

Research Article

Target Matching Recognition for Satellite Images Based on the Improved FREAK Algorithm

Yantong Chen,^{1,2} Wei Xu,¹ and Yongjie Piao¹

¹Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

²University of Chinese Academy of Sciences, Beijing 100049, China

Correspondence should be addressed to Yantong Chen; chenyantong1@yeah.net

Received 4 June 2016; Accepted 7 September 2016

Academic Editor: Yakov Strelniker

Copyright © 2016 Yantong Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Satellite remote sensing image target matching recognition exhibits poor robustness and accuracy because of the unfit feature extractor and large data quantity. To address this problem, we propose a new feature extraction algorithm for fast target matching recognition that comprises an improved feature from accelerated segment test (FAST) feature detector and a binary fast retina key point (FREAK) feature descriptor. To improve robustness, we extend the FAST feature detector by applying scale space theory and then transform the feature vector acquired by the FREAK descriptor from decimal into binary. We reduce the quantity of data in the computer and improve matching accuracy by using the binary space. Simulation test results show that our algorithm outperforms other relevant methods in terms of robustness and accuracy.

1. Introduction

Matching recognition has many practical applications, including target recognition by matching target features, which is an important issue in the pattern recognition field. Increasing the resolution of satellite remote sensing images generates detailed image features, thereby allowing target matching recognition [1]. Satellite image target matching recognition exhibits higher stability and accuracy than traditional methods [2].

However, satellite remote sensing image target matching recognition is a complex process. For example, the complex background of satellite images leads to poor robustness of feature extraction. A large amount of satellite images will also produce processing difficulties and matching errors. Therefore, a matching algorithm with high robustness and accuracy must be used. As a core step in satellite image target matching recognition, feature extraction may be classified into global and local invariant feature extraction based on the amount of utilised target information. Traditional target recognition methods frequently extract the target shape feature and other global invariant features, including invariant

moments [3] and transform field invariance [4]. In [5], the principal component analysis (PCA) algorithm is adopted to an image obtained via segmentation following the method of Otsu [6] to estimate the main direction and recognise the targets via template matching. In [7], Hu invariant moment is applied to recognise an aircraft from a satellite image, and then four features are extracted and combined to build the global invariant feature. However, these methods assume that the edges of the target can be perfectly extracted, which is difficult to achieve in practice. To address this problem, we use the local invariant feature that is robust to noise and partial occlusion.

The local invariant feature comprises feature detection and feature description, which have drawn the attention of many researchers. Introduced by Lowe [8], the scale-invariant feature transform (SIFT) is a new scale-invariant method that uses a difference of Gaussian (DoG) detector and a SIFT descriptor to extract a robust and high-class discriminative degree local invariant feature vector. Introduced by Bay et al. [9], the speeded up robust feature (SURF) uses a fast Hessian feature detector and applies Haar-like features to create a feature vector. Although simpler than SIFT, SURF exhibits

poor performance in extracting scale-invariant features. Ke and Sukthankar [10] proposed the PCA-SIFT descriptor, which outpaces SIFT in describing key points by using the PCA algorithm. Although the GLOH [11] and DAISY [12] algorithms apply linear dimension reduction to improve the SIFT descriptor, they both involve a complex computation process. Bit operation presents a favourable alternative to reduce the complexity of the computation process in a computer; however, this method requires the transformation of the vector into binary. In [13], the binary robust independent elementary feature (BRIEF) approach is proposed; this method uses a FAST detector and bit operation for matching and offers a considerably more suitable alternative for real-time applications. However, despite its obvious advantage in speed, BRIEF exhibits poor reliability and robustness because of its minimal tolerance to image distortions and transformations, particularly rotation and scale change. Despite simulating the human vision system using a simple binary transform method to accelerate the matching process, the fast retina key point (FREAK) [14] descriptor has a relatively low accuracy.

Given these limitations, we propose a new efficient method for satellite remote sensing image target matching recognition based on FREAK feature extraction algorithm. In particular, we extend the FAST detector by applying scale space theory to improve its poor robustness that results from a complex satellite image background. We also create a binary data space to transform the high-dimensional FREAK descriptor from a decimal feature vector to a binary feature vector to improve its accuracy. The rest of this paper is organised as follows. Section 2.1 presents the scale-invariant FAST feature detector. Section 2.2 describes the FREAK descriptor and the binary data space for transforming the feature vector into binary data. Section 3 presents the simulation results. Section 4 concludes the study and offers directions for future work.

2. Improved FREAK Algorithm

Feature extraction algorithm often comprises feature detection and feature description. And the FREAK feature extraction algorithm comprises FAST detector and FREAK descriptor. However, the FAST detector is not scale invariant. The FREAK descriptor suffers from large quantities of data storage. So we improved FREAK feature extraction algorithm as follows.

2.1. Scale-Invariant FAST Feature Detector

2.1.1. Scale Space Theory. Feature detection begins by identifying locations and scales that can be repeated under different views of the same target. Detecting locations involves the search for stable features at all possible scales. Therefore, confirming the invariant scale is the key step in feature detection. Scale space theory regards scale as a free parameter and adds it to the signal. Lindeberg [15] identified the Gaussian function as the only possible scale-space kernel under various reasonable assumptions. Therefore, the scale space of an image is defined as function $L(x, y, \sigma)$, which is

produced by the function of variant scale Gaussian $G(x, y, \sigma)$ with an input image $I(x, y)$ as follows:

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y), \quad (1)$$

where \otimes is the convolution operation, and

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

To detect features in the scale space effectively, we use the DoG function of Lowe [8] to convolve the image $D(x, y, \sigma)$ and then compute the difference between two nearby scale images using a constant multiplicative factor k as follows:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) \otimes I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma). \end{aligned} \quad (3)$$

Function L should compute all the scale space features of the images, and function D can be easily calculated via image subtraction. The difference of the Gaussian function is closed to the scale-normalised Laplacian of Gaussian $\sigma^2 \nabla^2 G$, which is a true yet complex scale invariance. In Mikolajczyk and Schmid [11], the Laplacian of the Gaussian function provided a more stable image feature than the gradient, Hessian, or Harris corner functions. The relationship between the difference of the Gaussian function and the Laplacian of the Gaussian function is expressed as follows:

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G, \quad (4)$$

where $\nabla^2 G^2$ is approximate to $\partial G / \partial \sigma$ at scales $k\sigma$ and σ ; that is,

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma}. \quad (5)$$

Therefore,

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k-1) \sigma^2 \nabla^2 G. \quad (6)$$

However, if k is equal to 1, then the approximate equation is incorrect. The factor $(k-1)$ in the equation is a constant and does not influence the location and stability of the features. Therefore, k can be defined as a constant value, such as $k = \sqrt{2}$.

2.1.2. Scale-Invariant FAST Detector. A FAST detector functions as the foundation of the smallest univalue segment assimilating nucleus (SUSAN) detector [16]. However, the FAST detector is faster and more accurate than the SUSAN detector, thereby making the former suitable for satellite remote sensing images. The SUSAN detector has been implemented using a circle window, and the centre of the circle window is defined as the nucleus. The univalue segment assimilating nucleus (USAN) area includes points with brightness levels that are equal to that of the nucleus, whereas the unassimilated area includes points that do not have the equivalent brightness level of the nucleus. The area near the

nucleus is divided into two. A larger USAN area is obtained if this area also covers the nucleus. Otherwise, the USAN area is only half the size of the circle window. The nucleus will be detected as a corner when the USAN area is small. Following this algorithm, the FAST detector can be expressed as follows:

$$c(r, r_0) = \begin{cases} 1 & \text{if } |I(r) - I(r_0)| \leq t \\ 0 & \text{if } |I(r) - I(r_0)| > t, \end{cases} \quad (7)$$

where r_0 is the position of the nucleus in the image, r is the position of any other point in the circle window, $I(r_0)$ is the brightness of the nucleus, $I(r)$ is the brightness of any other point, t is the brightness difference threshold, and c is the output result. The pixels in the image are compared with one another, and the total result n is calculated as follows:

$$n(r_0) = \sum_{\bar{r}} c(r, r_0), \quad (8)$$

which denotes the size of the USAN area and the number of pixels within this area. This result is minimised. t determines the minimum contrast of the features to be detected and the maximum amount of noise to be ignored.

The initial feature is then detected using the following rule:

$$R(r_0) = \begin{cases} g - n(r_0) & \text{if } n(r_0) < g \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

where $R(r_0)$ is the initial detected feature and g is the threshold compared with n , which is set to $3n_{\max}/4$, where n_{\max} is the maximum value of n . This value is calculated by analysing the response of noise.

Despite providing a favourable result, the aforementioned method still requires improvement. The FAST detector provides the following equation, which is more stable and reasonable than (7):

$$c(r, r_0) = e^{-((I(r) - I(r_0))/t)^6}. \quad (10)$$

The preceding equation is smoother than the SUSAN detector and allows slight changes in the brightness of each pixel, thereby permitting the new function to build a search decision tree for feature detection. The brightness degree of each pixel must be specified and compared with that of the eight nearby pixels. Consequently, the configuration space contains three states, namely, "darker," "brighter," and "similar." The searching tree is assigned as follows:

$$c(r, r_0) = \begin{cases} d, & I(r) \leq I(r_0) - t \text{ (darker)} \\ s, & I(r_0) - t < I(r) < I(r_0) + t \text{ (similar)} \\ b, & I(r_0) + t \leq I(r) \text{ (brighter)}. \end{cases} \quad (11)$$

Given its sensitivity to scale change, we add scale space to the FAST detector. We show the image at a Gaussian scale space through a convolution process. We mark the extreme point in a 3×3 template based on the computation results. If the corner point is not the extreme point, then we reinterpolate the image coordinates between the patches in the layers next to the determined scale. The scale-invariant FAST detector is expressed as follows:

$$\begin{aligned} S(r, r_0, \sigma) &= (c(r, r_0, k\sigma) - c(r, r_0, \sigma)) \otimes I(r, r_0) \\ &= L(r, r_0, k\sigma) - L(r, r_0, \sigma). \end{aligned} \quad (12)$$

Direction estimation is another important process in feature detection. The points are placed on the edge between two reigns of the USAN area that can consist of a line, which is the direction of the edge. Therefore, the edge direction is calculated by finding the longest axis of symmetry and taking the sum of the following equations:

$$\begin{aligned} \overline{(x - x_0)^2}(r_0) &= \sum_r (x - x_0)^2 c(r, r_0), \\ \overline{(y - y_0)^2}(r_0) &= \sum_r (y - y_0)^2 c(r, r_0), \end{aligned} \quad (13)$$

$$\overline{(x - x_0)(y - y_0)}(r_0) = \sum_r (x - x_0)(y - y_0) c(r, r_0).$$

We use the ratio of $\overline{(y - y_0)^2}$ to $\overline{(x - x_0)^2}$ to determine the orientation of the edge and then use the sign of $\overline{(x - x_0)(y - y_0)}$ to determine whether a diagonal edge has a positive or negative gradient. Finally, we estimate the feature direction in this manner.

2.2. Binary FREAK Descriptor

2.2.1. FREAK Descriptor. The previous operations have detected the features scale, location, and orientation of an image. We then compute a descriptor for the image that remains invariant and stable at different situations, such as scale change, viewpoint change, and illumination change.

The FREAK descriptor simulates the human retina vision system. The human vision system has a complex process. Firstly, the light stimulates the retina to excite the optic nerve cells. Secondly, the optic nerve transfers the information to the retinal lateral geniculate nucleus (LGN) for decoding. Thirdly, a central visual cell and a peripheral visual cell in the LGN extract the detail information and contour feature. Fourthly, the central nervous system transfers the processed information to the primary region of the human brain. The visual information is fully recognised by integrating the information obtained from different cortical regions.

The FREAK descriptor designs the fast retina key point sample model according to the human retina imaging principle. This model includes a concentric circle with seven floors as shown in Figure 1. Each concentric circle has six sampling points that imitate the relationship between central visual cells and peripheral visual cells. The central sample

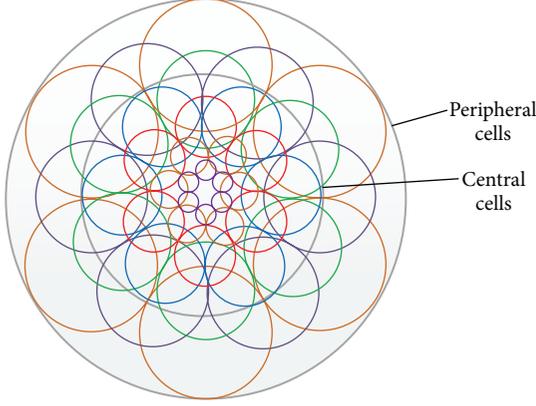


FIGURE 1: FREAK descriptor sampling model.

circle extracts the texture feature, whereas the peripheral cells extract the outline feature.

The FREAK descriptor is created using the grey threshold of the sampling points as follows:

$$\|F\| = \sum_{0 \leq a \leq k} 2^a J(S_a), \quad (14)$$

where $\|F\|$ is the value of the FREAK descriptor, S_a is the location of the sampling point, and $J(S_a)$ is the grey threshold value. Finally, we obtain the high-dimensional feature vector by training.

2.2.2. Binary Data Space. Traditional matching method matches two features by Euclidean distance. For example, these two features are $F_1 = m_1 m_2 \cdots m_{512}$ and $F_2 = n_1 n_2 \cdots n_{512}$. We compare them by Euclidean distance as follows:

$$\begin{aligned} \text{Euclidean distance}(F_1, F_2) = & \text{sqrt}((m_1 - n_1)^2 \\ & + (m_2 - n_2)^2 + \cdots + (m_{512} - n_{512})^2). \end{aligned} \quad (15)$$

It will cost a long time to compute the above equation. But if the features are binary, we can compute the features Hamming distance by bit operation.

$$\text{HM.distance}(F_1, F_2) = F_1 \oplus F_2, \quad (16)$$

and we can match two features quickly in this way.

To design an efficient method for target matching, we transform the high-dimensional feature vector into binary. Firstly, we apply the binary data space in [17] and mark the feature vector on the data space. We use two bits per each dimension image data $\{a, b, c, d, e, f\}$ as an example. Figure 2 shows the marked binary data space. Each dimension has four partitions according to this space, namely, 00, 01, 10, and 11, the ranges of which are (0–0.25), (0.25–0.5), (0.5–0.75), and (0.75–1), respectively. Therefore, the entire data space is divided into 16 groups ($2^2 \times 2^2$).

For example, the vector data item “a,” the real value of which is (0.1, 0.4), has been marked as “0001” as shown in Table 1.

TABLE 1: Marked data collection.

	Original data		Mark value	
a	0.1	0.4	00	01
b	0.2	0.9	00	11
c	0.3	0.7	01	10
d	0.4	0.6	01	10
e	0.7	0.7	10	10
f	0.9	0.2	11	00

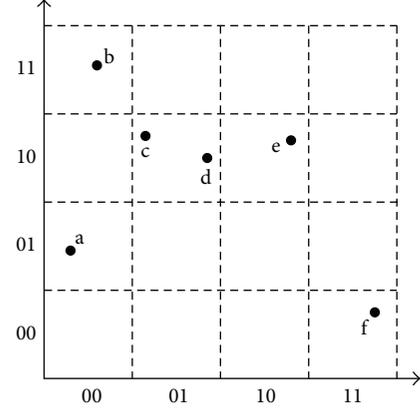


FIGURE 2: Binary system data space.

Different vector data items may have the same mark value, as shown by points c and d in Figure 2. Equation (15) presents the relationship amongst the number of feature dimensions (D), the number of bits used for signature (n), and P .

$$P = \left(\frac{1}{2^n}\right)^D. \quad (17)$$

However, satellite images always have high feature dimensions; hence, the probability that the aforementioned condition will occur is small.

3. Simulation Test

3.1. Parameter Setting. We conduct experiments to compare the results of our method with those of several state-of-the-art algorithms. Figure 3 shows our algorithm. Firstly, we use the scale-invariant FAST detector to detect feature points. Secondly, we describe the points to the high-dimensional feature vector using the FREAK algorithm. To achieve better performance, we transform the feature vector into binary using the binary data space instead of the FREAK conversion method. Thirdly, we compute the matching recognition results through a bit operation between the target template and the processed vector data. In our experiments, we analyse different satellite images from the company databases of Astrium (<http://www.astrium-geo.com/>) and Digitalglobe (<http://www.digitalglobe.com/products/basic-imagery>).

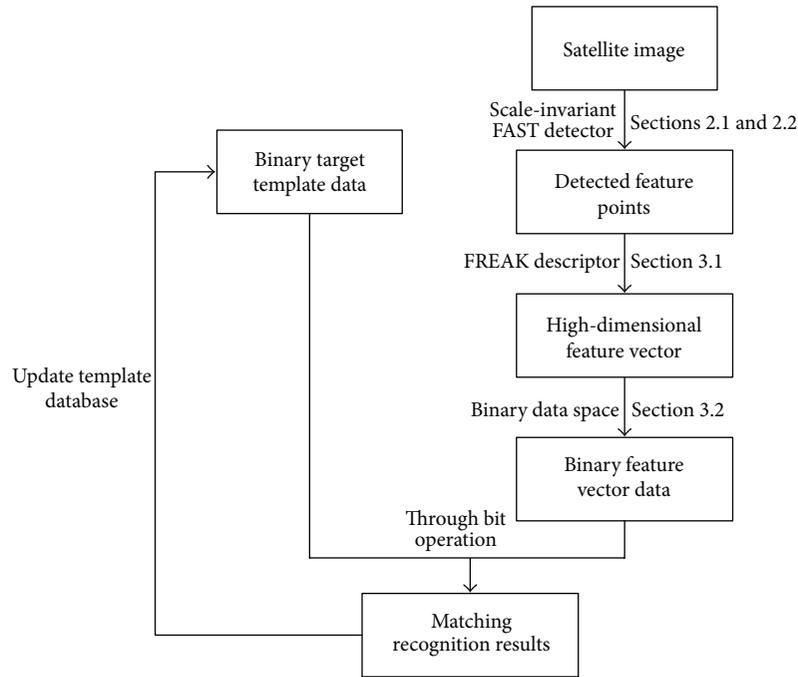


FIGURE 3: Flowchart of the algorithm.

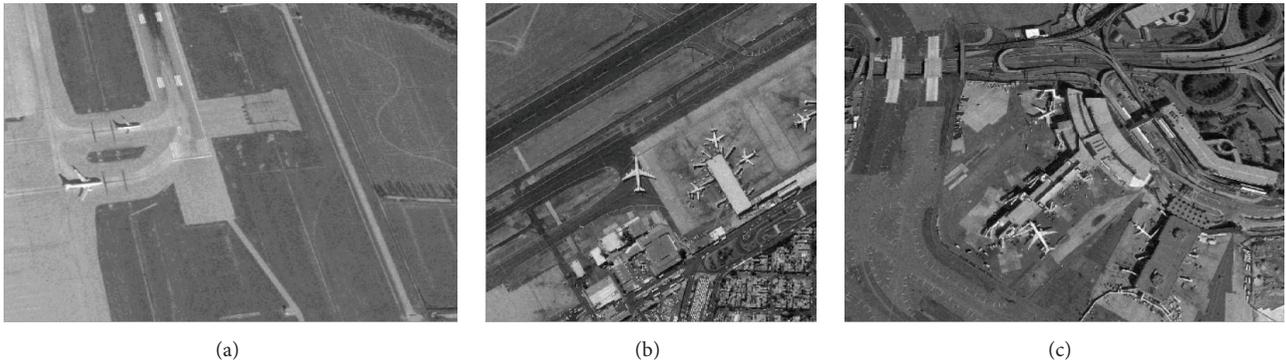


FIGURE 4: Satellite images of the (a) Beijing airport, (b) Mexico airport, and (c) New York airport.

Amongst the analysed images, 60 are from the SPOT6 satellite with a 1.5 m resolution, 60 are from the QuickBird satellite with a 0.61 m resolution, and 60 are from the WorldView-2 satellite with a 0.5 m resolution. Figures 4(a), 4(b), and 4(c) show the images from these three satellites, respectively.

We design two experiments to validate our method. In the first subsection, we compare our proposed method with various feature detectors. In the second subsection, we show our target matching recognition results and evaluate the accuracy of our method with the FREAK algorithm.

3.2. Performance of Feature Detectors. Scale and viewpoint change frequently according to the characteristics of the remote sensing image and illumination transfers with the

sun. Considering these factors, we conduct our repeatability experiments under different scale, viewpoint, and illumination conditions that adopt the Calonder et al. [13] change methods. By comparing our proposed detector with the DoG, FAST-Hessian, and FAST detectors, we test the performance of the feature detectors as shown in Figure 5.

Figure 5 shows the repeatability rate of the feature detectors. A higher repeatability indicates higher invariance and stability of the detectors. Figure 5(a) shows that our proposed detector is more robust under the scale change condition than the other detectors, which can be attributed to the addition of the Gaussian operator difference to our method. Figures 5(b) and 5(c) show the performance of the detectors under the viewpoint and illumination change conditions. Our detector behaves as favourably as the FAST

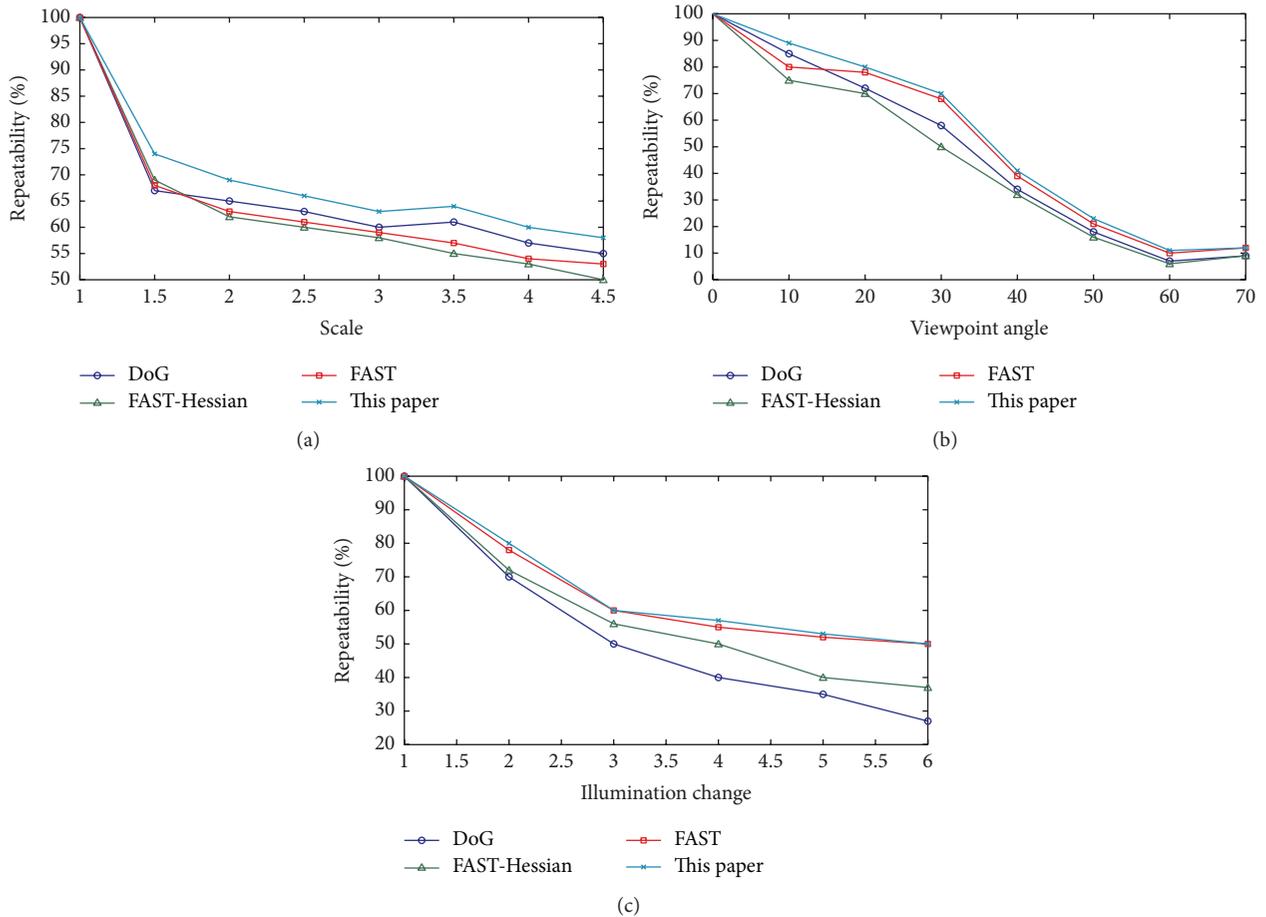


FIGURE 5: Repeatability rate of the feature detectors. (a) Scale change. (b) Viewpoint change. (c) Illumination change.

detector, which is more repeatable than the DoG and FAST-Hessian detectors. Therefore, our proposed detector inherits the characteristics of the FAST detector and extends its scale-invariant characteristic.

3.3. Matching Results and Analysis. Given that the algorithm is the key element in remote sensing image target matching recognition, we test the accuracy of different algorithms on various satellite images. We build the scale space to enhance the scale-invariant capability of the detector based on our proposed method as shown in Figure 6.

We then compare the matching accuracy of the FREAK algorithm with that of our method under simple, common, and complex image backgrounds. Figures 7, 8, and 9 show the matching results for these backgrounds, respectively.

Our method outperforms the FREAK algorithm regardless of the background. In our matching results, the small target matches more points than the larger ones because we apply the scale space in this target, thereby making the detectors scale invariant. The FREAK operator exhibits many matching errors as shown in Figures 7, 8, and 9. Therefore, instead of using the FREAK simple transformation

method, we create a binary data space to transform the high-dimensional FREAK descriptor from a decimal feature vector to binary.

In the objective evaluation, our method still outperforms the FREAK algorithm. Table 2 compares the matching accuracies of the two detectors. We can accurately judge the correct matching ratio of these two operators from the comparison results.

4. Conclusion

We proposed a new efficient feature matching method for satellite remote sensing image matching recognition. We extended the FAST detector with scale space to detect invariant key points and improved the FREAK descriptor to acquire highly accurate feature vectors at the feature description phase. We transformed the high-dimension feature vector into binary by using binary space. Our proposed method outperformed the other detection algorithms in terms of detector repeatability and accuracy. However, the matching time remains extremely long. Therefore, we plan to transform the nonlinear or fast search algorithm in our future work to

TABLE 2: Matching accuracy of the two detectors.

Operators	Image 1			Image 2			Image 3		
	Total	Correct	Rate	Total	Correct	Rate	Total	Correct	Rate
FREAK	35	31	88.6%	44	36	81.8%	72	58	80.6%
Proposed	43	39	90.7%	54	46	85.2%	84	69	82.1%

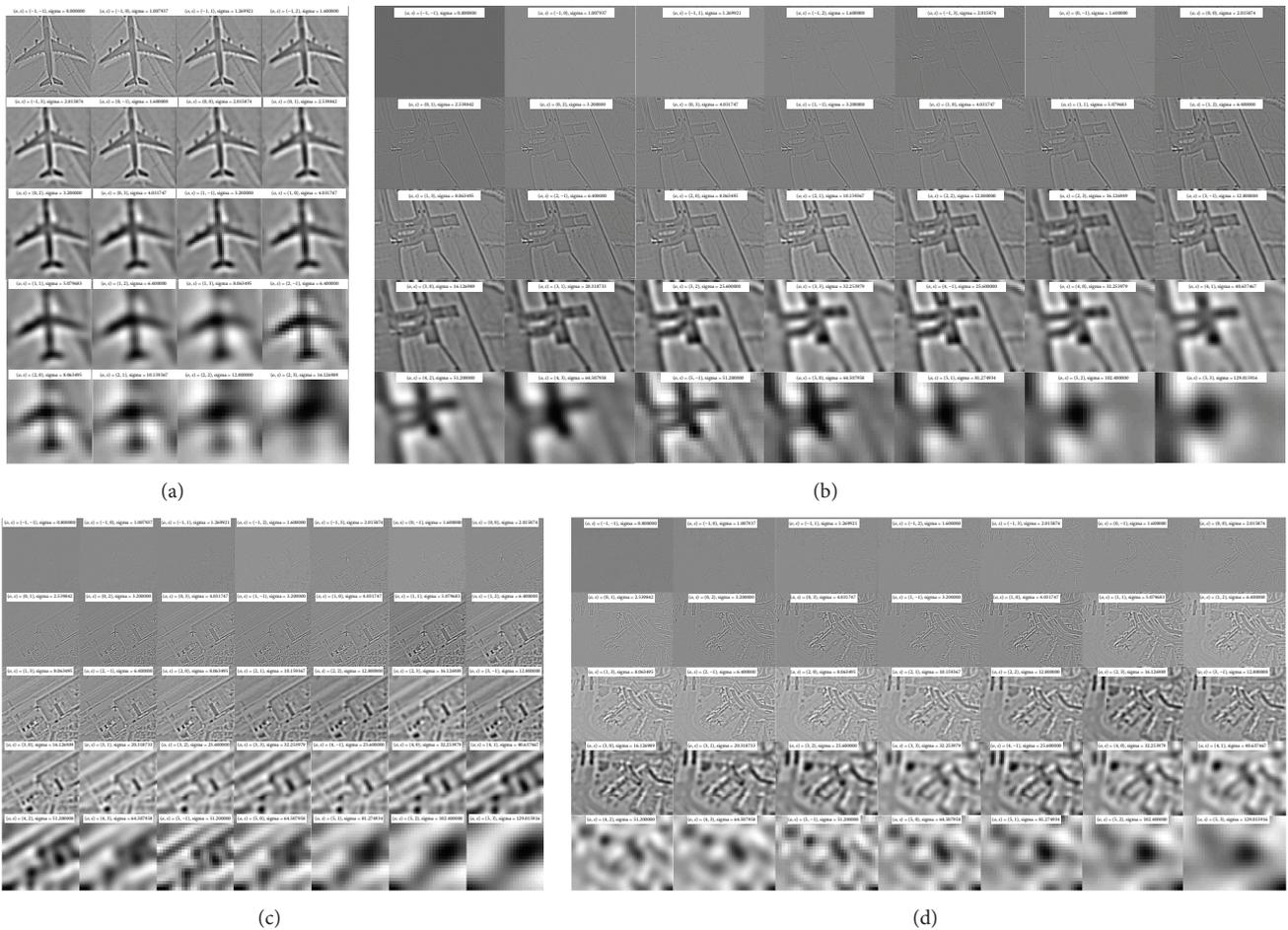


FIGURE 6: Scale space. (a) Target template. (b) Beijing airport. (c) Mexico airport. (d) New York airport.

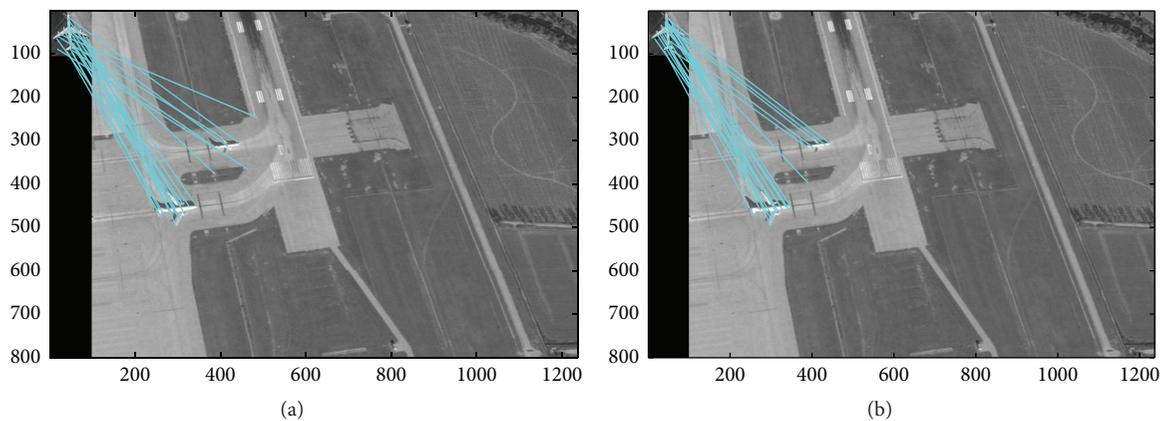


FIGURE 7: Simple background matching recognition. (a) FREAK result. (b) Our result.

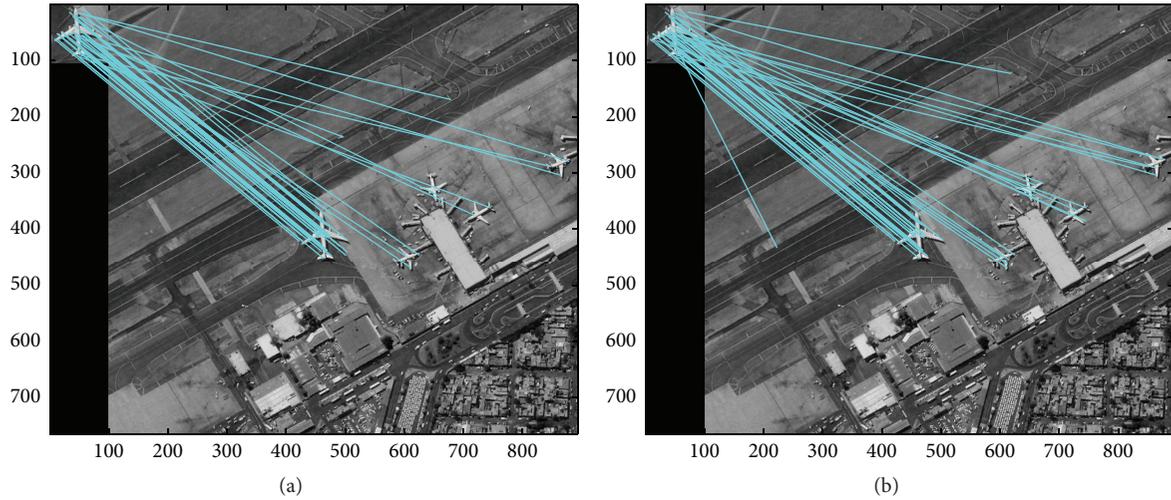


FIGURE 8: Common background matching recognition. (a) FREAK result. (b) Our result.

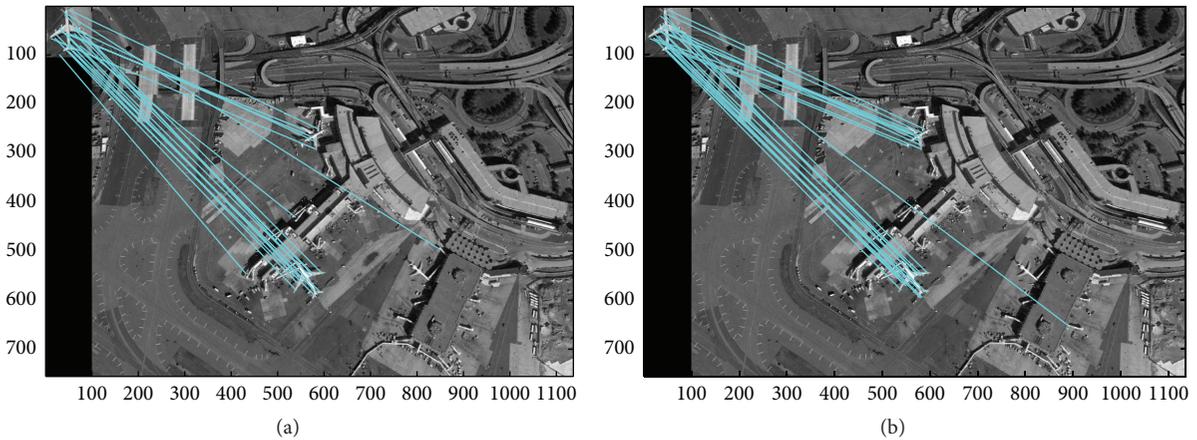


FIGURE 9: Complex background matching recognition. (a) FREAK result. (b) Our result.

create a new algorithm that can reduce matching recognition time.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work is partially supported by the National High-Tech R&D Program of China “863 Program” (no. 2012AA121502).

References

- [1] S. Korman, D. Reichman, G. Tsur, and S. Avidan, “FasT-match: fast affine template matching,” in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 2331–2338, Portland, Ore, USA, June 2013.
- [2] H. Jiang, T.-P. Tian, K. He, and S. Sclaroff, “Scale resilient, rotation invariant articulated object matching,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 143–150, IEEE, Providence, RI, USA, June 2012.
- [3] J. C. Yang and D. S. Park, “A fingerprint verification algorithm using tessellated invariant moment features,” *Neurocomputing*, vol. 71, no. 10–12, pp. 1939–1946, 2008.
- [4] C. Cyr and B. Kimia, “3D object recognition using similarity-based aspect graph,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (ICCV '01)*, pp. 254–261, IEEE, Vancouver, Canada, July 2001.
- [5] V. Deepu and S. Madhvanath, “Principal component analysis for online handwritten character recognition,” *Pattern Recognition*, vol. 2, no. 1, pp. 327–330, 2004.
- [6] N. Otsu, “A threshold selection method from gray-level histograms,” *Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [7] Q. Chen, “A comparative study of Fourier descriptors and Hu’s seven moment invariants for image recognition,” *Electrical and Computer Engineering*, vol. 1, no. 5, pp. 103–106, 2004.
- [8] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

- [9] H. Bay, A. Ess, and T. Tuytelaars, "A Ess, and T Tuytelaars, SURF: speeded up robust features," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [10] Y. Ke and R. Sukthankar, "PCA-SIFT: a more distinctive representation for local image descriptors," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, pp. 11506–11513, San Diego, Calif, USA, July 2004.
- [11] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [12] E. Tola, V. Lepetit, and P. Fua, "DAISY: an efficient dense descriptor applied to wide-baseline stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 815–830, 2010.
- [13] M. Calonder, V. Lepetit, and C. Strecha, "BRIEF: binary robust independent elementary features," in *Proceedings of the European Conference on Computer Vision (ECCV '10)*, pp. 778–792, Crete, Greece, 2010.
- [14] A. Alahi and E. Polytech, "FREAK: fast retina keypoint," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 510–517, IEEE, Providence, RI, USA, June 2012.
- [15] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.
- [16] D. Serby, Esther-Koller-Meier, and L. Van Gool, "Probabilistic object tracking using multiple features," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR '04)*, pp. 184–187, August 2004.
- [17] H.-J. Lee and J.-W. Chang, "Signature-based hybrid spill-tree for indexing high-dimensional data," in *Proceedings of the IEEE 9th International Conference on Computer and Information Technology (CIT '09)*, pp. 287–292, IEEE, Xiamen, China, October 2009.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

