

Image Stitching Based on Improved SURF Algorithm

Jinxian Qi¹, Gongfa Li^{1,2,*}, Zhaojie Ju³, Disi Chen³, Du Jiang, Bo Tao^{4,5}, Guozhang Jiang⁶ and Ying Sun^{4,5}

¹ Key Laboratory of Metallurgical Equipment and Control Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China

² Precision Manufacturing Research Institute, Wuhan University of Science and Technology, Wuhan 430081, China

³ School of Computing, University of Portsmouth, Portsmouth PO1 3HE, UK

⁴ Research Center of Biologic Manipulator and Intelligent Measurement and Control, Wuhan University of Science and Technology, Wuhan 430081, China

⁵ Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China

⁶ 3D Printing and Intelligent Manufacturing Engineering Institute, Wuhan University of Science and Technology
ligongfa@wust.edu.cn

Abstract. In order to solve the problem of uneven distribution of picture features and stitching of images, an improved SURF feature extraction method is proposed. Image feature extraction and image registration are the core of image stitching, which is directly related to stitching quality. In this paper, a comprehensive and in-depth study of feature-based image registration is carried out, and an improved algorithm is proposed. Firstly, the Heisen detection operator in the SURF algorithm is introduced to realize feature detection, and the features are extracted as much as possible. Secondly, the characteristics are described by BRIEF operator in the ORB algorithm to realize the invariance of the rotation change. Then, the European pull distance is used to complete the similarity calculation, and the KNN algorithm is used to realize the feature rough matching. Finally, the distance threshold is used to remove the matching pair with larger distance, and then the RANSAC algorithm is used to complete the purification. Experiments show that the proposed algorithm has good real-time performance, strong robustness and high accuracy.

Keywords: Image Stitching, SURF Algorithm, RANSAC.

1 Introduction

Image stitching mainly includes two modules: image registration technology and image fusion technology. Image registration is the basis of image fusion, and image fusion is an important branch and research hotspot of image processing. Image fusion technology extracts and synthesizes information from two or more source images to obtain a more accurate, comprehensive and reliable description of a region or target, and then captures similar regions in the image for stitching and fusion processing.

Complete panoramic image to meet the collection and understanding of image omnidirectional information and target detection, recognition and tracking. In response to this characteristic, image mosaic fusion technology has been widely used in urban transportation, airport scenes, military exercises and medical images. The technology has a basic architecture, and its main contents include image preprocessing, image stitching processing, image stitching evaluation and panoramic image results. The image stitching fusion module is the core technology, so the image fusion algorithm is directly related to the quality. The quality of the fusion after the entire image is stitched.

At present, the more commonly used image fusion algorithms are linear weighted fusion algorithms, such as the direct averaging method, but the pixel values of the overlapping regions are not simply superimposed but are first weighted and then superimposed and averaged, by selecting appropriate weights, the overlapping area can be smoothly smoothed during processing, and the splicing ghost is effectively avoided [1]. Therefore, the method is simple and fast, but is easily interfered by the transition bandwidth, which causes the smoothness of the spliced image and the problem of ghosting, resulting in poor stability of the algorithm. The multi-band fusion algorithm proposed in is based on the idea of decomposing images in different frequencies, using different transition band widths for weighted interpolation processing and then performing fusion processing, although the image quality after stitching and fusion is good [2, 3]. However, pictures with unevenly distributed features cannot be stitched together efficiently, and the algorithm has a large workload and a long calculation time, so the real-time performance is poor.

2 Related Work

For the fast extraction and matching algorithm of image features, many effective image feature extraction and matching algorithms have been proposed by scholars at home and abroad. The representative methods are Harris corner and SIFT feature points [4].

The image registration algorithm based on eigenvalue usually needs the following steps:

- Feature extraction is the most important and difficult step of eigenvalue-based registration algorithm. Feature extraction has a strong impact on the results. There are a lot of information that can be used to represent similarity in images, so there are various processing methods. Common image features include corners, high curvature points, straight lines, contours, closed areas, moment invariants, center of gravity, etc.
- Feature matching, comparing the features extracted from the image to be registered.
- Image registration is achieved by transforming the matched eigenvalues into corresponding images.

In 2006, Bay et al. improved the SIFT algorithm and proposed the SURF (Speeded Up Robust Features) algorithm [5]. The stitching speed was significantly improved. In 2011, Ruble et al. proposed the ORB (Oriented fast and Rotated BRIEF) algorithm based on the algorithm of Brief (Binary Robust Independent Elementary Features) [6]. The stitching speed is further improved than SURF, and the noise resistance is good. In 2004, an algorithm was proposed to automatically implement robust splicing [7]. The algorithm uses Harris to perform feature detection. Although the calculation speed is general, the image mosaic robustness is improved. In 2007, SIFT-based seamless splicing was proposed, which is robust to illumination changes and has scale invariance [8]. In 2010, the SIFT algorithm was used to splicing the remote sensing image of the drone, and good experimental results were obtained [9]. After the ORB algorithm was proposed, domestic research was also quickly integrated, and the image stitching method based on the ORB algorithm was continuously innovated and widely used [10,11]. In 2016, a method to improve the combination of ORB and symmetric matching was proposed [12]. In 2017, an improved algorithm for ORB feature image stitching was proposed [13].

3 Feature Extraction Algorithms Based on Optimized SURF and RANSAC

3.1 The Experimental Steps in This Paper

According to the registration process, this paper introduces the algorithm, and verifies the practicability of the algorithm by comparing the experiments in each link. In the process of image feature extraction, after the research and analysis of SURF and ORB algorithms, this paper introduces Heisen detection operator in SURF to feature detection, and BRIEF operator in ORB algorithm is selected for feature description.

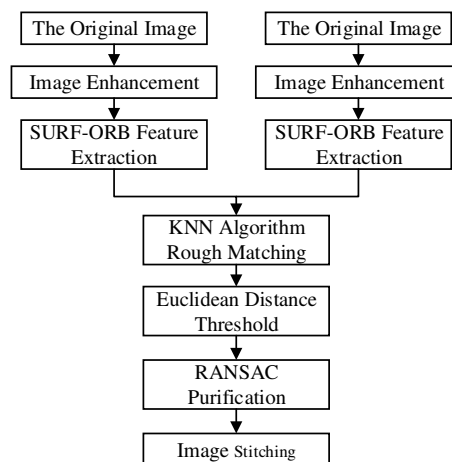


Fig. 1. The block diagram of the algorithm in this paper.

In the process of feature matching, after the research and analysis of and KNN (K-Nearest Neighbors) algorithm, this paper proposes to use Hamming distance to complete the similarity calculation of descriptors, using BF algorithm. The algorithm realizes rough matching of image features. In view of the mismatch of rough matching, an improved RANSAC algorithm is proposed to achieve accurate registration.

3.2 Feature Extraction Algorithm Based on SURF-ORB Algorithm

As an upgraded version of SIFT algorithm, SURF algorithm not only has the characteristics of high accuracy and robustness of SIFT algorithm, but also significantly improves its speed.

Feature Detection Based on SURF. The SIFT algorithm uses the DoG (Difference of Gaussian) operator for feature detection, and the SURF algorithm uses the Heisson detection operator to approximate the DoG operator, and the algorithm speed is improved by the integral graph [14].

Integral Map. The integral map of $p(x, y)$ in image I is the sum of all pixels from P to the origin of the image.

$$I_{\Sigma}(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y') \quad (1)$$

Among then, $I_{\Sigma}(x, y)$ is the integral map and $I(x', y')$ is the original image.

After calculating the integrated image, the sum of any rectangular partial pixels in the original image can be calculated by three additions. Calculate the sum of the pixels in the original image, just add 3 additions according to equation (2).

$$\Sigma = A - B - C + D \quad (2)$$

Among then, Σ is the sum of the rectangular partial pixels, and $A, B, C,$ and D are the pixel values of the four corners of the rectangular portion, respectively.

Therefore, the size of the filter has no effect on the calculation of the integral map. The use of a large-sized filter in the SURF algorithm can significantly increase the speed of the algorithm.

Constructing Scale Space. The SURF algorithm uses the Hessian matrix for feature detection, and uses the matrix determinant to calculate the extreme value as the image feature [15]. This image feature has a patchy structure called a speckle feature. Given an image $I(x, y)$, the Hessian matrix with scale σ is:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (3)$$

Among then, L_{xx}, L_{xy}, L_{yy} is the filtered second derivative.

Theoretically, the Gaussian function is the most ideal function for analyzing the scale space, but in practical applications, the Gaussian function needs to be interrupted and then discretized. In fact, the standard filter effect is sometimes not ideal. The SIFT algorithm uses the DoG operator to approximate the LoG (Laplacian of Gaussian) operator to achieve better results [16]. Therefore, in order to find the extreme

5

point, the SURF algorithm approximates the Hessian matrix, using square matrix filtering as a template to simulate Gaussian filtering.

The SURF algorithm groups the scale spaces when constructing the scale space, and the same group represents the response graph of the convolution of an image with different size square matrix filter templates magnified by the arithmetic progression. Figure 2 shows the pyramid of the constructed scale space. There are 3 groups of 4 layers. The 1st layer of the 1st group is a 9×9 square matrix filter. After that, the size of the first layer of the square matrix filter in each group is in the upper scale space.

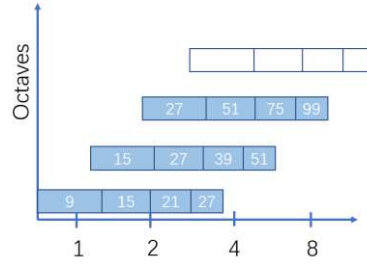


Fig. 2. SURF Algorithm Scale Space.

Finding Extreme Points. After constructing the scale space, the Hessian matrix is then used to calculate the extreme points. The result of convolving the image with the template is denoted by D_{xx}, D_{xy}, D_{yy} , respectively, to replace the L_{xx}, L_{xy}, L_{yy} to obtain a simplified version of the Hessian matrix determinant:

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (4)$$

By increasing the size of the square matrix filter template, the value of the approximate Hessian matrix determinant of the pixel at each scale can be obtained. This value is compared with the values of other pixels in the three-dimensional region of the pixel. If a pixel is an extreme value in its own scale and 26 pixels in its upper and lower scales, then the pixel is a feature point.

Looking for the main direction. The SURF algorithm defines the main direction for the feature points, making the algorithm invariant to the rotation. The SURF algorithm uses Harr wavelet to calculate the gradients in both horizontal and vertical directions. The response template is:

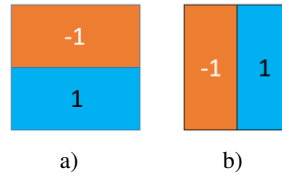


Fig. 3. Harr wavelet response template. a) is X-direction template and b) Y-direction template.

Feature Description Based on BRIEF of ORB. The feature points are used as the center of the square region with a side length of 20σ , which is divided into 16 small

6

squares of the same size, and the small square is divided into 25 smaller areas. The sampling of each small area is shown in the Figure 4.

Thus, the entire square area together form a 64-dimensional feature vector. As an expression of the feature, it embodies the basic information of the feature point, and can perform feature matching instead of the image pixel.

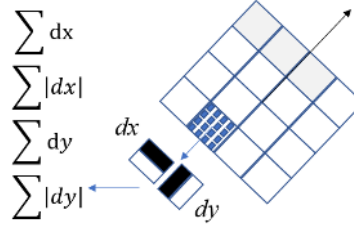


Fig. 4. SURF feature vector.

The ORB algorithm feature description is implemented by the BRIEF descriptor, and the descriptor obtained by the BRIEF is a binary string for easy operation. Take the pixel block P of $S \times S$ size and define the binary test as:

$$\tau(p; x, y) = \begin{cases} 1 & p(x) < p(y) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Where p_x is the gray value of pixel block P on x . The gray point centroid method is used to calculate the main direction of the feature point, and a feature descriptor of the n -dimensional binary string is generated. The transformed feature descriptor is:

$$g_n(p, \theta) = f_n(p) | (x_i, y_i) \in S_\theta \quad (6)$$

3.3 Feature Matching and Purification

Image Feature Matching Based on KNN. The two images will enter the feature matching link after the feature extraction is completed. Fast and accurate matching helps to obtain better transformation matrices and improve image stitching efficiency and quality. The image feature matching completes the search and matching process by calculating the similarity degree of the feature descriptors by a certain measurement method. The similarity calculation is done by using the Euclidean distance and the Hamming distance. The Euclidean distance can be used simultaneously for the case where the feature descriptor is a binary string and a feature vector. Commonly used matching algorithms include KNN algorithm and BF algorithm. In order to ensure the matching accuracy, this paper uses KNN algorithm to achieve image feature matching.

The KNN algorithm calculates the degree of similarity based on the Euclidean distance. The selection of K value is directly related to the effect of the algorithm. The selection of K value is too small, and the calculation of the algorithm becomes more complicated and cannot achieve the effect of approximate expression. If the K value is

7

selected too large, the point with less correlation is also taken into account. Affect the accuracy of the match, so the value of K is critical.

Image Matching Purification Based on RANSAC. RANSAC (Random Sample Consensus) is an abbreviation of the random sampling consensus algorithm. It is essentially a process of searching through continuous iterations and finding model parameters that can contain more interior points. The algorithm retains the "inner point" in the sample (that is, it can satisfy the model constraint within the error range) by setting a certain threshold, and eliminates the "outer point" or "noise point" (the model constraint cannot be satisfied at a given threshold). The algorithm first selects a small number of sub-point sets that are consistent with the original feature points in the feature points, uses the sub-point set to obtain the transformation matrix, and then detects the original feature point set according to the obtained transformation matrix, so that the error range can be obtained. And you can satisfy all the points of the model constraint and record it as an interior point. This process is continuously carried out. When the number of points obtained is large enough and the number does not change substantially, the transformation matrix obtained at this time is used as the final transformation matrix, and the outer points are eliminated, and all the obtained inner points are retained.

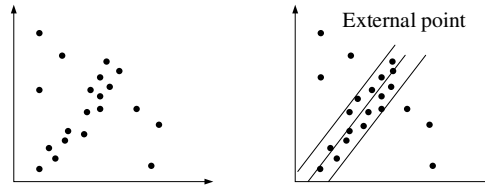


Fig. 5. Straight Line Fit.

Let q be the probability of the inner point in the feature point. At least n pairs of points are needed. p is the probability that the randomly selected feature point is the inner point. The iteration number is calculated as shown in Equation (7):

$$C = \frac{\log(1-p)}{\log(1-q^n)} \quad (7)$$

RANSAC uses a univariate matrix as a transformation model, as shown in Equation (7).

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \quad (8)$$

The implementation steps of RANSAC algorithm can be expressed as follows:

1) Four pairs of matching points are randomly selected from the matching point pairs, and the transformation matrix H is obtained according to the relationship between the pairs of points.

2) Set a threshold, calculate the distance between each point in the original feature set and the corresponding point according to the transformation matrix H obtained in

8

step 1. If the distance obtained is smaller than the given threshold, keep it and record the correct matching points.

3) Repeat the steps mentioned above. The iteration stops when the correct matching points are large enough and there is no change at all. In this case, the collocation point is the maximum internal point allowed in the threshold range, and the transformation matrix obtained is the transformation matrix of the original feature point set.

4 Analysis of Experimental Results

4.1 Data Acquisition and Platform Construction

In the experiment, indoor and outdoor images were collected by color CCD camera as shown in Figure 6, in which group A was outdoor and group B was indoor.

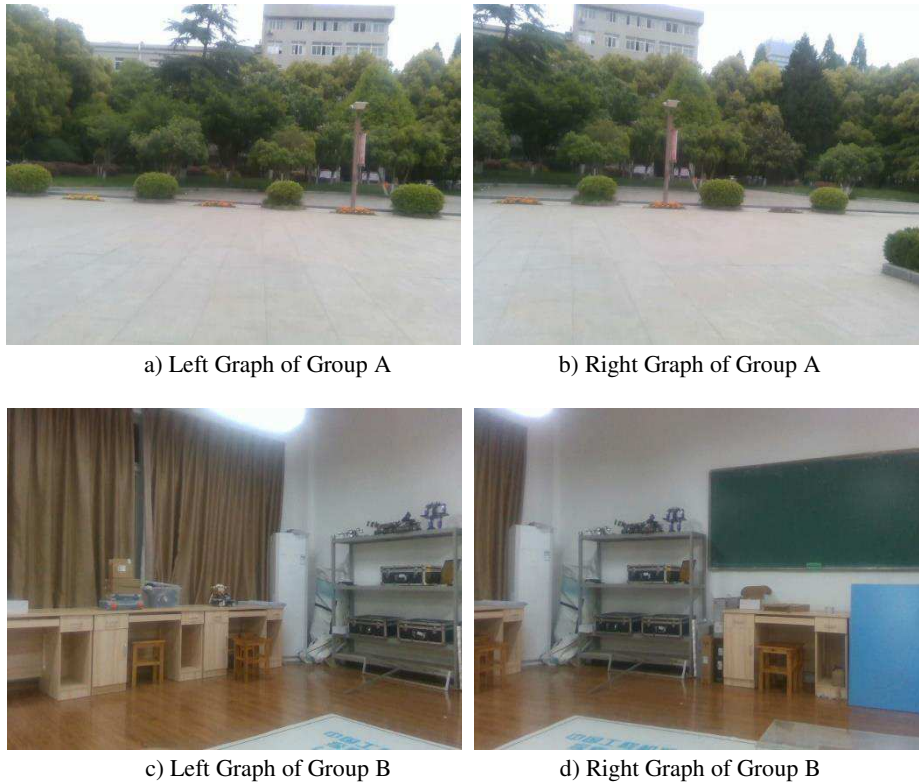


Fig. 6. They are divided into two groups, A and B, showing outdoor and indoor maps respectively.

Picture Mosaic Experiments. As shown in Figure. 7, the feature point map extracted for the SURF algorithm is improved. In order to better display the effect, the color picture is grayed out. It can be seen that the number of feature points extracted by the

algorithm is more and the distribution is more uniform. When feature extraction, the number of feature points, dimensions, time and average number of feature points extracted per second are shown in Table 1.

Table 1 Comparison of feature extraction effects

Group	algorithm	Number of feature points	Time/s
A	SURF	480/367	0.6250
	O-SURF	565/482	0.6563
B	SURF	485/432	0.7031
	O-SURF	607/556	0.6719

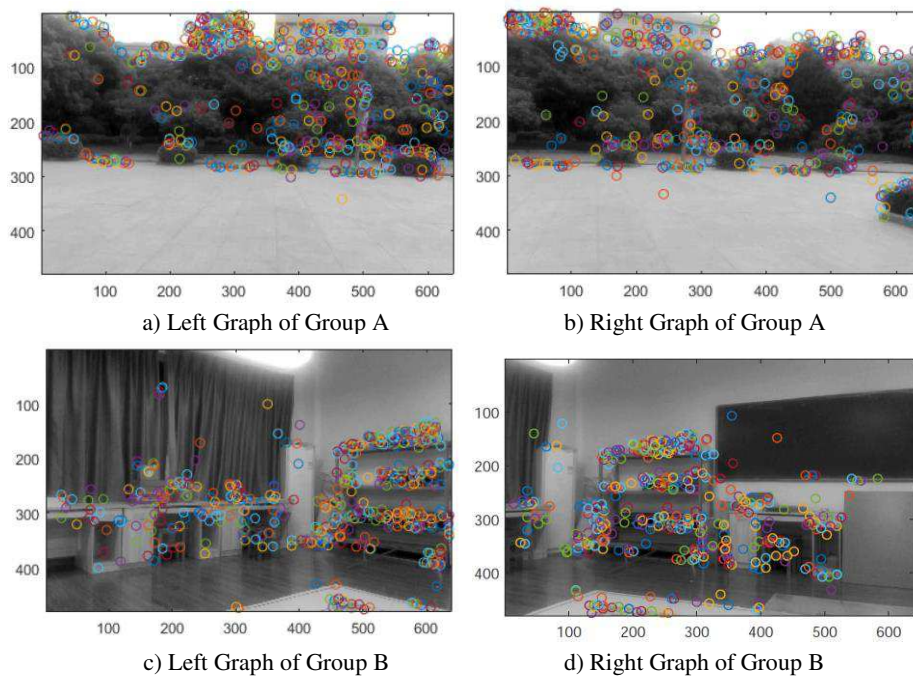


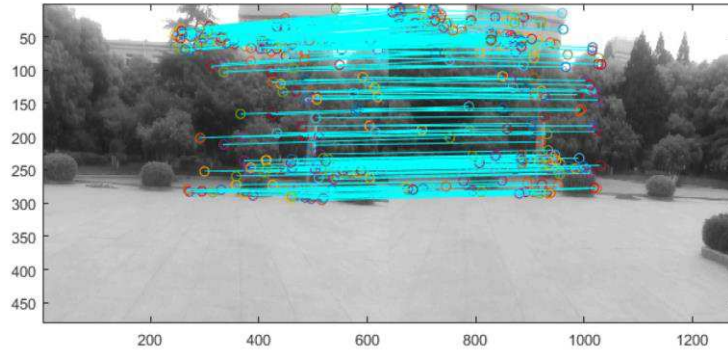
Fig. 7. Improved SURF Feature Extraction

As shown in Figure 8, for the image of feature matching effect after the purification of RANSAC algorithm, it can be seen from the graph that the matching effect after the purification of feature matching is significantly improved compared with that before purification. The connection lines between each feature are horizontally distributed, and there is no crossover. This also shows that the RANSAC algorithm can effectively improve the matching accuracy.

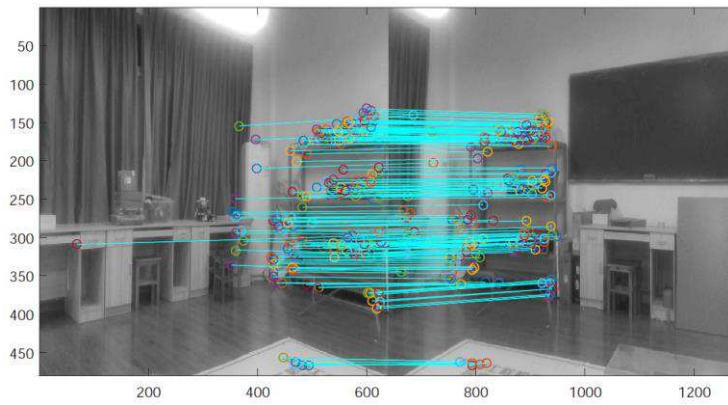
As shown in Figure 9, it is the effect diagram of the final stitching. The picture can be seamlessly docked without significant distortion. According to the splicing effect of the A outdoor and the B indoor, a good splicing can be achieved in some areas

10

without features. This shows that the algorithm has strong generalization ability. For images with uneven feature distribution, a good stitching function can also be achieved.



a) Feature matching effect of group A



b) Feature matching effect of group B

Fig. 8. Feature point matching effect based on RANSAC purification.



a) Group A outdoor stitching effect diagram



b) Group B indoor stitching effect diagram

Fig. 9. Resulting picture after stitching.

5 Conclusion

Image and video stitching technology has become a research hotspot in the field of digital image processing. The stitching technology can broaden the field of view and reduce redundancy while ensuring image resolution. In this paper, the splicing core technology is studied for the problem of uneven distribution of features.

The Heisson detection operator in the SURF algorithm is introduced in feature detection, and the feature description selects BRIEF operator in the ORB algorithm. In the feature matching process, through the research and analysis of KNN algorithm, this paper proposes to use the Euler distance to complete the descriptor similarity calculation, and use the KNN algorithm to perform image feature rough matching. And for the rough matching, there may be mismatches, and the improved RANSAC algorithm is proposed to complete the purification. Before the RANSAC algorithm, the matching of the Euclidean distance greater than a certain threshold is removed, the search range is reduced, the calculation amount is reduced, and the efficiency and accuracy of the feature matching are ensured.

Acknowledgements

This work was supported by Grants of National Natural Science Foundation of China (Grant Nos. 51575407, 51505349, 51575338, 51575412, 61733011), the Grants of National Defense Pre-Research Foundation of Wuhan University of Science and Technology (GF201705) and Open Fund of the Key Laboratory for Metallurgical Equipment and Control of Ministry of Education in Wuhan University of Science and Technology (2018B07).

References

1. Adwan, S., Alsaleh, I., & Majed, R.: A new approach for image stitching technique using dynamic time warping (dtw) algorithm towards scoliosis x-ray diagnosis. *Measurement*, 84, 32-46 (2016).
2. Suk, J. H., Lyuh, C. G., Yoon, S., & Roh, T. M.: Fixed homography-based real-time sw/hw image stitching engine for motor vehicles. *Etri Journal* 37(6), 1143-1153 (2015).
3. Li, G. F., Jiang, D. Zhou, Y. L., Jiang, G. Z., Kong, J. Y. Gunasekaran, M.: Human Lesion Detection Method Based on Image Information and Brain Signal. *IEEE Access* 7, 11533-11542 (2019).
4. An, J., Koo, H. I., Cho, N. I.: Unified framework for automatic image stitching and rectification. *Journal of Electronic Imaging* 24(3), 033007 (2015).
5. Bang, S., Kim, H., & Kim, H.: Uav-based automatic generation of high-resolution panorama at a construction site with a focus on preprocessing for image stitching. *Automation in Construction* 84, 70-80 (2017).
6. Holmes, G., Hale, M., Mcalindon, M. E., Anderson, S.: Pth-185 mapping the gastric mucosal surface: image mosaicking for capsule endoscopy. *Gut* 64(Suppl 1), A490.1-A490 (2015).
7. Sun, Y., Li, C. Q., Li G. F., Jiang, G. Z., Jiang, D., Liu, H. H., Zheng, Z. G., Shu, W. N.: Gesture Recognition Based on Kinect and sEMG Signal Fusion. *Mobile Networks and Applications* 23(4),797-805 (2018).
8. Sjudahl, M., Oreb, B. F.: Stitching interferometric measurement data for inspection of large optical components. *Optical Engineering* 41(2), 403-408 (2015).
9. Johnson, B. G.: Recommendations for a system to photograph core segments and create stitched images of complete cores. *Journal of Paleolimnology* 53(4), 437-444 (2015).
10. Cheng, W. T., Sun, Y., Li, G. F., Jiang, G. Z., Liu, H. H.: Jointly network: a network based on CNN and RBM for gesture recognition. *Neural Computing & Applications* 31(Supplement 1), 309-323 (2018).
11. Lee, C. O., Lee, J. H., Woo, H., Yun, S.: Block decomposition methods for total variation by primal--dual stitching. *Journal of Scientific Computing* 68(1), 273-302 (2016).
12. Berriman, G. B., Good, J. C.: The application of the montage image mosaic engine to the visualization of astronomical images. *Publications of the Astronomical Society of the Pacific* 129(975), 058006 (2017).
13. Chen, D. S., Li, G. F., Sun, Y., Kong, J. Y., Jiang, G. Z., Tang, H., Ju, Z. J., Yu, H., Liu, H. H.: An interactive image segmentation method in hand gesture recognition. *Sensors* 17(2), 253 (2017).
14. Frankl, A., Seghers, V., Stal, C., Maeyer, P. D., Petrie, G., Nyssen, J.: Using image-based modelling (sfm--mvs) to produce a 1935 ortho-mosaic of the ethiopian highlands. *International Journal of Digital Earth* 8(5), 421-430 (2015).
15. Vargiu, L., Rodrigueztomé, P., Sperber, G. O., Cadeddu, M., Grandi, N., Blikstad, V.: Classification and characterization of human endogenous retroviruses; mosaic forms are common. *Retrovirology* 13(1), 7 (2016).
16. Mort, R. L.: Quantitative analysis of patch patterns in mosaic tissues with clonaltools software. *Journal of Anatomy* 215(6), 698-704 (2015).