International Core Journal of Engineering
Volume 7 Issue 8, 2021
ISSN: 2414-1895
DOI: 10.6919/ICJE.202108_7(8).0072

Overview of Image Mosaic Technology based on SIFT Algorithm

Jinsong Zhang[1,2], Lei Tang[1], Zhiyao Sun[1], Jun Liao[1]

[1]School of Information Engineering, Sichuan University of Science & Engineering, Zigong 643000, China;

[2]Sichuan Provincial Key Laboratory of Artificial Intelligence, Sichuan University of Science & Engineering, Zigong 643000, China.

Abstract

Image mosaic technology has developed rapidly in recent years. The core of image mosaic is the processing method of image registration. SIFT algorithm has become a common method of image registration because of its simple and accurate characteristics. This paper reviews a large number of literatures and summarizes the general process of SIFT algorithm, which includes scale space construction, key point detection and elimination, key point direction assignment, key point description and key point matching. After these processing, the target image can be fused and reconstructed into a new image with a wide view angle and more information.

Keywords

Image Mosaic; Image Registration; SIFT Algorithm; Feature Extraction.

1. Introduction

In the early days, image mosaic technology was mainly used in drawing technology, which was used to process the photos from satellite to ground or aircraft aerial photography [1]. These pictures were spliced into a panoramic image, so that more information could be obtained. Although through continuous development, there are still some problems in image mosaic technology. The target image may be interfered by the external environment in the process of acquisition [2]. The noise can not be removed well after noise reduction, which leads to the problems of exposure difference and ghosting, which makes the image registration process difficult.

2. Image mosaic

Image mosaic technology usually has four steps [3]: image preprocessing, image registration, modeling and coordinate transformation, fusion and reconstruction. In these four steps, image registration and fusion reconstruction are the key links. For image fusion reconstruction, the current algorithm has been relatively mature. For image registration, the registration process will directly affect the success and speed of image mosaic.

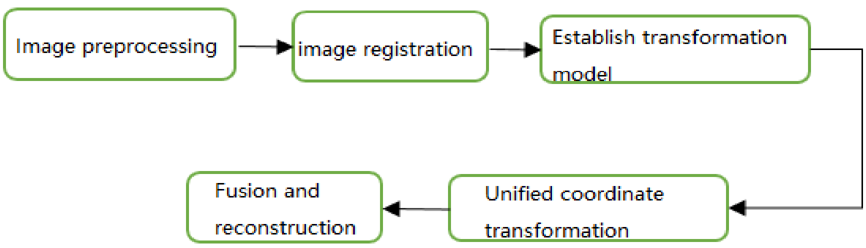

Figure 1. Splicing flow chart

ISSN: 2414-1895 DOI: 10.6919/ICJE.202108_7(8).0072

2.1 Research status at home and abroad

Image mosaic technology originated in foreign countries, and developed rapidly, it has made many achievements. Image mosaic algorithm is mainly used in feature extraction of image registration. LM algorithm [4] (Levenberg Marquardt iterative nonlinear minimization method) proposed by Richard Szeliski is a classical algorithm in this direction, which uses the geometric relationship between images for stitching. In addition, orb algorithm [5] (rotation invariant binary description algorithm), SIFT algorithm (scale invariant feature algorithm), surf algorithm (accelerated robust feature algorithm), etc. Orb algorithm [6] is fast, but it is not scale invariant; SIFT algorithm has a high success rate in key point matching, but the speed of real-time key point extraction is slow; Surf algorithm is fast, but it can't save the image edge and details.

Image mosaic technology appeared late in China, but developed rapidly. Through continuous research and improvement of the algorithm, domestic researchers have also achieved fruitful results. Zheng Yong bin proposed an image matching algorithm combining sift and rotation invariant LBP [7], which greatly improved the operation speed of SIFT algorithm; Liu Jia and Fu Wei ping proposed image matching based on Improved SIFT algorithm [8], adding wavelet change to the algorithm to process the image, so as to further improve the matching time and accuracy. With the continuous emergence of various technologies, image mosaic technology has become more mature. In recent years, SIFT algorithm has gradually become one of the most commonly used methods because of its good mosaic effect.

2.2 Research meaning

Through image mosaic technology, the target image is registered in Gaussian space [9], and then the image is sampled and fused again to obtain a new image with wide-angle scene, complete and high-definition information of each target image. This technology can remove redundant information, compress the information storage of the image, and make the information represented by the image more effective.

For example, in medical images, if you want to accurately observe the pathological parts of the patient's body, you need to obtain high-resolution images to observe the pathological parts. Because the use of larger instruments into the patient's body will cause discomfort, so the camera into the patient's body must be as small as possible, and the area that can be photographed is also very small. Doctors need to obtain a complete medical image to observe the whole lesion, which requires the splicing of multiple images.

2.3 Image registration algorithm

2.3.1 Phase correlation method

The core of this method is fast Fourier transform [10] (FFT), which transforms the target image from time domain to frequency domain, and then calculates the translation vector between the target images to achieve image registration.

2.3.2 Method based on time domain

Firstly, the feature points [11] (such as convex points and concave points) in the two images are found, and then the relationship between the key points in the image space is determined. Through the relationship between the two points, the transformation relationship between the target images is found. Another method is to select one of the template regions in the region of the target image, and then search for the image most similar to the selected template region in other images for matching.

2.4 Comparison of registration methods

The phase correlation method is simple and accurate, but the target image must have a large overlap area. If the overlap area is small, it is easy to cause calculation errors, which makes it difficult to achieve image registration.

The method based on time domain indirectly uses the gray information of the image [12], so it is not affected by the light and dark of the external environment, but it depends on the accuracy of the

ISSN: 2414-1895 DOI: 10.6919/ICJE.202108_7(8).0072

corresponding relationship of key points. This algorithm has high accuracy, but the amount of calculation is relatively large.

3. SIFT algorithm

SIFT (Scale-invariant feature transform) algorithm is proposed by David Lowe [13], and it is summarized and perfected. It is used to detect and describe the local characteristics of the image. It can find the extreme point in the spatial scale, and extract its position, scale and rotation invariants. The implementation steps of SIFT algorithm are generally divided into: scale space construction, key point detection and elimination, key point direction allocation, key point description, key point matching.

3.1 Building scale space

Scale space [14] is a concept and method of simulating human eyes to observe objects in the field of image. For the processing of visual information, the traditional method is to use a single scale, and the scale space is to put the information into a dynamic processing process to simulate the multi-scale features of image data, which is easier to obtain the essential features of the image. In order to preserve more details of the original image, the SIFT algorithm uses Gaussian kernel function to filter [15], and the scale space L(x, y, σ) It can be expressed as Gaussian function G(x, y, σ) Convolution with the target image I (x, y).

$$L(x,y,\sigma)= G(x,y,\sigma)* I(x,y)$$

$$G(x,y,\sigma)= \frac{1}{2\pi\delta^2}\cdot e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

Where (x, y) denotes coordinates, σ represents the spatial factor, σ from small to large represents the change of image from detail feature to general feature.

image I can be represented by an image pyramid after convolution with different Gaussian functions.

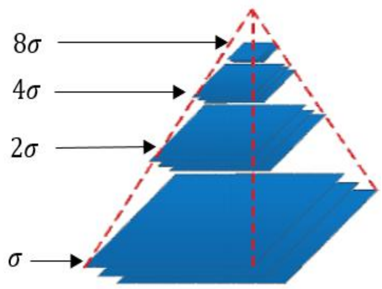

Figure 2. Gaussian image pyramid

The scale space operator Difference of Gauss(DoG) can be expressed as:

$$D(x,y,\sigma)= (G(x,y,k\sigma)- G(x,y,\sigma))* I(x,y)$$
$$= L(x,y,k\sigma)- L(x,y,\sigma)$$

Where D (x, y, σ) is the difference between two adjacent scale images, it means the difference between the adjacent Gaussian images of each order in Figure 1.

3.2 Key point detection and elimination

3.2.1 Key point detection

The key point is the extreme point in the scale space DoG [16]. To detect whether a point is a key point, it is necessary to compare it with the surrounding 8 points and the 18 points above and below the scale space in image domain and scale domain, so as to ensure that the comparison is carried out in the whole space. After comparison, if a point is the minimum or maximum of 26 points, this point is the key point.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202108_7(8).0072

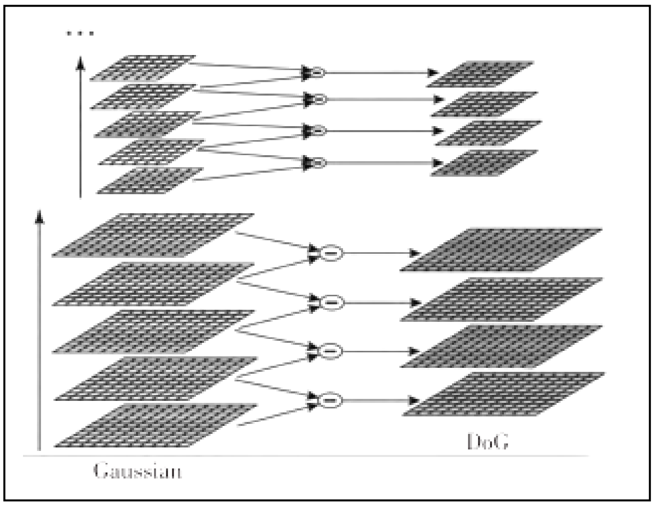

Figure 3. Gaussian difference pyramid

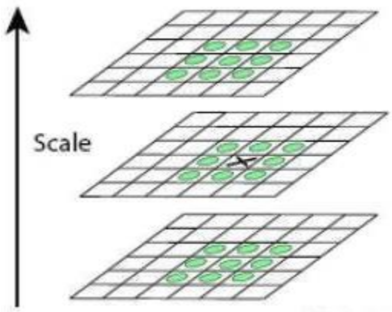

Figure 4. Key point detection

Because the detection of a key point requires three levels of scale space, it will lead to the first level and the last level of space can not carry out key point detection. The general method is to generate two images with Gaussian blur [17] and place them at the beginning and end of the scale space to ensure that the extreme points in the first layer and the last layer are not lost.

3.2.2 Eliminate the key points of disqualification

Because the values in the scale space dog are sensitive to noise and edge, it is necessary to select the key points, remove the unqualified points [18] (low contrast points and unstable edge points), and the remaining points are the feature points after accurate positioning. The steps are as follows.

(1) Remove low contrast points

The scale space function D (x, y, σ)open with Taylor formula:

$$D(X) = D + \frac{\partial D^T}{\partial X} + \frac{1}{2} \cdot X^T \cdot \frac{\partial^2 D}{\partial X^2} \cdot X$$

$$X = (x, y, \delta)^T$$

$$\frac{\partial D}{\partial X} = (\frac{\partial D}{\partial x}, \frac{\partial D}{\partial y}, \frac{\partial D}{\partial \delta})^T$$

$$\frac{\partial^2 D}{\partial X^2} = \begin{bmatrix} \frac{\partial^2 D}{\partial x^2} & \frac{\partial^2 D}{\partial xy} & \frac{\partial^2 D}{\partial x\delta} \\ \frac{\partial^2 D}{\partial yx} & \frac{\partial^2 D}{\partial y^2} & \frac{\partial^2 D}{\partial y\delta} \\ \frac{\partial^2 D}{\partial \delta x} & \frac{\partial^2 D}{\partial \delta y} & \frac{\partial^2 D}{\partial \delta^2} \end{bmatrix}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202108_7(8).0072

Key point $\hat{X}$ can be expressed as:

$$\hat{X} = -\frac{\partial^2 D^T}{\partial X^2} \cdot \frac{\partial D}{\partial X}$$

$$D(\hat{X}) = D + \frac{1}{2} \cdot \frac{\partial^2 D}{\partial X^2} \cdot \hat{X}$$

According to the literature, if $|D(\hat{X})| \geq 0.03$, which means that the key point is not a low contrast point and can be retained; On the contrary, if $|D(\hat{X})| \geq 0.03$, the key point will be removed.

(2) Remove unstable edge points

If the Gauss operator is not well defined, the maximum point will have a larger principal curvature on the edge of horizontal direction of scale space [19], and the main curvature will become smaller on the edge of vertical direction. This will have a great impact on the registration of images.

The principal curvature can be expressed as:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

H is Hessian matrix [20], and the principal curvature of function D and the eigenvalue of matrix H are expressed as proportional relations, and α is the larger eigenvalue in the matrix, β is a smaller eigenvalue. The following relationships can be obtained:

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta$$

$$Det(H) = D_{xx} \cdot D_{yy} - (D_{xy})^2 = \alpha \cdot \beta$$

set up $\alpha = r\beta$, There is the following derivation:

$$\frac{Tr(H)^2}{Det(H)} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(1+r)^2}{r}$$

Obviously, the ratio of trace square of H matrix to determinant is only related to r. According to the analysis, when $\alpha = \beta$ in the paper [21] of David Lowe, r is set to 10. Therefore, if you want to verify whether the key points are qualified, you only need to verify whether the following formula is true.

$$\frac{Tr(H)^2}{Det(H)} < \frac{(1+r)^2}{r}$$

3.3 Key direction assignment

After the above steps, the key point still does not have the property of rotation invariance. In this case, the gradient distribution [22] of each point field can be used to calculate the direction of the key points, so that the direction of the key points can be rotated unchanged. The gradient amplitude and gradient direction of the sampled pixels are as follows:

$$m(x,y) = ((L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2)^{\frac{1}{2}}$$

$$\theta(x,y) = arctan\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}$$

In the above formula, L(x, y) is the scale space value of the key points. Generally, the histogram method [23] is used to count the gradient direction. Samples are taken in the neighborhood window with the key point as the center. The sampling range is 0 ~ 360 degrees, in which there is a column every 10 degrees, and each square column represents the influence degree of the sampling points in the range of 10 degrees on the direction of the key points. The column with the largest peak value in the histogram can be regarded as the main direction of the gradient, it means the direction of the key points.

After a series of operations, each key point has position, scale and direction, through which a feature area can be determined.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202108_7(8).0072

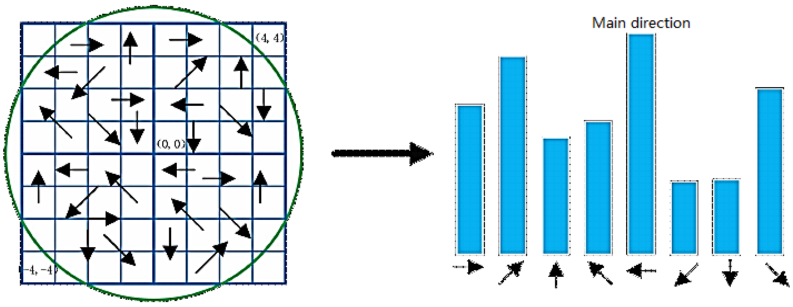

Figure 5. Gradient histogram

3.4 Key point description

In order to ensure the rotation invariance of the key points [24], the direction of the coordinate axis should be consistent with the direction of the key points. Take the key point as the center, select an 8 × 8. As shown in the figure below:

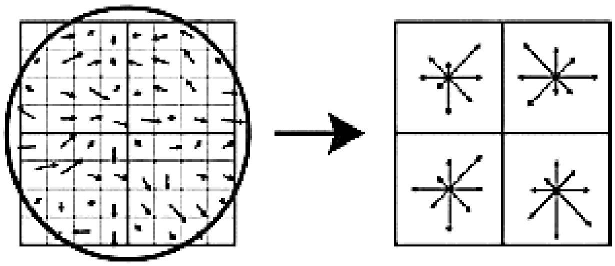

Figure 6. Feature description

The image on the left in Figure 5 shows the gradient of the image. Each small square represents a pixel in the key area, the arrow in the small square represents the gradient direction of the pixel point, and the length of the arrow represents the influence of this direction on the direction of the key point, while the circle represents the Gauss weighted range [25] (the gradient of pixels is closer to the key point, This shows that the greater the impact of this pixel on the key points. Calculate the histogram of gradient directions in 8 directions on each 4×4 square, and add up the length of arrows in each gradient direction to get a seed point [26], as shown in the picture on the right side of Figure5. A key point has 4 seed points, each seed point contains 8 directions information, so 32 data can be obtained, and a 32-dimensional feature vector can be generated. After normalization [27], the key points can be described.

$$W = (w_1, w_2, \cdots w_n)$$

$$L = (l_1, l_2, \cdots l_n)$$

$$l_j = \frac{w_j}{\sqrt{\sum_{i=1}^{n} w_i}} \qquad j = 1,2, \cdots n$$

Where W is the eigenvector before the gauge and L is the eigenvector after the gauge. David Lowe's paper proposes that 16 seed points are used to describe the key points [28], 128 data can be obtained and a 128 dimension eigenvector can be generated. After this change, the influence of scale change and rotation is eliminated, and then the gradient amplitude is normalized, so that the influence of light change can be reduced.

3.5 Key point matching

The key point description subsets are established for template graph and real-time graph respectively, and the key point description in the two sets is used for matching recognition [29]. There are two matching methods: one is the exhaustive method, which can identify the key points and other key points except itself. This method takes a lot of time.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202108_7(8).0072

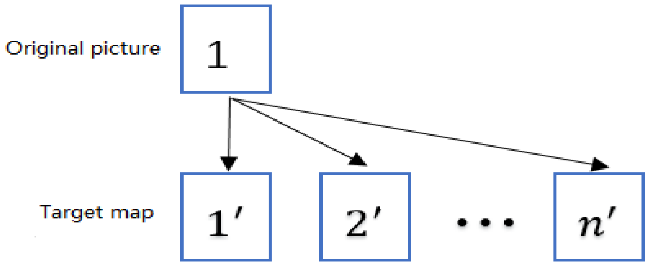

Figure 7. Exhaustive matching

Another method is the k-d tree algorithm [30-32]. This method searches through the data structure of the k-d tree, takes the key points of the image as the search target, finds the original image feature points closest to the processing image key points and the second closest original image key points, and then carries out recognition and matching.

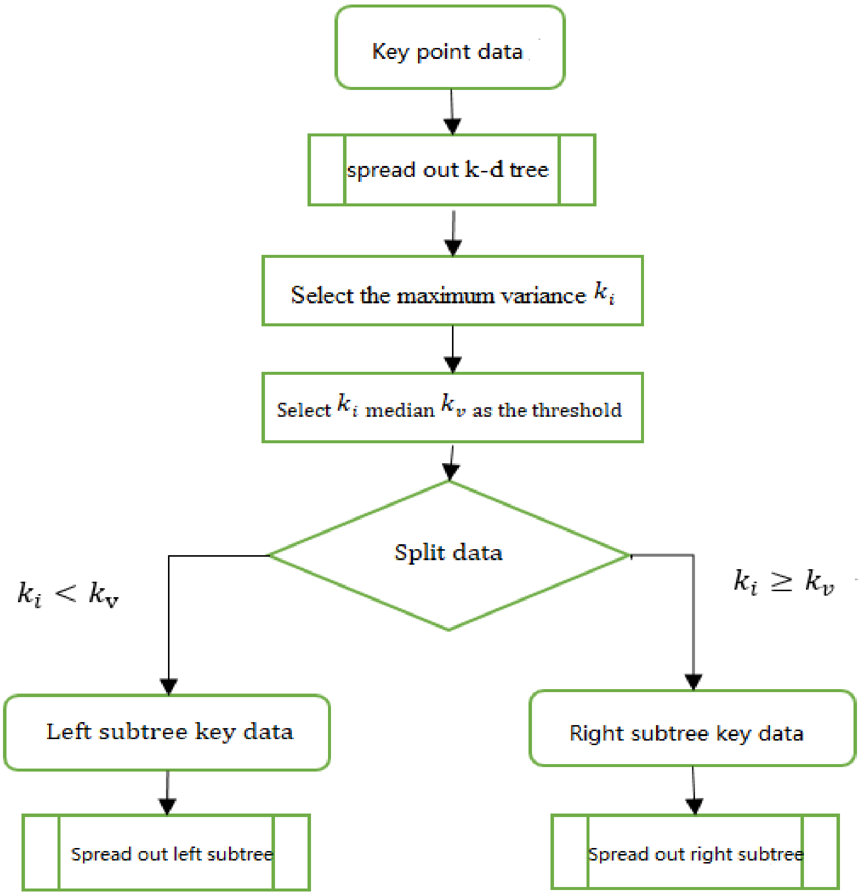

Figure 8. K-d tree flow chart

4. Conclusion

The key link of image mosaic is registration. This paper summarizes and analyzes the registration methods using SIFT algorithm, and concludes that the algorithm has good stability, invariance and discrimination, and also has multi quantity, high speed and scalability. However, the algorithm also has some shortcomings, such as low real-time performance, and the smooth edge of the target can not accurately extract the key points, which need further research and improvement. I believe that sift algorithm after continuous improvement and development, the future image mosaic technology will be more and more efficient.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202108_7(8).0072

Acknowledgments

Project of Sichuan Department of Science and Technology (2020JDJQ0075, 2021YFSY0027);
Project of Sichuan Provincial Key Laboratory of Artificial Intelligence (2020RZY02).

References

[1] Zhu Dalong, Yi Yunshan. Research on license plate location and parallelization based on SIFT algorithm[J]. Network Security Technology and Application, 2020(11):67-70.

[2] Sun Yanli, Zhou Wei, Zhang Yazhou, Liu Ye. Airborne reconnaissance image stitching algorithm based on SIFT[J]. Electronic Measurement Technology, 2020, 43(19): 112-116.

[3] Establishing'proof of concept' for a social cognition group treatment program (SIFT IT) after traumatic brain injury: two case studies.[J]. Brain injury,2020.

[4] Ding Sunan, Zhang Qiuju. Image matching method based on improved SIFT algorithm[J]. Sensors and Microsystems, 2020, 39(10):45-47+50.

[5] Gao Yu, Xu Zhongyu. SIFT block feature matching fundus image stitching [J]. Changchun University of Technology, 2020, 41(04): 395-398.

[6] He Huiyang, Han Jun. Research on image mosaic technology based on SIFT algorithm [J]. Automation and Instrumentation, 2020, 35(02): 57-60+79.

[7] Zheng Yongbin, Huang Xinsheng, Feng Songjiang. Image matching algorithm combining SIFT and rotation invariant LBP[J]. Journal of Computer Aided Design and Graphics, 2010, 22(02): 286-292.

[8] Liu Jia, Fu Weiping, Wang Wen, Li Na. Image matching based on improved SIFT algorithm[J]. Chinese Journal of Scientific Instrument, 2013, 34(05): 1107-1112.

[9] Earth Observations and Remote Sensing; Data on Earth Observations and Remote Sensing Discussed by Researchers at Zhengzhou University of Light Industry (Generating and Sifting Pseudolabeled Samples for Improving the Performance of Remote Sensing Image Scene Classification)[J]. Journal of Engineering, 2020.

[10] Zhang Heng. Research on text image mosaic technology based on SIFT algorithm [D]. Harbin Institute of Technology, 2020.

[11] Yang Zhifang, Yuan Jiakai, Huang Yaoyao. Indoor panoramic image stitching based on SIFT algorithm [J]. Automation and Instrumentation, 2020, 35(03): 58-62+87.

[12] Liu Jie, You Pinhong, Zhan Jianbin, Liu Jinfeng. Improved SIFT fast image stitching and ghosting optimization [J]. Optics and Precision Engineering, 2020, 28(09): 2076-2084.

[13] Hayder Dibs, Shattri Mansor, Noordin Ahmad, Biswajeet Pradhan, Nadhir Al-Ansari. Automatic Fast and Robust Technique to Refine Extracted SIFT Key Points for Remote Sensing Images[J]. Journal of Civil Engineering and Architecture, 2020,14(6).

[14] Du Yapeng. Research on key technologies of image stitching [D]. University of Electronic Science and Technology of China, 2020.

[15] Li Qin. Image mosaic method based on feature extraction [J]. Modern Computer, 2020(09): 69-72.

[16] Yu Jun. Research on the underwater image stitching algorithm based on SIFT [D]. Hangzhou Dianzi University, 2020.

[17] Gao Jing. Ship image stitching algorithm based on improved SIFT algorithm [J]. Ship Science and Technology, 2020, 42(02): 193-195.

[18] Ying Lan, Gaoyang Chen.Research on Workpiece Image Registration based on SIFT Algorithm[J]. International Core Journal of Engineering,2020,6(6).

[19] Tang Zetian, Ding Zhao, Zeng Ruimin, Zhong Minzhe, Zhu Dengwei, Wang Yuhao, Wang Yang, Yang Chen. SIFT quantum image mosaic based on dynamic threshold and global information[J]. Journal of Quantum Electronics, 2020, 37(01): 22- 28.

[20] Fu Xiaohui, Zhao Dequn, Zhang Junjie, Deng Qianhua. Design of low-altitude remote sensing image stitching system based on improved SIFT[J]. Foreign Electronic Measurement Technology, 2019, 38(08): 105-109.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202108_7(8).0072

[21] Han Weidong, Min Shitong. Image mosaic algorithm based on structure information[J]. Computer System Applications, 2019, 28(10): 120-129.

[22] Yang Siyan, He Guoqi, Liu Ruyi. Large scene video stitching algorithm and optimization based on SIFT algorithm[J]. Computer Science, 2019,46(07):286-291.

[23] Cao Guihong, Ouyang Ying, Guo Qing. Research on fabric image stitching technology based on SIFT algorithm[J]. Computer Knowledge and Technology, 2019, 15(20): 201-202+210.

[24] Pei Hongxing, Liu Jinda, Ge Jialong, Zhang Bin. Overview of image mosaic technology[J]. Journal of Zhengzhou University (Science Edition), 2019, 51(04):1-10+29.

[25] Andrew Zalesky, Tabinda Sarwar, Ramamohanarao Kotagiri.SIFT in pathological connectomes: Follow-up response to Smith and colleagues[J]. Magnetic Resonance in Medicine,2020,84(5).

[26] Zhang Fugui, Fu Jing, Meng Hui, Sun Xiaohe. A rapid splicing method of UAV smoke image based on SIFT algorithm[J]. Journal of Guizhou University (Natural Science Edition), 2019, 36(03): 62 -68.

[27] Liu Pengfei. Research on image registration method based on SIFT [D]. Harbin Institute of Technology, 2019.

[28] Feng Yanrong. Application of an improved SIFT algorithm in image stitching algorithm [J]. Computer Products and Circulation, 2019 (06): 86.

[29] Guodong. Research on Text Image Stitching Technology and Its Application [D]. North China University of Technology, 2019.

[30] Lu Qi. Research and evaluation of image stitching algorithm based on improved SIFT[D]. Shanghai Ocean University, 2019.

[31] Yao Wei, Hu Hong. Adaptive Rotating Image Seamless Mosaic Algorithm Based on SIFT[J]. Journal of Nanjing University of Science and Technology, 2019, 43(02): 250-254.

[32] Yu Shengju. Implementation of image stitching system based on feature matching[D]. Huazhong University of Science and Technology, 2019.