Acta Technica 62 No. 2A/2017, 475–484

The key technology of fast stitching of remote sensing images of small unmanned aerial vehicles

Kai Ma[1,2], Qingqing Huang[1], Tiebo Sun[1], Tingting Sui[1], Jinhao Liu[1,3]

Abstract. The purpose is to study the key technology of fast stitching of small unmanned aerial vehicle remote sensing image. Based on the characteristics of small UAV images, the rapid image stitching software is designed. It has the capability of on-site processing. At present, the light and small UAV platform cannot get accurate images when shooting outside the elements. Therefore, when the collinearity equation is used as the UAV image geometric correction model, the calculation of the exterior orientation elements of the image needs to be taken into account. For a large number of ineffective image matching relationships, the Euclidean space distance relationship of camera GPS information is used to quickly construct latent image matching pairs, so as to speed up post-image matching. The results show that the proposed method can improve the overall operational performance. Therefore, it can be concluded that this method can solve the problem of UAV image stitching.

Key words. UAV images, Ransac, quick stitching.

1. Introduction

Because of its convenience and flexibility, UAVs are attracting more and more attention from all walks of life [1]. UAV can efficiently carry out research and practical work. In the remote sensing and mapping industry, UAVs can display their characteristics of maneuvering and fast. At present, UAV flight platforms are used to carry sensors, including card cameras, micro single cameras, measuring cameras, small LIDAR and hyperspectral cameras [2]. However, compared to the rapid development of sensor hardware and platform, post-processing software efficiency and processing capacity is obviously insufficient, especially on-site processing capacity [3]. Image mosaic is a method that combines many images together to form a spa-

[1]Beijing Forestry University, Beijing, China, 100083, China
[2]Xinjiang Agricultural University, Xinjiang, China, 830052, China
[3]Corresponding author

tially consistent image. This image can provide more information in time and space. The core of the entire mosaic algorithm is the registration algorithm between multiple images. In the automatic image processing process of optical images, a key link is the geometric correction of the image without artificial participation. It is directly related to the process of pre-processing and splicing images between the geometric accuracy.

In order to meet the requirements of geometric accuracy of actual unmanned image, it is very important to study geometric correction based on strict imaging model. At present, in the field of photogrammetry, splicing is generally done with high-performance workstations or servers [4]. Day is the processing unit time. Its processing capacity cannot meet the needs of the rapid response characteristics of the business needs, such as disaster relief, site mapping guidance. There are many problems. The degree of automation and processing efficiency are low, and it relies on more measurement conditions. As computer performance is improved, multi-core CPU and image parallel capability GPU are included. At the same time, some new algorithms are proposed in the field of computer vision. It provides the possibility to solve the above problems in UAV image mosaic. Therefore, on the basis of the above background, the small UAV image is designed for the overall scheme of the low electric control system, and the process of image mosaic processing is discussed. The core link is concerned.

2. State of the art

2.1. The basic principle

On the basis of the Lindeberg signal spatial scale theory, the extraction of Sift feature points is done under the establishment of multi-scale space [5]. First of all, the Gaussian pyramid image is constructed by using a series of Gaussian functions of scale factors. Then, the Gaussian difference pyramid image is obtained by subtracting the adjacent two Gaussian images. The local extreme point is detected as a candidate extremum in the Gaussian difference pyramid image and the corresponding scale is recorded. A three-dimensional quadratic function is used to fit the feature points to precisely determine the position and scale of the feature points. At the same time, the low contrast and unstable edge feature points are eliminated to enhance the stability of the feature points. Finally, the main direction of the feature points is determined by using the pixel gradient information in the neighborhood of the feature points. According to the main direction, the feature point neighborhood is rotated to reconstruct the 128-dimensional eigenvector, so as to ensure the invariance of the rotation of the feature. By using the Sift algorithm, reliable matching features can be extracted from images of different viewpoints in the same scene. These local feature descriptors have the following characteristics. It has invariance to target scale and rotation. There is a strong robustness to the difference of the shooting angle of the target, the affine transformation, the illumination changes and so on [6]. It has the characteristic of locality. To some extent, it can reduce the influence of Sift feature descriptors such as occlusion and complex background. The uniqueness is high, and the amount of information is rich. It is suitable for accurate matching in large feature descriptor database. It can perform precise feature position localization, which can reach sub-pixel levels. In addition, it also has the characteristics of multi quantity, which can fully describe the image information [7].

2.2. *GPU-Sift feature extraction*

GPU (Graphic, Processing, Unit) is an arithmetic unit that different from CPU. Especially for graphics processing, it has its unique advantages. Today, CPU has entered the era of multi-computing nuclear. However, compared to GPU calculations, the number of cores is still very limited. For example, the Tesla C2050 of NVidia has 448 CUDA cores. Sinha [8] discusses all aspects of Sift feature extraction, which is distributed throughout the CPU-GPU computing framework. The basic process is depicted in Fig. 1.

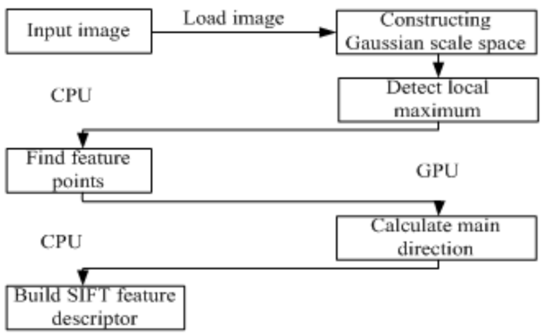

Fig. 1. GPU-Sift basic flow chart (GPU and CPU hybrid implementation)

Considering that all aspects of the Sift feature point extraction are not applicable to the parallel architecture of the GPU, the following steps are used to parallelize the GPU:

First, Gauss images and Gauss difference images are generated.

Second, feature point detection, which includes the extreme value of the scale space detection and feature points of the sub-pixel level positioning.

Third, calculate the main point of the feature point.

By using images of different sizes, the computing time of the main links in GPU-Sift is calculated. The computing platform is NVidia 8800GTX, 768 GB GPU memory, Win XP, Intel 3 GHz CPU.

3. Methodology

3.1. *GPU-Sift feature matching*

After determining the potential image matching pair, it is concerned that the Sift feature points between the two images are matched. Since the Sift feature matches the main operation time by retrieving the Eurasian distance from the KD tree to the nearest two character descriptors, this step only completes the parallel processing of the GPU. CPU is computed to construct the KD tree. The experimental results show that the GPU-Sift matching algorithm can improve the speed of image feature extraction. In the case of ensuring the image matching rate is stable, the operation speed is increased by about 3 times. By using GPU, computing efficiency can be improved to a certain extent. However, it makes a certain request for GPU memory. So, when the image feature matches, the CPU version of the matching algorithm is first considered. Unless the time performance of this session becomes the bottleneck of the entire process, the GPU version matching algorithm is considered.

3.2. Sift feature matching based on Ransac

Random Sampling Consistency Ransac (Random Sample Consensus) is an iterative algorithm for estimating the parameters of a mathematical model. The main idea of Ransac is to solve the parameters of the mathematical model that most samples (feature points) can satisfy by sampling and validating strategies. In the iteration, the minimum number of samples and the parameters of the model are needed to collect from the data set at a time. Then, the number of samples that match the parameters of the model is counted in the data set. The most sample parameters are considered to be the final model parameter values. The sample point of the model is called the Inliers, and the sample point that does not conform to the model is Outliers.

For Sift feature matching, the misaligned pair is Outliers. The algorithm flow is as follows:

First, select a pair of match pairs (5 points) randomly from the initial match point set.

Second, using the selected matching point to calculate the essential matrix model parameters between the two images.

Third, by using the distance threshold, the interior point that matches the model is calculated. The ratio of the number of points to the total number of points is counted, and the maximum number of iterations is updated accordingly.

Fourth, repeat first and second steps until the maximum number of iterations is reached, or enough of the Inliers points are found.

Fifth, through the above process, the essential matrix between the two images is determined. Based on this essential matrix, the matching pair of all Sampson errors is greater than the threshold, leaving the remaining Inliers pair. Eventually, the purification of the match point is completed.

3.3. The design of PLC module

The auxiliary electrical control system's lubrication system has two inputs, which are two alarm outputs. The cooling system also has two inputs for both alarm outputs. The automatic tool has 6 inputs and 6 alarm outputs. The lighting and semaphore system has an input and four outputs. The tool change system has 5 inputs and 5 outputs. Therefore, in the choice of PLC, taking into account the remaining 1/3 of the margin, Huichuan PLC was used. It has 24 inputs and 15 outputs. The scanning frequency is 1000 Hz, which fully meets the requirements.

4. Result analysis and discussion

4.1. Camera model

The camera's basic imaging model is often referred to as a small aperture imaging model. It is given by the three-dimensional space to the center of the plane projection transformation. The spatial point C is the projection center (the site), and the distance from the main point p is f (focal length). The projection of the spatial point C on the image plane is the intersection of the ray and the image plane at the point C and the point X. The results are shown in Figs. 2 and 3.

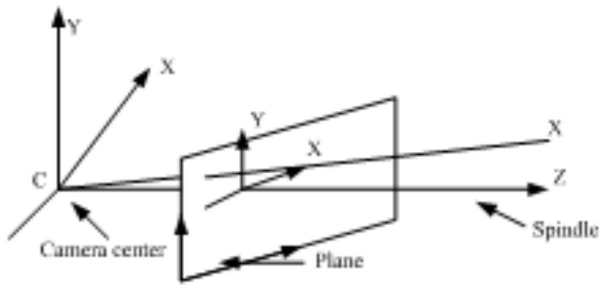

Fig. 2. Camera small hole imaging model 1

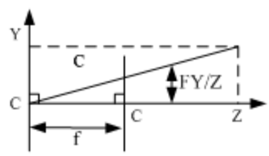

Fig. 3. Camera pinhole imaging model 2

To better describe the projective relation from algebra, three coordinate systems are established, which are camera space, Cartesian coordinate system, image plane coordinate system and pixel plane coordinate system. In practical applications, the upper left corner of the image (the upper left corner of the pixel in the upper left corner) is the pixel plane coordinate origin (0,0), and the horizontal line to the right and the vertical line down is the direct direction of the U axis and the V axis. On the image plane, the image plane coordinate system is established in the positive direction of the x-axis and the y-axis along the u-axis and v-axis positive directions of the pixel coordinate system, respectively, as the coordinate origin of the image plane coordinate system. In the space, the camera projection center C is the coordinate origin of the camera space Cartesian coordinate system, and the direction of the spindle pointing direction coincides with the positive direction of the Z_c axis. With the main axis as Z_c, the camera space Cartesian coordinate system (the right-hand space Cartesian coordinate system) is established in the positive direction of X_c and Y_c in the positive and negative directions of the plane coordinate system. According to the principle of triangular similarity, combined with the above three coordinate system, the basic imaging model of the camera is described as follows:

$$u = x + u_0, \ v = y + v_0 \,, \tag{1}$$

$$x = \frac{f x_c}{z_c}, \ y = \frac{f y_c}{z_c} \,. \tag{2}$$

The corresponding matrix of the above formula is

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f \ 0 \ u_0 \ 0 \\ 0 \ f \ v_0 \ 0 \\ 0 \ 0 \ 1 \ 0 \end{bmatrix} = \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{3}$$

In the above equations, (u, v), (x, y), (X_c, Y_c, Z_c) correspond to coordinates in the pixel plane coordinate system, image plane coordinate system and camera space Cartesian coordinate system. (u_0, v_0) is the coordinate value of the coordinate system as the main point p in the pixel plane coordinate system. f is the camera focus value. According to the triangular similarity principle, the above formula is obtained. It is also expressed in the case of adding a nonzero scale factor. Therefore, (X_c, Y_c, Z_c) in the above formula can be different from other variables.

4.2. Two-view geometry and three-view geometry

The two-view geometry is used to describe the polar geometric relationships between the two images, which express the inherent projective properties between the two images. It does not involve any geometric structure of the photographic scene. It is only related to the camera's inner and outer azimuth elements. The matching relationship between the two-view geometric description is shown in Fig. 4.

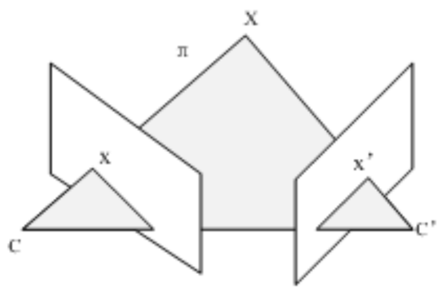

Fig. 4. Point matching relationships under two-view geometric descriptions

The three-view geometry is used to describe the polar geometries between three images, which express the inherent radio graphic properties of the three images. As with the two geometries, the three-dimensional geometry does not involve any geometric structure of the photographic scene. It is only related to the camera's inner and outer azimuth elements. The point matching relationship under the three-pole geometric description is shown in Fig. 5.

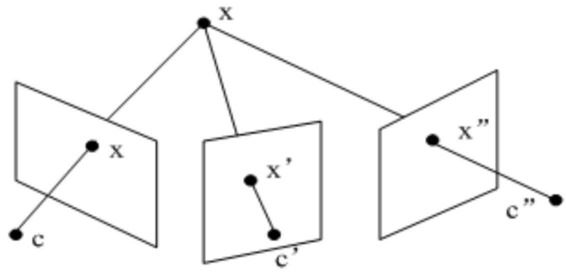

Fig. 5. Point matching relationship under three-view geometric descriptions

Two geometric geometries describe the geometric relationships between the two mating points using the basic matrix. However, when a point on one of the images is given, the specific position on the other image cannot be determined by the basic matrix. It can only be determined on a straight line through the pole. When an object point has three image points corresponding to it, the triplet tensor is used to describe it. When the object points are at known points, the corresponding image point position on the third image can be uniquely determined. Therefore, based on this strong constraint, we can remove some of the misplaced points in the two images, increasing the robustness of the existing model.

In the case of known azimuthal elements and camera lens distortion, there are many ways to solve the rotation matrix and the translation vector transformation between the two-view model and the three-view model. For example, through the relationship between points and points, triplet tensor is established. Then, the rotation matrix and the translation vector in the triplet are decomposed. According to the relative rotation and translation relationship between the two images, the rotation matrix and the translation vector between the triads are directly determined by matrix transformation. As with the two-dimensional geometric description, the amount of the three-dimensional geometry is also different from the overall scale factor.

4.3. Multi-view geometry

In the camera's inner azimuth elements and lens distortion, the problem is to use multiple two-dimensional images in Euclidean space to reconstruct the image corresponding to the three-dimensional point of the scene. Through the geometric characteristics of two views and three views, the rotation matrix, translation vector and projective depth of all three tuples can be computed from the image set by using the matching point pair relation between images. Although the relationship between the three tuples and the three tuple is uncertain, it effectively removes the external points and provides reasonable input for further multiview computation.

In photogrammetry, the nonlinear beam method is used to compute the exterior orientation elements of the camera and the coordinates of the points in the scene. Although this method relies heavily on reasonable initial values, it can obtain stable solutions in the case of good initial values. According to this, the larger part of the residual point can be removed so that the reliability of the subsequent treatment can be improved. At present, in computer vision, there are two main strategies for providing reasonable initial values for nonlinear beam method adjustment: one is incremental beam adjustment, and the other is global beam adjustment.

5. Conclusion

For low-level light and small UAV images, the stability of the platform is poor when acquiring images. It is susceptible to external weather conditions. The three rotation angles of the image are relatively large, and geometric deformation is particularly sensitive to ground elevation changes. When the size of the survey area image is particularly large, the geometric correction error generated by the approximate polynomial correction model is amplified. It is difficult to meet the requirement of image geometric correction precision when the image is spliced. At present, the light and small UAV platform cannot get accurate images when shooting outside the elements. Therefore, when using the colinearity equation as the geometric correction model of the UAV image, it is necessary to consider how to calculate the external azimuth elements of the image. On the basis of previous studies, aiming at the characteristics of remote sensing images of low - level light and small unmanned aerial vehicles (UAVs), the key technologies of UAV images are researched and analyzed. Combined with three parameters of Ransac's influence performance and iteration times, the parameters of Sift characteristic matching relation for high - resolution UAV images are given. It provides technical support for its efficient and fully automated process. Finally, the rapid processing software for remote sensing images of low altitude light and small UAVs is implemented.

References

[1] Y. Sun, K. Hara: *Fortran code of the projected shell model: Feasible shell model calculations for heavy nuclei.* Computer Physics Communications *104* (1997), Nos. 1 to 3, 245–258.

[2] Z. Zhou, M. Yan, S. Chen, Y. Lan, X. Luo: *Image registration and stitching algorithm of rice low-altitude remote sensing based on Harris corner self-adaptive detection.* Transactions of the Chinese Society of Agricultural Engineering *31* (2015), No. 14, 186–193.

[3] W. Yu, Y. Tao, A. X. Donghai: *Fast automatic stitching for images of unmanned aerial vehicle in emergency response.* Journal of Computer-Aided Design and Computer Graphics *25* (2013), No. 3, 410–416.

[4] E. Faleiro, U. Kuhl, R. A. Molina, L. Muñoz, A. Relaño, J. Retamosa: *Power*

spectrum analysis of experimental Sinai quantum billiards. Physics Letters A *358* (2006), No. 4, 251–255.

[5] Z. Li, Z. Chen, L. Wang, J. Liu, Q. Zhou: *Area extraction of maize lodging based on remote sensing by small unmanned aerial vehicle*. Transactions of the Chinese Society of Agricultural Engineering *30* (2014), No. 19, 207–213.

[6] I. Nitze, B. Barrett, F. Cawkwell: *Temporal optimisation of image acquisition for land cover classification with random forest and MODIS time-series*. International Journal of Applied Earth Observation and Geoinformation *34* (2015) 136–146.

[7] M. Z. Raqab: *Optimal prediction-intervals for the exponential distribution, based on generalized order statistics*. IEEE Transactions on Reliability *50* (2001), No. 1, 112 to 115.

[8] S. N. Sinha, J. M. Frahm, M. Pollefeys: *Feature tracking and matching in video using programmable graphics hardware*. Machine Vision & Applications *22* (2011), No. 1, 207 to 217.

Received May 7, 2017