A Robust Insulator Detection Algorithm Based on Local Features and Spatial Orders for Aerial Images

Shenglong Liao and Jubai An

***Abstract*—The detection of targets with complex backgrounds in aerial images is a challenging task. In this letter, we propose a robust insulator detection algorithm based on local features and spatial orders for aerial images. First, we detect local features and introduce a multiscale and multifeature descriptor to represent the local features. Then, we get several spatial orders features by training these local features, it improves the robustness of the algorithm. Finally, through a coarse-to-fine matching strategy, we eliminate background noise and determine the region of insulators. We test our method on a diverse aerial image set. The experimental results demonstrate the precision and robustness of our detection method, and indicate the possible use of our method in practical applications.**

***Index Terms*—Aerial image, insulator detection, local feature, point matching, spatial orders.**

I. Introduction

MONITORING the status of equipment on power lines with helicopters is one of the most important jobs in power system operation. To improve the efficiency of inspection, the traditional manual inspection technology is being replaced by new technology. Computer vision technology can quickly and efficiently detect defects on power lines, and greatly reduce the workload by automatically detecting and segmenting the power line insulators. The aerial images captured on helicopters often include various cluttered backgrounds such as mountains, rivers, grassland, and farmland; thus, the processing of original images is complicated, which will easily lead to a wrong result.

In the literature, Wu *et al.* [1] used the global minimization active contour model (GMAC) for insulator segmentation, resulting in good segmentation quality. However, the algorithm that uses global features requires a great deal of computation time, particularly for a high-resolution image. An object detection algorithm based on local features for high-spatial-resolution aerial images is more effective. Zhang and Yang [2] proposed an insulators recognition method with curved smooth and weak textured (CSWT) surfaces. Li *et al.* [3] used the improved MPEG-7 EHD (edge histogram method) technique to realize recognition of insulator. Oberweger *et al.* [4] presented a novel approach based on discriminative training of local gradient-based descriptors and a subsequent voting scheme for localization to detect insulators in aerial images. In addition, we found some useful object detection algorithm based local feature. Zhang *et al.* [5] conducted a comprehensive study of local features to classify texture and obtain object categories. Sırmaçek and Ünsalan [6] proposed an urban-area detection method based on local feature points and spatial voting. Weizman and Goldberger [7] applied a pixel-level variant of visual words to an urban-area extraction by adapting it to meet the demands of urban segmentation. Moreover, some methods based on the bag of words (BOW) model can build effective structural object descriptions using certain kinds of coding methods (see, e.g., [8]–[10]). Sun *et al.* [10] proposed a target detection algorithm for high-resolution remote sensing images using spatial sparse coding.

The contributions of this letter can be summarized as follows. First, we propose a novel insulator detection algorithm based on local features that can detect an insulator in a complex background. Second, we introduce a multiscale and multifeature (MSMF) descriptor. It not only effectively describes local features but also has the ability to handle rotation variations and scale variations. Third, we find several spatial orders features (SOFs) that improve the robustness of the algorithm. Moreover, our method adopts a coarse-to-fine matching strategy. The performance of feature point matching is significantly improved.

II. Object Detection Algorithm for Aerial Insulator Images

An object detection algorithm based on local features generally includes three main parts: local feature detection, local feature description, and feature matching. To improve the performance of the insulator detection algorithm for an aerial image, we optimize each part.

Fig. 1 shows the proposed insulator detection algorithm. In the training phase, a set of descriptors is obtained by feature detection and feature description from the training sample. With these features, we generate the visual vocabulary of insulator by the K-means algorithm, and find SOF of each visual word. In the detection phase, feature points of insulator are extracted by feature matching, and then the insulator is detected by using these target points.

Manuscript received April 15, 2014; revised September 2, 2014, October 17, 2014, and October 22, 2014; accepted November 6, 2014. This work was supported in part by the National Science Foundation of China under Grant 61201454 and the Special Funds in the Public Interest under Grant 2013418025. *(Corresponding author: Jubai An.)*

The authors are with the Information Science and Technology College, Dalian Maritime University, Dalian, 116026, China (e-mail: liaoshenglong7@163.com; jubaian@sohu.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2014.2369525

1545-598X © 2014 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

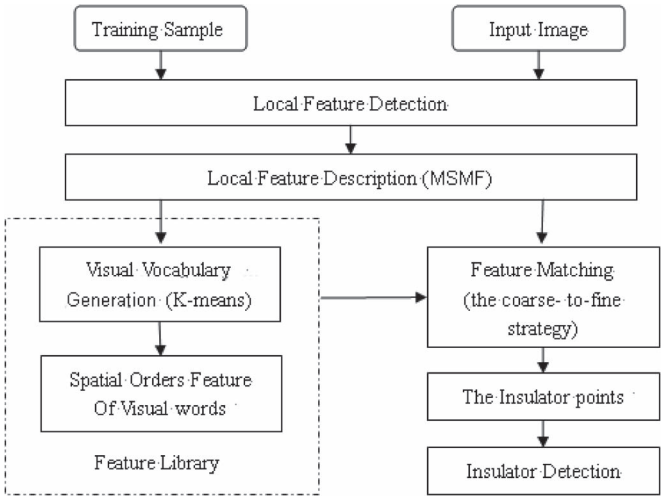

Fig. 1. Insulator detection algorithm for aerial images. (Left) Training phase. (Right) Detection phase. Feature library generation is shown in the dotted line box.

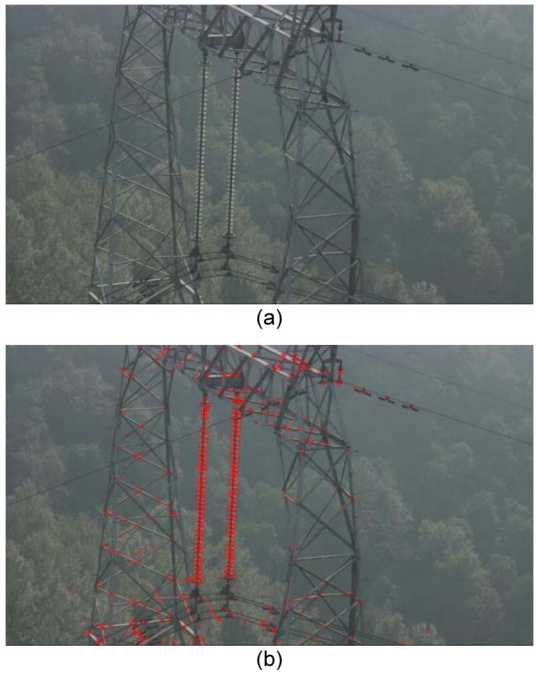

Fig. 2. (a) Aerial insulator image by using helicopters. (b) (Red) Local feature points that are detected.

A. Feature Detection and Description

By the term "feature detection," we mean the detection of local feature points or local patches. To reduce the complexity of the subsequent processing, we want to detect the feature points of insulator as many as possible (to improve the recall rate, we allow the existence of noise while forbid the missing of the insulator point). There are many methods to detect local feature points, such as the Harris affine region detector [11], and Lowe's difference-of-Gaussians detector [12]. We use an improved Harris corner selection strategy [13], which searches only the corners near enhanced edges by using a z-score normalization. Fig. 2 shows local feature points that are detected.

After detecting local feature points, we construct a descriptor to represent these local feature points. Over the past nearly

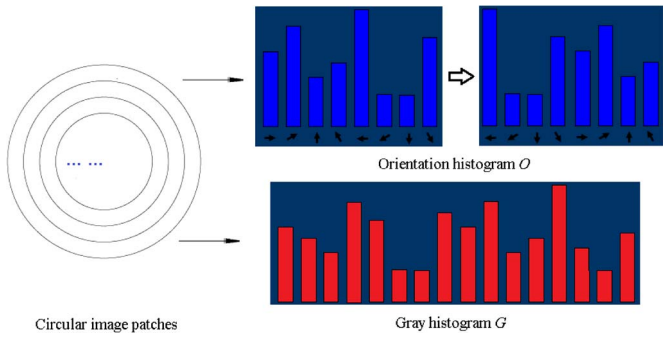

Fig. 3. Multiple scale circular image patch, orientation histogram O (calibration), and gray histogram G.

20 years, the scale-invariant feature transform (SIFT) descriptor [12] has been shown to outperform a set of existing descriptors. Recently, many new descriptors have been proposed, such as SPIN [14], RIFT [15], and DAISY [16]. In our method, we introduce an MSMF descriptor by improving the descriptor [15], which combines the SPIN descriptor and the RIFT descriptor in a unified framework. The term "multiscale" refers to the multiple regions of support (circular image patches around the feature point, as shown in Fig. 3). The term "multifeature" means including both gradient(orientation)-based and gray-level-based feature.

We can represent the descriptor of each feature point by a gradient (orientation) histogram vector O and a gray histogram vector G

$$O = [o_{\mathbf{1}}, o_{\mathbf{2}}, \ldots, o_{l-\mathbf{1}}, o_l, o_{l+\mathbf{1}}, \ldots, o_m] \tag{1}$$

$$G = [g_{\mathbf{1}}, g_2, \ldots, g_n] \tag{2}$$

where all histograms are normalized by using min–max normalization. o_l is the largest value in O. To maintain rotation invariance, we calibrate O as follows (as shown in Fig. 3):

$$O = [o_l, o_{l+\mathbf{1}}, \ldots, o_m, o_{\mathbf{1}}, o_{\mathbf{2}}, \ldots, o_{l-\mathbf{1}}]. \tag{3}$$

We used 8 bins for the scale, 16 for the gray level, and 8 for the gradient orientation, resulting in 192-dimensional descriptors.

B. Feature Library Generation (Training Phase)

The purpose of training is to obtain a set of insulator features (visual vocabulary) and relative parameters and to build a feature library. Each feature of the library consists of two parts: the visual vocabulary of insulator, and the SOF, which represents the spatial relations of visual words in the region of support (circular image patch).

During the training phase, we extract a large number of features from the training set. There are many redundant features. Therefore, the first step is to choose representative features (visual vocabulary) by cluster analysis. The cluster method we used is K-means, and we calculate the distance (the chi-square distance) as

$$d(H_1, H_2) = \frac{\mathbf{1}}{\mathbf{2}} \sum_i \frac{[H_{\mathbf{1}}(i) - H_{\mathbf{2}}(i)]^2}{H_{\mathbf{1}}(i) + H_{\mathbf{2}}(i)} \tag{4}$$

where H_1 and H_2, denote the histogram vector of two features (consists of orientation histogram O and gray histogram G), respectively, which are described in section A; i denotes the number of histogram bins.

We partition these features into k clusters. The choice of the k value is determined by the variance test. It can be represented as $f(k)$, which should be the minimum

$$f(k) = \frac{1}{k}\sum_{n=1}^{k}(d_n - d)^2 \tag{5}$$

where d_n denotes the minimum of the distance between nth cluster and other cluster, d denotes the proper distance (learn by experiment, d value of 0.15, k value is 11).

Finally, the visual vocabulary of insulator can be represented as

$$L_{vv} = \{(H_1, \mu_1), (H_2, \mu_2), \ldots, (H_k, \mu_k)\} \tag{6}$$

where H_k denotes the histogram vector of the kth visual word. It is calculated by taking the average of all feature vectors in the kth cluster, and we get the threshold μ_i $(i = 1, \ldots, k)$ as the matching threshold by training samples (consists of positive and negative samples).

The SOF is taken as the adjacent spatial order of each feature point. Liu [17] proposed a feature point matching algorithm based on restricted spatial order constraints. This algorithm obtained high precision and stability. Zhao *et al.* [18] also used spatial orders to improve feature matching in remote sensing image registration.

The main spatial relations between the points (all points in a training patch, as shown in Fig. 3) are direction and distance. By analyzing a large number of positive samples, we find several features of spatial orders. The SOF can be represented as a histogram vector of the Euclidean distance between the feature points, a histogram vector of the direction between the feature points and the density of the support region of the feature point (the number of points around the central point)

$$H_{\text{sof}} = [h_{\text{dis}}, h_{\text{dir}}, T_{\text{density}}] \tag{7}$$

where h_{dis} denotes the vector of the distance, and h_{dir} denotes the vector of the direction. T_{density} denotes the density of the support region of the feature point (obtained by training the positive sample).

Finally, the SOF of the feature library can be represented as

$$L_{\text{sof}} = \left\{\left(H^1_{\text{sof}}, \mu^1_{\text{sof}}\right), \left(H^2_{\text{sof}}, \mu^2_{\text{sof}}\right), \ldots, \left(H^k_{\text{sof}}, \mu^k_{\text{sof}}\right)\right\} \tag{8}$$

where all histograms are normalized, and the threshold μ^i_{sof} $(i = 1, \ldots, k)$ as the matching threshold is obtained by training positive and negative samples. To keep the rotation invariance, we calibrate these histograms by the method in section A and reference [10].

C. Feature Matching

Feature matching is a key step in the detection phase. To improve the efficiency of matching, we introduce a coarse-to-fine matching strategy. According to the feature library, the matching process is divided into two steps.

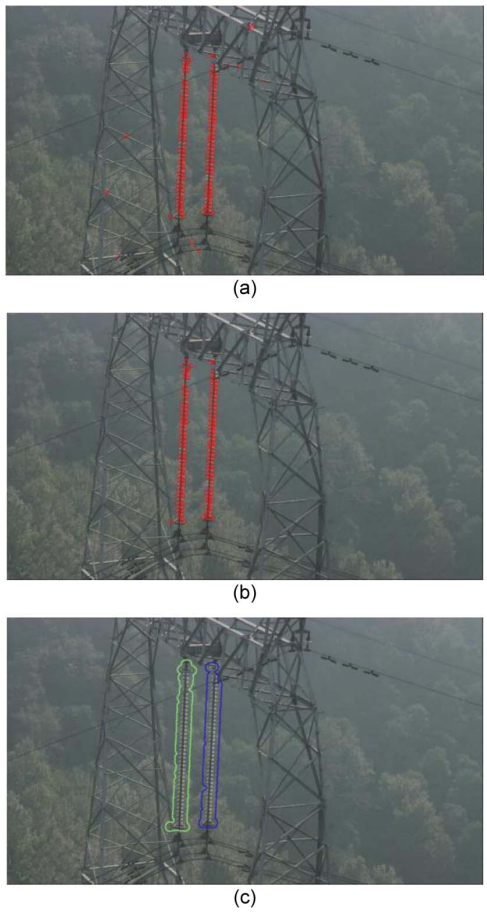

Fig. 4. (a) Result of the first matching. (b) Result of the second matching. (c) Result of object detection.

First, detect the features (with a lot of noninsulator features) from the input image and describe them by the method in section A, which are represented as $\{H^1_{\text{msmf}}, H^2_{\text{msmf}}, \ldots, H^i_{\text{msmf}}\}$. Then, these features will match with each visual word in L_{vv} using (4), and determine whether it is the feature of insulator by the corresponding threshold in L_{vv}. We can eliminate most of noninsulator features after the first matching. Fig. 4(a) shows the result of the first matching.

After the first matching, we extract the SOF from the rest of features (feature points) using the same way in section B, which are represented as $\{F^1_{\text{sof}}, F^2_{\text{sof}}, \ldots, F^j_{\text{sof}}\}$. Finally, these SOFs will match with corresponding SOF in L_{sof} using (4) and determine whether it is the feature of insulator by the corresponding threshold in L_{sof}. After the second matching, we can remove the rest of noninsulator features and leave all the insulator features (feature points). Fig. 4(b) shows the result of the second matching.

D. Object Detection

After feature matching, the location of the object is obvious. We define a voting matrix based on the extracted local feature points [5]. Around each extracted feature point, there is a high probability of the object occurring in the region. Therefore,

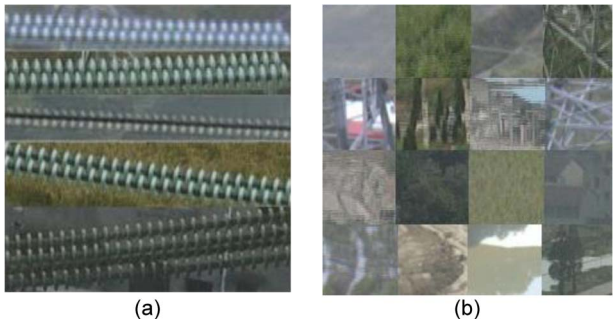

Fig. 5. (a) Positive training samples (the insulator image). (b) Negative training samples (the background image).

each extracted feature point has the highest vote at its spatial coordinate (x_i, y_i), and its votes decrease with spatial distance. We form the voting matrix as

$$V(x,y) = \sum_i \frac{\mathbf{1}}{\mathbf{2}\pi\sigma^2} \exp\left(-\frac{(x-x_i)^2 + (y-y_i)^2}{\mathbf{2}\sigma^2}\right) \quad (9)$$

where σ denotes the scale of the support region. Finally, we process the voting matrix using the threshold (the optimum threshold value is obtained by training, the maximum value of 0.7 times in the voting matrix), and thereby obtaining the region for the insulator. Fig. 4(c) shows the result of the insulator detection.

III. Results and Evaluation

We evaluate the proposed algorithm on 100 aerial remote sensing insulator images captured by helicopters in Shaoxing, Zhejiang, China, in November 2009. There are five types of insulators in this set of aerial images, as shown in Fig. 5(a). We get our training samples by cutting 100 image patches, each with an insulator in the center, and 200 image patches of background (there are three sizes of the training patches in pixels : 10, 20, and 30). Fig. 5 shows some training samples.

A. Evaluation Criterion

The values of *recall* and *precision* used as the criteria to evaluate the performance of the proposed algorithm, are defined as follows:

$$\text{Precision} = \text{T P}/(\text{T P} + \text{F P}) \quad (10)$$

$$\text{Recall} = \text{T P}/(\text{T P} + \text{F N}) \quad (11)$$

where TP is the number of correctly defined objects, $(\text{T P} + \text{F P})$ is the total number of detected objects, and $(\text{T P} + \text{F N})$ is the total number of actual objects.

B. Detection Result

We compare the proposed algorithm to the method-based on BOW and SIFT (SIFT-BOW), the method-based CSWT [2], the method-based MPEG-7 EHD technique (MPEG7-EHD) [3], and the method based on discriminative training of local gradient-based descriptors and a subsequent voting scheme

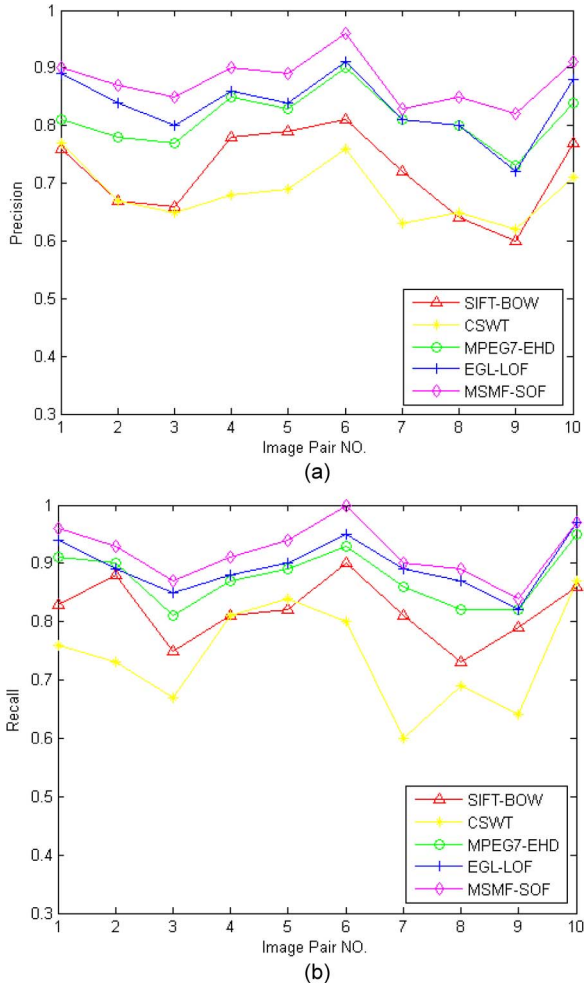

Fig. 6. Performance of five algorithms. (a) Precision. (b) Recall.

Elliptical GLOH-like descriptor and Local Outlier Factor (EGL-LOF) [4]. For convenience, the proposed algorithm is represented as MSMF-SOF. All methods exploit the same amount of training data. The results of the five algorithms are shown in Fig. 6.

Fig. 6 shows that the detection performance of the proposed algorithm is better than the other four algorithms both in recall and precision. Fig. 7 shows two detection results using the proposed algorithm. We can see from Fig. 7 that the proposed algorithm is effective in detecting different types of insulator from the complex background. In addition, by detecting and locating in advance, we can reduce about 80% of the workload than the artificial homework, and the error is within the allowable range.

C. Influence of the Number of Training Images

Here, we evaluate how the proposed algorithm varies in performance when the number of training samples varies. The result is shown in Fig. 8.

In this experiment, both the numbers of positive samples and negative samples increase from 10 to 100, and the number of training images is the only variable. Fig. 8 shows that the performance will be improved by adding more training images.

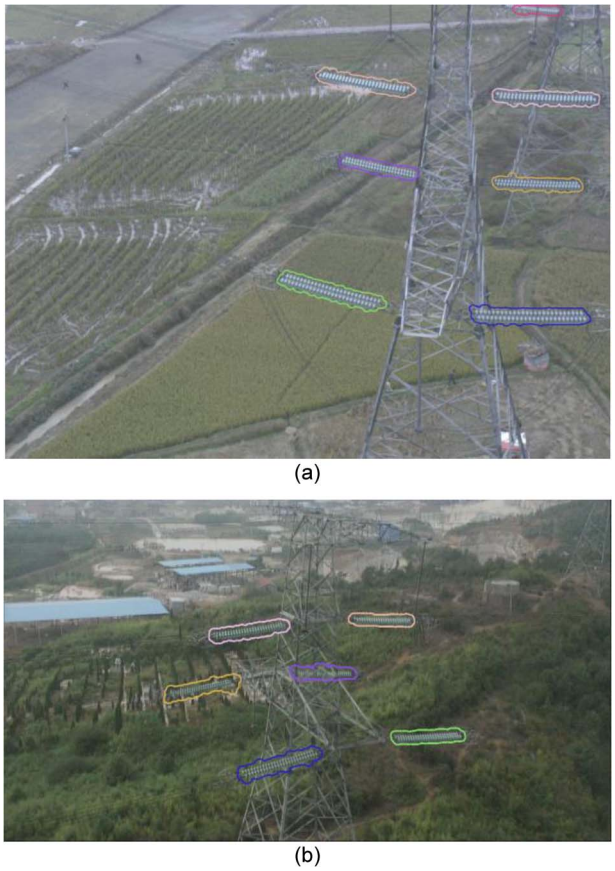

Fig. 7. Detection results on two aerial insulator image. (a) 1600 × 1200 pixels. (b) 1280 × 720 pixels.

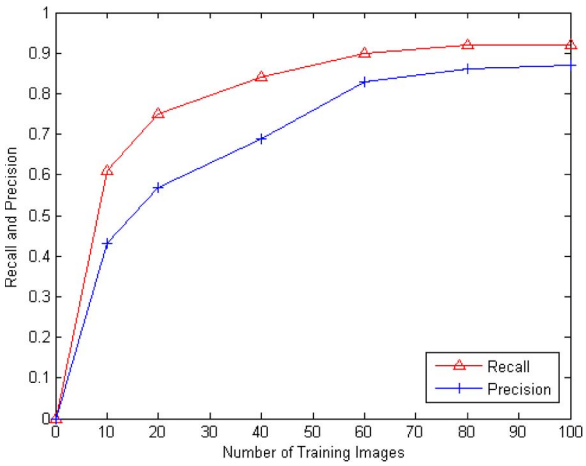

Fig. 8. Recall and precision with different numbers of training images.

The growth of recall and precision begins to slow when the number of training images is close to 30 and is almost stagnant when the number is over 60, which indicates that the number of training samples is well controlled in our algorithm.

IV. Conclusion

We have proposed a robust object detection algorithm based on local features and spatial orders for aerial insulator images. We introduce an MSMF descriptor, which effectively represents the local feature point, and we get several features of spatial orders for improving the robustness of the algorithm. Moreover, a coarse-to-fine strategy is used to improve the efficiency of feature matching. The results of the experiment show that the performance of the proposed method is good, and the results indicate that this method can improve the efficiency of operations in practical applications.

References

[1] Q. G. Wu, J. B. An, and B. Lin, "A texture segmentation algorithm based on PCA and GMAC for aerial insulator images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 5, pp. 1509–1518, Oct. 2012.

[2] J. Zhang and R. Yang, "Insulators recognition for 220 kV/330 kV high-voltage live-line cleaning robot," in *Proc. IEEE ICPR*, 2006, pp. 630–633.

[3] W. G. Li, G. Ye, F. Huang, S. Wang, and W. Z. Chang, "Recognition of insulator based on developed MPEG-7 texture feature," in *Proc. IEEE ICISP*, 2010, pp. 265–268.

[4] M. Oberweger, A. Wendel, and H. Bischof, "Visual recognition and fault detection for power line insulators," in *Proc. 19th CVWW*, 2014, pp. 1–8.

[5] J. G. Zhang, M. Marszałek, and C. Schmid, "Local features and kernels for classification of texture and object categories: A comprehensive study," in *Proc. IEEE CVPRW*, 2006, pp. 13–21.

[6] B. Sırmaçek and C. Ünsalan, "Urban area detection using local feature points and spatial voting," *IEEE Geosci. Remote Sens. Lett*, vol. 7, no. 1, pp. 146–150, Jan. 2010.

[7] L. Weizman and J. Goldberger, "Urban-area segmentation using visual words," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 388–392, Jul. 2009.

[8] S. Agarwal and D. Roth, "Learning a sparse representation for object detection," in *Proc. ECCV*, 2002, vol. 4, pp. 113–130.

[9] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. IEEE CVPR*, 2006, pp. 2169–2178.

[10] H. Sun, X. Sun, H. Wang, Y. Li, and X. J. Li, "Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 1, pp. 109–113, Jan. 2012.

[11] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vis. Conf.*, 1988, pp. 147–151.

[12] D. Lowe, "Object recognition with informative features and linear classification," in *Proc. IEEE ICCV*, 1999, pp. 1150–1157.

[13] F. Bellavia, D. Tegolo, and C. Valenti., "Improving Harris corner selection strategy," *IET Comput. Vis.*, vol. 5, no. 2, pp. 87–96, Jul. 2011.

[14] S. Lazebnik, C. Schmid, and J. Ponce, "Sparse texture representation using affine-invariant neighborhoods," in *Proc. IEEE CVPR*, 2003, pp. 319–324.

[15] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1265–1278, Aug. 2005.

[16] E. Tola, V. Lepetit, and p. Fua, "A fast local descriptor for dense matching," in *Proc. IEEE CVPR*, 2008, pp. 1–8.

[17] Z. X. Liu, J. B. An, and Y. Jing, "A simple and robust feature point matching algorithm based on restricted spatial order constraints for aerial image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 514–526, Feb. 2012.

[18] M. Zhao, B. An, and Y. P. Wu, "Bi-SOGC: A graph matching approach based on bilateral KNN spatial orders around geometric centers for remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1429–1433, Nov. 2013.