Feature Fusion Based Insulator Detection for Aerial Inspection

YAN Tiantian[1,2], YANG Guodong[2], YU Junzhi[2]

1. University of Chinese Academy of Sciences, Beijing 100049

2. State Key Laboratory of Management and Control for Complex Systems, Institute of Automation of, Chinese Academy of Sciences, Beijing 100190

E-mail: yantiantian2015@ia.ac.cn, guodong.yang@ia.ac.cn, junzhi.yu@ia.ac.cn

Abstract: This paper presents a detection method of insulator stings for aerial inspection based on feature-fusion. The local sub-images of insulator strings are firstly collected from aerial videos and tagged to establish a training dataset. The fusion feature is then composed by the histogram of oriented gradients (HOG) feature and local binary pattern (LBP) feature after the principal component analysis (PCA) dimension reduction separately. A training model is developed by SVM algorithm with the fusion feature. At the detection phase, threshold segmentation and morphological operation are adopted to preprocess the images. The sliding window method is then used to search the candidate region and the non-maximum suppression (NMS) method is adopted to fuse the candidate windows. Finally, the position of the insulator strings can be calculated by linear fitting. Both the efficiency and the effectiveness of the proposed method are verified through experiments on locating the multi-angle insulator strings under complex backgrounds.

Key Words: Insulator location, feature-fusion, power line inspection, HOG, LBP

1 Introduction

The insulator plays a critical role in power lines. Because of working for a long-term in the field, insulator strings need frequent inspection to prevent damaging from the sun or the rain. Inspection by helicopters or unmanned aerial vehicles (UAVs) is an advanced alternative, by carrying inspection equipments such as cameras. Currently, the image detecting work is carried out manually without efficiency. For the automatic detection of its damages, the position of the insulator should be recognized first. So the automatic localization of insulator strings is the premise on state monitoring and fault detection for the high voltage power line inspection by helicopters or UAVs.

There are two main kinds of methods to locate the insulator strings. The traditional way usually uses the image segmentation and then detects the targets according to the characteristics of the insulator strings. The edge extraction is the important premise of detecting insulator strings, so Zhao *et al.* proposed a method of aerial insulator image edge extraction based on Nonsubsampled Contourlet transform (NSCT)[1]; Jiang *et al.* achieved the recognition and fault diagnosis of insulator strings by Hough transform to detect the ellipsis[2]; Yu *et al.* firstly preprocessed the image by threshold segmentation and morphological operation, and got the possible region according to the shape feature, finally located the insulator strings based on texture characteristics of gray level co-occurrence matrix (GLCM)[3]; Jing recognized the insulator and detected the fault by the bottom-up method for the perceptual grouping of parallel lines according to the shape structure characteristics of insulator strings [4]; Zhao *et al.* got the insulator position by using threshold segmentation and binary shape feature description [5]; Zhao *et al.* could locate multiple insulators with different angles based on orientation angle detection and binary shape prior knowledge (OAD-BSPK)[6]. All these methods depend on the result of image segmentation, which are very inefficient when the insulator strings are similar with the backgrounds. At the same time, the characteristics are closely related to the angels of the insulators, resulting in lower reliability.

Target detection based on machine learning has been widely used in face detection, vehicle detection, and pedestrian detection. Chen *et al.* achieved human detection based on the fusion feature of the HOG and LBP feature[7]; Qu *et al.* detected pedestrian by integrating HOG with color frequent and skin color feature, the experiments showed the high detection rate by using the fusion feature[8]. Machine learning is developing in insulator detection. Zhai *et al.* optimized the Otsu threshold segmentation algorithm, realized the rough position based on the characteristic of the insulator string skeleton, and finally detected the insulators accurately by using Adaboost classifier[9]. Yang *et al.* preprocessed the images by the gauss filter and morphological operation, located the insulators coarsely, and finally achieved precise detection by machine learning[10]. Applications of machine learning in these methods are mostly combined with the traditional methods, which is ineffective and has certain dependence on the detection results by the traditional methods. Liao *et al.* proposed local features and adopted a coarse-to-fine machine strategy to detect insulators[11]; Liu *et al.* adopted deep learning algorithms, obtained detection model by convolution neural network (CNN), and obtained candidate regions by sliding window method, then fused windows by NMS, finally located insulators through linear fitting[12]. But these methods can not locate the multiple insulators with different angles. Besides, it is hard work to collect these training images while there is no standard insulator dataset. To solve the problems of the existing methods, a method of machine learning based on feature-fusion is proposed, which can locate multiple insulators with different angles.

The rest of the paper is organized as follows. Section 2 briefly introduces the training process. Section 3 describes how to locate the insulator strings by the training model. Section 4 offers the experimental results on the insulator strings with multi-angle and presents the accuracy of location based on the fusion feature. Finally, conclusions are drawn in Section 5.

*This work is supported by National Natural Science Foundation (NSFC) of China (Grant No. 61403374), State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (Grant No. LAPS16006), and National Key Technology Support Program (Grant No. 2015BAF01B01).

2 Model Training Based on Feature-Fusion

The accuracy of training model is of great importance to location results. In this paper, a model training process is designed as shown in Fig. 1. Firstly the local sub-images of insulator strings are collected from aerial videos and rotated randomly, which can locate the insulators with multiple angles. Since it is hard to get a satisfactory classification result with the single feature, a fusion feature combined by HOG and LBP is developed. Besides, the PCA algorithm is adopted to reduce the dimension, which saves the training time. Finally, the SVM classifier is used to train a model for its advantages on the nonlinear classification and small dataset.

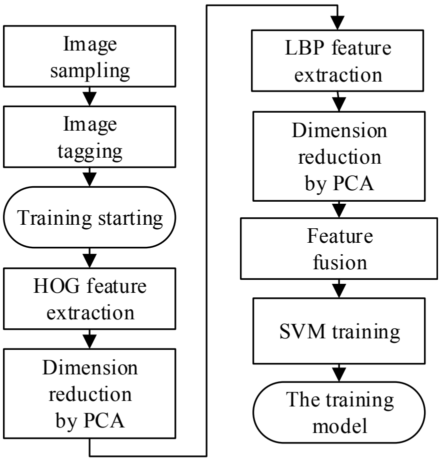

Fig. 1: The model training phase

2.1 Image Sampling and Tagging

Because of changing camera angle, insulators may have different scales, rotations and perspectives. To mark the location of the insulator in the image completely, the label box may be horizontal, vertical or tilted. Besides, the size of the training dataset in machine learning is required fixed. When stretching insulator images with different angles change to a fixed size, deformation will be caused, followed by bad training model. Taking the entire insulator strings with multi-angle as a sample, it will also be difficult to get an accurate training model due to the mass of background information.

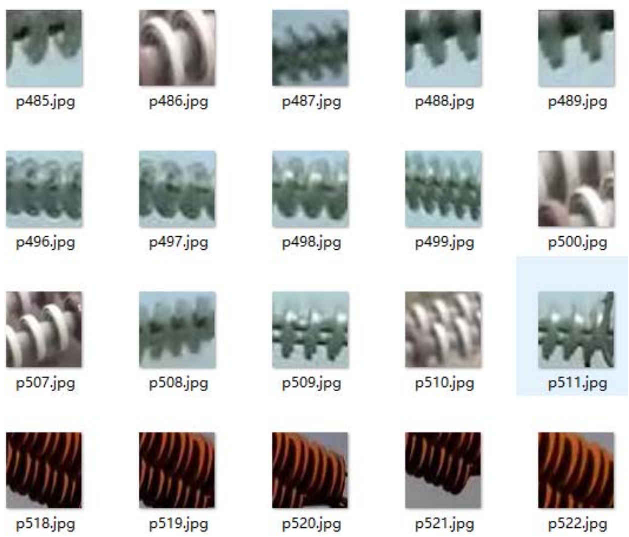

Fig. 2: Positive samples

To solve the above problems, we propose a new method to build the sample dataset. In the view that all the insulator strings are centrosymmetric and stacked by insulators with the same shape, we collect as many sub-images of insulators as possible from aerial videos, and rotate them randomly to form the positive dataset. The positive samples, which are shown in Fig. 2, should contain insulator images with various scales and perspectives to ensure the reliability. Negative samples are such as the pylon, grassland, trees, and so on.

2.2 Feature-Fusion Based on HOG and LBP

There are mainly color features, shape features and texture features in terms of feature description and feature representation. For different kinds of insulators, the shape and texture features are more representative, so we choose the upper two features to form the fusion feature. As one kind of shape feature, HOG feature with the ability of geometrical invariability is not sensitive to the change of illumination, and is suitable for the identification of rigid insulators under complex background[13]. As one kind of texture feature, LBP feature with the gray-scale invariance and the rotational invariance is also suitable. Because there are many miss and wrong detections while using the single feature, an efficient method is to adopt the fusion feature of the HOG and LBP features, which can improve the recognition accuracy.

The extracting process of HOG features includes:

1) Standardize the gamma space to reduce the effects of the partial shadow and illumination change.
2) Calculate the magnitude and direction of each pixel's gradient. Assuming the horizontal gradient is G_x and the vertical gradient is G_y, the magnitude and direction of gradient are as the formula (1) and (2):

$$G = \sqrt{G_x^2 + G_y^2} \tag{1}$$

$$\theta = \tan^{-1}(G_y / G_x) \tag{2}$$

3) Get the intracellular gradient direction histogram. The images are divided into cells of 8×8, where four cells form a block. According to the gradient magnitude, the gradient direction histogram can be calculated by weighted voting. Fig. 3 shows the gradient direction histogram structure;

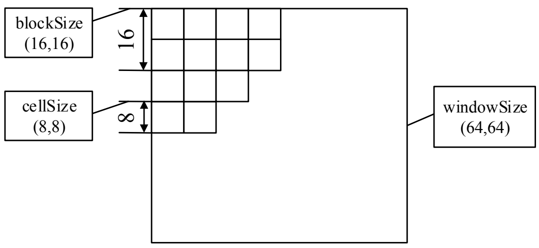

Fig. 3: The gradient direction histogram structure

4) Normalize the gradient direction histogram of each block to reduce further the impact of illumination and shadow;
5) Connect the block features in series to form the HOG feature vector. The dimension of the HOG feature vector can be calculated as follows:

$$D = \left(\frac{S_b}{S_c}\right)^2 \times \frac{(w - S_b + l_b)(h - S_b + l_b)}{l_b^{\,2}} \times D_c \tag{3}$$

In the formula (3), D_c represents the number of gradient direction; S_b is the block size; S_c is the cell size; l_b denotes the step length; w and h are the width and height of the image.

The fundamental of LBP feature is as follows:

In the formula (4), T defined as the texture is the joint distribution of the gray pixels in the neighborhood.

$$T = t\left(g_c, g_0, \ldots, g_{p-1}\right) \tag{4}$$

The sample radius of $g_0,\ldots,g_{P-1}$ is R, which is 10 in this paper. The sample points are the circular symmetry whose number P is 8. In Fig. 4, the neighbor gray values which do not locate in the pixel grid will be calculated by the quadratic linear interpolation

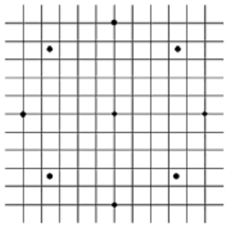

Fig. 4: Each represents two pixels (R=10, P=8)

Calculate the differences of gray values between the center pixel g_c and other points. Assuming that g_c and $g_p - g_c$ are independent, besides $t(g_c)$ does not provide useful information, the LBP values can be written as formula (5).

$$LBP_{(P,R)} = \sum\nolimits_{p=0}^{P-1} s(g_p - g_c)\times 2^P \tag{5}$$

When P=8, gray invariant LBP has 256 kinds of binary modes. While the image is rotating, the relative position of g_c and g_p will change, yielding different LBP values. Rotation invariant LBP are not affected by any angle through rotation invariant mapping, which has 36 kinds of patterns. The formula is as follows:

$$LBP_{P,R}^{ri} = \min\{ROR(LBP_{P,R}, i)\} \tag{6}$$

While calculating LBP feature vectors, the input image is divided into blocks. In the parallel function of parallel_for_, the boundaries of each block are padded according to the sample radius, rotation invariant LBP of each pixel are then calculated, and the LBP feature of the input image can finally be combined by connecting each block feature in serial.

In the paper, the size of our training image is 64×64. While calculating the HOG feature, the cell size is 8×8, the block size is 16×16, the step length is 8×8, and the gradient direction has 9 sections, so the dimension of HOG feature is 1764 according to the formula (3). While calculating the LBP feature, the block size is 16×16, the step length is 8×8, and the rotation invariant LBP has 36 kinds of binary modes, so the dimension of the rotation invariant LBP feature is 1764. Due to the high dimension, data has big correlation, which will waste the training time, thus the PCA algorithm is adopted to reduce the dimension before feature fusing.

At last the main steps of the feature-fusion are summarized as follows:

1) Normalize the LBP and HOG feature of the input image and then reduce the dimension separately by PCA;
2) Connect the LBP and HOG features in series to form the fusion feature;
3) Normalize the fusion-feature.

In this paper, the dimensions of LBP and HOG feature are reduced to 15, therefore the fusion-feature has the dimension of 30.

2.3 Classification Based on SVM

There are many classification algorithms, such as SVM classifier, AdaBoost, k neighbor, etc. The SVM algorithm is used because the number of the training dataset is small and there are only insulator and background two kinds of samples. Besides, SVM algorithm is suitable for both the linear and nonlinear classification.

For linear separable samples, there has a hyperplane $w^T x + b = 0$ to separate the samples completely.

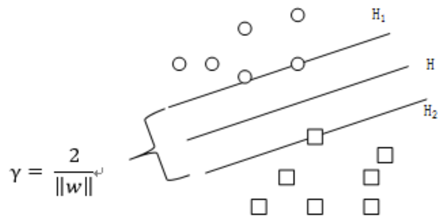

Fig. 5: SVM linear classification

Circles and squares represent two kinds of samples respectively in Fig. 5. By finding the parameters w and b which are calculated by formula (14), the hyperplane $H: w^T x + b = 0$, the straight $H_1: w^T x + b = 1$, and the straight $H_2: w^T x + b = -1$ can be found, which are parallel to each other.

$$\min_{w,b} \frac{1}{2} w^2 \ \text{s.t.}\ y_i\left(W^T x_i + b\right) \geq 1,\ i = 1,2,3,\ldots,m. \tag{7}$$

The optimal classification hyperplane can allow the fault samples. At this time the largest interval is equivalent to calculating the minimum of $\mathrm{F}(\mathrm{w},\xi) = \frac{1}{2} w^2 + C\left[\sum_{i=1}^{n} \xi_i\right]$, where C is the penalty factor.

For nonlinear classification, SVM will find the optimal separating hyperplane by using nonlinear transformation which can map the input variables with low dimension into high-dimensional feature space. Solving parameters w and b, will involve calculating the $\varphi(x_i)^T \varphi(x_j)$, which needs to calculate the inner product of sample x_i and x_j in high dimension. Because of the difficulty in calculating the inner product, the kernel function is introduced whose expression is as follows

$$\kappa\left(x_i, x_j\right) = \varphi\left(x_i\right)^T \varphi\left(x_j\right) \tag{8}$$

There are mainly linear kernel, polynomial kernel, Gaussian kernel (RBF kernel), Laplace kernel, Sigmoid kernel and so on. This method uses the RBF kernel, whose function is as follows:

$$\mathrm{K}\left(x_i, x_j\right) = \exp\left(-\gamma x_i - x_j^{\ 2}\right), \gamma > 0 \tag{9}$$

In training process, 700 local sub insulator images are as the positive samples and 2100 background images are as the negative samples. The performance of classification is mainly related to the C and γ , so the SVM parameters are optimized by using the K cross-validation. The optimal parameters based on HOG, LBP and fusion feature are close to each other, therefore C and γ are set to 4.0 and 1.0 uniformly. The test dataset of classification includes 300 positive images and 900 negative images. The results shown in Table 1 indicate that the classification accuracy based on fusion feature is better than that using the single feature.

Table 1: The classification accuracy of Hog, LBP, HOG+LBP

Feature	Accuracy
HOG	75.3%
LBP	76.9%
HOG + LBP	83.5%

3 Insulator Detection

For locating the insulator strings, the candidate sub-windows need to be extracted from the input image, and the training model is then used to judge whether the candidate windows contain insulators or not. If there contains insulator, the sub-windows will be saved. The redundant windows will then be removed by the window-fusing algorithm. The detection process is as shown in Fig. 6. The sliding window method is used to get the candidate sub-windows, which is a simple way. Before getting the candidate windows, the threshold segmentation and morphological operation are firstly used to preprocess the image, which can reduce the number of sub-windows that need to do feature extraction and prediction, which can save the detection time. Then, the model trained by SVM classifier is used to predict the candidate windows and the NMS algorithm is used to fuse the windows. Because of the local sub-images of insulators located firstly, the linear fitting is finally conducted to merge windows and the final location of the insulator strings can be marked.

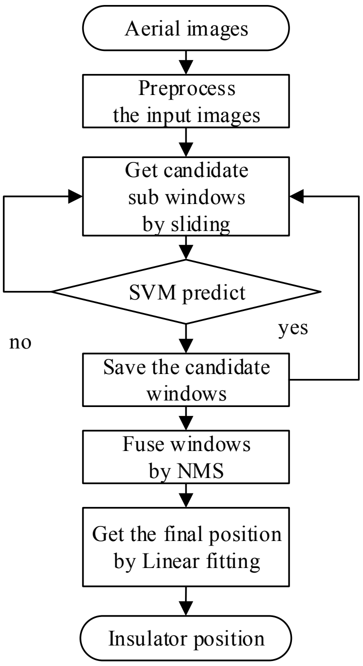

Fig. 6: The detection phase

3.1 Image Preprocess

There have two main advantages of preprocessing. The one is to save detecting time. When the candidate windows are obtained and all the values of the sub windows are equal to zero, it is unnecessary to implement feature extraction and prediction. While the images have many elements similar to sky or thin wires, this method is more effective. However, when the backgrounds are all elements such as grassland or pylons, this method does not work. It is necessary to preprocess the images firstly, which is less time-consuming and has no effect on the detection speed. Preprocessing can help find the sub-images which contain the elements such as the sky or thin wires, so it is unnecessary to collect these negative images. Thus the other advantage is that it can reduce the number of negative samples, reduce the complexity of the training model, speed up the classification, and at the same time improve the classification accuracy even though the glass insulators are similar to the sky.

In the preprocessing stage, the main steps are as follows: apply the Otsu threshold segmentation, make the morphological close operation, and get the mask image after reverse operation, finally multiply the mask image by the original image to get the final image to be detected. The preprocessing results are shown in Fig. 7.

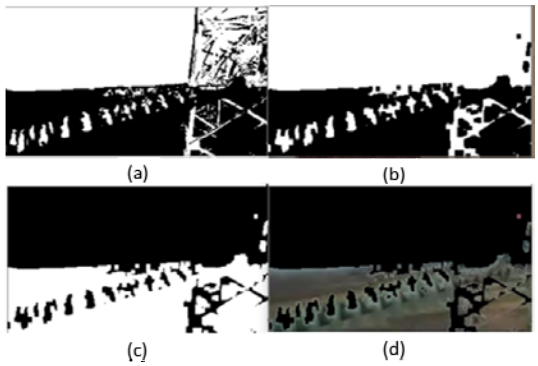

Fig. 7: Image preprocessing. (a) Otsu segmentation. (b) Morphological close operation. (c) Mask image by reverse operation. (d) Image to be detected

The threshold segmentation results in Fig. 7(a) show that, it is difficult to separate the insulators completely from the complex backgrounds by segmentation technology, and difficult to extract the edge or shape characteristics such as the ellipse characteristics, let alone detect the insulator strings when backgrounds and insulators are similar to each other. In this paper, the threshold segmentation and morphological technology are only to reduce the number of candidate sub-windows to be detected, and the whole detection process of insulator strings do not excessively rely on the results of segmentation.

3.2 Candidate Windows and Prediction

There are many methods to get the candidate windows, such as the selective search algorithm and the sliding window method. Because the candidate windows will contain the whole multi-angle insulators with much background information by the selective search method, which has conflicts with our training model by using the local sub images of insulator strings. Therefore, the sliding window method is adopted which is also a simple way. In the experiment, sliding windows with five scales are used because of the uncertain scales of insulators.

After saving the candidate sub-windows, the function of parallel_for_ is then used, which can get the classification probability of each candidate windows in parallel. The sub-windows, whose prediction probability is greater than the threshold, are saved to a global vector and then used to fuse. Owing to the dual-core CPU, the speed of prediction is increased about 1.5 times.

3.3 Image Fusion Based on NMS

The non-maximum suppression (NMS) is an algorithm which can search the local maximum values and restrain the non-maximum elements. Based on the classification probabilities of the candidate windows, the sub-windows with higher probability can be retained and the rests with lower probability are removed, which can reduce the number of final candidate windows to reduce the computational complexity of the window merging.

Its main steps are as follows: sort and traverse the sub-windows according to the possibilities if they contain insulators; retain the windows with bigger probability and remove the ones with smaller probability, where the overlaps between them are greater than the threshold.

The results of NMS algorithm are shown in Fig. 8, where the number of candidate sub-windows is 25 before NMS, and changes to 11 after the window fusion.

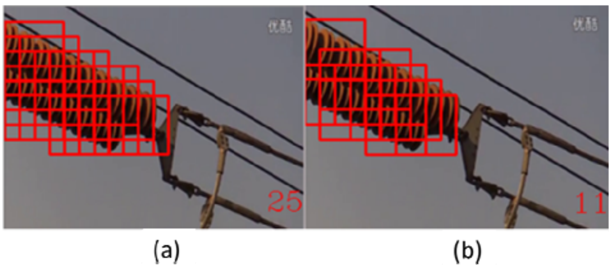

Fig. 8: (a) Before NMS. (b) After NMS

3.4 Image Mergence

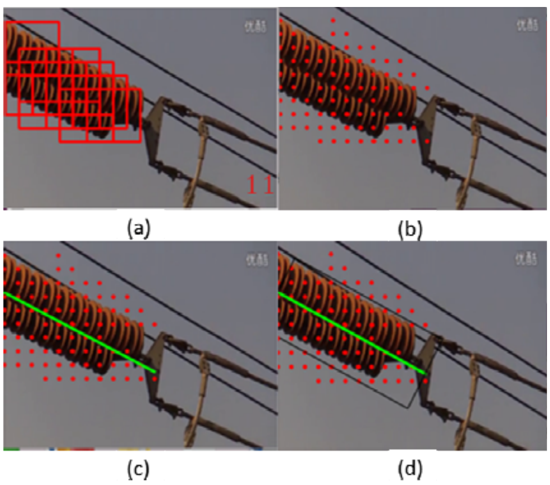

Fig. 9: Image mergence by linear fitting. (a) Final candidate windows after NMS. (b) The set of vertices. (c) The center line. (d)The location of the insulator

Finally, the final candidate sub-windows after NMS need to be merged to get the final location of insulator strings. The main steps are as follows:

1) Group the final candidate sub-windows. Adjacent windows are divided into the same group if having overlaps with each other. There may be many groups because the number of insulator strings is more than one in an image;
2) Storage the vertex coordinates of every sub-windows in the same group into a vector, while the number of windows in a group is greater than 1. The number of coordinate vectors is the same with groups';
3) Get the center line of each group by using linear fitting based on least squares method;
4) Do statistics for the distances of all points to the center line within the same group, and calculate the average and maximum of the distances, and then combine them linearly to get the width of the insulators;
5) Mark the location boxes in the image according to the center line and the width of insulators.

The results of window mergence are shown in Fig. 9.

4 Experimental Results

To verify the effectiveness of the proposed method in this paper, we locate insulators from three aerial videos of high voltage transmission lines. From the Fig. 10, it can be observed that three types of insulators (including glass insulators) with different angles are located accurately. For the insulators that are hard to segment and extract characteristics, this method can still detect them.

Fig. 10: The insulator detection

In the detection test, we choose 500 images from three videos, 400 of them contain insulators, others are absolute backgrounds. There are 430 insulators in total because of some images containing more than one insulator. For those insulators that are too far or too close in the image and their characteristics are barely distinguishable, they will be excluded as miss detections.

The final detection results are shown in Table 2. From the results, we can see that our fusion feature has better performance on detection accuracy, almost 90%, in comparison with the single feature based on HOG or LBP.

Table 2: The detection results of Hog, LBP, HOG+LBP

Feature	HOG	LBP	HOG + LBP
Right detection	366	352	383
Right detection rate	85.1%	81.8%	89.1%
Wrong detection	34	13	9
Wrong detection rate	7.9%	3.0%	2.1%
Miss detection	64	78	47
Miss detection rate	14.9%	18.1%	10.9%

The performance of the localization mainly depends on the accuracy of the classification, as shown in Fig. 11 giving some miss and wrong recognition results. Miss recognition is mainly because the HOG and LBP features has not rotation invariance. Besides samples can not contain insulator images with every angle and scale, resulting in bad influences on recognition results. Wrong recognition is because some backgrounds and insulators are similar on the shape and texture which are difficult to distinguish. So further improvement in classification accuracy is needed.

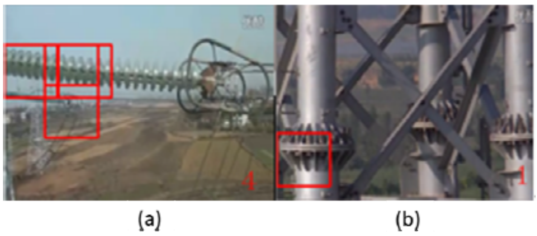

Fig. 11: (a) Miss recognition. (b) Wrong recognition

In addition, we test the running time of this method. For the images including some elements such as the sky, the detection time is about 800 ms because of the preprocessing stage. For the images with complex backgrounds, this method need to predict all candidate sub-windows, which is more time-consuming, about 2 s. The most time-consuming parts in the algorithm to predict a candidate window are HOG, LBP feature extraction, and the SVM classification. The LBP feature extraction consumes most among them and needs to be improved.

Table 3: The consuming time

Detecting an image	
No preprocessing	2 s
Preprocessing	800 ms~2 s
Main parts while predicting a sub-window	
HOG	5 ms
LBP	10 ms
SVM	5 ms

5 Conclusion

In this paper, a detection method of insulator stings for aerial inspection has been developed based on feature-fusion. In the training stage, the local sub images of insulator strings are collected and tagged. The fusion feature is then combined by the HOG feature and LBP feature after PCA dimension reduction separately. The SVM classification algorithm is adopted to get a training model. At the detection phase, threshold segmentation and morphological operation are used to preprocess the images, which can speed up the detection process. The candidate sub windows are extracted by the sliding window method, which are then predicted by SVM in parallel. The NMS algorithm is then used to fuse the candidate windows. The final positions of the insulator strings are finally obtained by the linear fitting. The results of the experiment indicate that the proposed method can locate multi-angle insulators under the complex backgrounds, and that the detection accuracy can be guaranteed satisfactorily.

References

[1] Z. Zhao, S Jin, Y. Liu. Aerial insulator image edge extraction method based on NSCT [J]. *Chinese Journal of Scientific Instrument*, 2012, 33(9): 2045-2052.

[2] H.R. Jiang, L.J. Jin, S.J. Yan. Recognition and fault diagnosis of insulator string in aerial images [J]. *Mechanical & Electrical Engineering Magazine*, 2015, 32(2): 274-278.

[3] L.Y Yu, B. Yao, W.H. Wu, H.L. Liu, W.Q. Tan. Insulator identification method based on multi-feature [J]. *Insulators and Surge Arresters*, 2016(3): 79-83.

[4] J. Zhang, J. Han, Y. Zhao, *et al.* Insulator recognition and defects detection based on shape perceptual [J]. J*ournal of Image and Graphics*, 2014, 19(8): 1194-1201.

[5] Z. Zhao, W. Le. Aerial insulator string image automatic location method [J]. *Chinese Journal of Scientific Instrument*, 2014, 35(3): 558-565.

[6] Z. Zhao, N. Liu, L. Wang. Localization of multiple insulators by orientation angle detection and binary shape prior knowledge [J]. *IEEE Transactions on Dielectrics and Electrical Insulation*, 2015, 22(6): 3421-3428.

[7] R. Chen, M. Wang, X. Chen. Human detection based on HOG-PCA and LBP characteristics [J]. *Information Technology*, 2015, 2: 101-105.

[8] Qu Y.Y, Liu Q, Guo J M, *et al*. HOG and color based pedestrian detection [J]. *Journal of Wuhan University of Technology*, 2011, 33(4): 134-138.

[9] Y.J. Zhai, D. Wang, Y. Wu, H.Y. Cheng. Two-stage recognition method of aerial insulator images based on skeleton extraction[J]. *Journal of North China Electric Power University*, 2005, 42(3): 105-110.

[10] W. Yang, C. Li, S.L. Yang, H. Zhou. Study on aerial image positioning of wide glass insulator [J]. *Electric Test*, 2016, (15):23-26.

[11] S. Liao, J. An. A robust insulator detection algorithm based on local features and spatial orders for aerial images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(5): 963-967.

[12] Y. Liu, J. Yong, L. Liu, *et al*. The method of insulator recognition based on deep learning [C]. *International Conference on Applied Robotics for the Power Industry*. 2016: 1-5.

[13] L. Y. Insulator location and recognition algorithm based on HOG characteristics and SVMD [J]. *Journal of Transportation Engineering and Information*, 2015, 13(4):53-60.