

Indoor Scene Splicing Based on Genetic Algorithm and ORB

Tao Zhang^{1,*} and Yi Cao²

¹School of Mechanical Engineering, North China University of Water Conservancy and Hydroelectric Power, Zhengzhou, 450045, China

²Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, 9B 3P4, Canada

*Corresponding Author: Tao Zhang. Email: ztncwu@126.com

Received: 10 January 2022; Accepted: 10 February 2022

Abstract: The images generated by the image stitching algorithm have false shadow and poor real-time performance, and are difficult to maintain visual consistency. For this reason, a panoramic image stitching algorithm based on genetic algorithm is proposed. First, the oriented fast and rotated brief (ORB) algorithm is used to quickly perform detection and description of feature, then the initial feature point pairs are extracted according to the Euclidean distance for feature point rough matching, the parallelism of genetic algorithm is used to optimize the feature point matching performance. Finally, the PROSAC algorithm is used to remove mismatched point pairs and get the transformation matrix to complete the image stitching.

Keywords: Indoor scene; image stitching; false shadow; oriented fast and rotated brief; genetic algorithm

1 Preface

Image stitching technology is a process of transforming multiple different images containing the same object into the same image with a larger field of view through a certain conversion algorithm. It is one of the research topics in the field of computer vision and graphics. Image stitching technology is widely used in many fields [1–4] such as security monitoring, battlefield situation awareness, virtual reality technology and medical images.

Image stitching mainly includes feature extraction, feature matching, removing mismatched points, and graphics conversion. The most widely used existing technologies are scale invariant feature transform (SIFT) algorithm [5,6] and speeded up robust features (SURF) algorithm [7–9]. The SIFT algorithm has good scale invariance, and good anti-interference ability for illumination, noise, rotation, scaling, and occlusion. However, the detection of feature points uses the scale space extreme value, and the feature descriptor uses the gradient histogram method. It leads to a large amount of calculation and low timeliness of the algorithm. The SURF algorithm, which based on the Hessian matrix, reduce the calculation time by acceleration of feature points detection in extreme space and feature descriptor dimensionality reduction, but still cannot meet the real-time requirements. In addition to traditional artificially annotated image matching methods, current image processing methods based on deep learning are also widely used to



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

solve image matching problems. Deep learning methods can extract more subtle feature information, but require complex structures and more computing time. Aiming at the real-time requirements of feature detection, Ethan et al. proposed the oriented fast and rotated brief (ORB) algorithm in 2011 [10], which is invariant to rotation changes and has been greatly improved the speed of feature extraction. In terms of feature matching, the classic random sampling consensus (RANSAC) algorithm has the problems of large amount of calculation and low accuracy. Literature [11] sorts the matching points obtained after rough matching according to the quality factor, and then samples the increasing interior points set, which saves the amount of calculation and improves the running speed. Literature [12] improved the accuracy of image registration by constructing block random sampling detection.

Therefore, this article proposes an image stitching algorithm based on ORB features. First, ORB algorithm is used for feature extraction, and then the parallel feature of genetic algorithm is used to achieve fast rough matching of features. Finally, PROSAC algorithm is combined to remove mismatched point pairs to complete the precise matching and calculate the transformation matrix.

2 Image Stitching Algorithm

Generally speaking, image stitching algorithms include feature point extraction, image matching and image transformation. The basic flow of the image stitching algorithm proposed in this paper is shown in Fig. 1.

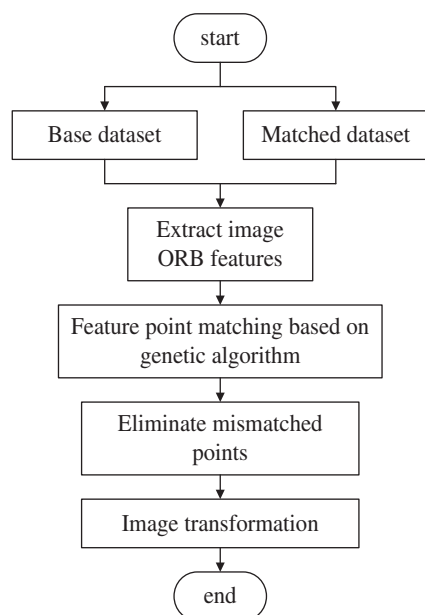


Figure 1: Image stitching process

2.1 Feature Extraction Algorithm

ORB algorithm [13–15] is an improved algorithm based on the fusion between FAST [13] feature point detection and BRIEF [14] feature description. The algorithm obtains the feature information of the image, uses the binary string to describe the feature information, and maintains invariance to changes in target translation, rotation, viewing angle and illumination.

FAST feature point detection algorithm is to select a pixel p from the picture as the center of the circle, and then make a circle with 3 pixels as the radius, set the threshold t , and compare the gray value of pixel

points p and the 16 pixels on the circle in turn. If the gray values of the n pixels appearing on the circle are all greater than $I(p) + t$ or less than $I(p) - t$, it can be judged that the selected pixel p is a feature point, where $I(p)$ represents the gray value of the pixel p , usually 9 or 12.

In order to make the obtained feature points have rotation invariance, the gray-centroid method is used to obtain the main direction of the feature points. First, it takes the feature point in the fast corner point set as the center, and calculate the gray centroid in the neighborhood of the feature point. Then a vector is constructed with the feature point as the start point and the corresponding gray centroid as the end point, and the direction of the vector can be judged as the main direction of FAST feature point.

The specific steps are:

(1) In an image block B , the local image block moment is defined as:

$$m_{pq} = \sum_{x,y \in B} x^p y^q I(x, y) \tag{1}$$

where $I(x, y)$ is the gray value of (x, y) .

(2) The centroid of the image calculated by the moment is:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \tag{2}$$

(3) Connect the geometric center O and the centroid C of the image block to obtain the direction vector \overrightarrow{OC} , and then define the direction angle of the feature point as:

$$\theta = \arctan(m_{01}, m_{10}) \tag{3}$$

Through the above method, the FAST corner point has the scale feature and the rotation feature, and its robustness is improved. The feature description in the ORB algorithm uses BRIEF, which is a binary descriptor whose description vector is composed of 0 and 1. Gaussian filtering is performed on the image first to reduce noise interference. In the pixel area centered on the feature point, random Select a point pair for gray value comparison, and perform the following binary assignment:

$$\tau(p; x, y) = \begin{cases} 1, & p(x) < p(y) \\ 0, & p(x) \geq p(y) \end{cases} \tag{4}$$

where $p(x)$ and $p(y)$ represent the pixel gray value of x and y separately, n binary code strings obtained by comparison to form a n -dimensional binary vector:

$$f_n(p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i, y_i) \tag{5}$$

In order to make the feature descriptor have directionality, the centroid direction information of the feature point is added to the descriptor. For a feature set containing n binary tests at the location (x_i, y_i) , the one forming $2 * n$ matrix Q is:

$$Q = \begin{pmatrix} x_1, \dots, x_n \\ y_1, \dots, y_n \end{pmatrix} \tag{6}$$

The characteristic point pair matrix, constructed by the direction θ of the characteristic point and the corresponding rotation matrix R_θ , is defined as:

$$Q_\theta = R_\theta Q \quad (7)$$

The feature descriptor with direction is:

$$g_n(p, \theta) = f_n(p) | (x_i, y_i) \in S_\theta \quad (8)$$

Greedy search is used to perform irrelevant tests on 256 pixel pairs to form the final feature descriptor BRIEF with direction.

2.2 Feature Point Matching and Optimization

The principle [15,16] of image feature point matching is to select a feature point in the reference image, and then search for feature points in the image to be matched according to a certain similarity measurement criterion. If the two feature points meet the criterion, these two points are considered to be matched. The commonly used method is to judge the similarity of the key points of the two images by using the Euclidean distance [17]. The matching process is: first select a feature point in the reference image, search in the image to be registered in turn, and find out the Euclidean distance from it. The first two closest key points; then divide the next closest distance by the closest distance. If the quotient is less than a certain threshold, the pair of points is considered to be a match. The Euclidean distance between two feature points is expressed as

$$d(i, j) = \sqrt{\sum_{m=1}^p [d_i(m) - d_j(m)]^2} \quad (9)$$

where d_i is the feature vector of the feature point in the reference image; d_j is the feature vector of the feature point in the image to be registered. According to the matching principle, it can be known that for the specified feature points p_1 in the reference image, to find the corresponding best matching point p_2 , it is necessary to calculate the Euclidean distance with the feature vectors of all the feature points in the image to be registered, and then obtain the minimum value. The algorithm is complicated High degree. From Eq. (9), it is known that only a pair of feature points must be matched to the p-th sub-square difference calculation, so it is necessary to optimize this process.

Genetic algorithm [18,19] (GA) is a general method for solving search problems, and it is a parameter optimization model that simulates genetic selection and natural elimination in the principle of biological evolution. Its search strategy and optimization algorithm are carried out independently and have little correlation with external information. It uses fitness function value to discriminate individuals, so it is widely used in various fields. The image matching optimization process based on genetic algorithm is as follows:

- (1) Encode the chromosome and initialize the population

Because the feature vector of the image feature point is obtained by accumulating the gradient values of the pixels in the neighborhood of 8 directions, the individual coding adopts binary number coding.

- (2) Construct fitness function

According to the algorithm principle, the Euclidean distance can be set as the evaluation function, and the fitness determines the criterion for whether a chromosome individual can be retained. In order to achieve the optimization effect, it is necessary to reduce the residual between the prediction and the expected value as much as possible. Therefore, the error matrix norm between the prediction and the expected value of the prediction sample is selected as the output of the fitness function.

(3) Regenerate chromosomes

According to the fitness function, the fitness value of each chromosome is calculated. Chromosomes with low fitness are selected as genetic seeds; chromosomes with high fitness are eliminated. Then some genes in the chromosome are exchanged, and a certain bit is randomly selected among these individuals for mutation operation to form a new generation of population.

(4) Decoding to obtain the best matching point

Decoding the chromosome with the smallest fitness value in the last generation is the coordinate of the matching point. The pixel point $p1$ corresponding to the pixel point $p2$ with the smallest Euclidean distance is the matching point with $p1$.

3 Image Transformation

The process of image transformation is to first find the transformation matrix of the reference image and the image to be registered, and then transform the image to be registered into the coordinate system of the reference image according to the transformation matrix to complete the image transformation. The existence of mismatched points will affect the quality of image stitching. In order to overcome the inaccuracy of the conversion matrix, RANSAC algorithm needs to be improved. PROSAC algorithm adopts a semi-random parameter estimation method [18,19], and ranks the sample points from the image matching according to the quality, and then extracts samples from the high-level sample point set to estimate the model. The algorithm improves the accuracy of the model while reducing the number of model iterations, greatly reducing the amount of calculation.

This paper considers to sort the matching points obtained after feature matching according to the results of the initial matching, and then uses PROSAC algorithm to solve the initial transformation matrix. The specific steps are:

(1) Set the initial value, upper limit of the number of iterations, threshold of the number of interior points, and error range. The initial value is set to 0.

(2) Arrange the matched feature points in descending order of quality, and then select n data with higher quality.

(3) Remove any m sample data from the extracted n sample data, and calculate the number of Euclidean distances between feature point pairs, which are less than the set error threshold, as the number of interior points.

(4) If the calculated number of interior points is greater than the set threshold, the number of interior points will be updated. Otherwise the number of iterations will be increased by 1, step (2) will be returned, and the above operations will be repeated until the conditions are met.

4 Experimental Results and Analysis

In order to verify the effectiveness and practicability of the algorithm in this paper, a splicing experiment of actual scenes is carried out, and the algorithm in this paper is compared with the existing splicing algorithm. The algorithm hardware environment is a high-speed visual processor (CPU i9-10900X, 3.7 GHz, 4.5 GHz Turbo, memory 64 GB DDR4, 32-bit Windows operating system), and the software platform is implemented by Matlab R2020a. The simulation experiment chooses to shoot multiple sets of images in the classroom and the laboratory, and compare the effects of the SIFT algorithm, the SURF algorithm, the ORB algorithm, and the algorithm in this paper. Parts of the image set is shown in [Fig. 2](#).



Figure 2: Test image set

This article compares the matching accuracy, relative error and feature point extraction time.

(1) In order to measure the accuracy of extracting feature points, the correct matching rate (CMR) is used for evaluation to detect all feature point pairs in two pictures. If the distance between the feature point pair descriptor sub-vectors is less than a certain threshold (generally set to 0.7), the point pair is a matching point, and then PROSAC is used to further match the matching point pair set. The formula is:

$$CMR = \frac{n_c}{N} \quad (10)$$

where CMR represents the ratio of the number of correct matching points to the logarithm of all matching points, n_c indicates the number of correct matching points. N is the number of all matching points. The larger the CMR value, the better the matching performance. Comparison of CMR is shown in [Tab. 1](#).

Table 1: Comparison of CMR

Algorithm	Data1	Data2	Data3
SIFT	93.1%	91.4%	92.4%
SURF	92.5%	91.7%	89.3%
ORB	94.2%	92.5%	90.2%
Improved algorithm in this paper	97.3%	96.5%	98.7%

(2) In order to reflect the effect of image splicing more accurately, the root mean squared error (RMSE) is introduced as evaluation standard. RMSE represents the expected value of the square of the image error, is expressed as

$$RMSE = \sqrt{\frac{1}{m * n} \sum_i^{m*n} (x_i - y_i)^2} \quad (11)$$

where x and y represent the original image and the stitched image, respectively, and m and n are the image size. Comparison of RMSE is shown in [Tab. 2](#).

Table 2: Comparison of RMSE

Algorithm	Data1	Data2	Data3
SIFT	0.15	0.11	0.13
SURF	0.13	0.12	0.16
ORB	0.069	0.078	0.067
Improved algorithm in this paper	0.012	0.023	0.031

(3) Feature point extraction time

In order to measure the speed of feature point extraction, this paper extracts 1000 feature points from the reference image and the image to be spliced. [Tab. 3](#) compares the feature point extraction time (ms) of various algorithms. The shorter the time, the faster the feature point extraction speed.

Table 3: Comparison of feature point extraction time

Algorithm	Data1	Data2	Data3
SIFT	1.15	1.11	1.13
SURF	0.63	0.72	0.66
ORB	0.13	0.38	0.27
Improved algorithm in this paper	0.12	0.23	0.21

The final stitching effect is shown in [Fig. 3](#). Compared with the traditional algorithm, the algorithm in this paper has a great advantage in the registration time, and the shortened time is proportional to the number of feature points. This is because the traditional SIFT algorithm first selects a feature point in the target image, and then calculates the Euclidean distance of the feature vector with each feature point in the image to be matched. All feature points are sorted after the Euclidean distance calculation is completed. Feature points can be matched only after obtaining the shortest Euclidean distance and the second shortest distance. If the number of feature points of the target image is m and the number of feature points of the image to be registered is n, the algorithm complexity is O(mn). With the introduction of genetic algorithm, the feature points with the smallest Euclidean distance can be searched in all feature point spaces at the same time, and the algorithm complexity is O(m). Because genetic algorithm has parallelism and does not fall into local optimum, the algorithm in this paper increases the accuracy of matching on the basis of improving the matching speed.

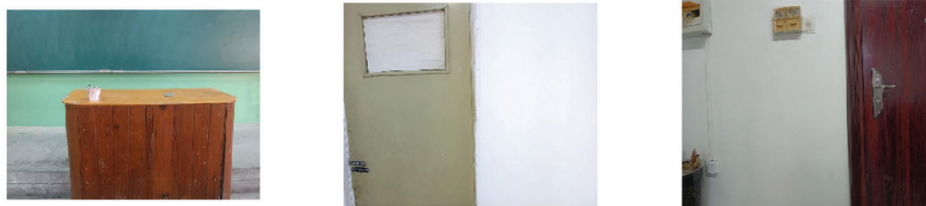


Figure 3: Splicing result

5 Conclusions

This paper mainly studies the indoor scene image mosaic technology based on ORB and genetic algorithm. The algorithm uses the image registration technology based on ORB feature point matching, and extracts the initial feature point pairs according to the Euclidean distance to perform rough matching of feature points. Aiming at the problem of long time-consuming and high false matching rate of image feature point matching, genetic algorithm is added to the process of feature point matching. The advantages of genetic algorithm, such as the global search characteristics and parallelism, make the search for optimal matching points faster. Finally, the PROSAC algorithm is used to remove the mismatched point pairs and obtain the conversion matrix, and complete the image stitching according to the conversion matrix. The test results show that the GA-ORB-based image stitching technology not only makes full use of the rotation and scale invariance of the ORB algorithm, but also makes full use of the powerful global search capabilities of the genetic algorithm, which shortens the matching time while ensuring the effect of image stitching.

Acknowledgement: The authors thank Dr. Jinxing Niu for his suggestions. The authors thank the anonymous reviewers and the editor for the instructive suggestions that significantly improved the quality of this paper.

Funding Statement: This work is funded by the Key Scientific Research Projects of Colleges and Universities in Henan Province under Grant No. 22A460022.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] Q. H. Zhu, Y. Y. Shang and Z. H. Shao, "Cylindrical panorama stitching algorithm based on local features and vision consistence," *Journal of Image and Graphics*, vol. 21, no. 11, pp. 1523–1529, 2016.
- [2] Y. J. Ren, F. J. Zhu, K. S. Pradip, T. Wang, J. Wang *et al.*, "Data query mechanism based on hash computing power of blockchain in Internet of things," *Sensors*, vol. 20, no. 1, pp. 1–22, 2020.
- [3] Y. W. Yang and B. L. Guo, "Algorithm for cylindrical panoramic imagemosaic," *Computer Engineering and Applications*, vol. 45, no. 9, pp. 171–173, 2009.
- [4] X. R. Zhang, W. F. Zhang, W. Sun, X. M. Sun, S. K. Jha *et al.*, "A robust 3-D medical watermarking based on wavelet transform for data protection," *Computer Systems Science & Engineering*, vol. 41, no. 3, pp. 1043–1056, 2022.
- [5] Z. Zhang, J. Feng, J. Yan, X. Wang and X. Shu, "Ground-based cloud recognition based on dense-sift features," *Journal of New Media*, vol. 1, no. 1, pp. 1–9, 2019.
- [6] Y. J. Ren, Y. Leng, J. Qi, K. S. Pradip, J. Wang *et al.*, "Multiple cloud storage mechanism based on blockchain in smart homes," *Future Generation Computer Systems*, vol. 115, no. 2, pp. 304–313, 2021.

- [7] Z. F. Yang and L. Yan, "Research on image mosaic technology based on improved SURF algorithm," *Journal of Wuhan Engineering University*, vol. 43, no. 2, pp. 223–226+231, 2021.
- [8] X. R. Zhang, X. Sun, X. M. Sun, W. Sun and S. K. Jha, "Robust reversible audio watermarking scheme for telemedicine and privacy protection," *Computers, Materials & Continua*, vol. 71, no. 2, pp. 3035–3050, 2022.
- [9] T. T. Leonid and R. Jayaparvathy, "Classification of elephant sounds using parallel convolutional neural network," *Intelligent Automation & Soft Computing*, vol. 32, no. 3, pp. 1415–1426, 2022.
- [10] E. Rublee, V. Rabaud and K. Konolige, "ORB: An efficient alternative to SIFT or SURF," in *Int. Conf. on Computer Vision, ICCV 2011*, Barcelona, Spain, pp. 2564–2571, 2012.
- [11] Z. Y. Li, Y. Tian and F. J. Chen, "An UAV aerial image stitching algorithm based on improved ORB and PROSAC," *Progress in Laser and Optoelectronics*, vol. 56, no. 23, pp. 91–99, 2019.
- [12] C. Zuo and C. J. Pang, "Fast image matching algorithm based on improved ORB anti-perspective change," *Journal of Sensor Technology*, vol. 31, no. 11, pp. 1714–1720, 2018.
- [13] T. A. Al-Shurbaji, K. A. AlKaabneh, I. Alhadid and R. Masa'deh, "An optimized scale-invariant feature transform using chamfer distance in image matching," *Intelligent Automation & Soft Computing*, vol. 31, no. 2, pp. 971–985, 2022.
- [14] Z. H. Xi, H. X. Wang and S. Q. Han, "Quick mismatch elimination based on ORB-SLAM2 system division algorithm and map construction," *Computer Applications*, vol. 40, no. 11, pp. 3289–3294, 2020.
- [15] T. Renukadevi, K. Saraswathi, P. Prabu and K. Venkatachalam, "Brain image classification using time frequency extraction with histogram intensity similarity," *Computer Systems Science and Engineering*, vol. 41, no. 2, pp. 645–460, 2022.
- [16] Y. Zhang and Y. Liu, "Research on fast image mosaic system based on FAST features," *Computer Engineering and Applications*, vol. 52, no. 10, pp. 167–170, 2016.
- [17] X. Z. Chen, R. J. Liu and S. Zhang, "Development of millimeter wave radar imaging and SLAM in underground coal mine environment," *Journal of China Coal Society*, vol. 45, no. 6, pp. 2182–2192, 2020.
- [18] Y. H. Zhang, X. Jin and Z. J. Wang, "A new modified panoramic UAV image stitching model based on the GA-SIFT and adaptive threshold method," *Memetic Computing*, vol. 9, no. 3, pp. 231–244, 2017.
- [19] Z. F. Xu, C. Shi and Y. F. Wang, "A study of visual SLAM based on ORB+PROSAC mismatch elimination algorithm," *Software Engineering*, vol. 22, no. 5, pp. 9–14, 2019.