Aerial Image Stitching Based on Fusion of Geographic Coordinates and Image Features

Liangchao Guo
College of Aerospace Science
National University of Defense Technology
Changsha, China
guoliangchao19@nudt.edu.cn

Huyang Zhu
College of Aerospace Science
National University of Defense Technology
Changsha, China
809897416@qq.com

Yuwei Liu
College of Aerospace Science
National University of Defense Technology
Changsha, China
897405846@qq.com

Xiaoliang Sun*
College of Aerospace Science
National University of Defense Technology
Changsha, China
alexander_sxl@nudt.edu.cn

Xichao Teng
College of Aerospace Science
National University of Defense Technology
Changsha, China
tengari@buaa.edu.cn

2022 IEEE International Conference on Unmanned Systems (ICUS) | 978-1-6654-8456-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICUS55513.2022.9986913

***Abstract*—With the rapid development of unmanned aerial vehicle technology, unmanned aerial vehicles have been widely used to obtain ground remote sensing images. Due to the limited field of view of a single remote sensing image, it is necessary to stitch the images to obtain a large-scale scene image to apply to various practical applications. This paper introduces a two-stage stitching algorithm combining geographic coordinate information and image features. The stitching of sequence images is divided into intra-line stitching and inter-line stitching. Firstly, we use geographic coordinates and image features to calculate transformation parameters between a row of images, respectively, and select an appropriate transformation matrix to stitch a single row of images based on the rotation parameters in the homography matrix. Then, we also calculate the transformation matrices between different rows of images based on the geographic coordinate information and image features, and select the appropriate transformation matrix based on the comparison of the rotation parameters in different transformation matrices. Stitching experiments are carried out for various scenarios. Compared with traditional stitching methods, the proposed method integrates information from different sources, has higher reliability, and can adapt to stitching of various types of scenarios.**

Keywords—aerial images, geographic coordinates, image features, stitching

I. Introduction

With the continuous improvement of camera resolution and the rapid development of unmanned aerial vehicle (UAV) control systems, Ground-based remote sensing images obtained by UAVs is becoming more common. Aerial images captured by UAVs are widely utilized in environmental monitoring, land planning, disaster prediction, 3D scene reconstruction, inspection in heritage, archeological applications and other fields[1][2]. For obtaining an intricate image of a large area, it is necessary to stitch multiple UAV images to generate a high-resolution single image that can display the geographic information of the whole area.

Image stitching is mainly divided into two stages: image registration and image combination. The image registration operation is to obtain the rotation and translation transformation parameters between related images. The image composition transforms a series of images into a common

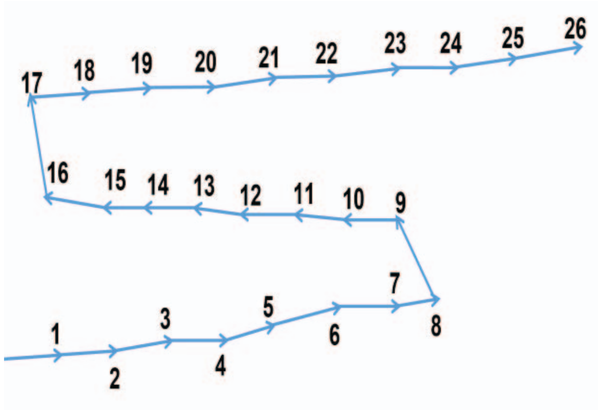

Fig. 1. Flight path diagram of aerial photography aircraft（Numbers such as 1, 2 and 3 indicate the image number obtained at the current position, and the arrow indicates the flight direction of the aircraft）

coordinate system for eliminating redundant information between images to generate a large image.

The selection and design of image registration algorithm have a key impact on the result of image stitching. Selecting an appropriate registration algorithm can effectively improve the accuracy and visual effect of image stitching result, and simplify the difficulty of image composition steps. Currently, the image registration algorithms mainly include direct registration method and feature-based registration method. The direct registration method mainly calculates the gray information similarity between the images to be stitched, and finds the transformation parameters that maximize the similarity. The feature registration method is mainly to calculate the transformation parameters between images by registering the specified types of features extracted from the reference image and the target image. The commonly used features include Harris feature[3], scale-invariant feature transform (SIFT) feature[4], speeded up robust feature (SURF) [5][6] and oriented fast and rotated brief (ORB) feature[7]. Image registration based on SURF features has the characteristics of high accuracy and long time, while the image registration time using ORB feature points is usually only one tenth of that based on surf feature points, but the accuracy is slightly lower.

*Corresponding Author.

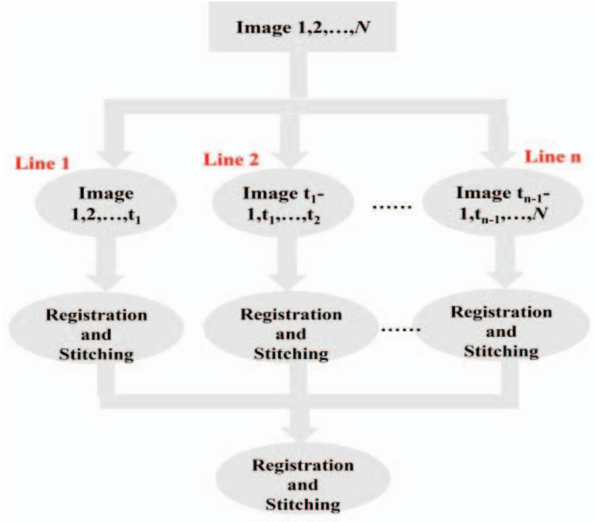

Fig. 2. Flowing diagram showing the proposed method

The main task of the image combination stage is to splice the target image and the reference image based on the transformation parameters obtained from the registration. The combination task involves two steps: the selection of image unified coordinate system and the selection of image fusion mode. There are usually two methods about the coordinate system selection. One is to convert the target image to the reference image coordinate system which is relatively simple and easy to implement, and the disadvantage is that the image with a long distance from the reference image may appear distorted, which will make the image stitching effect worse. The other is to transform the images to be spliced to the spherical coordinate system or cylindrical coordinate system. This method can effectively reduce the deformation of the image and make the global image spliced have a more regular shape to improve the splicing effect. However, this splicing method has a high error rate and is complex to realize. The methods of image fusion mainly include direct replacement of overlapping areas, weighted fusion at seams, etc. the direct replacement method is faster, but due to the influence of illumination changes, obvious stitching traces will be generated. The weighted fusion method can make the seams transition naturally and effectively eliminate the stitching traces, but it will slow down the stitching speed and is not conducive to the real-time performance of stitching.

For the image sequence obtained by the flight photography of UAV according to the path shown in Fig. 1, a two-stage stitching method based on fusion of geographic coordinates and image features, as shown in Fig. 2, is proposed to realize the mosaic of aerial images, which decomposes the global image stitching task into single line image stitching and interline image stitching. In the process of single line image stitching, taking the middle frame image as the reference image, the transformation parameters of between each image in a single line and the reference image are calculated by using the geographic coordinate data and SURF feature points respectively, and the appropriate transformation parameters are selected by comparison. Then, each image is transformed into the coordinate system of the reference image, and the combination and splicing of single line images are realized by direct replacement. The interline image stitching process is similar to the single line image stitching and the geographic coordinate data and SURF feature points are also used to calculate the transformation parameters of each line of image

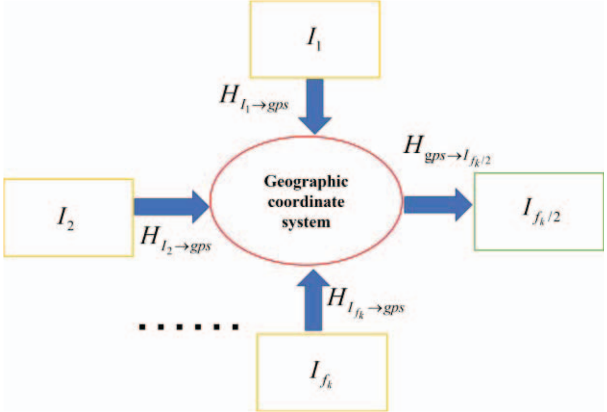

Fig. 3. Calculating perspective transformation matrix using geographic coordinate information of image

and the first line of image, then the accurate transformation parameters are selected by comparing the rotation parameters. Finally, the splicing of the whole sequence of images is completed.

This paper is organized as follows: Section II introduces the process and algorithm details of single line image stitching and interline image stitching in detail. In Section III, three sets of image data are used to verify the stitching algorithm proposed in Section II, and compared with the global stitching algorithm based on OpenCV. Section IV summarizes the advantages and disadvantages of the proposed algorithm and introduces the future work.

II. Materials and Methods

Based on the flight path of the UAV shown in Fig. 1, the stitching algorithm proposed in this paper is mainly divided into intra line stitching and interline stitching. Multiple images input in sequence are firstly divided into multi-line images according to the given line ID, and then solves the inter image transformation parameters by combining the geographic coordinate information and image feature points to splice each line of images. In the inter row stitching step, multiple rows of images are registered with a similar way to the intra-row stitching, and finally multiple remote sensing images in a sequence are stitched into a large-scale image.

A. Inline Image Stitching

As shown in Fig. 1, since the UAV acquires the ground remote sensing image according to the s-shaped flight path, this paper first stitches the remote sensing images belonging to the same line. Taking the k-th row of images as an example, the number of images is . The intermediate frame image in one row is set as the reference image, and the transformation parameters between the other images in the row and the reference image are solved. Considering that the scene in the remote sensing image may lack the feature points required for image matching or the matching features are not enough, this paper, as shown in Fig. 3 and (1), gives priority to using the geographic coordinates of the four vertices of a single image obtained by the aircraft GPS equipment to solve the required transformation parameters.

$$H_{I_c \to I_{fk/2}} = H_{gps \to I_{fk/2}} H_{I_c \to gps} \tag{1}$$

There is a large deviation in the accuracy of the geographic coordinate data given by a single picture. In the study, we find that when the obtained remote sensing image contains rich feature information, the accuracy of the transformation parameters obtained by image feature matching is much higher than that obtained by geographic coordinate information. However, when the UAV takes a scene image with few available image features, such as the sea surface, the transformation parameters calculated from the image features will no longer be trusted, and even the matching may fail. Therefore, it is necessary to find a judgment standard to make a reasonable choice between the transformation parameters obtained from geographic coordinate and the transformation parameters obtained from image features. Through experiments, we found that the smaller the rotation angle in the calculated homography matrix, the more trustworthy the transformation matrix is. The reason is that the images we want to splice are roughly distributed near a straight line, and a large amount of rotation is not required during stitching.

Algorithm1: Transformation Parameter Selection Algorithm

Input: Homography matrices set H_IMG based on image features, homography matrices set H_GPS based on geographic coordinate, Second-order identity matrix $BASE\_H$

Output: H_GPS

mid is the index of the intermediate frame image

$$
\begin{aligned}
& for\ k = mid-1:0:-1 \\
& \quad H\_img_{k\to mid} = H\_GPS_{k+1\to mid} H\_IMG_{k\to k+1} \\
& \quad diff\_gps = sum(abs(H\_GPS_{k\to mid}(:2,:2) - BASE\_H))/4 \\
& \quad diff\_img = sum(abs(H\_img_{k\to mid}(:2,:2) - BASE\_H))/4 \\
& \quad if\,(diff\_gps \geq 0.1\ \&\&\ diff\_img \geq 0.1\ \&\&\ diff\_gps > diff\_img) \\
& \qquad H\_img_{k\to mid}.copyTo(H\_GPS_{k\to mid}) \\
& \quad if\,(diff\_img < 0.1) \\
& \qquad H\_img_{k\to mid}.copyTo(H\_GPS_{k\to mid}) \\
& end
\end{aligned}
$$

$$
\begin{aligned}
& for\ k = mid+1: I.size()-1 \\
& \quad H\_img_{k\to mid} = H\_GPS_{k-1\to mid} H\_IMG_{k\to k-1} \\
& \quad diff\_gps = sum(abs(H\_GPS_{k\to mid}(:2,:2) - BASE\_H))/4 \\
& \quad diff\_img = sum(abs(H\_img_{k\to mid}(:2,:2) - BASE\_H))/4 \\
& \quad if\,(diff\_gps \geq 0.1\ \&\&\ diff\_img \geq 0.1\ \&\&\ diff\_gps > diff\_img) \\
& \qquad H\_img_{k\to mid}.copyTo(H\_GPS_{k\to mid}) \\
& \quad if\,(diff\_img < 0.1) \\
& \qquad H\_img_{k\to mid}.copyTo(H\_GPS_{k\to mid}) \\
& end
\end{aligned}
$$

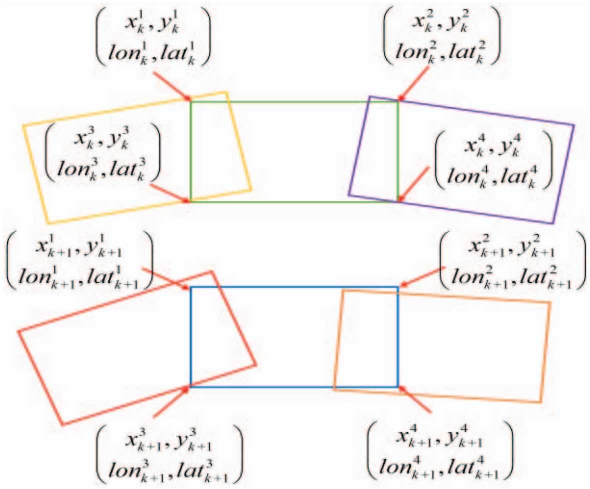

Fig. 4. Calculation of transformation matrix between adjacent images by geographic coordinate information（x,y is the pixel coordinate, lon is the longitude, lat indicates the latitude, and k is the row index)

According to this principle, we designed a transformation parameter selection algorithm as shown in algorithm 1.

B. Stitching Between Multiple Lines of Images

The inter row stitching algorithm also uses the combination of geographic coordinate and image features to calculate the transformation parameters between images. During the stitching process of intra row images, the intermediate frame images are not deformed. Therefore, we use the geographic coordinate information and pixel coordinate information of the intermediate frame image vertices as shown in Fig. 4 to solve the transformation matrix between adjacent rows of images, use SURF feature points +RANSAC for feature matching and then set reasonable criteria to select and replace the two transformation matrices see algorithm 2 for detailed selection process.

Note that in order to ensure that the two adjacent lines of images have enough overlapping region for stitching, in the

Algorithm 2: Selection of Perspective

Transformation Parameters in Interline Image Stitching Algorithm

Input: Homography matrix $H\_IMG_{k\to k-1}$ between adjacent lines based on image features, homography matrices $H\_GPS_{k\to 0}$ based on geographic coordinate information, Second-order identity matrix $BASE\_H$

Output: $H\_GPS_{k\to 0}$

k is the row image index

$$
\begin{aligned}
& H\_IMG_{k\to 0} = H\_GPS_{k-1\to 0} H\_IMG_{k\to k-1} \\
& diff\_gps = sum(abs(H\_GPS_{k\to 0}(:2,:2) - BASE\_H))/4 \\
& diff\_img = sum(abs(H\_IMG_{k\to 0}(:2,:2) - BASE\_H))/4 \\
& if\,(diff\_gps > diff\_img) \\
& \quad H\_IMG_{k\to 0}.copyTo(H\_GPS_{k\to 0})
\end{aligned}
$$

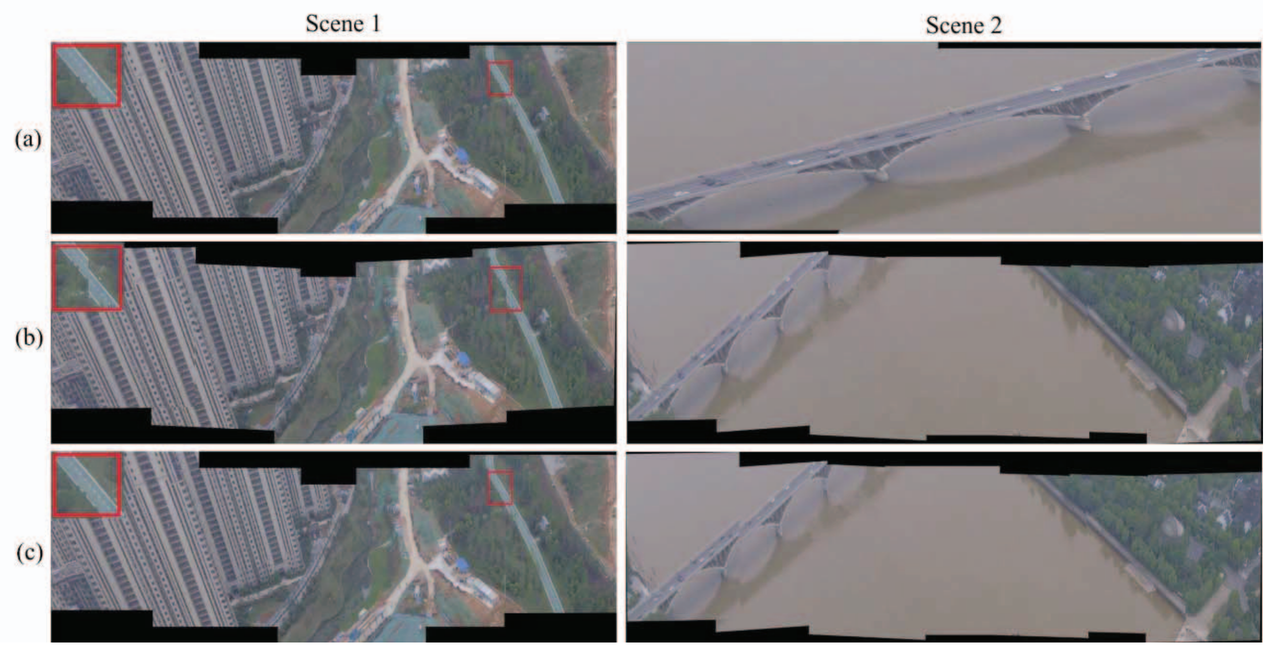

Fig. 5. Comparison of ablation experiments of single-line spliced output maps in two scenarios：(a) Image feature stitching, (b) Geographic coordinate stitching, (c) The proposed method

intra-row stitching stage, we stitch the last two images of the previous line of images with the current line of images to generate a large image of the current line.

III. EXPERIMENTS

In this paper we used three aerial image datasets for experiments. It includes three three-channel color visible light image data sets. See Table 1 for details. The scenarios covered by the dataset cover buildings, vegetation and rivers.

A. Data Acquisition

In this study, we captured three ground image scenes in a wild area in in the urban area of Changsha using UAV. Note that due to the fast flight speed of the UAV and the limited level of the equipment, there may be large errors in the GPS information of the images captured by the UAV.

B. Experimental Settings

To verify the effectiveness of the proposed algorithm, we compare the proposed algorithm with the classical image stitching method which completely adopts SURF image feature matching+bundle adjustment optimization+spherical projection. For the proposed method, the details of the algorithm are set as follows: using SURF feature points + RANSAC for image matching, where the Hessian matrix threshold is set to 100, and BestOf2NearestRangeMatcher is used to match the effective feature points of different images. The feature extraction and matching parameters of the comparison algorithm are the same as those of the proposed algorithm to facilitate a fair comparison of the stitching effects of the two algorithms.

The simulations were performed on a Ubuntu 20.0 computer equipped with an AMD R7-6800H CPU, 16GB RAM, and an NVIDIA GeForce RTX 3060 GPU.

C. Ablation Studies

The algorithm reasonably selects and fuses the transformation parameters calculated from image features and geographic coordinate data respectively, which ensures the reliability of the algorithm. In order to test the comprehensive effect of the algorithm proposed in this paper on the above two splicing methods, we set up an ablation experiment, using image features, geographic coordinate data and the proposed algorithm to splicing single-line images on two different scenes, respectively. Compare the stitching effects of these three algorithms.

Fig. 5 shows an example of the aerial stitching effect of three different stitching methods in two scenes. In the stitching of scene 1 images, due to the deviation of the geographic coordinates of the images, there is a local misalignment in the images stitched with geographic coordinates as shown in the enlarged box in the figure, and because the features in the image scene are relatively rich, the image feature stitching is adopted that can better stitch together single-line images. The results in Fig. 5 show that the proposed algorithm correctly selects the transformation parameters obtained from the image feature data in scene 1 stitching, and performs image stitching accordingly. In the stitching of scene 2, the water surface lacks image features, so the method using image features cannot extract enough features, resulting in the failure of image stitching and it can only stitch the first two images in a row of images. In contrast, the stitching method using geographic coordinate information can still obtain a better and complete stitching result despite some errors. Compared with Scenario 1, in Scenario 2, the proposed method selects the transformation matrix obtained by trusting the geographic coordinate information.

TABLE I. DETAILED INFORMATION OF THREE AERIAL IMAGE SCENES

Image Type	Number of Images	Image Resolution(pixel)	Number of Images Per Line
Scene 1: Color image	15	640×480	5
Scene 2: Color image	23	640×480	5~10
Scene 3: Color image	20	640×480	5~10

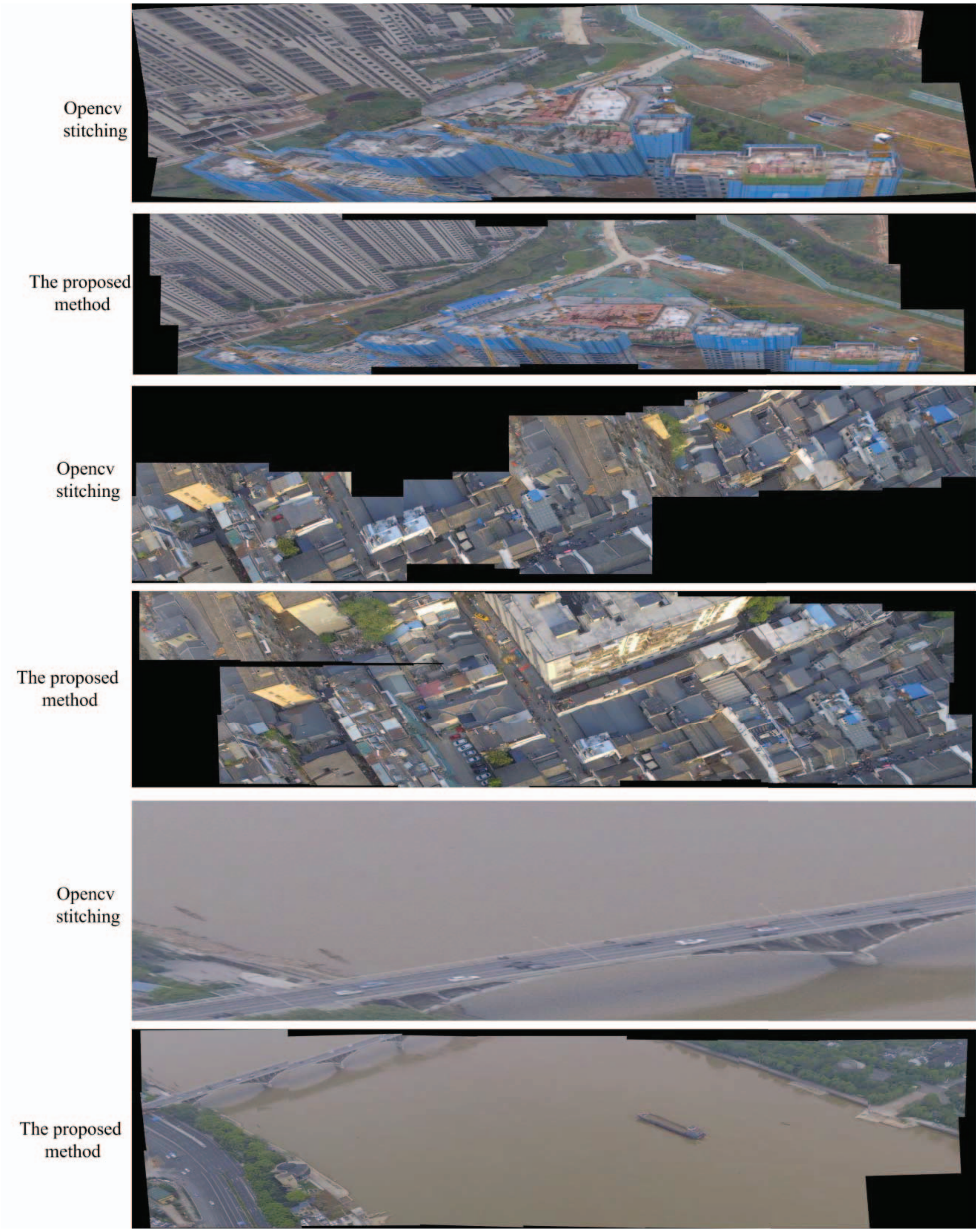

Fig. 6. Comparison of three line image output mosaic between traditional opencv mosaic method and the proposed algorithm in three scenarios. The highlighted red and blue boxes in each image are zoomed in (shown by the corresponding boxes on the right) for a detailed visual comparison.

The comparison results of ablation experiments in two scenarios show that the proposed algorithm can reasonably choose between different image transformation parameters obtained from image features and geographic coordinate data to get a large stitched image with high accuracy. Compared with stitching methods that use image features or geographic coordinate data alone, the proposed algorithm has higher reliability, which is particularly important in practical engineering.

D. Experimental Results

Fig. 6 shows a comparison of the image stitching results of three-line images in three different scenes between the proposed method and the traditional OpenCV stitching method. In the stitching of scene 1, the aerial image is a color three-channel image, which has rich features and is easy to extract. The stitching method based on OpenCV optimizes multiple camera motion matrices through the beam adjustment method, projects all images into a spherical coordinate system, and finally obtains a complete stitched large image. However, due to the stitching of multiple rows of images at one time, the relative position of each image the spliced large image cannot be accurately located, so that the sequence of multiple original images in the spliced large image obtained as shown in Fig. 6 is dislocated. The method proposed in this paper firstly accurately selects the transformation parameters obtained based on image features to ensure the alignment of local details, and the two-stage stitching ensures that the order of the original images in the obtained stitched large image will not be wrong. The stitching of scene 2 is similar to scene 1: the stitching method of OpenCV has a serious image sequence error, and all images are stitched on one line. While the method proposed in this paper also chooses to use image features for image stitching to obtain more accurate stitching results. In the stitching of scene 3, due to the lack of image features, the stitching method based on OpenCV cannot calculate the correct transformation parameters, and the interaction between images causes the stitching to fail. The proposed method uses the transformation parameters calculated based on geographic coordinate data to obtain ideal splicing results.

It can be seen from the above comparative experiments that the proposed method can comprehensively utilize two stitching methods based on image features and geographic coordinate information to ensure that the stitching map with the highest accuracy can be obtained in the process of image stitching in different scenes. Both in scene stitching with rich image features and in image stitching with lack of image features, the stitching accuracy and reliability of the proposed method are better than traditional stitching methods.

IV. CONCLUSION

In this paper, we propose an algorithm that can combine two stitching methods based on geographic coordinate and based on image features to more reliably stitch a large number of aerial images. The method proposed by us judges the accuracy of the transformation parameters obtained by the two splicing methods of latitude and longitude and image features based on the rotation amount in the transformation matrix, and independently selects more accurate transformation parameters to stitch the images. For adapting to the serpentine aerial photography route of the aircraft and speeding up the stitching speed, we divide the stitching process into in-line stitching and inter-line stitching, so that the relevant operators can quickly and comprehensively analyze the captured scene.

However, the stitching accuracy of the stitching method proposed by us still needs to be improved. In-line stitching has the problem of large deformation of some images, and inter-line stitching has the problem that lines and textures in some areas of the image are difficult to align. Future research will focus on addressing these issues.

ACKNOWLEDGMENT

Special thanks to Jinghao Wang, Qinwei Wang et al. for their selfless image data sharing.

REFERENCES

[1] N. T. Pham, S. Park and C. -S. Park, “Fast and Efficient Method for Large-Scale Aerial Image Stitching,” IEEE Access, vol. 9, pp. 127852-127865, September 2021.

[2] Gomez-Reyes, JK, Benitez-Rangel, JP, Morales-Hernandez, LA, Resendiz-Ochoa, E, Camarillo-Gomez, KA, “Image Mosaicing Applied on UAVs Survey,” Applied Sciences, vol. 15, pp. 2729, March 2022.

[3] C. G. Harris, M. Stephens, “A combined corner and edge detector,” Proceedings of the 4th Alvey Vision Conference. Manchester, Britain, 1988:147-151.

[4] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” International Journal of Computer Vision, 2004, 60(2):91-110.

[5] H. Bay, T. Tuytelaars, G. L. Van, “Surf: Speeded up robust features,” European conference on computer vision 2006 May 7, 404-417.

[6] H. Bay, A. Ess, T. Tuytelaars, et al, “Speeded-up robust features (SURF),” Computer vision and image understanding. 2008, 110(3):346-59.

[7] E. Rublee, V. Rabaud, K. Konolige, et al, “ORB:An efficient alternative to SIFT or SURF,” International conference on computer vision, 2011, Nov 6, 2564-2571.