International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.7, No.3 (2014), pp.395-406
http://dx.doi.org/10.14257/ijsip.2014.7.3.32

A Sub-Harris Operator Coupled with SIFT for Fast Images Matching in Low-altitude Photogrammetry

Haiqing He[1*], Xiaoyong Chen[1], Bo Liu[1] and Zhiyong Lv[2]

[1]*Faculty of Geomatics, East China Institute of Technology, Nanchang, China*
[2]*School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China*
**Corresponding author: hyhqing@163.com*

Abstract

Fast and robust images matching are of vital importance in low-altitude photogrammetric process. Among the most popular features of images matching are currently SIFT and Harris etc. For time-critical applications such as disaster monitoring, the SIFT features extraction are too slow, and the location accuracy of SIFT and Harris is insufficient in photogrammetric process. In this paper, we present a sub-Harris operator coupled with SIFT for fast images matching in low-altitude photogrammetry. Firstly, the original stereopair is down-sampled to small scaling images, in which rough relative orientation is computed by the corresponding points obtained from SIFT matching. Then sub-pixel level precise Harris corners are extracted within original scale images. Finally, the corresponding points are found in the sets of sub-Harris corners consistent with epipolar geometry obtained from rough images matching. Experimental results show that the proposed method can achieve more excellent performances in accuracy than SIFT and Harris operator based method in the relative orientation, and significantly improve the computational efficiency compared with SIFT.

Keywords: *low-altitude image, SIFT matching, sub-Harris operator, epipolar geometry, corresponding points*

1. Introduction

In the recent few years, Low-altitude photogrammetry using UAV, airship or balloon has been applied extensively in the civil field, while most of the applications are disaster monitoring, resource investigation, city planning and surveying. Compared with satellite or traditional airplane photogrammetry, low-altitude photogrammetry has many advantages: high resolution, high efficiency, high flexibility and low cost. Considering these advantages, low-altitude photogrammetry has become a hot topic throughout the world and a popular field of research.

Compared with traditional aerial photography, low-altitude photography has great differences in platform, flight altitude, sensor, pose *etc*. A single low-altitude photographic image only covers a small area, so it demands hundreds of images to have the whole interesting area be covered completely. And image has irregular overlap, big rotation angle, *etc.*, Therefore, the robustness and high efficiency of images processing are of great significance for fast applications such as disaster monitoring. Among the data processing of low-altitude photogrammetry, images matching are a key and time-consuming step, so fast and robust images matching is of vital importance. In the photogrammetry field, an image matching mainly includes template-based and feature-based approach, the latter is more common. At present, the most widely used method is SIFT (Scale-Invariant Feature

ISSN: 2005-4254 IJSIP

Transform) algorithm proposed by Lowe [1], which has been used for the applications of photogrammetry and computer vision. Moreover, SIFT has also been applied in low-altitude applications such as visual real time mapping for unmanned aerial vehicles [2], automatic feature extraction and matching [3], visual global positioning system for unmanned aerial vehicles [4], UAV images mosaic [5].

However, in low-altitude photogrammetry, only using SIFT for images matching, there are many disadvantages as follows: 1) As a result of constructing multi-layers image pyramid structure, it needs a lot of memory to extract and match features, which is one of the most time-consuming processing parts; 2) A large number of features can be extracted by SIFT, so features matching may be ambiguous, and the corresponding points may be not evenly distributed; 3) In SIFT algorithm, the optimal location of key points are determined gradually by moving half of one pixel in each iteration, so the location precision of the matched corresponding points is not better than traditional least square images matching method (the precision may reach 1/10 pixel matching level) proposed by Prof. Ackermann [6, 7]. In [8], a speeded up robust features algorithm named SURF is presented, feature extraction using SURF is faster than SIFT, but it is still too slow for tens of millions of pixels low-altitude images matching. And SURF reduces the feature location accuracy [3], SIFT is better than SURF in invariance to scale and rotation [9]. In [3], the accuracy of SIFT matching in photogrammetric process can be improved by using the least square matching algorithm. To accelerate feature extraction and matching, an approach of combining Harris interest points and the SIFT descriptor for fast scale-invariant object recognition is presented in [10], which extracts interest points using Harris corner detector instead of SIFT. However, the location accuracy of Harris corner is pixel-level, which is insufficient in photogrammetric process.

To achieve images matching invariance to translation, scaling, rotation, and illumination changes, and take into account precision and efficiency, in this paper, we present a sub-pixel level precise Harris operator (sub-Harris operator) coupled with SIFT for fast images matching in low-altitude photogrammetry. Firstly, the original stereopair is down-sampled to small scaling images, in which rough relative orientation is computed by the corresponding points obtained from SIFT matching. Then sub-Harris corners are extracted within original scale images. Finally, the corresponding points are found in the sets of sub-Harris corners consistent with epipolar geometry obtained from rough images matching.

This paper is organized as follows. In Section 2, the rough images matching by SIFT is described. Section 3 and Section 4 introduces sub-Harris operator, descriptor and matching method respectively. Section 5 shows the experimental results in low-altitude images matching. Finally the conclusions are given in Section 6.

2. Rough Images Matching

In order to accelerate low-altitude images matching, original images are down-sampled to small scaling images (as shown in Figure 1), which are implemented by SIFT to extract and match features, so it can be suitable for image scale, rotation and illumination changes.

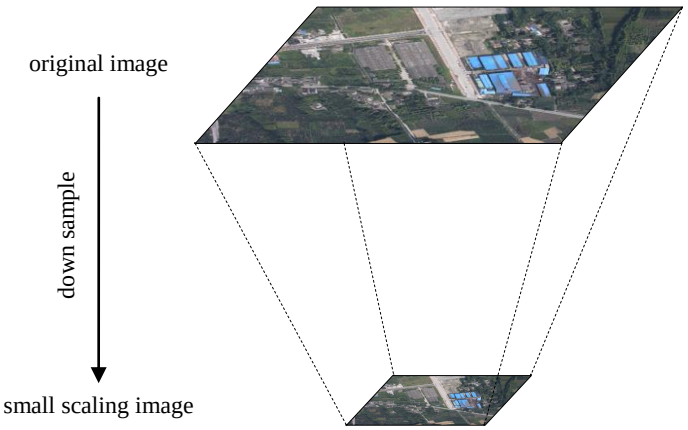


Figure 1. The Down-sampled Image

After SIFT matching, there may be false matched points, which are eliminated by RANSAC scheme combining with epipolar geometry (refer to [11]), in which the coplanarity constraint of two homogeneous image points q and q' is represented with a fundamental matrix F:

$$q'Fq = 0 \tag{1}$$

A valid F is not full rank, it satisfies the following single cubic constraint:

$$\det(F) = 0 \tag{2}$$

Generally, camera is fully-calibrated in low-altitude photogrammetry, so the intrinsic matrix K is constant. Then, the fundamental matrix F is transformed to an essential matrix E, which can be represented as:

$$E = K^{\mathrm{T}}FK \tag{3}$$

A real non-zero 3×3 matrix E is an essential matrix if and only if it satisfies the following equation:

$$2EE^{\mathrm{T}}E - trace(EE^{\mathrm{T}})E = 0 \tag{4}$$

The essential matrix E can be computed by several equations established by the constraints of equations (1), (2) and (4), which require a finite number of the corresponding points in two views. The solutions mainly include five-point method, six-point method, seven-point method and eight-point method. Obviously, the number of the extracted features using SIFT in small scaling image is less than original image, and there may be a corresponding drop of the number of the matched points. So we implement five-point algorithm (given in [12]) to compute relative orientation, as it calls for a minimal number of solutions.

3. Sub-Harris Operator

In the previous section, the original stereopair is down-sampled to small scaling images, in which rough relative orientation is computed by the corresponding points obtained from SIFT matching. In this section, to extract high-precision features for original images matching, sub-Harris operator is presented.

Harris corner detector presented by Harris and Stephens in [13] is one of the most commonly used interest point detectors in photogrammetry field. A Harris corner is determined based on autocorrelation matrix of the image gradients in horizontal and vertical directions. The mathematical formulation of the autocorrelation matrix $M(x, y; I)$ of a gray-level image I at a pixel location (x, y) can be represented as:

$$M(x, y; I) = \begin{bmatrix} I_x^2 & I_{xy} \\ I_{xy} & I_y^2 \end{bmatrix} \tag{5}$$

Where $I_x = \frac{\partial I}{\partial x}$ and $I_y = \frac{\partial I}{\partial y}$ computes the derivatives or gradients in horizontal and vertical directions, respectively. If the two eigenvalues λ_1 and λ_2 of autocorrelation matrix $M(x, y; I)$ are large, then the pixel $I(x, y)$ is considered as a corner. In [13], the corner identification can be measured by the corner response function R, which is given as follows:

$$\begin{aligned} &R = \det(M) - k(\mathrm{tr}(M))^2 \\ &\det(M) = \lambda_1 \lambda_2, \mathrm{tr}(M) = \lambda_1 + \lambda_2 \end{aligned} \tag{6}$$

Where det and tr is the determinant and trace of autocorrelation matrix M, respectively. k is a parameter, which is empirically determined in the interval [0.04, 0.06] (given in [14]). If the corner response function R of a pixel at the location $(\hat{x}, \hat{y})$ is local maximum within a certain size window (non-maximum suppression), which can be represented as:

$$[\hat{x}, \hat{y}] = \arg local \max_{x,y} |R(x, y)| \tag{7}$$

Where if R is greater than the threshold value set manually, then the pixel $I(\hat{x}, \hat{y})$ can be considered as a corner.

The significant advantages of Harris corner operator are simple and efficient, while the location accuracy of Harris corner is pixel-level, which is insufficient in photogrammetric process. So we present an improved Harris operator with sub-pixel level precise.

Harris operator detects corner by the local non-maximum suppression of the corner response function, ignoring the contribution of adjacent similar corners cluster to the location, which may deviate the ideal corner location. By contrast, sub-Harris operator gives consideration to the contribution of adjacent similar corners cluster in a certain radius r (as shown in Figure 2), and least square method is introduced to optimize the location of Harris corner as follows:

$$J = V^{\mathrm{T}} P V \tag{8}$$

or

$$J = \sum_{j=1}^{n} p_j s_j, \left(s_j = (\hat{x} - x_j)^2 + (\hat{y} - y_j)^2, p_j = \frac{R_j}{\sum_{i=1}^{n} R_i} \right) \tag{9}$$

Where $V = \begin{bmatrix} \hat{x} - x_1 \\ \hat{y} - y_1 \\ \hat{x} - x_2 \\ \hat{y} - y_2 \\ \vdots \\ \hat{x} - x_n \\ \hat{y} - y_n \end{bmatrix}$, n is the number of similar corners adjacent to Harris corner in the radius r. s_j and p_j is the squared Euclidean distance and weight of jth similar corner adjacent to Harris corner, respectively. And the weight matrix

$$P = \begin{bmatrix} p_1 & & & & & & \\ & p_1 & & & & & \\ & & p_2 & & & & \\ & & & p_2 & & & \\ & & & & \ddots & & \\ & & & & & p_n & \\ & & & & & & p_n \end{bmatrix}$$

. To achieve the optimal corner location, the derivatives computed in $\hat{x}$ and $\hat{y}$ are based on least square method as follows:

$$\begin{cases} \dfrac{\partial J}{\partial \hat{x}} = 2\sum_{j=1}^{n} p_j \left(\hat{x} - x_j\right) = 0 \\ \dfrac{\partial J}{\partial \hat{y}} = 2\sum_{j=1}^{n} p_j \left(\hat{y} - y_j\right) = 0 \end{cases} \tag{10}$$

then the optimal location $(\hat{x}, \hat{y})$ of sub-Harris corner can be computed as follows:

$$\begin{cases} \hat{x} = \sum_{j=1}^{n} p_j x_j \\ \hat{y} = \sum_{j=1}^{n} p_j y_j \end{cases} \tag{11}$$

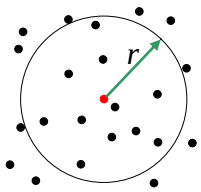


Figure 2. The Adjacent Similar Corners Cluster of Harris Corner in Radius r

Meanwhile, in order to accelerate sub-Harris detection, some non-corner pixels can be excluded before implementing sub-Harris operator. If the gray of a pixel in the center of 3×3 window is completely similar or dissimilar with the pixels in eight neighborhood (as shown in Figure 3 and Figure 4, the pixel in the center of 3×3 window

is denoted by circle), we can assume that the pixel is obvious not a corner according to the definition of Harris corner. The gray similarity in 3×3 window can be measured by the following mathematical formulation:

$$SimCount(i,j)=\sum_{(m,n)} cnt(i+m,j+n), \quad (-1\le m\le 1,-1\le n\le 1, and\ m\ne 0, n\ne 0) \tag{12}$$

$$cnt(i+m,j+n)=\begin{cases}1, & |g_{i,j}-g_{i+m,j+n}|\le t\\ 0, & |g_{i,j}-g_{i+m,j+n}|>t\end{cases} \tag{13}$$

Where $g_{i,j}, g_{i+m,j+n}$ is the gray value at pixel location (i,j) and $(i+m,j+n)$,respectively. t is a threshold of gray difference. If $cnt=1$, then it indicates similar; if $cnt=0$, it indicates dissimilar. When the gray similar number $SimCount$ of a pixel and its eight neighborhood pixels is equal to 0 or 8, we assume that the pixel is completely similar or dissimilar with the pixels in eight neighborhood, so the pixel is not considered as a corner.

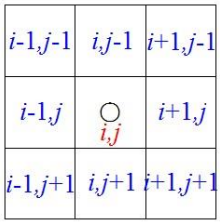


Figure 3. Completely Similar Gray in 3×3 Window

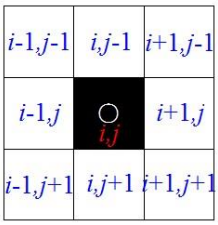
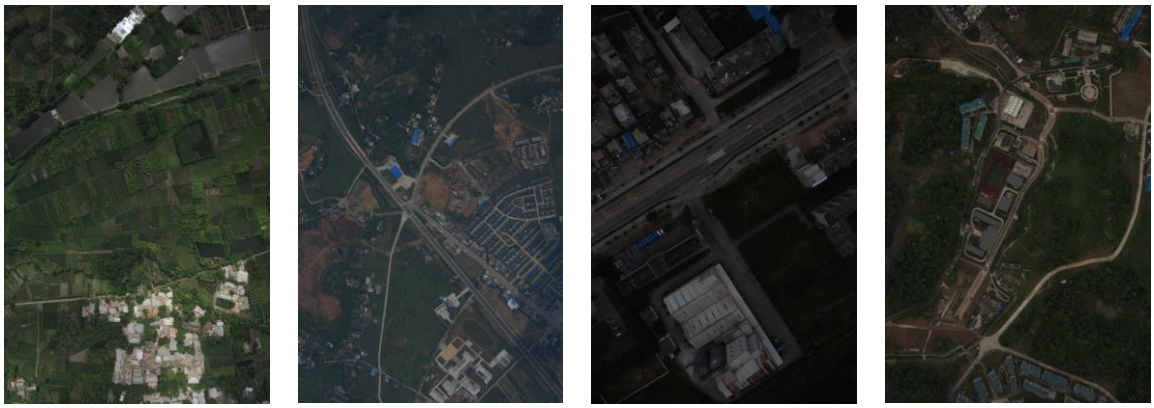


Figure 4. Completely Dissimilar Gray in 3×3 Window

We take four low-altitude images for example to detect sub-Harris corner (as shown in Figure 5), noting that sub-Harris corners are not evenly distributed as a result of inconsistent image texture characteristic. In order to acquire evenly distributed corners, as is shown in Figure 6, images are partitioned to regular grids, in which a certain number of sub-Harris corners are detected. Then it will be conducive to acquire evenly distributed corresponding points.



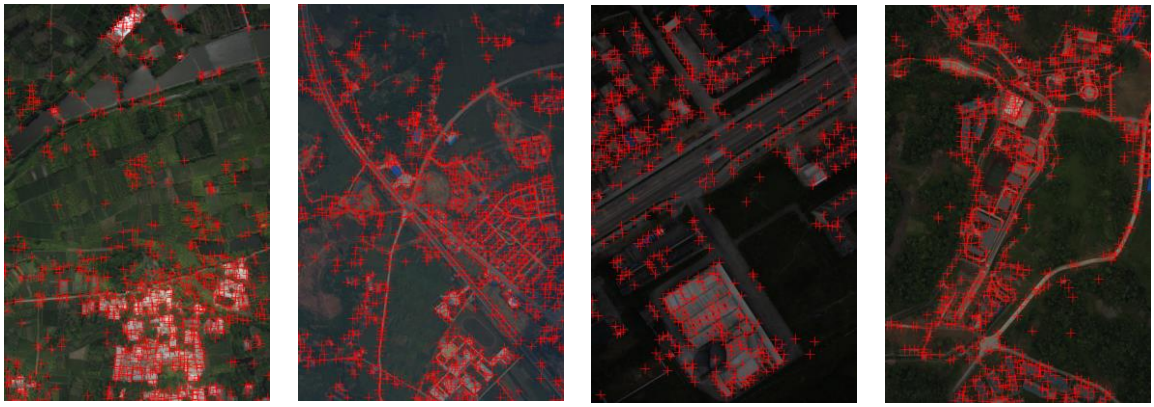


Figure 5. Examples of sub-Harris Corners Detection in Low-Altitude Images

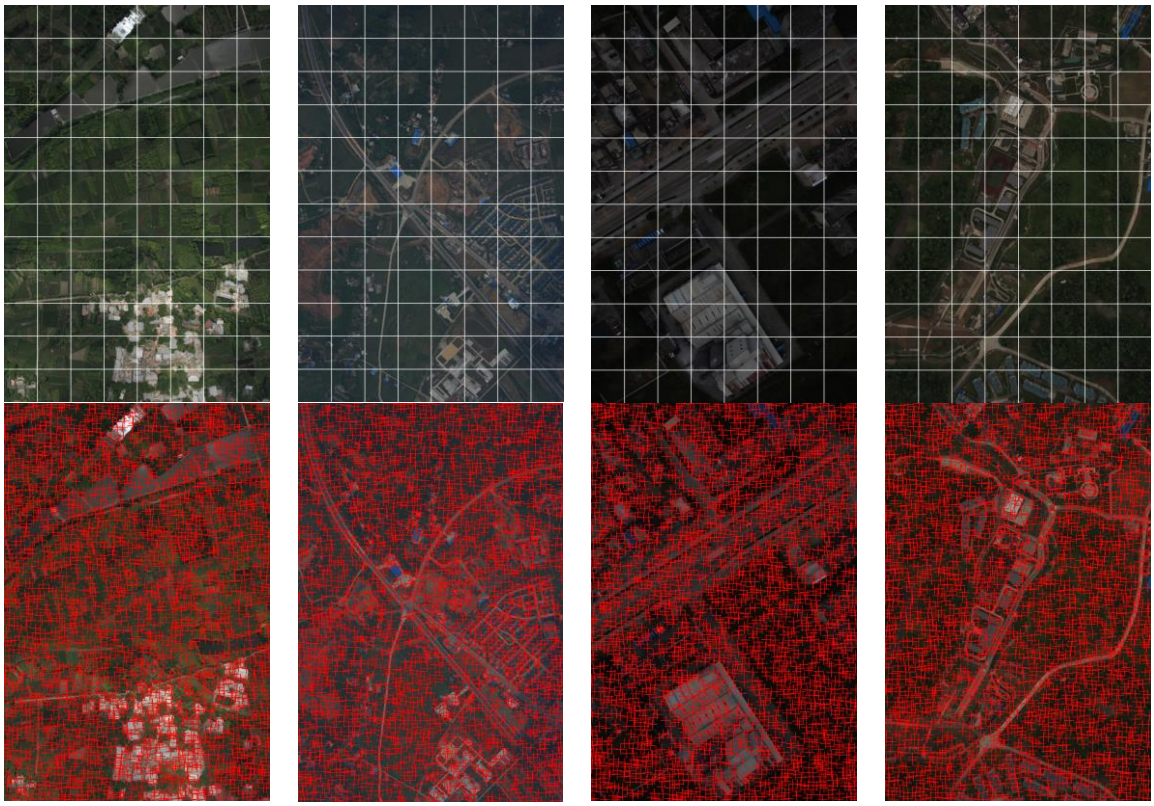


Figure 6. Examples of sub-Harris Corners Detection in Low-Altitude Images Partitioned to Regular Grids

4. Descriptor and Matching

The two sets of corners can be detected by sub-Harris operator in a stereopair, and then we require matching them. In order to achieve reliable descriptor for patch matching, SIFT feature descriptor is adopted to describe sub-Harris corner. At each sub-Harris corner $c_{\hat{x},\hat{y}}$, gradient magnitude $m(\hat{x},\hat{y})$ and orientation $\theta(\hat{x},\hat{y})$ are computed based on the following formula:

$$\begin{cases} m(\hat{x},\hat{y}) = \sqrt{(I(x+1,y)-I(x-1,y))^2 + (I(x,y+1)-I(x,y-1))^2} \\ \theta(\hat{x},\hat{y}) = \arctan(I(x,y+1)-I(x,y-1))/(I(x+1,y)-I(x-1,y)) \end{cases} \quad (14)$$
$$(x = Round(\hat{x}), y = Round(\hat{y}))$$

Similar to SIFT descriptor, a sub-Harris corner can form a 128 dimensional vector. Then the corresponding points can be acquired by matching sub-Harris corners consistent with epipolar geometry, which is computed in rough images matching (section Ⅱ).

5. Experiments

In this section, the experiments focus on sub-Harris operator coupled with SIFT for images matching, employing low-altitude images sized 3744×5616 from Cannon EOS 5D Mark Ⅱ camera. Moreover, the proposed images matching approach is compared to common techniques (including SIFT, Harris operator based image matching (using Harris operator instead of sub-Harris operator of the proposed method)), considering the efficiency and accuracy of relative orientation estimated by the corresponding points. Experiments using the typical stereopairs were implemented to validate the proposed method. The results are shown in Figure 7 (two images of evenly distributed texture), Figure 8 (two images of dissimilar illumination) and Figure 9 (two images of dissimilar scale and rotation). The corresponding points are denoted by red cross in Figure 7, 8 and 9. 1154, 297, 289 corresponding points were matched by SIFT, Harris operator based method, the proposed method, respectively. Other corresponding experimental results are shown in Table 1, Table 2, Table 3 respectively.

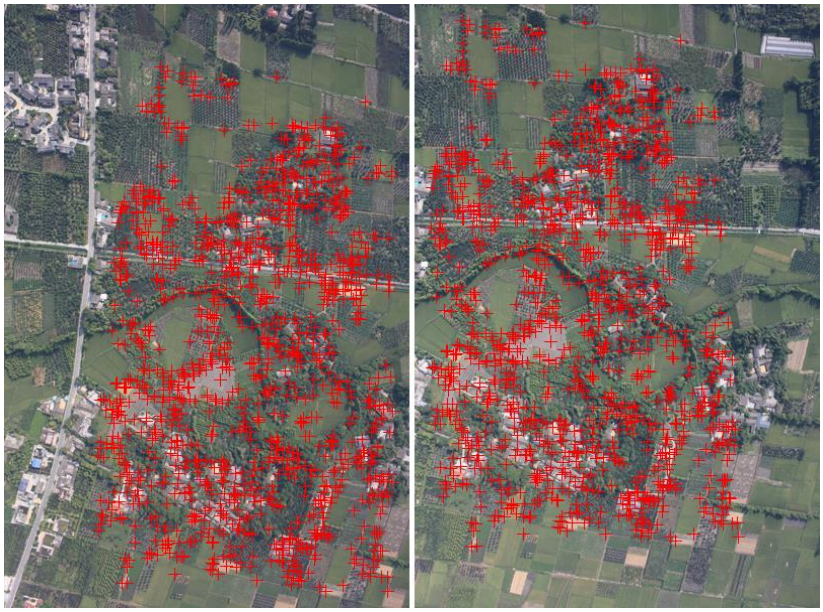


Figure 7. Matched Corresponding Points of a Stereopair with Evenly Distributed Texture

Table 1. Comparison of Low-Altitude Images Matching

	images matching algorithm		
	SIFT	Harris operator based method	the proposed method
the maximum residual vertical parallax in the pixel	0.612	1.077	0.293
the RMS of vertical parallax in the pixel	0.273	0.455	0.107
time [s]	47.3	13.4	14.5

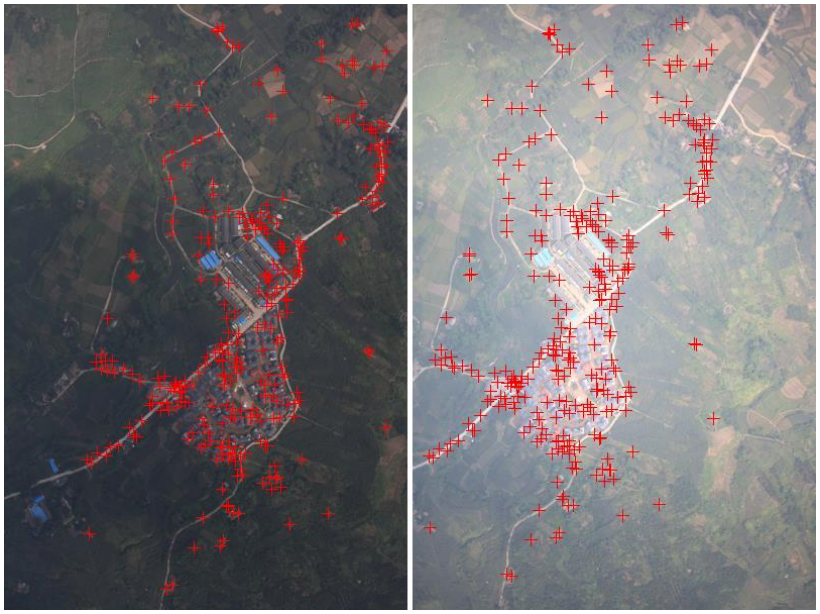


Figure 8. Matched Corresponding Points of a Stereopair with Dissimilar Illumination

Table 2. Comparison of Low-Altitude Images Matching

	images matching algorithm		
	SIFT	Harris operator based method	the proposed method
the maximum residual vertical parallax in the pixel	0.573	0.992	0.204
the RMS of vertical parallax in the pixel	0.261	0.431	0.102
time [s]	38.7	10.6	11.9

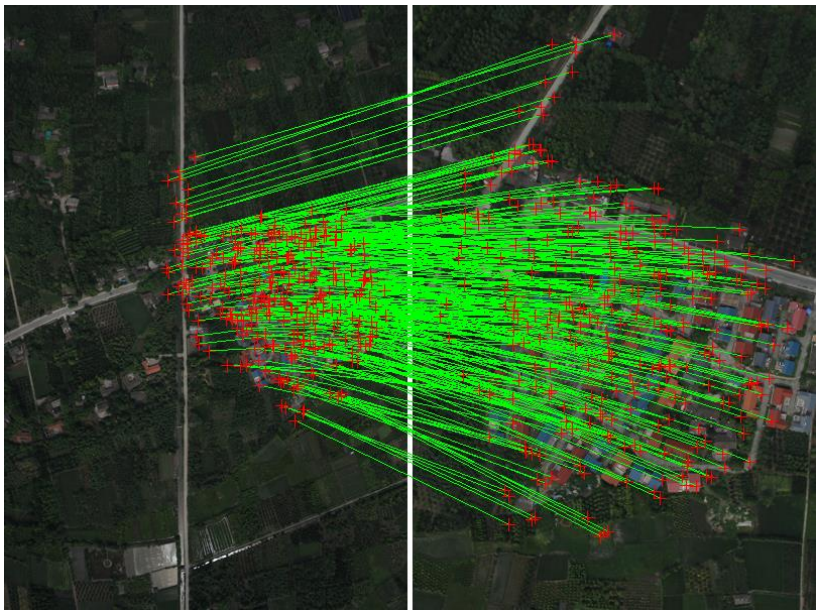


Figure 9. Matched Corresponding Points of a Stereopair with Dissimilar Scale and Rotation

Table 3.Comparison of Low-Altitude Images Matching

	images matching algorithm		
	SIFT	Harris operator based method	the proposed method
the maximum residual vertical parallax in the pixel	0.527	0.944	0.265
the RMS of vertical parallax in the pixel	0.294	0.516	0.113
time [s]	40.1	12.2	13.3

As can be seen in the three tables, the maximum residual vertical parallax of the relative orientation obtained by the proposed method is the smallest one, indicating that the proposed method can also achieve more excellent performances in accuracy. And the processing time of images matching using Harris operator based method or the proposed method is less than one-third of using SIFT. Then, we assume that the proposed method can achieve better performances in low-altitude photogrammetry by comparison of the three images matching method.

6. Conclusion

In this paper, we present a sub-Harris operator coupled with SIFT for fast images matching in low-altitude photogrammetry. The approach can optimize the location accuracy of Harris corner by least square method, and improve the efficiency of extracting and matching features in two ways: 1) some non-corner pixels can be

excluded before implementing sub-Harris operator by pixels similarity analysis in eight neighborhood; 2) the matched corresponding points are found in the sets of sub-Harris corners consistent with epipolar geometry obtained from rough images matching. Experimental results show that low-altitude images matching using the proposed method is invariance to image translation, scaling, rotation, and illumination changes. Furthermore, the proposed method can obtain better precision than SIFT and Harris operator based method in the relative orientation, and significantly improve the computational efficiency compared with SIFT.

Acknowledgements

This research was financially supported by the Open Research Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing ((13)04); the Open Research Fund of Jiangxi Province Key Lab for Digital Land (DLLJ201402); the Doctoral Scientific Research Fund of East China Institute of Technology (DHBK2013204).

References

[1] D. G. Lowe, "Object recognition from local scale-invariant features", IEEE International Conference on Computer Vision, **(1999)**, pp. 1150-1517.
[2] R. Steffen and W. Förstner , "On visual real time mapping for unmanned aerial vehicles", Proceedings of XXI ISPRS Congress, Beijing, China, **(2008)**, pp. 57-62.
[3] A. Lingua, D. Marenchino and F. Nex, "Performance analysis of the SIFT operator for automatic feature extraction and matching in photogrammetric applications", Sensors, vol. 9, no. 5, **(2009)**, pp. 3745-3766.
[4] A. Cesetti, E. Frontoni, A. Mancini, A. Ascani, P. Zingaretti and S. Longhi, "A visual global positioning system for unmanned aerial vehicles used in photogrammetric applications", Journal of Intelligent & Robotic Systems, vol. 61, no. 1, **(2011)**, pp. 157-168.
[5] C. Xing, J. Wang and Y. Xu, "A method for building a mosaic with UAV images", International Journal of Information Engineering and Electronic Business, vol. 2, no. 1, **(2010)**, pp. 9-15.
[6] F. Ackermann, "High precision digital image correlation", Proceedings 39th Photogrammetric Week, **(1983)**.
[7] F. Ackermann, "Digital image correlation: performance and potential application in photogrammetry", The Photogrammetric Record, vol. 11, no. 64, **(1984)**, pp. 429-439.
[8] H. Bay, T. Tuytelaars and L. V. Gool, "SURF: Speeded Up Robust Features", European Conference on Computer Vision, vol. 3951, **(2006)**, pp. 404-417.
[9] L. Juan and O. Gwun, "A comparison of SIFT, PCA-SIFT and SURF", International Journal of Image Processing, vol. 3, no. 4, **(2009)**, pp. 143-152.
[10] P. Azad, T. Asfour and R. Dillmann, "Combining Harris interest points and the SIFT descriptor for fast scale-invariant object recognition", International Conference on Intelligent Robots and Systems, **(2009)**, pp. 4275-4280.
[11] R. Hartley and A. Zisserman, "Multiple view geometry in computer vision", Second Edition, Cambridge University Press, Cambridge, UK, **(2004)**.
[12] H. Li and R. Hartley, "Five-point motion estimation made easy", International Conference on Pattern Recognition, Hong Kong, **(2006)**, pp. 630-633.
[13] C. Harris and M. Stephens, "A combined corner and edge detector", Proceedings of the 4th Alvey Vision Conference, **(1988)**, pp. 147-151.
[14] U. Orguner and F. Gustafsson, "Statistical characteristics of Harris corner detector", Proceedings of the IEEE/SP 14th Workshop on Statistical Signal Processing, **(2007)**, pp. 571-575.

Author

Haiqing He, he received his PhD in Geodesy and Surveying Engineering (2013) from Wuhan University. Now he is a lecturer at Faculty of Geomatics, East China Institute of Technology. His current research interests include photogrammetry and remote sensing, surveying data processing. E-mail: hyhqing@163.com