

An improved SIFT algorithm in the application of close-range Stereo image matching

Xuehua Zhang¹, Xiaoqing Wang¹, Xiaoxiang Yuan¹, Shumin Wang¹

¹Institute of Earthquake Science, China Earthquake Administration, Beijing, China 100036

E-mail: zhangxhua1990@163.com wangxiaoq517@163.com yuanxx1983@163.com wmcnu@163.com

Abstract. As unmanned aerial vehicle (UAV) remote sensing is applied in small area aerial photogrammetry surveying, disaster monitoring and emergency command, 3D urban construction and other fields, the image processing of UAV has become a hot topic in current research. The precise matching of UAV image is a key problem, which affects the subsequent processing precision directly, such as 3D reconstruction and automatic aerial triangulation, etc. At present, SIFT (Scale Invariant Feature Transform) algorithm proposed by DAVID G. LOWE as the main method is, is widely used in image matching, since its strong stability to image rotation, shift, scaling, and the change of illumination conditions. It has been successfully applied in target recognition, SFM (Structure from Motion), and many other fields. SIFT algorithm needs the colour images to be converted into grayscale images, detects extremum points under different scales and uses neighbourhood pixels to generate descriptor. As we all know that UAV images with rich colour information, the SIFT algorithm improved through combining with the image colour information in this paper, the experiments are conducted from matching efficiency and accuracy compared with the original SIFT algorithm. The results show that the method which proposed in this paper decreases on the efficiency, but is improved on the precision and provides a basis choice for matching method.

1. Introduction

Unmanned aerial vehicle (UAV) is a kind of integrated the technologies of UAV remote sensing system, flight control system, remote sensing and GPS navigation and positioning, and other new technology. Because its characteristics, such as high efficiency, less cost and low labour intensity, increasingly cause the attention of many fields. UAV remote sensing has been widely applied in precision agriculture^[1, 2], fire detection^[3] and characteristics of forest^[4, 5], especially in the building 3D modelling of digital city^[6]. In the process of images processing, the corresponding points matching is a key problem, as its result will directly affect the final subsequent precision, such as image matching, 3D reconstruction and automatic aerial triangulation^[7]. Now in order to solve this problem there are two main methods: corresponding points matching based on grey correlation and corresponding point matching based on features. Based on grey correlation matching method has the advantages in calculation, but the accuracy are greatly influenced by SNR(Signal-to-Noise Ratio) of window, and there will be a low matching success rate if the grey level information is poor or SNR is small, and it is difficult to guarantee the stability of the matching accuracy^[8]. By contrast, the method of based on the features have a stable matching precision, Mikolajczyk compared some representative feature matching algorithms, such as moment invariants, SIFT, etc., through the experiments and



analysis of the pictures that in the cases of image geometric changes, compression and illumination change and so on, the SIFT algorithm show the best extracting result^[9]. SIFT algorithm has strong stability to image scaling, rotation, shade, light and other conditions change^[10], and has been successfully applied in target recognition, SFM (Structure from Motion), and many other fields. But the SIFT algorithm is based on image grey value to seek extreme value point as characteristics points and convert colour images to grayscale images, not use the rich colour information of the UAV images. Many scholars to research the SIFT algorithm in order to improve the precision, such as Jijia Xu in combination with Harris operator, this paper proposes a fast images matching based on improved Harris - SIFT operator method^[11]; Ping Xie proposed the method to improve the matching accuracy through improving the corresponding points matching^[12]. Therefore, this paper will combine the images colour information to improve the SIFT algorithm. Experimental results show that the method in matching efficiency is slightly lower, but the precision is better than the original SIFT algorithm.

2. Extracting and matching feature points

2.1. Feature points extraction

The number and the precision of the control points eventually used to matching determines the accuracy of images matching^[13]. For the characteristics of the UAV images such as high resolution, flight attitude instability and different light conditions, this paper uses the SIFT algorithm to detect the points. The first step is to convert colour images to grey images when the feature points are extracted, grey images are sampled fall under the different levels of images structure. In order to retain the image colour information of original image, the same degree of reducing the sampling method is adopted in this paper, so that we can get feature points corresponding to the colour information. Then using a different scale factor of the Gaussian function to each layer (Eq. 1), forming the pyramid structure under different scale space (σ). This method simulates the human eye observation from coarse to fine scenery, the outline to the details of mode of observation, to satisfy the visual invariance. In order to more effectively detect the stability of the extreme value points, we use the pixel value of adjacent two layers that after filtering to subtract to get difference of Gaussians pyramid images (DoG)(Eq. 2). Extract the extreme value points of stability of the DoG scale space, must compare the pixel value between each pixel point and all of its adjacent points, including the same scales around eight points and adjacent scale up and down in the corresponding 9×2 points, ensure that point for the image space and scale space under the local minimum or maximum (as shown in Fig. 1). At the same time, record the extreme value point, scale, and the coordinates of the pyramid.

$$L(x, y, \sigma) = F(x, y, \sigma) * I(x, y); \quad (1)$$

Among: (x, y) —— the coordinates of the pixel

$$F(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \text{ —— Gaussian blur function}$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma); \quad (2)$$

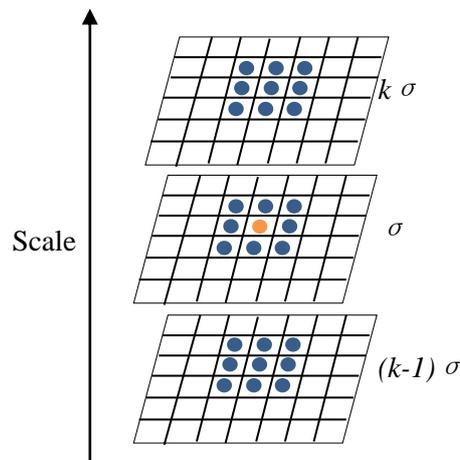


Figure 1 Extreme point in the DoG

Then get the extreme value point location is in discrete space under the rough position. Lowe David using three dimensional quadratic function to correct the point's position. In order to remove the extreme points of low contrast, we set a threshold value that the point's position correction is greater than the threshold will be deleted. In addition, DoG operator will produce strong edge effect, and need to use the eigenvalues of the Hessian matrix ratio to weed out the points which edge is not stable. The Hessian matrix of the H (Eq. 3), maximum and minimum eigenvalue is α 、 β respectively, and $\alpha = r\beta$, $\text{Tr}(H)$ (trace of H , Eq. 4), $\text{Det}(H)$ (determinant of H , Eq. 5):

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (3)$$

$$\text{Tr}(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (4)$$

$$\text{Det}(H) = D_{xx}D_{yy} - D_{xy}^2 = \alpha\beta \quad (5)$$

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r+1)^2}{r} \quad (6)$$

Set a threshold value, if the ratio is greater than the threshold is determined that point as the edge points, delete them. After the above steps extreme value points of quality was improved, identified as the final feature points.

2.2. To determine the main direction

In order to guarantee the algorithm of rotation invariance, we need to rotate selected feature points and its local characteristics, make sure all feature points generated by the descriptor in the same direction vector. Different feature points with different local grey-level gradient distribution, so it can be determined a principal direction as the direction of the stability of local structure according to gradient distribution around the point. In the first step we have determined the feature points in the DOG pyramid, to collect all the pixels value which located the Gaussian pyramid image within 3σ neighbourhood window of the point, press type calculation of the gradient of each pixel (direction Fig. 8 and amplitude Fig. 7).

$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2} \quad (7)$$

$$\theta(x, y) = \tan^{-1} \left[\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right] \quad (8)$$

After computing feature points and neighbourhood of gradient, using histogram method to statistics. Gradient will be divided by direction in the histogram graph, the interval is 10 degrees, a total of 36 columns. Histogram peak direction represents the main direction of the point. Lowe David thought only keep peak is greater than the principal direction of the direction of the peak of 80% as the auxiliary direction, this can set up multiple different directions at the same location and scale of feature points to increase the stability of the match.

2.3. Generate the descriptor

After the above processing each feature points have three information: position, scale and direction. SIFT algorithm is to establish a set of feature points vector as a descriptor, including this point and its surrounding point's features. Lowe David was used the feature points in the scale space of $4 * 4$ window to calculate eight direction of gradient information, a total of $4 * 4 * 8 = 128$ dimensional, and proved that the $4 * 4$ window as the best detection window. This paper uses Lowe David's conclusion, but need to extract RGB information according to the feature points location, which makes the 128 dimensional descriptor increased to 131 dimensional, makes the feature points matching is not only has the characteristics of the original SIFT algorithm, but considered the rich colour information of the images, finally get a stable matching results.

3. Experimental analysis

3.1. Experimental data

This paper used four rotorcraft UAV for data acquisition, flight radius about 5 km, the SONY A6000 16-50 mm lens, the image below for the UAV.



Figure 2 The photo of the UAV

Its wheelbase diagonal motor is 105cm, there are many types of flight mode, including artificial remote autonomous hover, independent route and go around the points of interest, the biggest task load can reach 8kg, maximum take-off weight of 14kg, battery life varies according to the load, when the load is zero it can flying 90 minutes. The cruising speed for 3-50km/h, 6 class wind resistant effect. Navigation method have three modes: GPS navigation, beidou navigation and combine GPS with beidou navigation.

The data acquisition area is Beizhen city, Liaoning province. The flight course overlap rate can reach 74%, side overlap rate was 40%, and the average speed is about 15m/s. We get a series of high quality images because the weather is good and no wind. The aircraft sorties flying height is about 120m, the images spatial resolution is 0.02m. The images like Fig. 3.



Figure 3 the sample pictures and range that we acquired

3.2. Experimental results and analysis

In this paper, we will use VS2010 compiler as a platform, with the aid of using Opencv2.4.10 function library to implement the traditional SIFT matching algorithm and SIFT matching algorithm with colour information. Purpose research discussed with colour information SIFT and effectiveness and efficiency of the traditional SIFT algorithm, therefore not use all flight images. By using two images to contrast the differences of the two methods, in the aspect of time may not be much difference, therefore, decided to use the ratio of time as a measure. Used in the experiment process parameters are considered as the best value in the Lowe David's experiments.

Through use of the two images that has the overlap area of UAV images as experiment analysis data. Below pictures are the effect pictures respectively by using traditional SIFT method and the SIFT which combined with colour information and corresponding points matching. In Fig. 4 is the images that used a traditional method of SIFT under the different scales of Gaussian pyramid and DoG. From the pictures we can see the different objects in different layers of the pyramid, conforms to the human eye to observe the features of the object. In DoG, image feature points and line is more apparent and can be good to extract the feature points in the hierarchy. The extraction of feature points set is used in the original method of matching eventually and get the matching results of the method. By the method of computing the match to get 5219 points, with the same consistency by using random sampling algorithm (RANSAC) resulting from the processing of the same number of points to 3298, this experiment shares 26.284 seconds.

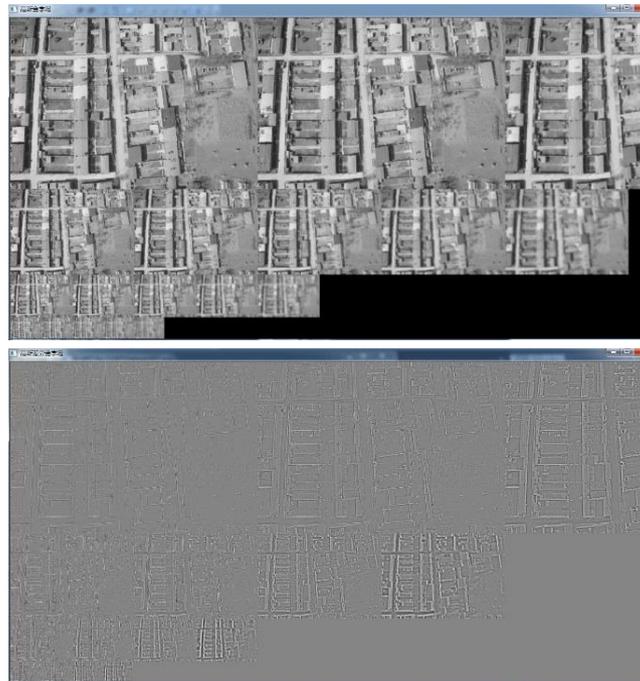


Figure 4 The Gaussians and the DoG

On the basis of using OpenCv function library to improve the program, through using the function in the library's to generate colour description operator to get the corresponding points, a total of 4805(Fig. 5), after removal of gross error and two-way matching strategy received 3547 valid points, a total of 46.204 seconds.



Figure 5 Key Feature Points of the two Pictures

Through the experiment we can get the comparison of the accuracy and efficiency of two methods, experimental results are integrated in the table below for two methods.

Table 1 Comparison of two methods

method	Total corresponding points	Effective corresponding points	Elapsed time (s)
Traditional SIFT	5219	3298	26.284
SIFT which combined with colour information	4805	3547	46.204

From the above statistics can be seen that the improved SIFT algorithm have more effective corresponding points, and the ratio is 63.19% for total corresponding points. The effective points of the traditional SIFT algorithm is less, the ratio is 73.82%. For the same images for both effective points was similar, with the ratio is about 0.93. In terms of the efficiency of the improved SIFT algorithm is more elapsed time, it is about 1.7 times of the former, the reason is that with the colour information of the SIFT algorithm in matching dimension is more, and it is also possible that has certain influence in the program optimization.

Comprehensive the above analysis when using SIFT algorithm for UAV images matching, we can according the need to choose a different method. When to process a large amount of data, we can give priority to efficiency to choose the traditional matching SIFT algorithm. When the amount of data is small, we can consider the precision priority, SIFT algorithm with colour information is a good choice.

4. Conclusion

In this paper, we represented the traditional SIFT algorithm and the colour information SIFT algorithm, using the same UAV images to perform experiment and analyses, and found out that the colour information of SIFT algorithm is higher effective corresponding points ratio than the traditional SIFT algorithm, but lacking in efficiency, and time consuming in this experiment is about 1.7 times that of traditional SIFT algorithm, especially in a large amount of data. The reason of time consuming, on the one hand is the dimensions of the colour information of SIFT algorithm is higher when corresponding points matching, on the other hand may be the program optimization problem. This experiment will make a certain guiding role for UAV image matching using the SIFT algorithm.

Acknowledgments

The research was supported by the Fundamental Research Funds of Institute of Earthquake Science, China Earthquake Administration (Grant No. 2013IES0204).

References

- [1] Valente, J. A., Sanz, D., Del Cerro, J., Barrientos, A. and de Frutos, M., 2013. Near-optimal coverage trajectories for image mosaicking using a mini quad-rotor over irregular-shaped fields [J]. *Precision Agriculture* 14, pp. 115–132.
- [2] Zhang, C. and Kovacs, J., 2012. The application of small unmanned aerial systems for precision agriculture: a review [J]. *Precision Agriculture* 13, pp. 693–712.
- [3] Merino, L., Caballero, F., Martinez-de Dios, J., Maza, I. and Ollero, A., 2012. An unmanned aircraft system for automatic forest fire monitoring and measurement [J]. *Journal of Intelligent & Robotic Systems* 65, pp. 533–548.
- [4] Tao, W., Lei, Y. and Mooney, P., 2011. Dense point cloud extraction from UAV captured images in forest area. In: *Spatial Data Mining and Geographical Knowledge Services [C], 2011 IEEE International Conference*, pp. 389–392.
- [5] Wallace, L. O., Lucieer, A. and Watson, C. S., 2012. Assessing the feasibility of UAV-based Lidar for high resolution forest change detection [C]. *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 499–504.

- [6] Tang min, Li yongshu, Lu heng. 2012. Correspondence Points Matching of UAV Images [J]. *Opto-Electronic Engineering*. 39(3): 19-24.
- [7] Liu you. 2012. Research on Matching and Mosaicking Methods for UAV Image Sequence [D]. CENTRAL SOUTH UNICERSITY. pp3.
- [8] Liu qingyuan, et, al. 2010. Research on key technology of digital photogrammetry based on UAV remote sensing image [J]. *Science of Surveying and Mapping*. 35(1): 28-30.
- [9] Mikolajczyk K, Schmid C. Scale & Affine Invariant Interest Point Detectors [J].2004. *International Journal of Computer Vision*, 60 (1):63-86.
- [10] Lowe David G. 2004. Distinctive Image Features from Scale-Invariant Keypoints [J]. *International Journal of Computer Vision*. 2(60): 91-110.
- [11] Xu jiajia, Zhang ye, Zhang he. 2015. Fast image registration algorithm based on improved Harris-SIFT descriptor [J]. *Journal of Electronic Measurement and Instrumentation*. 1(29): 48-54.
- [12] Xie ping, Zhou zhengrong, Xiao qi. 2012. A close-range image matching based on Harris corners and SIFT features [J]. *Science of Surveying and Mapping*, 37(4): 107-110.
- [13] Schmid C , Mohr R , Bauckhage C. 2000. Evaluation of interest point detectors [J]. *International Journal of Computer Vision (S0920-5691)*, 37 (2): 151-172.