

A Sequence Image Matching Method Based on Improved High-Dimensional Combined Features

Leng Xuefei^{*}, Gong Zhe, Fu Runzhe, Liu Yang

College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P. R. China

(Received 17 November 2017; revised 19 May 2018; accepted 10 October 2018)

Abstract: Image matching technology is theoretically significant and practically promising in the field of autonomous navigation. Addressing shortcomings of existing image matching navigation technologies, the concept of high-dimensional combined feature is presented based on sequence image matching navigation. To balance between the distribution of high-dimensional combined features and the shortcomings of the only use of geometric relations, we propose a method based on Delaunay triangulation to improve the feature, and add the regional characteristics of the features together with their geometric characteristics. Finally, k -nearest neighbor (KNN) algorithm is adopted to optimize searching process. Simulation results show that the matching can be realized at the rotation angle of -8° to 8° and the scale factor of 0.9 to 1.1, and when the image size is 160 pixel \times 160 pixel, the matching time is less than 0.5 s. Therefore, the proposed algorithm can substantially reduce computational complexity, improve the matching speed, and exhibit robustness to the rotation and scale changes.

Key words: sequence image matching; navigation; Delaunay triangulation; high-dimensional combined feature; k -nearest neighbor

CLC number: TN967.2

Document code: A

Article ID: 1005-1120(2018)05-0820-09

0 Introduction

Image matching aided navigation, as one autonomous navigation technology, is of great theoretical value and has been widely used in many fields such as robot, unmanned aerial vehicle and cruise missile^[1-2]. However, the imaging difference between images usually leads to large calculation in the matching procession^[3]. Even though the speed of image matching has been greatly improved, its real-time performance is nevertheless hard to meet the requirements in many scenarios^[4]. Therefore, image matching navigation can only assist other navigation systems, forming aided navigation systems to realize high-precision and real-time navigation^[5].

To tackle the shortcomings of traditional image navigation, sequence image matching navigation has prominent advantages and has attracted an increasing attentions. Since two adjacent

frames are obtained at the same time under the same condition, sequence images just have the same noise distribution and small geometric deformation, which usually surface as translation and small rotation^[6]. So matching between sequence images can evidently reduce the time costs and improve the accuracy and speed of image matching substantially.

At present, the studies on image matching navigation mainly focus on matching algorithm, search strategy, filtering algorithm^[7], adaptation region analysis^[8], and so on. Among them, sequence image matching navigation has been thoroughly studied. Bodensteiner et al. proposed a method matching video streams to lidar data accurately and efficiently using key frames, and introduced the process of simultaneous localization and map building in detail^[9]. Ref. [10] proposed a positioning method based on monocular image,

^{*} Corresponding author, E-mail address: lengxuefei@nuaa.edu.cn

computing six-freedom degrees camera pose estimates continuously by matching natural features from input video and three-dimensional points reconstructed by structure from motion. In the algorithm, the fast key point tracker based on binary feature descriptor was used to realize the matching of two-dimensional to three-dimensional data, so the online scale invariant feature extraction was avoided, and the delay of each frame was reduced. Ref. [11] presented an improved Radon transform method to extract the line segments from the edges of the image, which were represented by four dimensional vectors. Time consumption of vector matching was greatly reduced. Ref. [12] adopted a novel algorithm, which constructed high-dimensional combined features based on stable branch points, and determined the position relation of two adjacent frames according to the geometric constraint relationship between the features. The algorithm combined the advantages of point features and line features, and eliminated the adverse effect of inaccuracy line segments extraction on matching results. However, there might be uneven distribution of features and redundant calculation if using the traversal to construct high-dimensional combined features. The requirements for image quality are too strict to meet.

A fast sequence image matching algorithm based on improved high-dimensional combined features is proposed in this paper. In the process of constructing high-dimensional combined features, Triangulation is introduced to achieve better dispersion of features, less repeated calculation, and improved matching performance. Meanwhile, the regional attributes of the features are added, so the matching based on the geometric as well as regional characteristics can greatly improve the robustness of the algorithm. Finally, this paper adopts k -nearest neighbor (KNN) algorithm based on KD-tree structure to optimize the search process and to further improve the efficiency of the algorithm. Simulation results show that the proposed algorithm can improve the real-time performance of image navigation, and meet the

requirements of accuracy and robustness of the navigation system.

1 Sequence Image Matching Method

1.1 Improved high-dimensional combined feature

The high-dimensional combined feature is a pair of intersecting lines which is composed of three stable branch points^[13]. These branch points are extracted from the edged binary images after denoising and thinning, and each branch point has more than two neighborhood points.

Each high-dimensional combined feature contains three branch points, which are defined as vertex $\mathbf{z}(z_x, z_y)$, the first endpoint $z_1(z_{1x}, z_{1y})$, and the second endpoint $z_2(z_{2x}, z_{2y})$, where the y coordinate value of \mathbf{z} is the smallest, and the first endpoint z_1 is the largest, if the y coordinate values are equal. The point with the smallest x coordinate value is defined as \mathbf{z} , and the point with the largest x coordinate value is the first endpoint z_1 ^[12-14]. In fact, this feature is not the basic feature directly detected from the image, but constructed by basic features according to some rules. Therefore, it is defined as a high-dimensional combined feature.

To balance the distribution of high-dimensional combined features, Delaunay triangulation algorithm is introduced to divide the branch points into a uniform triangle network before constructing the high-dimensional combined features.

Delaunay triangulation, widely used in practice^[15], has two important criteria: Empty circle and maximizing the minimum angle^[16]. These criteria guarantee that Delaunay triangulation network is unique. No matter where the feature construction starts, the final result will be consistent and adding, deleting or moving a point will only affect adjacent triangles.

Delaunay triangulation is implemented by incremental insertion method in this paper whose time efficiency and space efficiency are relatively higher. The process of constructing improved high-dimensional combined feature is shown in Fig.1, where only one high-dimensional com-

bined feature is marked in the diagram. After the extraction of stable branch points, Delaunay triangulation for the points set is done. Then three vertices of each triangle in the triangle network are used to construct a high-dimensional combined feature.

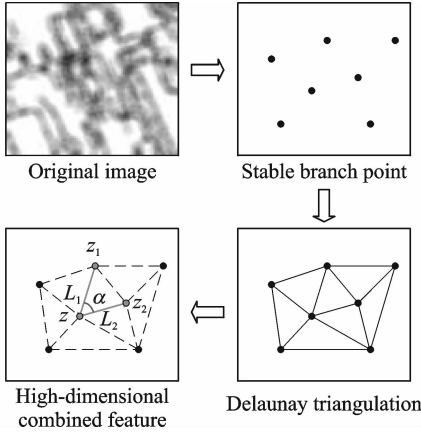


Fig. 1 Construction of improved high-dimensional combined features

In Fig. 1, L_1, L_2 represent the lengths of the two segments of the pair and α represents the angle between two line segments, which are described as

$$L_1 = \sqrt{(z_x - z_{1x})^2 + (z_y - z_{1y})^2} \quad (1)$$

$$L_2 = \sqrt{(z_x - z_{2x})^2 + (z_y - z_{2y})^2}$$

$$\alpha = \arccos\left(\frac{L_1^2 + L_2^2 - [(z_{1x} - z_{2x})^2 + (z_{1y} - z_{2y})^2]}{2L_1L_2}\right) \quad (2)$$

Usually, the features are not obvious enough and the information provided is limited when the branch points are too close. Contrarily, the features are not stable enough and large matching errors are easy to occur when the branch points are far away. Therefore, eliminating the un conspicuous and unstable features can further improve the matching performance. The eliminating rule is as follows

$$\begin{cases} \tau_1 \leq L_1 \leq \tau_2 \\ \tau_1 \leq L_2 \leq \tau_2 \end{cases} \quad (3)$$

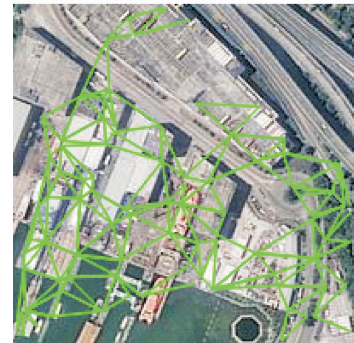
where τ_1, τ_2 are boundary constants whose value can be modified according to the specific experimental environment. The number of high dimensional combined features can be adjusted based on different τ_1 and τ_2 , which directly affects the accuracy and speed of matching. When imaging condi-

tions or image size change, the matching time can be controlled by adjusting τ_1 and τ_2 , so the algorithm is flexible.

Figs. 2(a), (b) show the results of construction of features without triangulation and the improved high-dimensional combined features, respectively.



(a) Features without triangulation



(b) Improved features

Fig. 2 Distribution of features

It is observable that the improved high-dimensional combined features retain the pros of unimproved features, with the advantages of point features as well as line features. Further, the number of features is obviously reduced. The distribution is more uniform. And no features aggregation exists.

1.2 Image matching based on geometric and regional characteristics

The improved high-dimensional combined features have some geometric characteristics, and the image matching can be greatly simplified according to the geometric constraints between features, so the vertex angle and the length information can be used to describe the feature.

The vertex angle α and the proportion of two lengths $L_1/(L_1 + L_2)$ are selected as the geomet-

ric information descriptor of the feature, which is a two-dimensional vector \mathbf{d}_g

$$\mathbf{d}_g = (d_{g1}, dg_2) = (a_1\alpha, a_2L_1/(L_1 + L_2)) \quad (4)$$

where a_1, a_2 are weight coefficients.

For more accurate matching relation between the two features, the regional information is added on the basis of the geometric information of the feature^[17]. Vertex \mathbf{z} of the high-dimensional combined feature is chosen as the center of the region, and the reference direction is pointed from \mathbf{z} to the first endpoint z_1 , as shown in Fig. 3.

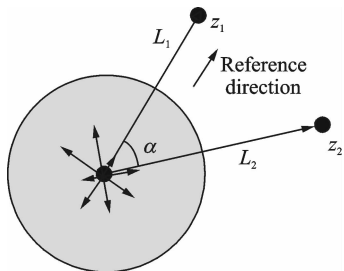


Fig. 3 Construction of regional information descriptor

The gray scale gradient $m(x, y)$ and gradient direction $\theta(x, y)$ of the pixels in a circular neighborhood centered on \mathbf{z} with a radius of $L_1/2$ can be calculated

$$m(x, y) =$$

$$\sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2} \quad (5)$$

$$\theta(x, y) = \arctan[(L(x+1, y) - L(x-1, y)) / [L(x, y+1) - L(x, y-1)]] \quad (6)$$

The gray scale gradients are divided into 8 intervals according to the deviation of the gradient direction relative to the reference direction, so the accumulated value of gray scale gradients in each interval $d_{r1}, d_{r2}, \dots, d_{r8}$ can be used as the regional information descriptor, which is an eight-dimensional vector \mathbf{d}_r

$$\mathbf{d}_r = a_3 (d_{r1}, d_{r2}, \dots, d_{r8}) \quad (7)$$

Similarly, a_3 is the weight coefficient. After normalization, it can be calculated as $a_3 =$

$$1 / \sum_{i=1}^8 d_{ri}.$$

Therefore, the descriptor of the improved high-dimensional combined feature \mathbf{d} is a 10-dimensional vector

$$\mathbf{d} = (d_1, d_2, \dots, d_{10}) = (d_g, d_r) = (a_1\alpha, a_2L_1 / (L_1 + L_2), a_3d_{r1}, a_3d_{r2}, \dots, a_3d_{r8}) \quad (8)$$

Considering that each dimension information descriptor should have similar weight coefficient, a_1, a_2 must satisfy the following relations

$$\begin{cases} (2\pi a_1 + a_2) / 2 = 1/8 \\ 2\pi a_1 = a_2 \end{cases} \quad (9)$$

It can be calculated that $a_1 = 1/(8\pi), a_2 = 1/4$. Finally, whether the two features are matched can be estimated according to whether the Euclidean distance between two features equals to a certain threshold value.

When Eq. (10) is satisfied, these two features are considered to be matched, and then the position relation of them can be calculated

$$\sqrt{\sum_{i=1}^{10} (d_i - d'_i)^2} \leq \epsilon \quad (10)$$

where ϵ is a threshold whose value can be modified according to the specific experimental environment.

The method integrating geometric information descriptor and regional information descriptor can improve the real-time performance and robustness of the algorithm. Moreover, the radius of the region is not a fixed value, but determined by the length of the line segment of the given feature, which ensures that the algorithm has certain scale invariance.

1.3 Search strategy based on KNN

Since the ergodic searching can hardly meet the real-time requirements, we choose KNN algorithm, a supervised learning classification algorithm based on KD-tree structure^[18], to optimize the search process and reduce the matching time.

KD-tree is a data structure that divides data points in k -dimensional space, and is suitable for fast nearest neighbor searching and approximate nearest neighbor searching in high dimensional space. In the proposed algorithm, the features of the last frame image is stored in KD-tree structure, that is, the features set of the last frame image is selected as the training sample set, so the sample data is 10-dimensional data.

The clustering strategy of KNN is used to quickly find two features, the nearest and the

next nearest to the feature to be matched, and obtain two corresponding distances l_{\min}, l_2 . If Eq. (11) holds, the closest feature in the previous frame picture is considered to match this feature, σ is a threshold, and the value of ϵ and σ directly affects the accuracy of the matching results

$$\begin{cases} l_{\min} \leq \epsilon \\ (l_2 - l_{\min})/l_{\min} \geq \sigma \end{cases} \quad (11)$$

After all the features in current frame image appear in the matching judgment, the matching positions can be counted.

If the matching points concentrated in a certain area, the position is considered as the correct matching point between the current frame image and the last frame image. Then the confidence factor Ψ can be introduced to evaluate the credibility of the matching result.

$$\Psi = \frac{\lambda_{\max} - \bar{\lambda}_R}{\lambda_{\max}} \quad (12)$$

where λ_{\max} is the maximum counting value of the accumulator, and $\bar{\lambda}_R$ is the sum of rest accumulator values more than $0.25\lambda_{\max}$. When the value of Ψ is larger than a certain threshold, the matching result is believed to be reliable.

The position relation of the current frame image with the reference image can be derived on the basis of sequence image matching, but certain cumulative errors would inevitably appear after continuous matching of adjacent frames sequence images. Therefore, so the matching of the real time image with the reference map at regular intervals is needed, based on the improved sequence image matching method.

The flow chart of the algorithm is shown in Fig. 4.

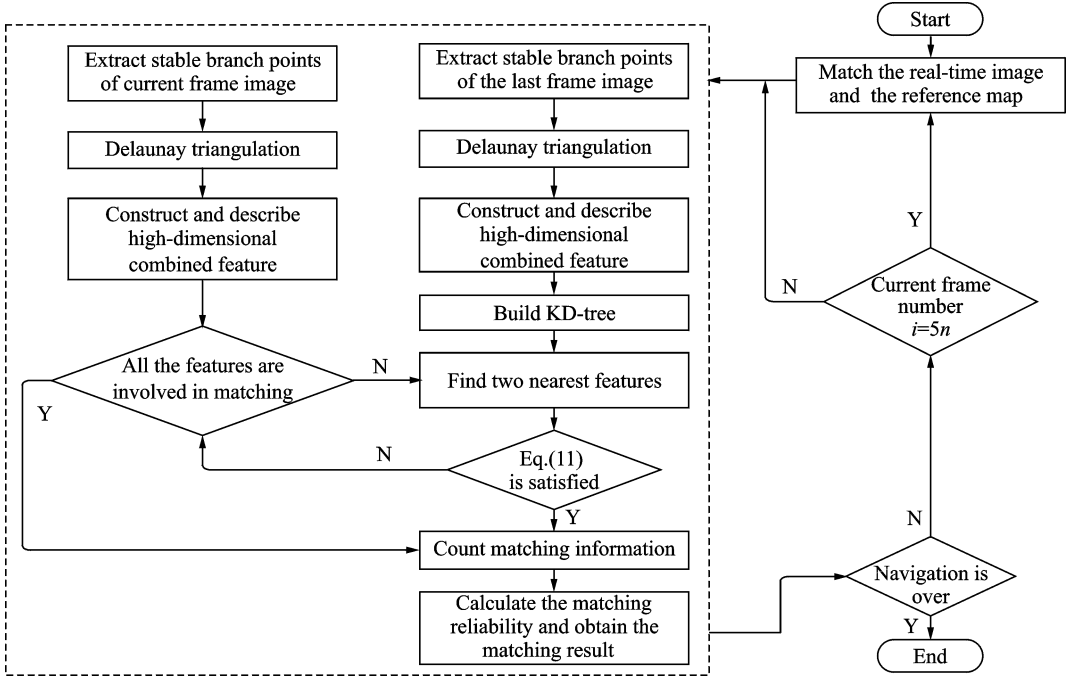


Fig. 4 Flow chart of the algorithm

2 Simulation

A series of testing experiments to verify the proposed algorithm were performed on a computer with a CPU of Intel CORE(TM) i3-2330M, 2.2 GHz dominant frequency and 4.00 GB memory storage, and the programming language was Visual C++ 6.0.

Firstly, the numbers of branch points, unimproved high-dimensional combined features and the high-dimensional combined features improved by triangulation of different images captured by different sensors were counted, as shown in Table 1, and the results of the distribution of features using this algorithm are shown in Fig. 5. It can be seen that the number of high-dimensional combined features is closely related to images

quality and sensor types. The features number of synthetic aperture radar (SAR) images is much more than that of optical satellite images. Since Delaunay triangulation algorithm divided the branch points into a triangle network, and the three points of each improved high-dimensional combined feature could only be located on a triangle in the network, the number of high-dimensional combined features was reduced obviously after improvement, which can effectively avoid the aggregation of high-dimensional combined features caused by the aggregation of point features. Therefore, the using of triangulation improves the dispersion of features, reduces the

number of features, and then reduces the computational costs and space occupation, which is beneficial to improve the performance of image matching navigation.

Several sets of tests were conducted to verify the effectiveness of the proposed algorithm based on improved high-dimensional combined feature in the presence of translation, small scale change and rotation of images. In the experiment, the size of sequence image was $160 \text{ pixel} \times 160 \text{ pixel}$; boundary constants $\tau_1 = 5, \tau_2 = 40$; thresholds $\epsilon = 0.02, \sigma = 0.2$. Considerable experiments have verified this is the best condition. The simulation results are shown in Table 2.

Table 1 Comparison of the number of features

Image classification	Name	Size/(pixel \times pixel)	Number of branch point	Number of unimproved feature	Number of the improved feature
Optical satellite image	Picture 1	300×300	81	167	50
	Picture 2	300×300	106	301	56
SAR image	Picture 3	160×160	127	1 457	184
	Picture 4	160×160	92	577	137

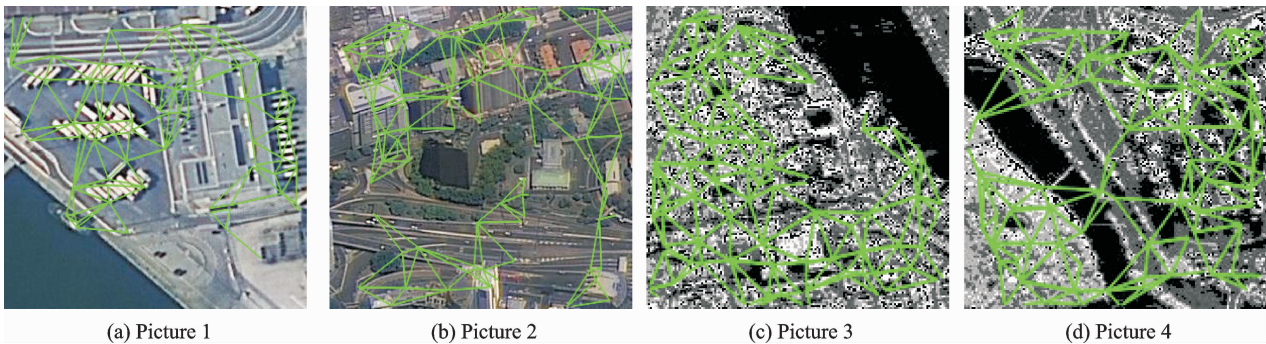


Fig. 5 Results of the distribution of features by the proposed algorithm

Table 2 Results of matching

Number of features in current frame image	Number of features in last frame image	The correct matching point	Rotation angle/($^{\circ}$)	Scale factor/time	Matching results			
					Matching point	Rotation angle/($^{\circ}$)	Scale factor/time	Running
100	111	(20,20)	0	1.00	(20,20)	0	1.00	0.219
77	56	(12,9)	-2	1.10	(12,9)	-2.33	1.07	0.269
125	138	(17,15)	3	0.96	(17,15)	3.09	0.95	0.397
146	184	(7,13)	-8	1.05	(7,12)	-7.52	1.05	0.331
137	126	(9,6)	-5	0.90	(9,6)	-5.21	0.92	0.305
183	196	(7,11)	8	1.06	(7,11)	8.34	1.05	0.384

In Table 2, it can be seen that the matching time was within 0.5 s at the rotation angle of -8° to 8° and the scale factor of 0.9 to 1.1, so the speed of the algorithm was relatively fast. This

fast sequence matching algorithm proposed can meet the requirements of navigation system of real-time performance, robustness and accuracy.

In order to further verify the effectiveness

and superiority of the proposed algorithm, we chose the images from different sensors, and compared the proposed algorithm with the algorithm without KNN, the algorithm based on unimproved high-dimensional combined features and the algorithm based on SIFT feature. Given that only translations and rotations of small degrees u-

sually exist in reality, no scale change was adopted in each group experiment. The rotation angles were -5° and -2° , and their correct matching points were (13, 7) and (12, 10), respectively. Figs. 6, 7 show the matching results of two group images using the proposed algorithm. Table 3 shows the matching results of different algorithms.

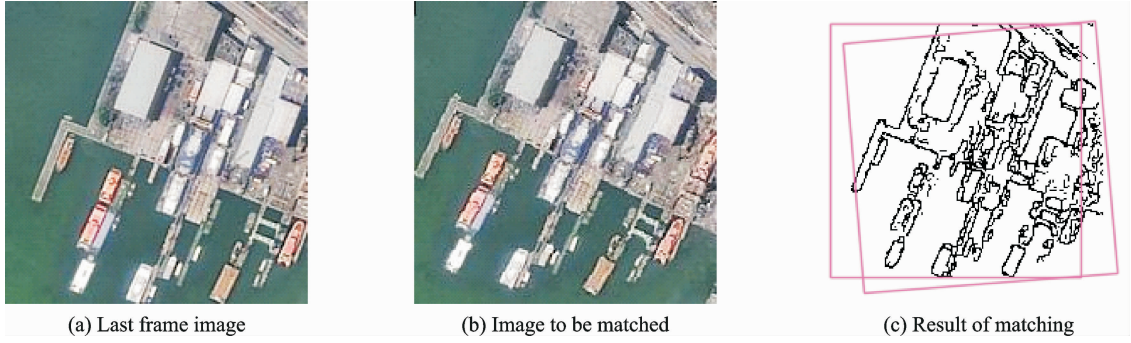


Fig. 6 Optical satellite image

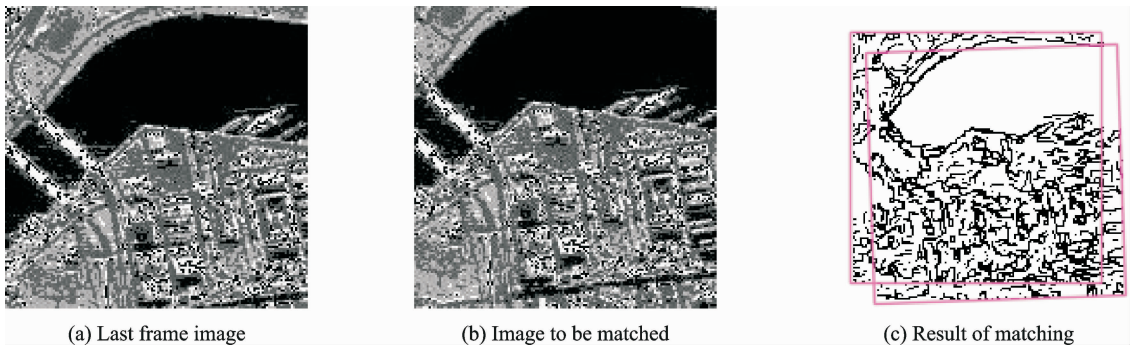


Fig. 7 Synthetic aperture radar image

Table 3 Results of different algorithms

Matching object	Size/(pixel \times pixel)	The correct matching point	Algorithm	Matching point	Rotation angle/ $^\circ$	Scale factor/time	Running time/s
SAR images	200 \times 200	(13,7)	The proposed algorithm	(13,7)	-4.61	1.00	0.242
			Algorithm without KNN	(13,7)	-4.61	1.00	0.406
			Algorithm based on unimproved feature	(14,6)	-3.9	1.02	0.533
			Algorithm based on SIFT feature	(13,7)	-4.63	1.00	1.989
The optical satellite images	160 \times 160	(12,10)	The proposed algorithm	(12,10)	-1.79	1.00	0.267
			Algorithm without KNN	(12,10)	-1.79	1.00	0.513
			Algorithm based on unimproved feature	(12,10)	-1.80	1.00	0.482
			Algorithm based on SIFT feature	(12,10)	-1.79	1.00	1.571

Table 3 implies that time consumption of the proposed algorithm is the shortest. Since the number of the features constructed by this method was reduced, compared with the method of constructing high-dimensional combined features by searching all the branch points, and a large number of redundant computations were avoided,

the speed of the algorithm was improved. Although the number of features was reduced, the matching accuracy did not decline because in essence a small amount of more dispersive features which may better represent the image information were selected to participate in the matching. When the rotation angle of optical satellite image

was 5° , the algorithm using unimproved features could hardly find the matching position accurately, but our algorithm could still achieve the matching. That means the robustness of the proposed algorithm is further improved, resulting from the adding of regional characteristics. Meanwhile, the descriptor dimension of the feature proposed in this paper was far less than the SIFT feature, so the proposed algorithm has a better speed performance.

3 Conclusions

Based on sequence image matching navigation, the algorithm is introduced in three aspects: Feature detection, matching rules and search strategy. Simulation experiments verified its superiority. The following conclusions can be obtained.

(1) When the high-dimensional combined feature is improved by Delaunay triangulation, the distribution of features is improved, and redundant calculation is reduced.

(2) In the process of matching, after the geometric as well as regional characteristics of the feature are taken into account, the algorithm is more robust to rotation and scale changes.

(3) KNN algorithm is introduced to optimize the search process, which can improve the speed of the algorithm further.

In conclusion, the proposed algorithm can meet navigation requirements, and exhibit good robustness, accuracy, and a promising engineering application. And based on this algorithm, we will conduct further study on key frame extraction and adaptation research.

Acknowledgements

This work was supported by the National Natural Science Foundations of China (Nos. 51205193, 51475221).

References:

- [1] LUITPOLD B. Flight path planning for unmanned aerial vehicles with landmark-based visual navigation [J]. *Robotics and Autonomous Systems*, 2014, 62 (2): 142-150.
- [2] RAMIREZ-TORRES J G, LARRANAGA-CEPEDA A. Real-time reconstruction of heightmaps from image taken with an UAV: Robotics and mechatronics [M]. U. K. : Springer International Publishing, 2016.
- [3] DAI Tao, ZHU Changren, HU Shuping. Overview of image matching technology[J]. *Digital Technology and Applications*, 2012(3): 174-177.
- [4] CHEN Tianhua, ZHOU Aide, LI Huixi, et al. Vision enhancement technology of drivers based on image fusion[J]. *Transactions of Nanjing University of Aeronautics and Astronautics*, 2015, 32 (5): 495-501.
- [5] KAJAL S, INKYU M. Improved scale-invariant feature transform feature-matching technique-based object tracking in video sequences via a neural network and Kinect sensor[J]. *Journal of Electronic Imaging*, 2016, 9(6): 9-24.
- [6] LIU Hongliang, CHEN Weiyi, XU Zhongsheng. Research on vehicle-borne velocity measurement method based on sequence image matching[J]. *Systems Engineering and Electronics*, 2015, 37(4): 964-968.
- [7] RU Jiangtao, LENG Xuefei, GONG Zhe. Study on filtering algorithm with asynchronous measurement information for ins/sar integrated navigation system [J]. *Journal of Nanjing University of Aeronautics and Astronautics*, 2017, 49(2): 276-282.
- [8] SHEN Lincheng, BU Yanlong, XU Xin, et al. Research on matching-area suitability for scene matching aided navigation[J]. *Acta Aeronautica Et Astronautica Sinica*, 2010, 31(3): 553-563.
- [9] BODENSTEINER C, BULLINGER S, LEMAIRE S, et al. Single frame based video geo-localisation using structure projection [C]//IEEE International Conference on Computer Vision Workshop. [S. l.]: IEEE, 2015: 1036-1043.
- [10] LIM H, SINHA S N, COHEN M F, et al. Real-time monocular image-based 6-DoF localization[J]. *The International Journal of Robotics Research*, 2015, 34(4-5): 476-492.
- [11] MANZAR A, CHEEMA T A, QURESHI I M. New scale invariant template matching technique using hyper space image representation[J]. *Pattern Analysis and Applications*, 2008, 12(3): 201-214.
- [12] GONG Zhe, LENG Xuefei, LIU Yang. A fast image matching method based on high-dimensional combined features[C]//The 36rd Chinese Control Conference. Dalian, China: IEEE Conference Publications, 2017: 5993-5999.

- [13] LENG Xuefei. Research on the key technology for scene matching aided navigation system based on image features [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2007.
- [14] WANG Bihui, LENG Xuefei, WU Songsen, et al. A subpixel precision image matching-aided navigation method[C]//The 33rd Chinese Control Conference. Nanjing, China: IEEE Conference Publications, 2014: 556-561.
- [15] GUO Xiaojie, CAO Xiaochun. Good match exploration using triangle constraint[J]. Pattern Recognition Letters, 2012, 33(7): 872-881.
- [16] CAO Jianchao. Fast image matching algorithms based on new feature description[D]. Dalian: Dalian University of Technology, 2015.
- [17] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.
- [18] WU Xindong, VIPIN K, ROSS J Q, et al. Top 10 algorithms in data mining[J]. Knowledge and Information Systems, 2008, 14(1): 1-37.

Dr. **Leng Xuefei** is currently an associate professor at College of Astronautics, Nanjing University of Aeronautics and Astronautics, China. She received her Ph. D. degree from Nanjing University of Aeronautics and Astronautics, China, in 2007. Her research interests include image matching, navigation, inertial devices.

Ms. **Gong Zhe** received her Master degree from Nanjing University of Aeronautics and Astronautics China in 2018. Her research interests include inertial technology and multiple information fusion integrated navigation.

Mr. **Fu Runzhe** received his Bachelor's degree from Nanjing University of Aeronautics and Astronautics and National Aerospace University named after. M. E. Zhukovskiy, Ukraine in 2017. He is currently a postgraduate at Nanjing University of Aeronautics and Astronautics. His research interests include stereo image matching and navigation.

Mr. **Liu Yang** graduated from Nanjing University of Posts and Telecommunications in 2016 and is currently a postgraduate at Nanjing University of Aeronautics and Astronautics. His research interests are image matching and navigation.

(Production Editor: Zhang Bei)