

## ROBUST LOW-ALTITUDE IMAGE MATCHING BASED ON LOCAL REGION CONSTRAINT AND FEATURE SIMILARITY CONFIDENCE

Min Chen<sup>a, b, c, \*</sup>, Qing Zhu<sup>a, b, c</sup>, Shengzhi Huang<sup>a, b, c</sup>, Han Hu<sup>a</sup>, Jingxue Wang<sup>a</sup>

<sup>a</sup> Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, 611756, China –  
minchen@home.swjtu.edu.cn; zhuq66@263.net; shzhhuang3@163.com; huhan@whu.edu.cn; xiaoxue1861@163.com

<sup>b</sup> State-province Joint Engineering Laboratory of Spatial Information Technology for High-speed Railway Safety, Chengdu, 610031,  
China

<sup>c</sup> Collaborative Innovation Center for Rail Transport Safety, Chengdu, 610031, China

### Commission III, WG III/1

**KEY WORDS:** Low-altitude remote sensing, Image matching, Local region, Feature similarity confidence

#### ABSTRACT:

Improving the matching reliability of low-altitude images is one of the most challenging issues in recent years, particularly for images with large viewpoint variation. In this study, an approach for low-altitude remote sensing image matching that is robust to the geometric transformation caused by viewpoint change is proposed. First, multiresolution local regions are extracted from the images and each local region is normalized to a circular area based on a transformation. Second, interest points are detected and clustered into local regions. The feature area of each interest point is determined under the constraint of the local region which the point belongs to. Then, a descriptor is computed for each interest point by using the classical scale invariant feature transform (SIFT). Finally, a feature matching strategy is proposed on the basis of feature similarity confidence to obtain reliable matches. Experimental results show that the proposed method provides significant improvements in the number of correct matches compared with other traditional methods.

#### 1. INTRODUCTION

In recent years, comprehensive low-altitude remote sensing platforms, such as unmanned aerial vehicle (UAV), have provided new possibilities for high-resolution image acquisition and have been extensively used in many applications (Bulatov et al., 2011; Choi et al., 2011; Colomina et al., 2014; Wallace et al., 2014a, 2014b; Goncalves et al., 2015; Zhou et al., 2015). Compared with traditional approaches (e.g., satellite remote sensing and aerial photogrammetry), low-altitude remote sensing platforms have the following inherent advantages: first, the work mode is flexible, efficient, and less affected by weather, and they can take off any time as tasked; second, they are able to obtain large-scale and high-precision remote sensing images; and third, the overlapping degree between images is relatively large. It can enhance the reliability of the subsequent processing. The cost of platform construction, maintenance, and operation is also low. Although low-altitude remote sensing platforms have many advantages in image acquisition, the image processing technology is still unable to meet the needs of many applications. Therefore, investigating automatic low-altitude remote sensing image processing technology, in which robust image matching is a fundamental issue, is necessary.

Image matching is a key component of many tasks in photogrammetry and remote sensing, which is extensively used in many applications, such as image registration (Song et al., 2010; Gruen et al., 2012; Cheng et al., 2013; Chen et al., 2014; Li et al., 2014), stitching (Brown et al., 2007), and 3D reconstruction (Kratochvil et al., 2010). Image matching methods can be mainly classified into two categories, namely,

intensity-based methods and feature-based methods (Xiong et al., 2010). In intensity-based methods, many approaches are based on cross-correlation (Zhao et al., 2006; Karna et al., 2008), which is a simple concept and easy to implement. The main problem with cross-correlation is that it is difficult to deal with image deformations except for shift transform. Therefore, cross-correlation-based methods are usually used to match epipolar images. Many matching frameworks, such as the least squares (Gruen et al., 2005) and relaxation matching techniques (Zhang et al., 2007), have been proposed to improve the reliability of remote sensing image matching. In addition to cross-correlation, researchers have proposed a class of frequency domain matching methods based on Fourier spectrum (Zitova et al., 2003). This type of method searches the best match by using image frequency domain information. Compared with the cross-correlation-based methods, the frequency domain matching methods based on Fourier spectrum can obtain better matching results under image illumination change and noise interference. Other well-known intensity-based matching methods adopt mutual information to find matches (Viola et al., 1997; Thevenaz et al., 1998). The basic idea of mutual-information-based matching methods is to keep moving the target window in the search area. When the mutual information between the target window and the search window achieves a maximum value, the target window center and the search window center are regarded as one pair of match. Therefore, the matching problem can be converted into the computation of the mutual information maximum value. In image matching, the number of feature-based methods far exceeds that of intensity-based methods. In feature-based methods, those based on local invariant feature descriptor are

\* Corresponding author

the most well-known. In general, the framework of such methods includes three steps, namely, feature detection, description, and matching. In the field of computer vision and pattern recognition, researchers have proposed many well-known local invariant feature detectors (Harris et al., 1988; Smith et al., 1997; Lowe, 1999; Matas et al., 2004; Mikolajczyk et al., 2004) and descriptors (Belongie et al., 2002; Mikolajczyk et al., 2005; Bay et al., 2008; Calonder et al., 2010; Rublee et al., 2011). One of the most well-known methods is the scale invariant feature transform (SIFT) (Lowe, 2004). The SIFT method has been extensively used in many applications because of its robustness for image rotation, scale change, and a certain degree of viewpoint and illumination change. In addition to investigating invariant descriptors, researchers have enhanced image matching performance by improving the matching strategy (Morel et al., 2009; Yu et al., 2012). Given that feature-based methods have made remarkable achievements in computer vision and pattern recognition, they are used to match remote sensing images (Li et al., 2009; Lingua et al., 2009; Huang et al., 2010; Sedaghat et al., 2011; Wang et al., 2012).

The characteristics of low-altitude platforms, i.e., bumpy and low height, result in failure of the commonly used matching methods to match low-altitude remote sensing images. Although many methods can solve image rotation and scale change well, they cannot obtain satisfactory results when large viewpoint change exists, as shown in Figure 1.

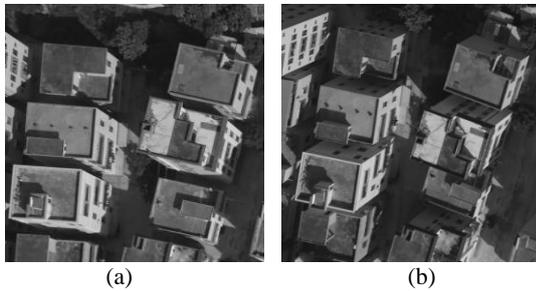


Figure 1. Low-altitude images with viewpoint change: (a) reference image. (b) test image.

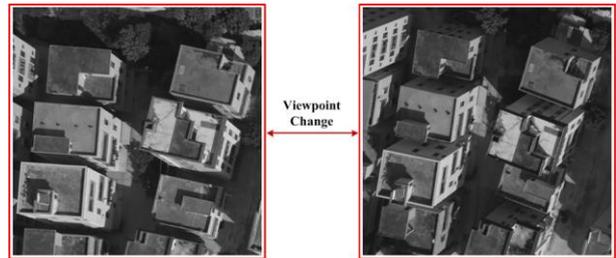
A robust matching method is proposed in this study to improve the matching performance of low-altitude remote sensing images, particularly for those images with large viewpoint change. First, multiresolution local regions are extracted from the whole image, and a geometric transform is implemented on the extracted regions. Based on this transformation, the viewpoint change between images is converted into rotation and scale change. Second, point features are detected inside and outside of the local regions. Then, the viewpoint invariant feature area is computed for each interest point based on the constraint of local regions. In this procedure, the size of the feature area is determined according to image resolution instead of feature scale value. Then, the SIFT method is adopted to calculate the descriptor. Finally, feature similarity confidence is defined and a matching strategy based on it is presented to find feature correspondences.

The remainder of this paper is organized as follows: Section 2 presents the overall methodological consideration of how the proposed algorithm will improve the matching performance. Section 3 presents the proposed low-altitude remote sensing image matching algorithm in detail. The experimental results, along with the detailed analysis and discussion, are presented in Section 4. The final section concludes the paper by discussing

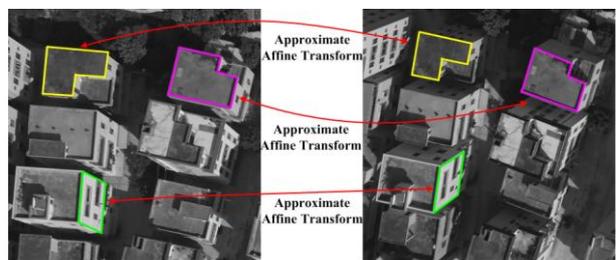
the advantages and disadvantages of the proposed method and further improvements that can be made.

## 2. OVERALL METHODOLOGICAL CONSIDERATION

Projection transformation between different pixel pairs is different because of the large object depth variation in low-altitude remote sensing images. All the pixel pairs within the whole images do not meet the same affine transformation. We divided the images into local regions to overcome this problem. The transformation between the pixels in the corresponding regions can be approximated to an affine transform because the depth variation in the local region is small (see Figure 2).



(a) transformation between the whole images



(b) transformation between local regions

Figure 2. Image geometric transformation.

Based on the previously presented analysis, the proposed matching method is conducted through the following three steps: First, the local regions are extracted from the input images and are normalized based on a transformation. Second, the interest points are detected and described based on the local region constraint. Third, descriptor similarity confidence is defined and a matching strategy is proposed based on the concept. The procedure is summarized in Figure 3.

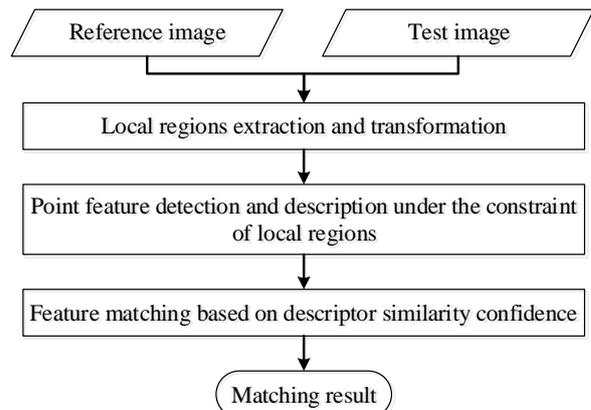


Figure 3. Flowchart of the proposed method.

### 3. THE PROPOSED IMAGE MATCHING METHOD

#### 3.1 Local Regions Extraction and Transformation

Region detectors are adopted to obtain the local regions. In this study, given that local regions are used to constrain the subsequent different viewpoint image matching, the selected region detector should be robust to image viewpoint change. In the study conducted by Mikolajczyk et al. (2005b), they demonstrated that the maximally stable extremal region (MSER) detector (Matas et al., 2004) is more robust than other region detectors for image viewpoint change. Therefore, the MSER detector is adopted in this study to extract the local regions. The initial irregular regions extracted by using the MSER detector have been fitted to elliptical regions based on the region second-order moment to facilitate the subsequent processing.

The original MSER detector is not scale invariant. In this study, we improved the multiresolution strategy (Forssen et al., 2007) and integrated it with MSER to obtain multiresolution local regions. To do this, a Gaussian scale pyramid is constructed by image blurring and subsampling with a series of Gaussian kernels. Then, local regions are detected separately at each resolution image by using the original MSER detector. Finally, duplicate regions from different scale images are removed by eliminating fine-scale MSERs. The following three criteria are used to distinguish duplicate MSERs:

- 1) Distance between the centroids of the two MSER elliptical regions should be smaller than 4 pixels.
- 2) The value of  $abs(S_1 - S_2) / \max(S_1, S_2)$  should be less than 0.2, where  $S_1$  and  $S_2$  are the sizes of the two regions.
- 3) The value of  $abs(\alpha_1 - \alpha_2) \cdot \max(L_1, L_2) / \pi$  should be less than 4, where  $\alpha_1$  and  $\alpha_2$  are the directions of the major axis of the two elliptical regions and  $L_1$  and  $L_2$  are the perimeters of the two ellipses.

The transformation between corresponding local regions is approximated to affine transform. In this study, a transform is proposed to map the affine change into rotation and scale change to improve the feature repeatability rate.

We assume that  $EAp$  is the local elliptical region,  $l$  is the length of the major axis,  $w$  is the length of the minor axis,  $\alpha$  is the orientation of the major axis, and  $CAp$  is the circular region transformed from  $EAp$ . The radius of  $CAp$  is  $r = \sqrt{l \cdot w}$ . If matrix  $\mathbf{T}$  denotes the transformation between  $EAp$  and  $CAp$ , then  $\mathbf{T}$  satisfies the following equation:

$$\left[ \mathbf{T}(\mathbf{X} - \mathbf{X}_c) \right]^T \left[ \mathbf{T}(\mathbf{X} - \mathbf{X}_c) \right] = r^2, \quad (1)$$

where  $\mathbf{X}$  is a point on the ellipse and  $\mathbf{X}_c$  is the region center. Since  $\mathbf{X}$  is located on the ellipse, we obtain the following equation:

$$(\mathbf{X} - \mathbf{X}_c)^T \boldsymbol{\mu}^{-1} (\mathbf{X} - \mathbf{X}_c) = 1, \quad (2)$$

where  $\boldsymbol{\mu} = [\mu_{20}, \mu_{11}; \mu_{11}, \mu_{02}]$  is the second-order moment of  $EAp$ . Matrix  $\mathbf{T}$  can be calculated from Equations (1) and (2) as follows:

$$\mathbf{T} = \frac{r}{\left[ \mu_{20} (\mu_{20} \cdot \mu_{02} - \mu_{11}^2) \right]^{1/2}} \begin{bmatrix} \left( \mu_{20} \cdot \mu_{02} - \mu_{11}^2 \right)^{1/2} & 0 \\ -\mu_{11} & \mu_{20} \end{bmatrix} \quad (3)$$

Then, the local elliptical region can be transformed to a circular region based on the following equation:

$$CA_p = \mathbf{T} \cdot EA_p. \quad (4)$$

The affine transformation between the elliptical regions has been converted into the scale and rotation transformation after the aforementioned processes:

If the transformation between the corresponding transformed regions is  $\mathbf{B}$ , then we obtain the following equation (Hartley et al., 2000):

$$\boldsymbol{\mu}'_L = \mathbf{B}^T \boldsymbol{\mu}'_R \mathbf{B}, \quad (5)$$

where  $\boldsymbol{\mu}'_L$  and  $\boldsymbol{\mu}'_R$  denote the second-order moment of the corresponding transformed regions. For the two transformed regions, we derive the following expression:

$$\begin{cases} \boldsymbol{\mu}'_L = \lambda_L \mathbf{E} \\ \boldsymbol{\mu}'_R = \lambda_R \mathbf{E} \end{cases} \quad (6)$$

where  $\mathbf{E}$  denotes an identity matrix. From Equations (5) and (6), we can derive the following equation:

$$(\lambda_L / \lambda_R) \mathbf{E} = \mathbf{B}^T \mathbf{B}. \quad (7)$$

where  $\lambda_L$  and  $\lambda_R$  are scale factors.  $\mathbf{B}$  is a matrix including scale and rotation factors. Therefore, only rotation and scale change exist between the corresponding transformed regions. The transformation between the transformed regions can be expressed as follows:

$$\mathbf{B} = \mathbf{H}_{\lambda'} \mathbf{R}_1(\psi'). \quad (8)$$

where  $\mathbf{H}_{\lambda'}$  denotes a scale transformation.  $\mathbf{R}_1(\psi')$  denotes a rotation transformation.

#### 3.2 Point Feature Detection and Description

Considering the rotation and scale change between the transformed local regions, the difference of Gaussians (DoG) detector is adopted to extract the point features from the local regions. In practical applications, the resolution of the inputted low-altitude remote sensing image is known. Therefore, for each extracted point feature, the feature area size is determined based on the image resolution instead of the feature scale value.

We assume that the resolution of the two inputted images is  $\lambda_1$  and  $\lambda_2$  ( $\lambda_1 \geq \lambda_2$ , unit: meter). The feature area radius of all the features in the image with resolution  $\lambda_1$  is set as follows:

$$r_1 = \begin{cases} 16, & \lambda_1 \geq 2 \\ 24, & 1 \leq \lambda_1 < 2 \\ 32, & 0.2 \leq \lambda_1 < 1 \\ 64, & \lambda_1 \leq 0.2 \end{cases} \quad (9)$$

The feature area radius of all features in another image with resolution  $\lambda_2$  is set as follows:

$$r_2 = r_1 \cdot \lambda_1 / \lambda_2 \quad (10)$$

This strategy can overcome the unreliability of feature scale computation. After point feature detection, the well-known SIFT method is adopted to compute the feature descriptor.

Given that the extracted local regions cannot cover the whole image, the point features from the outside area of local regions should be detected and described. In this study, the following approach is proposed:

1) The DoG detector is used to find interest points and obtain feature coordinates from the outside area of the local regions.

2) The detected interest points are clustered into local regions. The distances between each interest point to all local region centers are computed. If the ratio between the smallest distance  $d_1$  and the second smallest distance  $d_2$  is less than a threshold  $t$ , then the interest point will be clustered into the local region corresponding to the smallest distance  $d_1$ . Otherwise, the interest point is regarded as two features and they are clustered into the nearest two regions, as shown in Figure 4.

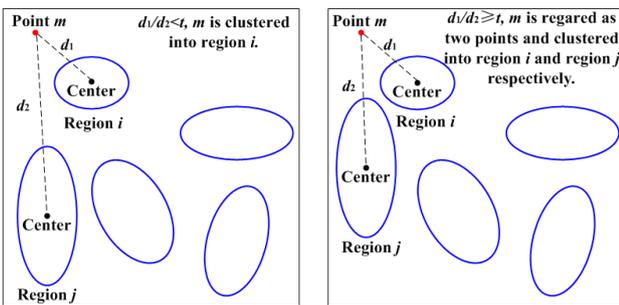


Figure 4. Interest points clustering rule.

3) The feature area of each interest point is computed according to the elliptical parameters of the region in which the interest point is clustered. Then, the SIFT method is adopted to describe each feature. The flowchart of the feature area computation is summarized in Figure 5.

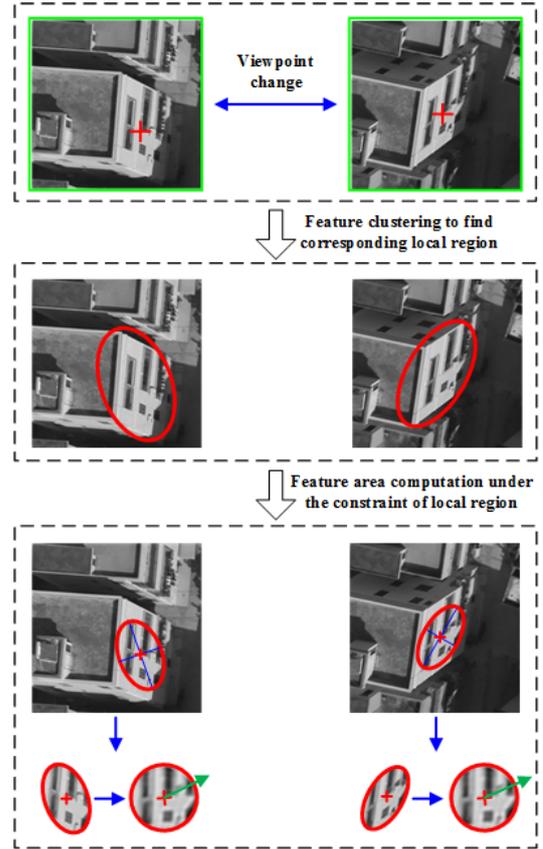


Figure 5. The flowchart of feature area computation.

### 3.3 Feature Matching Based on Similarity Confidence

In this paper, the similarity confidence is defined as follows:

For a feature  $p$  with descriptor  $D_p$  on the reference image, we compute the Euclidean distance between  $D_p$  and the descriptors of all the features ( $q_1, \dots, q_n$ ) on the test image. We find the closest feature  $q_i$  and the second closest feature  $q_j$ , and their descriptors are  $D_{q_i}$  and  $D_{q_j}$ , respectively. Then, the similarity confidence  $C$  between feature  $p$  and  $q_i$  is expressed as follows:

$$C(p, q_i) = 1 - \text{dist}(d_p, d_{q_i}) / \text{dist}(d_p, d_{q_j}), \quad (11)$$

where  $\text{dist}(\ )$  is a function used to compute the Euclidean distance between two descriptors.

From the previously presented definition, the similarity confidence is only computed between each feature on the reference image and its closest feature on the test image.

Based on the similarity confidence, a robust feature matching method is proposed in this study, which is implemented as follows:

Step 1: A set used to save final matches is marked as *SetFinal*, which is initially an empty set.

Step 2: The Euclidean distance between all the features on the reference image and all the features on the test image are computed.

Step 3: For each feature on the reference image, its corresponding closest feature on the test image is determined. These feature pairs are regarded as candidate matches and are saved into another set  $SetAll$ . Meanwhile, the similarity confidence of each feature pair is computed.

Step 4: A threshold  $T_c$  is set. Those feature pairs with similarity confidence conforming to  $C > T_c$  are saved into a set  $SetC$ .

Step 5: An affine transformation  $H$  is estimated in  $SetC$  by using the random sample consensus (RANSAC) algorithm. In  $SetC$ , those feature pairs conforming to  $H$  are saved into sets  $SetCfitH$  and  $SetFinal$  at the same time. Then, we reset  $SetC$ , such that  $SetC = SetC - SetCfitH$ .

Step 6: The affine transformation  $H$  is adopted to verify the feature pairs in  $SetAll$ . Those pairs conforming to  $H$  are saved into sets  $SetAllfitH$  and  $SetFinal$  at the same time. Then, we reset  $SetAll$ , such that  $SetAll = SetAll - SetAllfitH$ .

Step 7: An iteration algorithm is implemented for Steps 5 and 6 until the following expression is true:

$$setAll = \emptyset \text{ or } Num(setC) < 4 \text{ or } Num(iteration) > 5, \quad (12)$$

where  $Num(setC)$  denotes the number of feature pairs in  $SetC$  and  $Num(iteration)$  denotes the number of iterations.

Step 8: All matches in  $SetFinal$  are outputted as initial matching result.

In the initial matching result, there are several outliers inevitably. In traditional methods, the RANSAC algorithm is used to estimate the affine transform between images. Those matches that do not conform to this transform will be eliminated as outliers. The algorithm performs well for middle- or low-resolution remote sensing images because the transformation between all pixels of the whole inputted images can be approximated to an affine transform. However, for low-altitude remote sensing images, several correct matches are eliminated because the transformation between the whole inputted images cannot be approximated to an affine transform. An epipolar constraint based on the fundamental matrix is used to eliminate outliers with the RANSAC algorithm to overcome this problem.

## 4. EXPERIMENT

### 4.1 Experimental Data

In the experiments, the three pairs of low-altitude remote sensing images shown in Figures 6 to 8 were used to evaluate the performance of the proposed method. We compared the proposed method with traditional methods by using the DoG (Lowe, 1999), HarAff (Harris-Affine) (Mikolajczyk, 2004),

HesAff (Hessian-Affine) (Mikolajczyk, 2004) and MSER (Matas, et al., 2004) detectors combined with the most well-known SIFT descriptor (Lowe, 2004). The proposed method was also compared with the state-of-the-art Affine-SIFT (ASIFT) (Morel et al., 2009) and iterative SIFT (ISIFT) (Yu et al., 2012) methods.



Figure 6. Dataset1: low-altitude UAV images. Image size is 1024×1024 pixels. Image resolution is 0.05m.



Figure 7. Dataset 2: low-altitude oblique images. Image size is 1024×1024 pixels. Image resolution is 0.08m.



Figure 8. Dataset 3: low-altitude oblique images. Image size is 1024×1024 pixels. Image resolution is 0.08m.

### 4.2 Matching Results

In the experiments, parameters of all the aforementioned methods in our experiments were set according to their recommended values in the original references. Table 1 summarizes the performance of each method in terms of number of correct matches. In Table 1, the four methods that combined the HarAff, HesAff, MSER, and DoG detectors and SIFT descriptor obtained fewer correct matches than those methods that improved the matching framework, namely, ASIFT, ISIFT, and the proposed method. More specifically, in all the three group experiments, the proposed matching method obtained more correct matches than all the other methods. Although the proposed method is based on MSERs, it detects and matches features from the whole images rather than MSER matches. Therefore, the proposed method performs well when the correct matches of the MSER method are few.

Matching methods	Experimental data		
	Dataset 1	Dataset 2	Dataset 3
HarAff	71	19	158
HesAff	34	13	163
MSER	19	7	38
SIFT	48	4	165
ASIFT	274	45	532
ISIFT	144	4	222
<b>Proposed</b>	<b>750</b>	<b>352</b>	<b>865</b>

Table 1. Number of correct matches of different methods on the three datasets.

The performance of the ISIFT method is better than that of the SIFT method because the ISIFT method is an iterative method. The ISIFT method estimates the transformation between images and builds a simulation of one of the inputted images. Then, feature detection and matching are conducted between another inputted image and the simulated image based on the original SIFT method. However, in low-altitude remote sensing image matching, the transformation between the whole inputted images cannot be approximated to an affine transform. The simulation of one image generated in the iteration of the ISIFT method can only correspond to a partial area of another image. Therefore, its performance improvement relative to the SIFT method is limited.

The ASIFT method performs generally well (only ranking behind the proposed method). The method improves the view invariance of the SIFT method by simulating and building the affine space of inputted images. For image viewpoint variation, several images in the affine space of one inputted image have a pose similar to the images in the affine space of another inputted image. Good matching results can be obtained between these simulated images. Although the ASIFT method can improve the matching performance under view change, its application is limited because the transformation between the inputted images is not considered and an exhaustive strategy is needed for the feature search. Thus, the time efficiency of the ASIFT method is an issue of concern. The simulated affine space in the ASIFT method is discontinuous, and the matching result can be improved by decreasing the sample interval. For low-altitude remote sensing images, several regions cannot be covered in the affine space because of the topographic relief, which is why the number of correct matches generated by the ASIFT method is less than that of the proposed method in the experimental results.

Similar to the ASIFT and ISIFT methods, the proposed method also improves the matching performance by simulating images. The differences are as follows:

(1) The ASIFT method simulates images without considering the transformation between images. In our proposed method, each local image region corresponds to a simulated circular image area. It can avoid detecting and matching features in a mass of useless images (images without correspondence). Besides, the proposed method simulates each region with different transform. These regions can cover the input image better and facilitate more matches.

(2) The ISIFT method simulates the whole image with one transformation. In our proposed method, the simulation process is implemented to local regions. The transformation between the

local regions can be approximated to an affine transform whether the transformation between the whole images conform to an affine transform or not. Thus, the proposed method can obtain satisfactory matching results when the ISIFT method fails.

Some matching results in our experiments are shown in Figures 9 to 11. In the results, matches are linked with white lines.

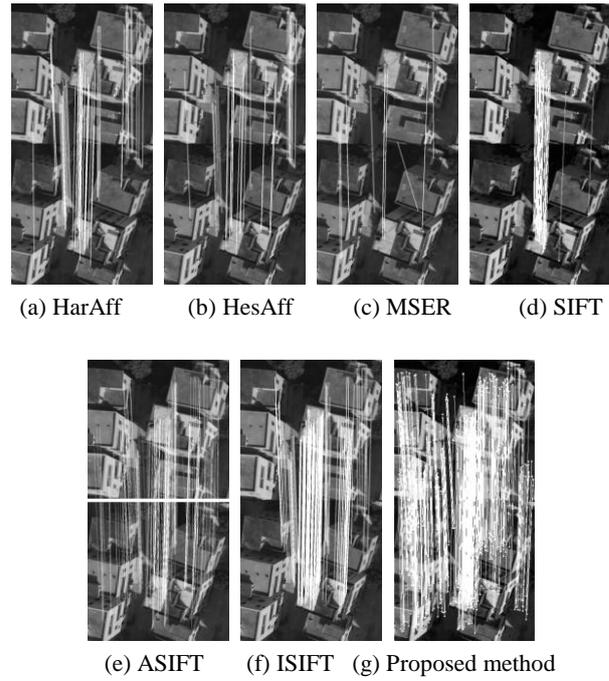


Figure 9. Matching results based on Dataset 1.

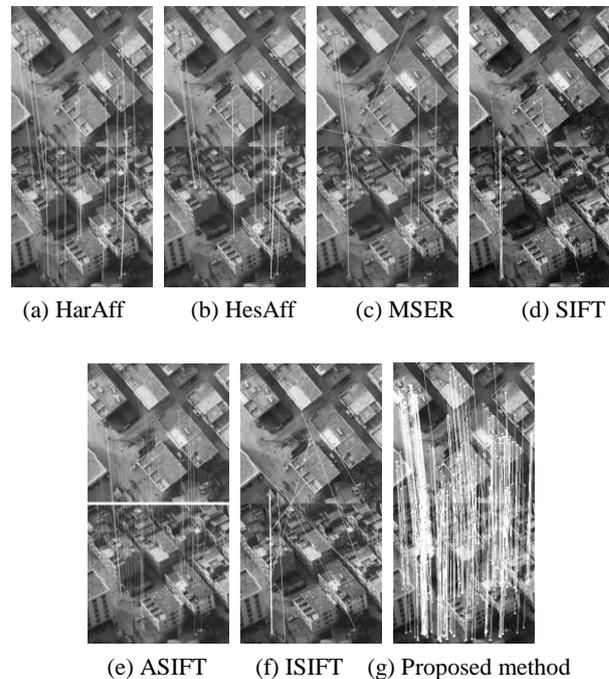


Figure 10. Matching results based on Dataset 2.

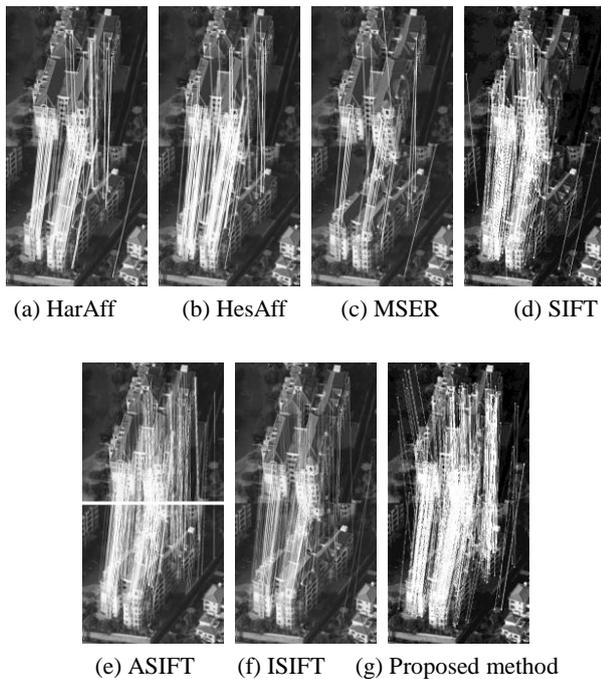


Figure 11. Matching results based on Dataset 3.

## 5. CONCLUSION

In this study, we proposed a novel point feature matching method for low-altitude remote sensing images based on the analysis of image geometric transformation. The contribution of this study lies in the following aspects:

(1) A new matching framework for low-altitude remote sensing images is proposed based on the local region constraint and feature similarity confidence. The proposed method can effectively match images with large viewpoint difference.

(2) Compared with the state-of-the-art matching method ASIFT, only one geometric transformation is implemented in our method for each local region, ensuring satisfactory results and time efficiency.

The experimental results showed that the proposed method performs better than the other methods for low-altitude remote sensing images with viewpoint variation. However, the local region extraction in the proposed method highly depends on the image content, which indicates that the proposed matching method works better for structured images. A possible future work is to improve the local region extraction method and make the proposed matching method perform well in structured and textured areas.

## ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (No. 41501492 and 41471320), the Science and Technology Program of Sichuan, China (No. 2015SZ0046), the Open Research Fund by Sichuan Engineering Research Center for Emergency Mapping & Disaster Reduction (No. K2015B006), the Open Research Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (No. (14) Key 03), and the Opening Research Fund of National Engineering

Laboratory for Surface Transportation Weather Impacts Prevention (No. NELXX201601).

## REFERENCES

- Bay H., Ess A., Tuytelaars T. et al., 2008. Speeded-up robust features (SURF). *Comput. Vis. Image Und.*, 110(3), pp. 346-359.
- Belongie S., Malik J., and Puzicha J., 2002. Shape Matching and Object Recognition Using Shape Contexts. *IEEE T. Pattern Anal.*, 24(4), pp. 509-522.
- Brown M., and Lowe D. G., 2007. Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vis.*, 74(1), pp. 59-73.
- Bulatov D., Solbrig P., Gross H., et al., 2011. Context-based urban terrain reconstruction from UAV-videos for geoinformation application. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-1/C22, pp. 75-80.
- Calonder M., Lepetit V., Strecha C., et al., 2010. BRIEF: Binary Robust Independent Elementary Features. In *Proceedings of European Conference on Computer Vision*. 05-11, September, Heraklion, Greece, pp. 778-792.
- Chen Q., Wang S., Wang B., et al., 2014. Automatic Registration Method for Fusion of ZY-1-02C Satellite Images. *Remote Sens.*, 6(1), pp. 157-179.
- Cheng L., Tong L., Li M., et al., 2013. Semi-Automatic Registration of Airborne and Terrestrial Laser Scanning Data Using Building Corner Matching with Boundaries as Reliability Check. *Remote Sens.*, 5(12), pp. 6260-6283.
- Choi F., and Lee I., 2011. A UAV-based close-range rapid aerial monitoring system for emergency responses. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-1/C22, pp. 247-252.
- Colomina I., and Molina P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm.*, 92, pp. 79-97.
- Goncalves J. A., and Henriques R., 2015. UAV photogrammetry for topographic monitoring of coastal areas. *ISPRS J. Photogramm.*, 104, pp. 101-111.
- Forssen P. E., and Lowe D. G., 2007. Shape descriptors for maximally stable extremal regions. In *Proceedings of IEEE International Conference on Computer Vision*, pp. 1-8.
- Gruen A., and Akca D., 2005. Least squares 3D surface and curve matching. *ISPRS J. Photogramm.*, 59(3), pp. 151-174.
- Gruen A., 2012. Development and Status of Image Matching in Photogrammetry. *Photogramm. Rec.*, 27(137), pp. 36-57.
- Harris C., and Stephens M., 1988. A combined corner and edge detector. In *Proceedings of the 4th Alvey Vision Conference*. pp. 147-152.
- Hartley R., and Zisserman A., 2000. Multiple View Geometry in Computer Vision, Cambridge university press.

- Huang L., Li Z., 2010. Feature-based image registration using the shape context. *Int. J. Remote Sens.*, 31(8), pp. 2169-2177.
- Karna D. K., Agarwal S., and Nikam S., 2008. Normalized Cross-Correlation Based Fingerprint Matching. In *Proceedings of the Fifth International Conference on Computer Graphics, Imaging and Visualisation*. Penang, Malaysia, 26-28 August, pp. 229-232.
- Kratochvil B. E., Dong L. X., Zhang L., et al., 2010. Image-based 3D reconstruction using helical nanobelts for localized rotations. *J. Microsc.*, 237(2), pp. 122-135.
- Li A., Cheng X., Guan H., et al., 2014. Novel Image Registration Method Based on Local Structure Constraints. *IEEE Geosci. Remote S.* 11(9), pp. 1584-1588.
- Li Q., Wang G., Liu J., et al., 2009. Robust Scale-Invariant Feature Matching for Remote Sensing Image Registration. *IEEE Geosci. Remote S.*, 6(2), pp. 287-291.
- Lingua A., Marenchino D., and Nex F., 2009. Performance Analysis of the SIFT Operator for Automatic Feature Extraction and Matching in Photogrammetric Applications. *Sensors.*, 9(5), pp. 3745-3766.
- Lowe D. G., 1999. Object Recognition from Local Scale-Invariant Features. In *Proceedings of IEEE International Conference on Computer Vision*. 20-27, September, Kerkyra, Greece, pp. 1-8.
- Lowe D. G., 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.*, 60(2), pp. 91-110.
- Matas J., Chum O., Urban M., et al., 2004. Robust wide-baseline stereo from maximally stable extremal regions. *Image Vis. Comput.*, 22(10), pp. 761-767.
- Mikolajczyk K., and Schmid C., 2004. Scale and affine invariant interest point detectors. *Int. J. Comput. Vis.*, 60(1), pp. 63-86.
- Mikolajczyk K., and Schmid C., 2005. A Performance Evaluation of Local Descriptors. *IEEE T. Pattern Anal.*, 27(10), pp. 1615-1630.
- Mikolajczyk K., Tuytelaars T., Schmid C., et al., 2005. A comparison of affine region detectors. *Int. J. Comput. Vis.*, 65(1-2), pp. 43-72.
- Morel J. M., and Yu G., 2009. ASIFT: A New Framework for Fully Affine Invariant Image Comparison. *SIAM J. Imaging Sci.*, 2(2), pp. 1-31.
- Rublee E., Rabaud V., Konolige K., et al., 2011. ORB: an efficient alternative to SIFT or SURF. In *Proceedings of IEEE International Conference on Computer Vision*, 06-13, November, Barcelona, Spain, pp. 2564-2571.
- Sedaghat A., Mokhtarzade M., and Ebadi H, 2011. Uniform Robust Scale-Invariant Feature Matching for Optical Remote Sensing Images. *IEEE T. Geosci. Remote*, 49(11), pp. 4516-4527.
- Smith S., and Brady J., 1997. Susan: a new approach to low-level image-processing. *Int. J. Comput. Vis.*, 23(1), pp. 45-78.
- Song Z., Li S., and George T. F., 2010. Remote sensing image registration approach based on a retrofitted SIFT algorithm and Lissajous-curve trajectories. *Opt. Express*, 18(2), pp. 513-522.
- Thevenaz P., and Unser M., 1998. An efficient mutual information optimizer for multiresolution image registration. In *Proceedings of IEEE International Conference on Image Processing*. 04-07, October, Chicago, USA, pp. 833-837.
- Viola P., Wells W. M., 1997. Alignment by maximization of mutual information. *Int. J. Comput. Vis.*, 24(2), pp. 137-154.
- Wallace L., Lucieer A., and Watson C. S., 2014a. Evaluating Tree Detection and Segmentation Routines on Very High Resolution UAV LiDAR Data. *IEEE T. Geosci. Remote*, 52(12), pp. 7619-7628.
- Wallace L., Musk R., and Lucieer A., 2014b. An Assessment of the Repeatability of Automatic Forest Inventory Metrics Derived From UAV-Borne Laser Scanning Data. *IEEE T. Geosci. Remote*, 52(11), pp. 7160-7169.
- Wang L., Niu Z., Wu C., et al., 2012. A robust multisource image automatic registration system based on the SIFT descriptor. *Int. J. Remote Sens.*, 33(12), pp. 3850-3869.
- Xiong Z., and Zhang Y., 2010. A critical review of image registration methods. *Int. J. Image Data Fusion*, 1(2), pp. 137-158.
- Yu Y., Huang K., Chen W., et al., 2012. A novel algorithm for view and illumination invariant image matching. *IEEE T. Image Process.*, 21(1), pp. 229-240.
- Zhang C.S., and Fraser C. S., 2007. Automated Registration of High-Resolution Satellite Images. *Photogramm. Rec.*, 22(117), pp. 75-87.
- Zhao F., Huang Q., and Gao W., 2006. Image matching by normalized cross-correlation. In *Proceedings of the 31st IEEE International Conference on Acoustics, Speech and Signal Processing*. Toulouse, France, 14-19 May, pp. 1977-1980.
- Zhou H. L., Kong H., Wei L., et al., 2015. Efficient Road Detection and Tracking for Unmanned Aerial Vehicle. *IEEE T. Intell. Transp.*, 16(1), pp. 297-309.
- Zitova B., Flusser J., 2003. Image registration methods: a survey. *Image Vision Comput.*, 21(11), pp. 977-1000.