. The track lines and swath width are employed to automatically confirm the overlapping areas between adjacent strips. With regards to the image matching within the overlapping area, the geographic coordinates are used as a constraint to improve the initial matching accuracy. Otherwise, k-means cluster algorithm is used to objectively segment the overlapping area into several regions where the corresponding coordinate transformation models are established.

The proposed method was successfully applied for SSS image mosaic in water areas of Shenzhen and Bohai sea, providing great objectivity and automaticity in terms of overlapping area determination and segmentation. Experimental results showed that the coordinate deviations of the same features in the overlapping areas were reduced. The same targets were associated with unique positions in the mosaicked image. Such a mosaicked image exposes correct substrate characteristics and surface features, which is meaningful for understanding and interpreting the seafloor.

References

[1] M. Prampolini, P. Blondel, F. Foglini, and F. Madricardo, "Habitat mapping of the maltese continental shelf using acoustic textures and bathymetric analyses," *Estuarine Coastal Shelf Sci.,* vol. 207, pp. 483–498, 2018.

[2] G. Rousakis, A. P. Karageorgis, and P. Georgiou, "Geological structure and seabed morphology of the stoupa submarine groundwater discharge system, Messinia, Greece," *Environ. Earth Sci.,* vol. 71, no. 12, pp. 5059–5069, 2014.

[3] J. Petrich, M. F. Brown, J. L. Pentzer, and J. P. Sustersic, "Side scan sonar based self-localization for small autonomous underwater vehicles," *Ocean Eng.,* vol. 161, pp. 221–226, 2018.

[4] B. Liu *et al.*, "Underwater hyperspectral imaging technology and its applications for detecting and mapping the seafloor: A review," *Sensors,* vol. 20, no. 17, pp. 4962–4982, 2020.

[5] P. Davis and J. Brockhurst, "Subsea pipeline infrastructure monitoring: A framework for technology review and selection," *Ocean Eng.,* vol. 104, pp. 540–548, 2015.

[6] J. Zhao, X. Shang, and H. Zhang, "Side-Scan sonar image mosaic using couple feature points with constraint of track line positions," *Remote Sens.,* vol. 10, no. 6, pp. 953–974, Jun. 2018.

[7] C. Chailloux, J. M. Le Caillec, D. Gueriot, and B. Zerr, "Intensity-Based block matching algorithm for mosaicing sonar images," *IEEE J. Ocean. Eng.,* vol. 36, no. 4, pp. 627–645, Oct. 2011.

[8] P. Cervenka and C. De Moustier, "Sidescan sonar image processing techniques," *IEEE J. Ocean. Eng.,* vol. 18, no. 2, pp. 108–122, Apr. 1993.

[9] P. S. Chavez *et al.*, "Processing, mosaicking and management of the monterey bay digital sidescan-sonar images," *Mar. Geol.,* vol. 181, no. 1–3, pp. 305–315, 2002.

[10] D. T. Cobra and A. V. Oppenheim, "Geometric distortions in side-scan sonar images: A procedure for their estimation and correction," *IEEE J. Ocean. Eng.,* vol. 17, no. 3, pp. 252–268, Jul. 1992.

[11] J. E. H. Clarke, "Dynamic motion residuals in swath sonar data : Ironing out the creases," *Int. Hydrograph. Rev.,* vol. 4, no. 1, pp. 6–23, 2003.

[12] A. Burguera and G. Oliver, "High-resolution underwater mapping using side-scan sonar," *PLOS ONE,* vol. 11, no. 1, 2016, Art. no. e0146396.

[13] Y. Tian, A. Sun, N. Luo, and Y. Gao, "Aerial image mosaicking based on the 6-DoF imaging model," *Int. J. Remote Sens.,* vol. 41, no. 1, pp. 74–89, 2020.

[14] X. Li, N. Hui, H. Shen, Y. Fu, and L. Zhang, "A robust mosaicking procedure for high spatial resolution remote sensing images," *ISPRS J. Photogrammetry Remote Sens.,* vol. 109, pp. 108–125, 2015.

[15] X. Li *et al.*, "Generating high-quality and high-resolution seamless satellite imagery for large-scale urban regions," *Remote Sens.,* vol. 12, no. 1, pp. 81–105, 2020.

[16] J. Zhao *et al.*, "Rapid mosaicking of unmanned aerial vehicle (UAV) images for crop growth monitoring using the SIFT algorithm," *Remote Sens.,* vol. 11, no. 10, pp. 1226–1244, 2019.

[17] J.-I. Kim, H.-C. Kim, and T. Kim, "Robust mosaicking of lightweight UAV images using hybrid image transformation modeling," *Remote Sens.,* vol. 12, no. 6, pp. 1002–1019, 2020.

[18] W. Zhang, X. Li, J. Yu, M. Kumar, and Y. Mao, "Remote sensing image mosaic technology based on SURF algorithm in agriculture," *EURASIP J. Image Video Process.*, pp. 85–93, 2018.

[19] R. Aktar *et al.*, "Robust mosaicking of maize fields from aerial imagery," *Appl. Plant Sci.,* vol. 8, no. 8, 2020, Art. no. e11387.

[20] X. Li, R. Feng, X. Guan, H. Shen, and L. Zhang, "Remote sensing image mosaicking: Achievements and challenges," *IEEE Geosci. Remote Sens. Mag.,* vol. 7, no. 4, pp. 8–22, Dec. 2019.

[21] S. Daniel and F. Le Leannec, "Side-scan sonar image matching," *IEEE J. Ocean. Eng.,* vol. 23, no. 3, pp. 245–259, 1998.

[22] P. Schwind, S. Suri, P. Reinartz, and A. Siebert, "Applicability of the SIFT operator to geometric SAR image registration," vol. 31, no. 7/8, pp. 1959–1980, 2010.

[23] J. Zhao, A. Wang, H. Zhang, and X. Wang, "Mosaic method of side-scan sonar strip images using corresponding features," *IET Image Process.,* vol. 7, no. 6, pp. 616–623, 2013.

[24] X. F. Ye, P. Li, J. G. Zhang, J. Shi, and S. X. Guo, "A feature-matching method for side-scan sonar images based on nonlinear scale space," *J. Mar. Sci. Technol.,* vol. 21, no. 1, pp. 38–47, 2016.

[25] W. Tao and Y. Liu, "Combined imaging matching method of side scan sonar images with prior position knowledge," *IET Image Process.,* vol. 12, no. 2, pp. 194–199, 2018.

[26] Z. Tang, G. Ma, J. Lu, Z. Wang, and Y. Wang, "Sonar image mosaic based on a new feature matching method," *IET Image Process.,* vol. 14, no. 10, pp. 2149–2155, 2020.

[27] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up robust features (SURF)," *Comput. Vis. Image Understanding,* vol. 110, no. 3, pp. 404–417, 2008.

[28] J. Zhao, X. Wang, H. Zhang, and A. Wang, "A comprehensive bottom-tracking method for sidescan sonar image influenced by complicated measuring environment," *IEEE J. Ocean. Eng.,* vol. 42, no. 3, pp. 619–631, Jul. 2017.

[29] P. Blondel, *The Handbook of Sidescan Sonar*, Berlin Germany: Springer, 2009.

[30] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM,* vol. 18, no. 9, pp. 509–517, 1975.

[31] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM,* vol. 24, no. 6, pp. 381–395, 1981.

[32] K. Ismail, V. A. I. Huvenne, and D. G. Masson, "Objective automated classification technique for marine landscape mapping in submarine canyons," *Mar. Geol.,* vol. 362, pp. 17–32, 2015.

[33] I. Barrodale, R. Kuwahara, R. Poeckert, and D. Skea, "Side-scan sonar image processing using thin plate splines and control point matching," *Numer. Algorithms,* vol. 5, no. 2, pp. 85–98, 1993.

[34] D. Buscombe, "Shallow water benthic imaging and substrate characterization using recreational-grade sidescan-sonar," *Environ. Model. Softw.,* vol. 89, pp. 1–18, 2017.

[35] J. Y. Li, Q. W. Hu, and M. Y. Ai, "Robust feature matching for remote sensing image registration based on L_q-Estimator," *IEEE Geosci. Remote Sens. Lett.,* vol. 13, no. 12, pp. 1989–1993, Dec. 2016.

[36] C. A. Zala and I. Barrodale, "Warping aerial photographs to orthomaps using thin plate splines," *Adv. Comput. Math.,* vol. 11, pp. 211–227, 1999.

[37] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image matching from handcrafted to deep features: A survey," *Int. J. Comput. Vis.*, vol. 129, pp. 23–79, 2020.

[38] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," 2012.

[39] D. Tamsett, J. McIlvenny, and A. Watts, "Colour sonar: Multi-frequency sidescan sonar images of the seabed in the inner sound of the pentland firth, Scotland," *J. Mar. Sci. Eng.*, vol. 4, no. 1, pp. 26–43, 2016.
[40] B. Misiuk, C. J. Brown, K. Robert, and M. Lacharité, "Harmonizing multi-source sonar backscatter datasets for seabed mapping using bulk shift approaches," *Remote Sens.,* vol. 12, no. 4, pp. 601–636, 2020.
[41] J. Li, Q. Hu, and M. Ai, "RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform," *IEEE Trans. Image Process.,* vol. 29, pp. 3296–3310, Dec. 2019.
[42] Y. Ye, L. Shen, M. Hao, J. Wang, and Z. Xu, "Robust optical-to-SAR image matching based on shape properties," *IEEE Geosci. Remote Sens. Lett.,* vol. 14, no. 4, pp. 564–568, Apr. 2017.

Xiaodong Shang received the B.S. degree in in surveying and mapping engineering from Shandong Agricultural University, Tai'an, China, and M.S. degree in Geodesy and Survey Engineering from Wuhan University, Wuhan, China, where he is currently working toward the Ph.D. degree at the School of Geodesy and Geomatics.

His research interests include side scan sonar data processing, seafloor classification, and computer vision.

Jianhu Zhao received the B.S. and the M.S. degrees in surveying and mapping from University of Wuhan surveying and mapping, Wuhan, China, in 1994 and 1998, respectively, and the Ph.D. degree in in geodesy and geomatics from Wuhan University, Wuhan, China, in 2002.

He is currently a Professor with the School of Geodesy and Geomatics, and Institute of Marine Science and Technology, Wuhan University. He is also the Deputy Director of the Institute of Marine Science and Technology, Wuhan University. His research interests include underwater topography, seafloor sediment classification, marine subbottom profile data processing, marine hydrology, and underwater navigation.

Hongmei Zhang received the B.S. and the M.S. degrees in physics from Wuhan University of Hydraulic and Electric Engineering, Wuhan, China, in 1992 and 1995, respectivel, and the Ph.D. degree in hydraulics and river dynamics from Wuhan University, Wuhan, China, in 2003.

She is currently a Professor with the Department of Automation, Wuhan University, Wuhan, China. Her research areas focus on computer detection and control technology, signal processing, fault diagnosis, underwater integrated navigation technology and other aspects of teaching and scientific research. Her research interests include image processing and computer version.