Journal of Imaging Science and Technology® 66(3): 030502-1–030502-8, 2022.

A Fast Underwater Image Stitching Algorithm with Adaptive Adjustment of Attitude Angle

Lizhou Jiang, Zhijie Tang, Zhihang Luo, and Chi Wang
School of Mechatronic Engineering and Automation, No. 99 Shangda Road, Shanghai University, Shanghai, China
E-mail: tangzhijie@shu.edu.cn

Abstract. *In underwater image acquisition process, due to the impact of water currents and other disturbances, the movement posture of the underwater machine will be unstable, which could lead to unusual problems such as twisting of underwater image capture. These factors will increase the error rate of feature point matching and lead to the failure of panoramic image mosaic. In this regard, we propose a new, highly applicable underwater image stitching algorithm. Firstly, the posture angle adjustment link is added to the underwater image processing, and the angle deflection problem of the underwater image is effectively improved by using the posture angle information. Secondly, the feature points of underwater images are extracted based on the accelerated robust feature (SURF) algorithm. Then, the reference image is matched with the feature points of the image to be registered, and effective feature point pairs are obtained by screening. Finally, the images are stitched based on OpenCV to obtain a good panoramic image. After experimental analysis and comparison, our method can increase the number of matching feature point pairs between images. In addition, the Euclidean distance is significantly shortened during the matching process, which further makes the matching of feature points more accurate. Our method satisfactorily overcomes the adverse effects of actual underwater operations and has a better application prospect.*

[DOI: 10.2352/J.ImagingSci.Technol.2022.66.3.030502]

1. INTRODUCTION

Oceans accounts for about 71% of the earth's surface area. These vast oceans contains abundant resources for survival of mankind. The development and utilization of oceans has become one of the basic factors that determine the rise and fall of a country. It is very important to use modern scientific and technological means to explore and research water resources. Among them, underwater image processing technology has received extensive attention in the field of marine science and technology, and has shown significant results. The method of acquiring underwater pictures and target scene video sequences by using Remote Operated Vehicle (ROV) has become an important apparatus of acquiring underwater information. However, the field of view of the underwater camera is limited, the target scene range obtained by a single underwater picture or video sequence is small, the target information is less, and multiple images have the problem of poor continuity. A successful method to accurately complete underwater image stitching has become the main subject of the current research on underwater image stitching algorithms [1, 2].

Richard Szeliski [3] proposed a motion-based panoramic image stitching model in 1996. This technology uses an iterative nonlinear minimization method and the algorithm can stitch images through translation, rotation, and affine transformations. In 2000, Shmuel Peleg et al. [4] proposed an adaptive image stitching model. This technology can adaptively select the stitching model according to the different movement modes of the camera. Brown [5] and D.G. Lowe proposed an image mosaic method based on Scale Invariant Feature (SIFT) in 2003. In 2006, a new scale rotation invariant feature detection operator was proposed by Herbert Bay [6] et al., that is, the Speed-Up Robust Features operator (SURF), which uses integral image and box convolution methods to improve the algorithm speed and robustness. Zhao Xiangyang [7] et al. proposed an automatic stitching algorithm based on Harris corner matching. The algorithm uses the Harris corner detection method to extract features and has good stability when processing images. In addition, Zeng Luan [8] et al. proposed an improved SIFT [9] feature extraction algorithm in 2011, so that image stitching can be completed well in the case of low overlap.

In this paper, the new underwater image stitching method we proposed is implemented based on the SURF feature detection algorithm. The problem of image distortion can be effectively solved by using inertial sensor to obtain attitude angle information and using this information to correct the angle of underwater image. In this regard, the corrected image is more conducive to the extraction and matching of feature points, reducing matching error information. The final experimental results prove that this method can effectively improve the speed and accuracy of underwater image matching. The overall framework flow chart of the system is shown in Figure 1.

2. THEORETICAL BASIS

2.1 *Introduction to SURF Algorithm*

The SURF [10, 11] algorithm is used in the field of computer vision, such as object recognition or three-dimensional reconstruction. The algorithm uses integral image, Haar [12] wavelet transform, and approximate Hessian matrix operation to improve time efficiency, among which Haar wavelet transform increases the robustness of the algorithm.

There are always special pixels in an image. These points can be regarded as the characteristics of this image,

Received Apr. 18, 2021; accepted for publication Sept. 6, 2021; published online Oct. 1, 2021. Associate Editor: Steven Simske.
1062-3701/2022/66(3)/030502/8/$25.00

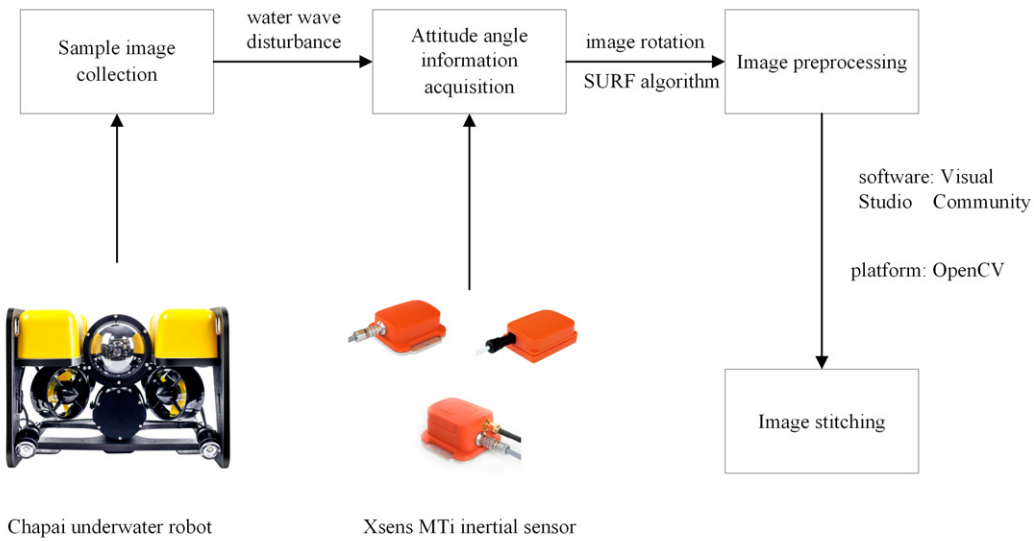

Figure 1. System overall framework flow chart.

called feature points. In the field of computer vision, the concept of feature points has been widely used, including target recognition, image registration, visual tracking, and 3D reconstruction [13, 14]. The principle of this concept is to select certain feature points from the image and analyze the image locally instead of observing the entire image. As long as there are enough detectable feature points in the image, and these interest points are different and have stable features, they can be accurately located [15, 16].

2.2 *SURF Algorithm Feature Extraction*

2.2.1 *Construct Hessian Matrix to Generate All the Feature Points for Feature Extraction*

Hessian Matrix [17] is a square matrix composed of the second-order partial derivatives of a multivariate function, which describes the local curvature of the function. The purpose of constructing the Hessian matrix is to generate stable edge points of the image, which is similar to the effect of Laplacian edge detection, and provides a basis for subsequent feature extraction.

For an image $f(x, y)$, its Hessian matrix is as follows:

$$H(f(x,y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}. \tag{1}$$

Before constructing the Hessian matrix, it is necessary to perform Gaussian filtering on the image. The filtered Hessian matrix is expressed as:

$$H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix}, \tag{2}$$

where L_{xx}, L_{xy}, L_{yy} are the second derivative of the image $g(\sigma)$ in each direction after Gaussian filtering and σ stands for the scale space factor.

When the discriminant of the Hessian matrix obtains the local maximum value, it is determined that the current point is a brighter or darker point than other points in the surrounding neighborhood, so as to locate the position of the feature point.

The $f(x, y)$ in the Hessian matrix discriminant is the Gaussian convolution of the original image. Since the Gaussian kernel obeys the normal distribution, the coefficient is getting lower and lower from the center point. SURF algorithm uses box filter to approximate instead of Gaussian filter, which can improve the calculation speed. In order to reduce the error caused by the box filter, multiply L_{xy} by a weighting coefficient of 0.9:

$$\det(H) = L_{xx} * L_{yy} - (0.9 * L_{xy})^2. \tag{3}$$

2.2.2 *Build a Scale Space*

The scale space is expressed as an image pyramid, the input image function is repeatedly convolved with the Gaussian kernel, and it needs to be sub-sampling continuously, as shown in Figure 2(a). The scale space of SURF is composed of M groups and N layers, and the image size between different layers is the same. It allows multi-layer images in the scale space to be processed at the same time, only changing the size of the filter, and improving the calculation speed, as shown in Fig. 2(b).

2.2.3 *Feature Point Positioning*

Use a $3 \times 3 \times 3$ template to find feature points in a three-dimensional scale space. Excluding the central point, there are 26 points left in the space around the point as illustrated in Figure 3. Calculate the Hessian matrix

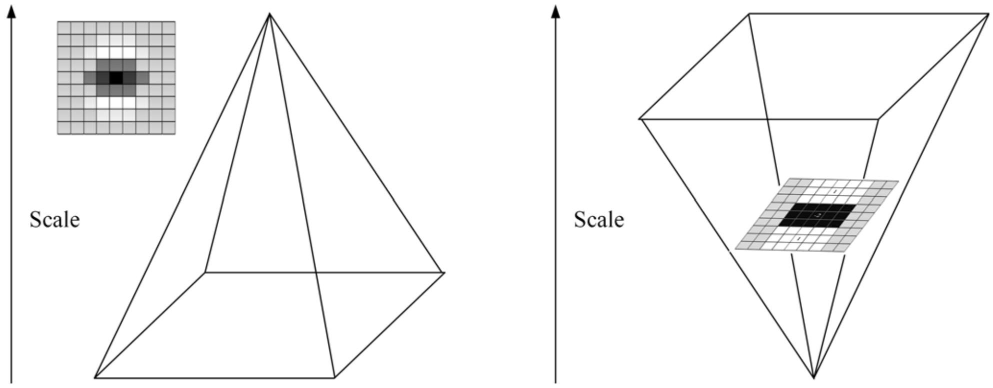

Figure 2. Schematic diagram of the SURF algorithm scale space.

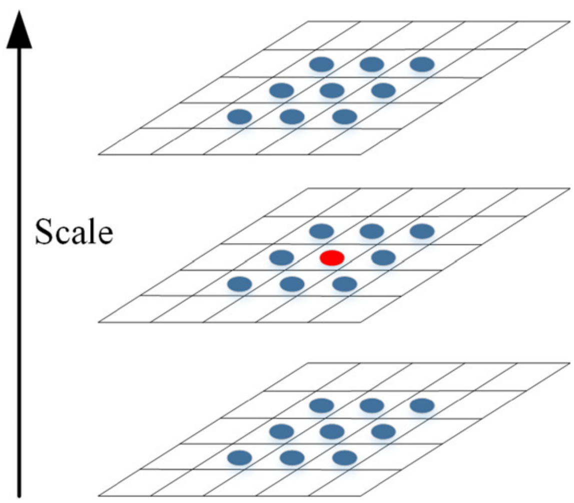

Figure 3. Examples of 3D-scale feature points.

determinant value of these 27 points. When the response value of the central point is greater than the response values of the 26 surrounding points, the central point can be determined as a characteristic point. This method is used to traverse all pixels and screen out the feature points with weak energy and incorrect positioning, and all stable feature points can be obtained.

2.2.4 *Select the Main Direction of the Feature Point*

In SURF, this method is used to count the Harr wavelet features in the circular neighborhood of feature points. In the circular neighborhood of the feature points, the sum of the horizontal Haar and vertical Harr wavelet features of all points in the 60-degree sector is counted, and then the sector is rotated at intervals of 0.2 radians and counted again. Finally, the direction of the sector with the largest statistical value is taken as the main direction of the feature point, as shown in Figure 4.

2.2.5 *Generate Feature Point Descriptors*

Take a 4×4 rectangular area block around the feature point, and the direction of the taken rectangular area is along the main direction of the aforementioned feature point. Each sub-region uses Haar wavelet template to calculate its response value. Each sub-region has 25 sampling pixels, which can accumulate the corresponding horizontal and vertical directions. Obtain the sum of values $\sum d_x$ in the horizontal direction, the sum of values $\sum d_y$ in the vertical direction, the sum of absolute values $\sum |d_x|$ in the horizontal direction, and the sum of absolute values $\sum |d_y|$ in the vertical direction. Considering these four values as the feature vector of each sub-block region, there are a total of 64-dimensional feature vectors as the descriptor of SURF features.

2.2.6 *Feature Point Matching*

For the feature points on the image to be registered, the degree of matching is determined by calculating the Euclidean distance between them and the feature points on the reference image. The shorter the Euclidean distance, the better the matching degree between the two feature points. At the same time, SURF also added the judgment of the Hessian matrix trace. If the matrix traces of the two feature points have the same sign, it means that the two features have the contrast change in the same direction. If they are different, it indicates the contrast change direction of the two feature points. On the contrary, even if the Euclidean distance is 0, it is directly excluded.

2.3 *Image Post-Processing*

During the shooting process, the underwater robot uses its own inertial sensor to obtain the attitude angle information of each image. The zoom factor R can change the size of the image to prevent the loss of edge information after the image is rotated. The posture angle information is used to rotate the image, and then interpolate the image to make it have an appropriate size.

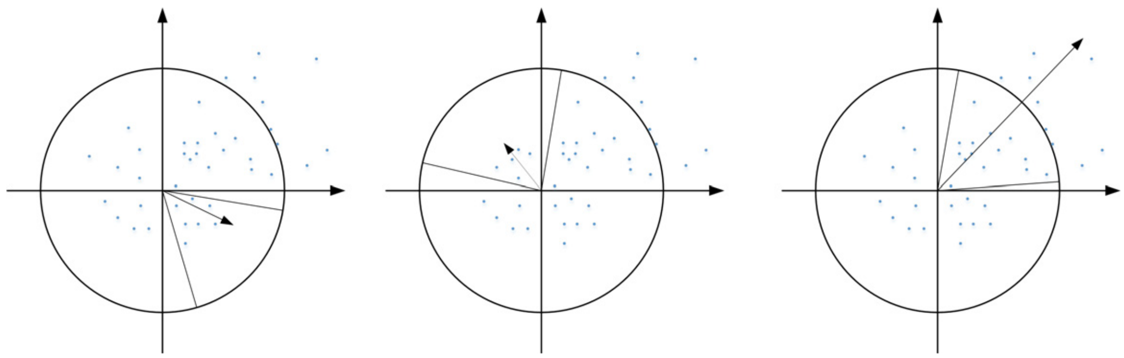

Figure 4. Determination of the main direction of feature points.

2.3.1 *Image Rotation*
Taking the upper left corner of the image as the origin, any pixel and the origin form a vector. Suppose the obtained attitude angle is θ, any pixel is $P_0(x_0, y_0)$, the rotated point is $P'(x', y')$, the image width is W, the height is H, the rotated width is W', and the rotated height is H'. The matrix expression is shown in Eq. (4).

$$\begin{bmatrix} x' & y' & 1 \end{bmatrix} = \begin{bmatrix} x_0 & y_0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ -0.5W & 0.5H & 1 \end{bmatrix} \times \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0.5W' & 0.5H' & 1 \end{bmatrix}. \quad (4)$$

After the image is rotated, the reverse mapping method is used to scan each output pixel in order, and then inversely transform it to the corresponding position of the corresponding input image, and interpolate it to the nearest input pixel.

2.3.2 *Bilinear Interpolation*
Suppose the size of the source image is $m * n$ and the target image is $a * b$. Then the side length ratios of the two images are: m/a and n/b. However, this ratio is usually not an integer, and a floating point type should be used when programming and storing. The pixel points in the ith row and jth column of the target image can be mapped back to the source image through the side length ratio, and the corresponding coordinates are $(i * \frac{m}{a}, j * \frac{n}{b})$. The corresponding coordinates are generally not integers, and non-integer coordinates cannot be used on discrete data such as images. Bilinear interpolation calculates the value of the point by finding the four pixels closest to the corresponding coordinate, which can be a gray value or an RGB value. If the image is a grayscale image, $f(i, j)$ represents the grayscale value of point (i, j), then bilinear interpolation can be performed by the following interpolation formula (i, j) can be calculated by the following formula:

$$\begin{aligned} f(i+v, j+u) &= (1-u)(1-v)f(i,j) + u(1-v)f(i,j+1) \\ &\quad + v(1-u)f(i+1,j) + uv * f(i+1,j+1), \end{aligned} \quad (5)$$

where u is the deviation in the column direction and v is the deviation in the row direction.

3. EXPERIMENTAL ARGUMENT

3.1 *Experimental Process*

A 100 cm × 100 cm porphyritic stone slab is placed in the pool as the object of underwater image data collection. The experiment used Chapai underwater robot, equipped with SONY 8 million pixel camera and 1080 low light chip. A fixed distance was kept between the robot and the slate so that the robot can slowly move parallel to the brick to take pictures. The leftmost end of the first image block is defined as the zero point of the coordinate system to facilitate the description of the specific position of the underwater robot.

In order to verify the algorithm, we add a certain degree of water wave disturbance to simulate the actual underwater environment during the robot movement. When the robot is disturbed, the captured image will have a certain angle offset, and the robot will record the angle offset information. The experimental images were put into OpenCV environment, the attitude angle was used to correct the image, and then the image information was restored by inverse mapping and bilinear interpolation. Finally, the SURF method is used to extract the feature points, and the effective feature points of the reference image and the image to be registered are matched, and the panoramic image stitching is finally completed. The software flow chart is shown in Figure 5.

3.2 *Experimental Results Display*

It can be seen from the captured images that the shooting angle of view occupies about 60% of the width of the brick. From this, the position of each picture can be obtained. The first picture is not disturbed. The experimental results are shown in the following charts.

In Figure 6, the characteristic points of the captured image and the corrected image are shown:

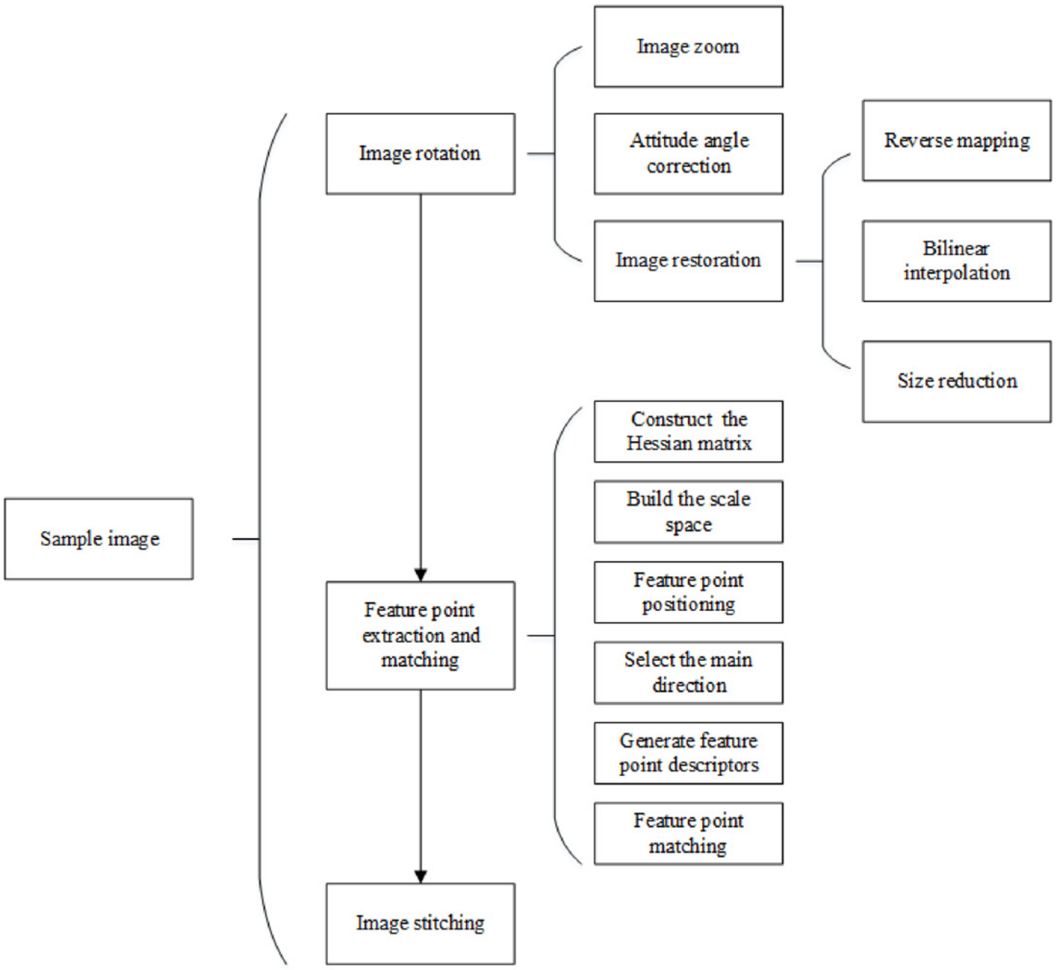

Figure 5. System overall framework flow chart.

Table I. Location and attitude angle information of the photographs.

	(a)	(b)	(c)	(d)
Position	0 cm	15 cm	22 cm	41 cm
Angle	0°	6°	11°	−5°

In Figure 7, the captured image and the corrected image are divided into two groups (1) and (2). Group (1) is the result without image correction, and group (2) is the result after correction algorithm.

The data in Tables I and II can clearly show the matching effect of the image.

When the image is twisted, Pan Jianping's [18] algorithm will form a large number of false matches, Yuan Liying's [19] algorithm will cause the output image to have overlapping areas, and Jin Binying's [20] algorithm will cause the panoramic image to fail to be stitched successfully. These scholars did not use the location and attitude angle info of the image. Therefore, it leads to poor image stitching. This also shows that our algorithm can help image mosaic have a better visual effect.

The images before and after correction are stitched separately, and the following results are obtained. In Figure 8, the splicing effect of the corrected image is much better than that of the image before correction, and there is almost no image distortion and distortion.

3.3 *Analysis of Experimental Results*

It can be seen from Fig. 7 and Table II that the corrected image has more matching feature point pairs, which makes image stitching have a better visual effect. The data in Table II shows the longest Euclidean distance and the shortest Euclidean distance. After the image is corrected, the Euclidean distance has been significantly reduced, the longest Euclidean distance has been reduced by 7%–19%, and the shortest Euclidean distance has been reduced by 30%–90%. These data prove that the proposed algorithm can improve the accuracy of feature point matching. The final stitching result is shown in Fig. 8.

After the posture angle correction, the image stitching result has no obvious distortion, the edge information is not lost, and the stitching effect is greatly improved. The impact of underwater images from water wave interference will be greatly reduced, which will in turn reduce the difficulty

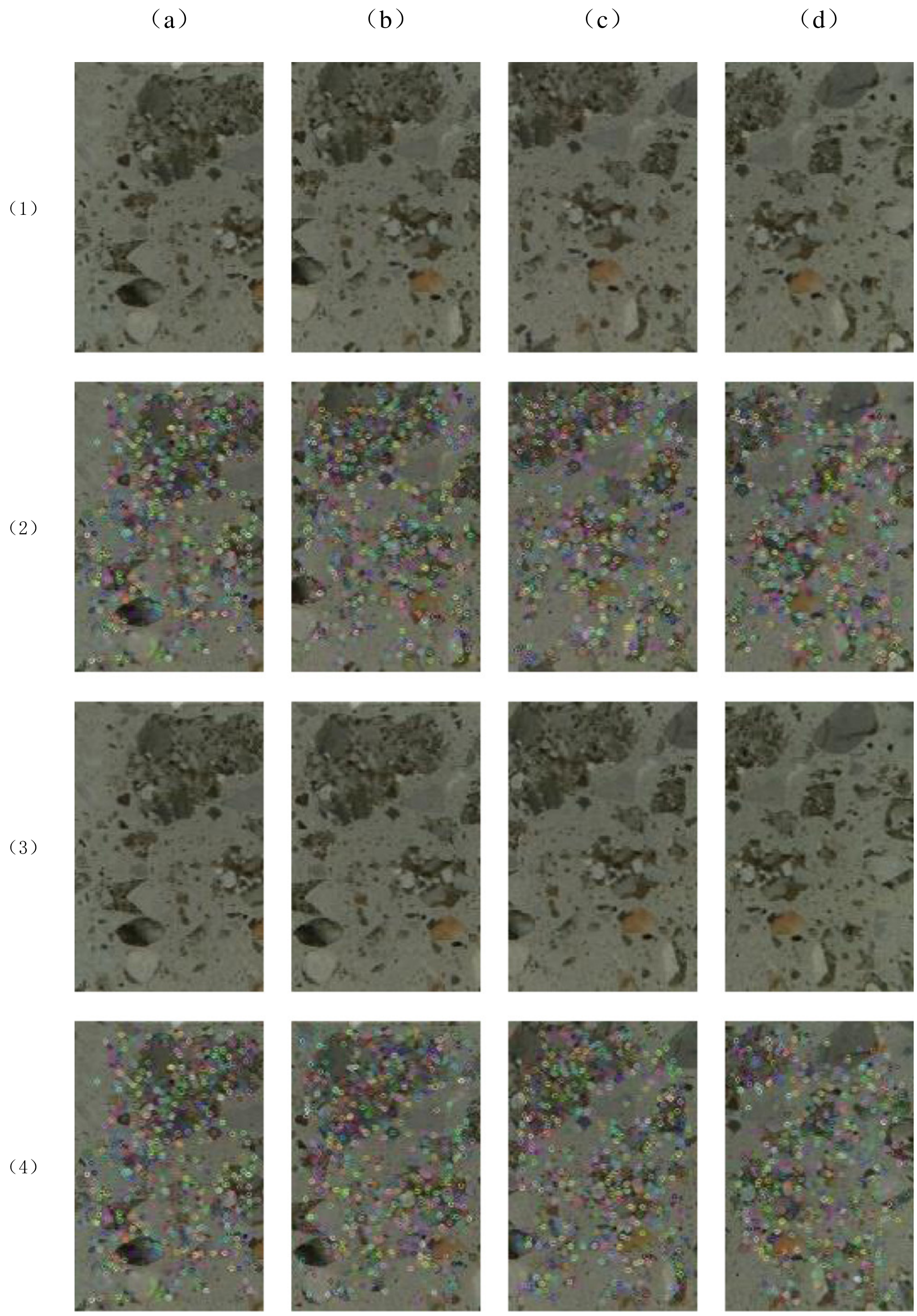

Figure 6. (a), (b), (c) and (d) respectively correspond to the photographed images at four positions; (1) represents the water wave disturbance image, (2) represents the feature point extraction of the image after the disturbance, (3) represents the corrected image, and (4) represents the feature point extraction of the corrected image.

of actual underwater operations, increase flexibility and adaptability, and have good application prospects.

4. CONCLUSION

In this paper, we have developed a new underwater image stitching with the help of SURF, OpenCV and attitude angle information processing to deal with the difficulty of acquiring underwater images in complex underwater environments. Problems such as surge impact disturbance and unstable motion attitude of underwater machines can cause difficulties in image processing. For example, the feature point matching error information increases, and the corresponding matching point pairs decrease, and finally panoramic stitching cannot be realized. We use inertial sensors to obtain real-time attitude information of underwater robots, and use this attitude angle information

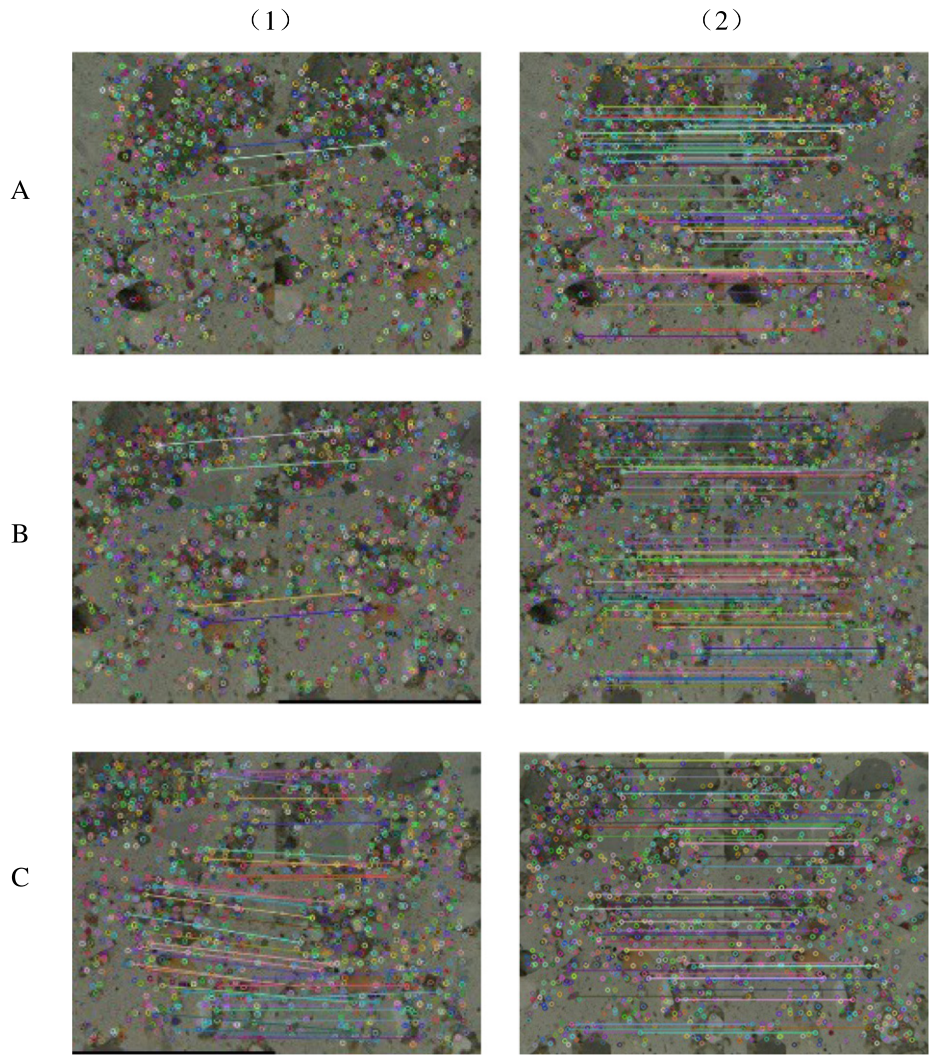

Figure 7. A represents image (a) and image (b) matching, B represents image (b) and image (c) matching, C represents image (c) and image (d) matching.

Table II. Number of matching points and Euclidean distance between each matching graph.

	Disturbance graph			Correction diagram		
	Number of matches	Longest Euclidean distance	Shortest Euclidean distance	Number of matches	Longest Euclidean distance	Shortest Euclidean distance
A	4	0.533989	0.0226876	66	0.451012	0.0159466
B	7	0.520612	0.0240225	97	0.485568	0.0044199
C	42	0.566464	0.0488156	61	0.457259	0.0045621

to scale, rotate and alternately fill underwater images. This step can help us to have a good image stitching foundation for underwater images under adverse effects. Experiments show that our algorithm can significantly improve the matching number and accuracy of feature points. The acquisition and effective use of attitude angle information can eliminate the trouble caused by the harsh underwater environment. The visual effectiveness, accuracy and stability of underwater images have been improved.

ACKNOWLEDGMENT

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation of China (no. 51005142), the Innovation Program of Shanghai Municipal Education Commission (no. 14YZ010), and the Natural Science Foundation of Shanghai (nos 14ZR1414900,19ZR1419300).

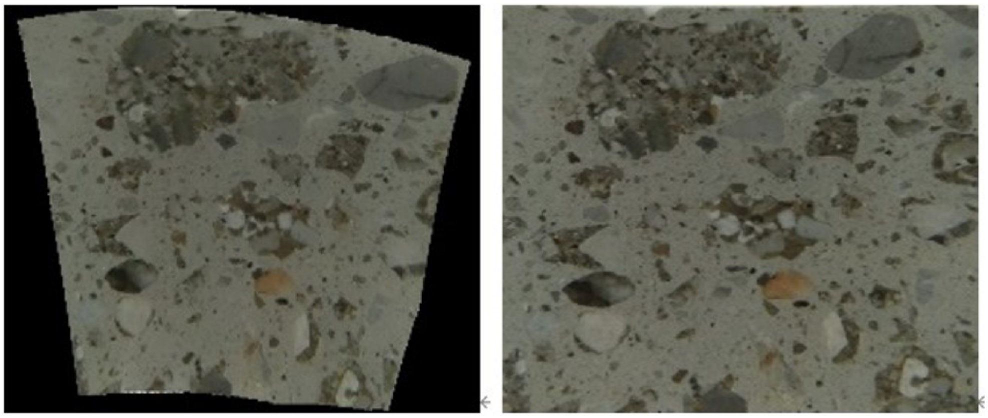

Figure 8. Output mosaic before and after posture angle correction.

REFERENCES

1 M. Sheng, S. Tang, Z. Cui, W. Wu, and L. Wan, "A joint framework for underwater sequence images stitching based on deep neural network convolutional neural network," Int'l. J. Adv. Robotic Syst. **17.2**, 172988142091506 (2020).

2 R. Rajendran, S. P. Rao, K. Panetta, and S. S. Agaian, "Adaptive alpha-trimmed correlation based underwater image stitching," *IEEE Int'l. Symp. on Technologies for Homeland Security* (IEEE, Piscataway, NJ, 2017).

3 R. Szeliski, "Video mosaics for virtual environments," IEEE Comput. Graphics Appl. **16.2**, 22–30 (1996).

4 S. Peleg and B. Rousso, "Mosaicing on adaptive manifolds," IEEE Trans. Pami **22.10**, 1144–1154 (2000).

5 M. Brown and D. G. Lowe, "Recognising panoramas," *Proc. Int'l. Conf. Computer Vision, 2003* (IEEE, Piscataway, NJ, 2003).

6 H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: speeded up robust features," Comput. Vis. Image Understanding **110.3**, 404–417 (2006).

7 M. Teke, M. Firat Vural, A. Temizel, and Y. Yardimci, "High-resolution multispectral satellite image matching using scale invariant feature transform and speeded up robust features," J. Appl. Remote Sensing **5**, 053553 (2011).

8 P. P. R. Filho, F. D. L. Moreira, F. G. D. L. Xavier, S. L. Gomes, J. C. D. Santos, F. N. C. Freitas, and R. G. Freitas, "New analysis method application in metallographic images through the construction of mosaics via speeded up robust features and scale invariant feature transform," Materials **8**, 3864 (2015).

9 M. Chen, R. Nian, B. He, S. Qiu, X. Liu, and T. Yan, "Underwater image stitching based on SIFT and wavelet fusion," *Oceans* (IEEE, Piscataway, NJ, 2015).

10 M. Hu, Y. Peng, and Y. Xu, "Fast image registration algorithm based on improved SURF," Sensors Microsyst. **36**, 151–153 (2017).

11 G. Ju, L. Yuan, X. Liu, and H. Yue, "Research on real-time image registration method of moving target based on improved SURF algorithm," J. Commun. **38**, 177–186 (2017).

12 Q. Tang and X. Zhang, "An improved image denoising method based on wavelet transform and You-Kaveh model," Electronic Design Eng. **27**, 120–124 (2019).

13 E. Zhang, J. Ma, and X. Wang, "An improved SURF color remote sensing image registration algorithm," Liquid Crystal Display **32**, 144–152 (2017).

14 S. Li, X. Wang, and K. Yang, "Improved SURF color remote sensing image registration algorithm," Comput. Meas. Control **25**, 209–212+216 (2017).

15 Q. Zhang, L. Sun, J. Chen, M. Zhou, M. Hu, Y. Wen, and Q. Li, "Speeded up robust features based image mosaic method for large scale microscopic hyperspectral pathological imaging," Meas. Sci. Technol. **32** (2020).

16 Z. Pan and H. Zhang, "A fast image registration algorithm based on improved feature point detection," Sensors Microsyst. **37**, 148–150+160 (2018).

17 Z. Zhuang and H. Wang, "A novel nonuniformity correction algorithm based on speeded up robust features extraction," Infrared Phys. Technol. **73**, 281–285 (2015).

18 J. Pan, J. Hao, and J. Zhao, "Improved image registration algorithm based on SURF," Remote Sensing Land Resources **29**, 110–115 (2017).

19 L. Yuan, J. Liu, and F. Wang, "Improved image registration algorithm based on SURF," J. Detection Control **42**, 65–70+78 (2020).

20 B. Jin, "Research on improved SURF image registration algorithm," Comput. Meas. Control **27**, 228–232 (2019).