International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.8, No.11 (2015), pp.365-374
http://dx.doi.org/10.14257/ijsip.2015.8.11.33

Research on UAV Remote Sensing Image Mosaic Method Based on SIFT

Yinjiang Jia[1], Zhongbin Su[1*], Qi Zhang1, Yu Zhang[1] Yunhao Gu[1]
and Zhongqiu Chen [2]

[1]*College of Electrical and Information, Northeast Agricultural University, Harbin 150030, China,*
[2]*China mobile communication technology engineering company, Harbin 150000, China*
jiayinjiang@126.com

Abstract

UAV remote sensing, as a new method of remote sensing, has the characteristics of higher spatial resolution, fine timeliness and high flexibility. It is widely used in the field of natural disaster monitoring, urban planning, resource investigation, and has become one of the indispensable method of remote sensing data acquisition. However, because the UAV remote sensing platform is limited by the flight height and focal length of camera, the acquired image size is smaller, single image can't cover the entire target area. Therefore, image mosaic has become a key technology to solve the problem. Image matching and image fusion are the key techniques of image mosaic. Due to the good robustness of image scaling, translation and rotation, this paper uses the SIFT algorithm to realize image matching of UAV. Since the feature extraction may produce false matches, RANSAC algorithm is applied to the feature point purification points. According to the seam-line in jointing overlap region, weighted fusion algorithm is applied to realize the image seamless splicing.

Keywords: *UAV, image mosaic, image registration, image fusion, SIFT algorithm*

1. Introduction

UAV is Unmanned Aerial Vehicle for short. With the technology development and the needs of civil field, UAV has gradually penetrated into many industries. UAV remote sensing has the characteristics of high flexibility, low cost, large scale and high accuracy, which is unparalleled for the traditional remote sensing satellites and manned aircraft [1-2]. UAV remote sensing is widely used in agricultural production, including land resource survey, crop monitoring, pest monitoring and early warning, disaster monitoring and evaluation. So it has broad application prospects [3-5].

However, since the UAV remote sensing platform is limited by the flight height and focal length of camera, the acquired image size is smaller and single image can't cover the entire target area. Therefore, in order to obtain the entire target area image, image mosaic has become a key technology to solve the problem [6]. Meanwhile, due to small size of Micro UAV, its poor stability and wind resistance, even if equipped with autopilot and stability augmentation gyro equipment, flight attitude tilt and shake phenomena are difficult to be avoided, and image distortion is serious. Therefore, it is necessary to correct and process the remote sensing images before jointing panorama mosaic [7].

UAV remote sensing image mosaic is a process of automatically splicing some certain image sequences with the overlapping area into larger field of view panorama [8]. The core technology during the process is the image registration. Image registration refers to how to make a corresponding image of the same object at different times, with different resolutions, different light, and different circumstances pose, which is directly related to

ISSN: 2005-4254 IJSIP

the image mosaic efficiency and quality [9]. Among the current image mosaic algorithm, SIFT algorithm is invariant in tilt, light, pose and scale, more widely applied in low altitude UAV image mosaic. This paper is based on the SIFT algorithm to realize UAV image mosaic.

2. UAV Remote Sensing Image Mosaic Process

Image mosaic is composed of image preprocessing, image registration and image fusion [10]. The basic process of image mosaic is as shown in Figure 1. Among them, image registration and image fusion is the key technology of image mosaic [11].

Figure 1. Basic Process of Image Mosaic

2.1. Image Preprocessing

The main task of image preprocessing is to solve the distortion problem of remote sensing image, as far as possible to reduce the impact due to image distortion, to ensure the precision and accuracy of image registration.

2.2. Image Registration

Image registration is one of the key technologies of image mosaic. Its accuracy largely determines the quality of image mosaic [12]. Image registration is at, two or more images acquired at different angles for optimal matching process. Image registration is the best matching processing of two or more images from the different times with different sensors [13].

Image registration process generally includes the four key steps. The first step is feature extraction. Feature extraction is the first step, but also is the most critical step. Feature extraction directly determines the image registration follow-up speed and accuracy. Its main task is to extract similar features in the reference image and registration image, the most representative feature is point, curve, and contour [14-16]. Image feature point extraction algorithm includes SUSAN, Harris and SIFT, *etc.* SIFT algorithm has strong robustness to image scaling, rotation, light and noise [17-18]. This paper uses SIFT algorithm for feature extraction.

The second step is feature matching. Feature matching is obtained by extraction of the reference image, and matching the two feature sets of the image through a series of similarity measure, which establish a corresponding relation between the similar characteristics.

The third step is model transformation and parameter estimation. The image registration process is geometrical transformation between different images, and transformation model is a mathematical description of this transformation process. Estimate model parameters are defined by feature matching obtained matching features. The most commonly used model transformations are rigid transformation, affine transformation, projection transformation and nonlinear transformation.

The forth step is image interpolation and transformation. According to the model of model transformation and parameter estimation, registration image make the corresponding geometric transform by reference image, and achieve non-integer coordinates gray image interpolation of registration image.

2.3. Image Fusion

Image fusion is an indispensable step in image mosaic. Due to the different shoot condition and the effect of illumination, after image matching, there exists obvious mosaic trace. The purpose of image fusion is to eliminate the image intensity or discontinuity of color, so as to achieve a smooth transition [19]. The most commonly used image fusion method is average method, weighted fusion algorithm, median filtering fusion algorithm.

3. SIFT Algorithm Principle

SIFT algorithm is the characteristic scale of the image on the basis of ideological choice, to create a multi-scale space images, which can get the Gaussian pyramid [20-21]. Then subtract the adjacent Gaussian image to get the DOG (difference-of-Gaussian) pyramid. And scale in DOG that each point scale with point and the adjacent position to the next adjacent one by one space for the resulting local extreme point which is the keypoint candidate. In scale space and image space interpolation, so that the keypoint position accuracy is improved, with a 128 dimensional vector representation of the gradient region near the keypoint that is SIFT feature descriptors. And make the subsequent feature matching capability stronger. SIFT keypoint extraction consists of 4 steps.

3.1. Construction of Scale Space

The first step is construction of scale space, detection of extreme points. The scale space theory thought is using Gauss transform to check the original image scaling. Multi-scale image obtained under the scale space representation of the sequence, then make these sequences of the scale spatial feature extraction. Gaussian convolution kernel proved to be the only nuclear achieve linear scaling. The definition of two-dimensional Gaussian convolution can be expressed by equation:

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

Scale space image is as shown in equation as follows:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{2}$$

Where L represents a scale space, (x, y) is the pixel coordinates of the image point, $I(x, y)$ is the image data, σ is the scale space factor and the variance of the Gaussian normal distribution, which reflects the extent of the image is smoothed. The smaller the value characterizing the degree of smoothing the image is smaller, the smaller the appropriate scale.

In order to efficiently extract stable keypoint in scale space, Lindeberg proposed a differential Gaussian scale space DOG (Difference of Gaussian) to detect the extreme point. Differential function represents two differential Gaussian kernel functions at different scales, DOG operator is defined as follows:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{3}$$

Where k is the scaling factor for the scale space adjacent and a constant, which generally take $\sqrt{2}$. SIFT algorithm to generate a DOG by constructing a pyramid image that shown in Figure 2. Gaussian pyramid is divided into multiple groups. Each group was divided into multiple layers, a difference of scale between adjacent layers of a scaling factor k. The Gauss image down-sampled by a factor of twice, and each group to generate the next set of scale space images, each set of adjacent layers image subtraction, a difference image is generated Gaussian pyramid.

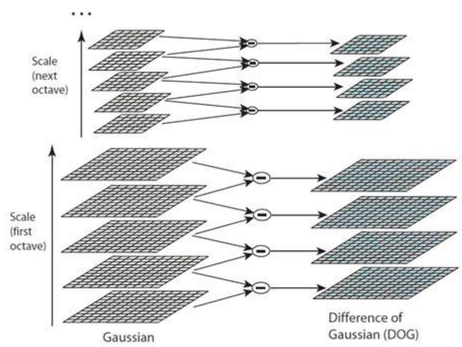

Figure 2. Gaussian Pyramid and Difference of Gaussian Pyramid

To search for extreme point of DOG, each point and all its neighboring points should be compared, with the layer 8 and the adjacent points corresponding to the upper and lower adjacent layers 18 points which a total of 26 points. If the value of sample is the minimum or the maximum point that is considered an extreme point.

3.2. Keypoint detection

The second step is to determine the keypoint. Extreme point of the generated keypoint are not all stable, which should be excluding those edge points that relatively low contrast and unstable. The Difference of Gaussians used of the Taylor expansion at key points can be expressed by equation:

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \tag{4}$$

On both sides of the equation and set the derivative is zero, we can get extreme points:

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \tag{5}$$

DOG operator at extreme point is:

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \hat{X} \tag{6}$$

Here $\left| D(\hat{X}) \right| \geq 0.3$, that retains the keypoint, otherwise remove the keypoint.

Because of the keypoint at the edges is not stable enough, so we should eliminate the unstable edge response point. Hessian matrix can be used to eliminate the edge of instability point. Hessian matrix can be expressed by equation:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{7}$$

Where D_{xx}, D_{xy}, D_{yy} is the second derivative. For matrix H, the maximal eigenvalue is α, the minimal eigenvalue is β, and $\alpha = \gamma\beta$, there are:

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

$$Det(H) = D_{xx} D_{yy} - (D_{xy})^2 = \alpha\beta \tag{8}$$

Where $Tr(H)$ is the tracks of matrix H, $Det(H)$ is the value of the determinant of matrix H.

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\alpha\beta + \beta)^2}{\alpha\beta^2} = \frac{(r+1)^2}{r} \tag{9}$$

Therefore, the follows equation can be used to determine whether the principal curvature is less than some threshold.

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \tag{10}$$

3.3. Determination of the Keypoint

The third step is to determine the main direction of the keypoint. In order to have a rotation invariant of SIFT operator, we need to determine the maximum gradient direction of the keypoint. For each point of the Gaussian image, the gradient magnitude $m(x, y)$ and orientation $\theta(x, y)$, is can be obtained by the equation:

$$\begin{cases} m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \\ \theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y))) \end{cases} \tag{11}$$

Gradient direction of the keypoint is statistics from the gradient histograms. The highest peak in the histogram representative keypoint main direction, then any other peak that is greater than 80% of the highest peak is used to an auxiliary direction of the keypoint.

3.4. Generation of Keypoint Descriptor

The forth step is to generate the keypoint descriptor. SIFT descriptor is a description of the SIFT keypoint area. In the configuration descriptor, taken the keypoint as the center of the size of the 16× 16 neighborhood window, then evenly divided into 4× 4 sub regions, as shown on the left in Figure 3, the gradient direction calculating eight histogram of each area map, as shown on the right in Figure 3, which constitutes 4× 4× 8=128 dimensional vector which is the SIFT descriptor. As shown on the right in Figure 3, the length of each arrow is gradient magnitude and the direction of the arrow is the gradient direction.

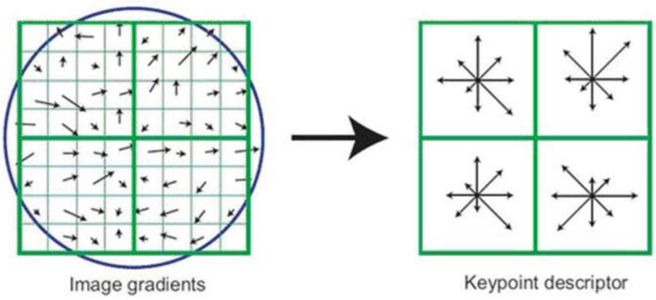

Figure 3. SIFT Keypoint Descriptor

4. The Experimental Environment and Data

In this paper, the selected experimental environment is as follows, CPU is Intel (R) Core (TM) i7-3770 3.40GHz, HP desktop with 8GB of memory. UAV remote sensing platforms is TY-6 fixed-wing UAV, which fuselage length 2.6 meters, 4 meters wingspan, flight level is 300-1000 meters altitude, flight speed is 70-110 kilometers per hour, the control mode is automatically controlled program, record sensor device is Canon EOS 6D, focal length is 50mm, the exposure time is 1/1600 of a second, 20 million effective pixels, the maximum resolution images is 5472 * 3648.

Research group conducted a number of UAV experiments in Heilongjiang province Fangzheng County Rice Research Institute during July to August 2014, and obtained a lot of experimental data. Figure 4 shows the UAV route plans on July 16, 2014. Among them, the route is designed as serpentine route, heading overlap is set to 80%, next to the overlap is set to 50%, the figure scale is 1: 500, with a ground resolution of 0.05 meters, shooting image area is the rice experimental fields belong to Fangzheng County Rice Research Institute.

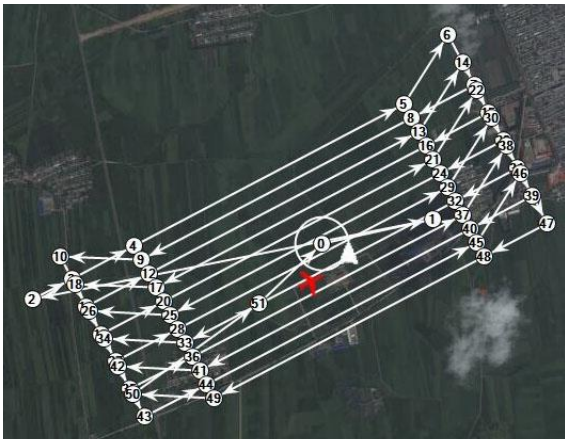

Figure 4. The UAV Route

5. The Experimental Results and Analysis

Experiment is conducted in strict accordance with image mosaic process. This paper mainly analyzes and discusses the image matching and image fusion.

5.1. Image Read

Reference image and registration image is as shown in Figure 5.

Figure 5. Reference Image and Registration Image

5.2. Gaussian Pyramid Build

Gaussian pyramid building first is Gaussian-smoothed, then is down-sampled. Figure 6 shows a forth order five layer Gaussian pyramid. As can be seen from the figure, with the number of layers increases, the image is gradually blurred so as to reduce image noise.

Figure 6. Gaussian Pyramid

5.3. Difference of Gaussian Pyramid Build

Adjacent Gaussian images are subtracted to produce the difference of Gaussian images based on the Gaussian pyramid. Figure 7 shows Difference of Gaussian pyramid. It is can be seen that the DOG image is a description of the target image contour.

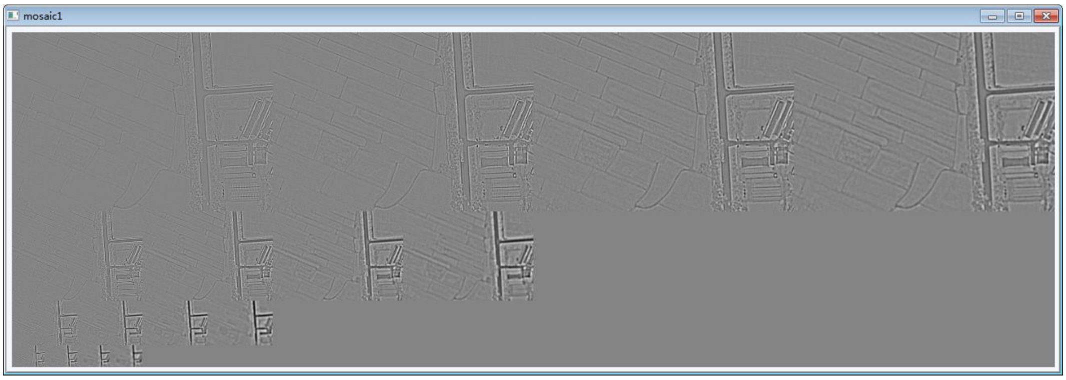

Figure 7. Difference of Gaussian Pyramid

5.4. Feature point extraction

Feature point extraction is to find extreme value in scale space. Feature point extraction image is as shown in Figure 8. In order to improve the image processing speed, as the difference of Gaussian pyramid construction, image equal scaling is to be done. The feature point extraction and the elapsed time are as shown in Table 1.

Figure 8. SIFT Feature Vectors

Table 1. Feature Point Extraction

Image size	Feature points	Time consuming（s）
480×320	189	0.846
640×426	313	1.513
800×533	489	2.345
1024×683	824	4.052

5.5. Feature point matching

This paper adopts Nearest Neighbor method for coarse feature point matching. The current image feature points with the ratio of the Euclidean distance between another image the nearest feature points and its nearer feature points is less than a predetermined threshold value, consider that the recent feature point is the current feature point matching point. Image of feature point matching is as shown in Figure 9.

Figure 9. Feature Point Rough Matching

5.6. Feature Point Purification

As can be seen from Figure 10, some matching points is wrong. It is necessary to deal with coarse matching feature point that obtained from the previous step for purification processing. RANSAC (Random Sample Consensus) algorithm is used to delete false matching points. Its idea is to design an objective function according to the specific problem, and then extract the minimum set of points by repeatedly estimating the initial value of the function parameters, all the data are divided into interior points and external points by using these initial values. In the end, interior points are used to recalculate function parameters [22]. The image by using RANSAC algorithm is as shown in Figure 10.

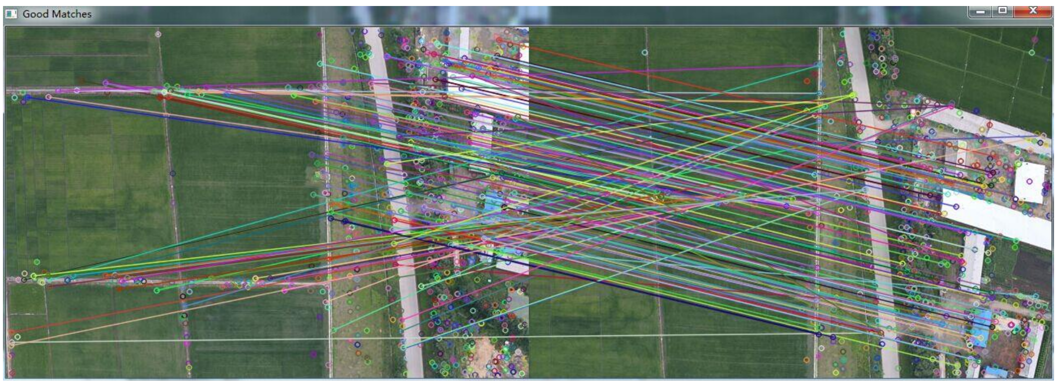

Figure 10. Feature Point Precise Matching

5.7. Image fusion

If the two images in adjacent, there will be obvious mosaic trace, we call it seam or ghosting, as shown in Figure 11 (a). This paper adopts weighted fusion algorithm to eliminate mosaic trace, the algorithm idea is to the pixel values of the overlapping area, multiplied by the respective weights and then superimposed on average, as shown follows:

$$f(x,y) = \begin{cases} f_1(x,y) & (x,y) \in f_1 \\ w_1(x,y) f_1(x,y) + w_2(x,y) f_2(x,y) & (x,y) \in (f_1 \cap f_2) \\ f_2(x,y) & (x,y) \in f_2 \end{cases}$$

Where w_1 and w_2 are the two images overlap region corresponding pixel weights, and satisfy $w_1 + w_2 = 0$, $0 \le w_1, w_2 \le 1$. Weighted fusion algorithm achieves a smooth transition of two images overlapping regions, as shown in Figure 11 (b).

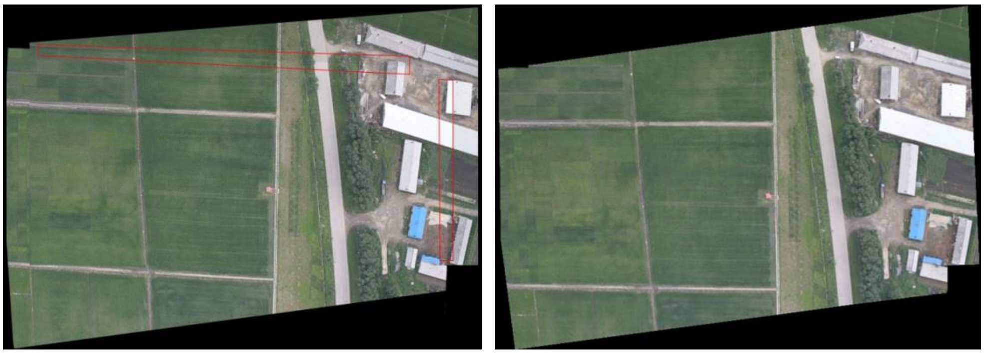

Figure 11. Image Fusion

6. Conclusions

With the rapid development of UAV and remote sensing technology, UAV remote sensing has been widely used, because single remote sensing images can not cover the entire target area, and it is necessary to splice multiple remote sensing images. This paper focuses on image matching and image fusion of image mosaic technology research, using SIFT algorithm for UAV remote sensing image mosaic. Since the feature extraction may produce false matches, RANSAC algorithm is applied to the feature point purification point. According to the seam-line in jointing overlap region, weighted fusion algorithm is applied to realize the image seamless splicing.

This paper mainly studies the image mosaic algorithm based on feature point. As to image feature extraction, besides the point features, there are line features, facial features, edge features, etc to be considered. In addition, for the oversized image, due to the large number of feature points, high program memory usage, large amount of calculations, system processing effect is not ideal. These will be further studied in the future work.

Acknowledgments

The authors would like to acknowledge the financial support from the National High-tech R&D Program of China (863 Program) (2013AA102303) and Key Technologies R&D Program of Harbin (2013AA6BN010).

References

[1] Y. Hong, J. Gong, S. Hu and M. Huang, "Discussiong on UAV Remote Sensing Image Capture and Disposal", vol. 23, no. 4, **(2008)**, pp. 462-466.
[2] J. Zhang, "The Research of UAV Remote Sensing Image Fast Splicing Method[J]", Urban Geotechnical Investigation and Surveying, vol. 5, **(2013)**, pp. 73-75.
[3] J. Su, "Mosaicing of Multiple Spectrum Images Acquired from Unmanned Airship with SIFT Feature Matching [J]", Journal of Applied Sciences, vol. 28, no. 6, **(2010)**, pp. 616-620.
[4] H. Yu and B. Kong, "Automatic Seamless Mosaic of UAV Remote Sensing Image [J]", Remote Sensing Technology and Application, vol. 27, no. 3, **(2012)**, pp. 347-352.
[5] B. Yang, X. Li, L. Yan and Y. Yu, "Mosaic Algorithm for Aerial Remote Sensing Image and Its Accuracy Evaluation [j]", Computer Engineering, vol. 40, no. 3, **(2014)**, pp. 253-257.
[6] M. Li, H. Liu and X. Zhu, "Approach to Fast Mosaic UAV Images of Disaster Emergency [J]", Journal of Catastrophology, vol. 27, no. 3, **(2012)**, pp. 139-144.
[7] J. Wang, J. Shi and X. Wu, "Survey of Image mosaics Techniques[j]", Application Research of Computers, vol. 25, no. 7, **(2008)**, pp. 1940-1944.
[8] L. Cai, Y. Liao and D. Guo, "Study on Image Stitching Methods and Its Key Technologies[j]", Computer Technology and Development, vol. 18, no. 3, **(2008)**, pp. 1-4.
[9] X. Liu, J. Yang and J. Sun, "Image Registration Approach Based on SIFT[j]", Infrared and Laser Engineering, vol. 2, no. 1, **(2008)**, pp. 25-29.
[10] H. Liu and H. Shen, "Application of Improved SIFT Algorithm on Stitching of UAV Remote Sensing Image[j]", Semiconductor Optoelectronics, vol. 35, no. 1, **(2010)**, pp. 108-112.
[11] R. Yin and S. Chen, "Technical Research Based on ASIFT of Low Altitude Remote Sensing Mosaic[j]", Journal of Jimei University, vol. 17, no. 2, **(2012)**, pp. 153-160.
[12] X. Zhao and L. Du, "An Automatic and Robust Image Mosaic Algorithm[j]", Journal of Image and Graphics, vol. 9, no. 4, **(2004)**, pp. 417-422.
[13] Y. Cheng, Y. Zhou, H. Lin and H. Pan, "Automatic Registration and Mosaic of Remote Sensed Imagery Based on SIFT Feature[J]", Remote Sensing Technology and Application, vol. 23, no. 6, **(2008)**, pp. 721-728.
[14] L. Yu and J. Xie, "A Remote Sensing Image Classification Method Combining Multiple Features[J]", Computer Application and Software, vol. 31, no. 11, **(2014)**, pp. 183-185.
[15] Z. Liu and J. Chu, "Feature Extraction and Template Matching Combined with Image Stitching Method[j]", Micro-Computer Information, vol. 26, no. 1, **(2010)**, pp. 117-118.
[16] D. Li and Y. Wang, "Research of the Image Mosaic Method Based on Corner Match[j]", Chinese Journal of Electron Devices, vol. 31, no. 3, **(2008)**, pp. 919-922.
[17] C. Wang, C. Guo, X. Duan and Q. Zhang, "Grid Affine Image Mosaic Algorithm based on SIFT[J]", Computer Engineering and Applications, vol. 47, no. 36, **(2011)**, pp. 187-190.
[18] T. H. Zhou and J. Zhang, "Robust Virtual Image Stitching Method by Image Warping[j]", Journal of Wuhan University of Technology, vol. 30, no. 7, **(2008)**, pp. 130-133.
[19] H. Wen and J. Zhou, "Improved Algorithm for Remote Sensing Image Mosaics[J]", Electronics Options and Control, vol. 16, no. 12, **(2009)**, pp. 34-37.
[20] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints[j]", International Journal of Computer Vision [J]", vol. 60, no. 2, **(2004)**, pp. 91-110.
[21] M. Brown and D. G. Lowe, "Automatic Panoramic Image Stitching Using Invariant Features[j]", International Journal of Computer Vision[J], vol. 74, no. 1, **(2007)**, pp. 59-73.
[22] Y. Wu, T. Yu and D. Xie, "Fast Automatic Stitching for Images of Unmanned Aerial Vehicle in Emergency Response[j]", Journal of Computer-Aided Design and Computer Graphics, vol. 25, no. 3, **(2013)**, pp. 410-416.

Author

Yinjiang Jia (1977), male, born in Harbin, Heilongjiang Province, China. he is a D.E. student in the college of Electrical and Information at northeast agricultural university. His research direction is agricultural remote sensing and information technology.

