COMPUTER VISION-BASED DETECTION AND STATE RECOGNITION FOR DISCONNECTING SWITCH IN SUBSTATION AUTOMATION

Hongkai Chen,* Xiaoguang Zhao,* Min Tan,* and Shiyong Sun*

Abstract

State recognition in disconnecting switches is important during substation automation. Here, an effective computer vision-based automatic detection and state recognition method for disconnecting switches is proposed. Taking advantage of some important prior knowledge about a disconnecting switch, the method is designed using two important features of the fixed-contact facet of such disconnecting switches. First, the Histograms of Oriented Gradients (HOG) of the fixed-contact are used to design a Linear Discriminant Analysis (LDA) target detector to position the disconnecting switches and distinguish their loci against a usual cluttered background. Then a discriminative Norm Gradient Field (NGF) feature is used to train the Support Vector Machine (SVM) state classifier to discriminate disconnecting switch states. Finally, experimental results, compared with other methods, demonstrate that the proposed method is effective and achieves a low miss rate while delivering high performance in both precision and recall rate. In addition, the adopted approach is efficient and has the potential to work in practical substation automation scenarios.

Key Words

Computer vision, substation automation, disconnecting switch, state recognition, histograms of oriented gradients, norm gradient field

1. Introduction

The disconnecting switch (also referred to as a disconnector or isolator) is one of the most important items of electrical equipment in a substation. A disconnector comprises five parts (see Table 1 and Fig. 1). The circuit disconnection point can be found by driving the operating mechanism to form its open and closed states. Although the structure and operational principle of a disconnecting switch is relatively simple, it plays a critical role in substation design, construction, and safe operation. For example, it is used in switching circuits to change the operating mode of power systems and is used to disconnect high-voltage (HV) maintenance equipment from charged equipment to ensure safe maintenance works.

In HV, or extra-high-voltage (EHV) substations, switching of electrical equipment usually needs to be performed when the operating mode of connected power systems, or the state of any item(s) of electrical equipment, needs to be changed. Although switching operations are being changed to remote automation, operatives have to go to working sites to confirm whether the state of any disconnecting switches is correct. Only after a state confirmation message from a disconnecting switch has been received by an operator in the control room can subsequent operations be executed. This kind of mode of operation has the disadvantages of being labour-intensive and slow. Therefore, automatic state recognition for disconnecting switches has become a pressing need in substation automation.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conductive part</td>
<td>Includes: moving-contact, fixed-contact, disconnecting link, and connector base; its main role is to carry current and change the state of the circuit</td>
</tr>
<tr>
<td>Operating mechanism</td>
<td>To provide energy to the disconnecting switch by manual, electric, pneumatic action, etc.</td>
</tr>
<tr>
<td>Transmission mechanism</td>
<td>Includes: crank arm, link rod, and the operating insulator</td>
</tr>
<tr>
<td>Insulating part</td>
<td>Includes: support insulator and operating insulator</td>
</tr>
<tr>
<td>Support pedestal</td>
<td>Connects conductive parts, operating mechanism, transmission mechanism, and insulating parts</td>
</tr>
</tbody>
</table>
It benefits the development of one-touch sequence control [1] which refers to substation automation systems which can perform sequential assigned tasks in a so-called smart substation, provided that certain previous operations have been completed and the corresponding switch states have been confirmed. One-touch sequence control includes multiple control steps where the state recognition of a disconnecting switch might be automatically recognized by a computer centre. Automatic state recognition also provides a verification method, and technical support, for unattended substation operations managers.

Here, the focus is on the automatic state recognition problem of a disconnecting switch which is widely used in substations. Based on computer vision, an effective detection and state recognition method for disconnecting switches was proposed. By analysis of the prior knowledge of an image (Fig. 1), it was found that the disconnector had the following characteristics:

1. The switch, and its fixed-contact, have a one-to-one relationship, that is to say we can position the disconnecting switch as long as we detect a fixed-contact in an image.

2. Compared with the whole body of the disconnecting switch, the shape characteristic of the fixed-contact is more stable and less noisy.

3. The fixed-contact is fixed to the busbar and exhibits bilateral symmetry. Moreover, the shape of a fixed-contact image is distinctive.

Based on the above findings, two important features of the fixed-contact were used to design the computer vision-based detection and state recognition method. The first feature was the histograms of the oriented gradients (HOG) of the fixed-contact of the disconnector. Combined with an LDA, the LDA detector was designed such that the high-dimensional HOG feature was projected onto a one-dimensional space, to generate candidates. The fixed-contact had a relatively stable shape characteristic and was readily distinguished from non-disconnector samples, so the second feature used was a proposed NGF that was used to discriminate different states of the switch. Compared with some other methods, experimental results showed that the proposed method was effective, efficient, and had a low miss rate while delivering high-performance precision and recall rate.

The main contributions of this work are:

1. An automatic detection and state recognition method is proposed for a disconnecting switch for substation automation, which makes use of shape information for higher performance.

2. The state recognition problem of disconnector is transformed into the issue that requires only the recognition of the state of the fixed-contact of the disconnector.

3. A detection-to-recognition scheme is adopted to recognize the status of a disconnecting switch where it is first proposed to use HOG for detecting the fixed-contact and a new discriminative feature (the NGF) is proposed for classification.

The structure of the rest of this manuscript is as follows. An introduction to related work is given in Section 2. Section 3 presents an overview of the proposed method. Sections 4 and 5 describe the proposed detection and state recognition method for a disconnecting switch. Section 6 summarizes the experimental results. Conclusions, with recommendations for future research, are given in Section 7.

2. Related Work

Using computer vision methods for substation automation tasks has many advantages. For example, one can make use of the powerful data processing capabilities of modern computers for rapid analysis of massive data in video images, then remove irrelevant information, leaving only essential information for operatives. It can not only reduce labour demands, but also improve the ability to deal with unexpected problems timeously. Hence, computer vision methods used for substation automation tasks are becoming increasingly popular, such as: electrical equipment recognition [2], [3], state monitoring for electrical equipment [4]–[8], security monitoring tasks [9], [10], etc.

Zhou and Zhao [2] presented an image recognition method for substation equipment based on SIFT feature matching: they proposed the use of the RANSAC algorithm and the relative distance ratio to perform SIFT feature matching between the equipment, and template, images. To promote the development of one-touch sequence control techniques, Wang et al. [3] proposed an image recognition-based state recognition method for an isolator and breaker. Their approach is based on the SIFT feature matching algorithm, the KNN algorithm, and the Hough algorithm. The purpose of their research is to use a robot to identify the state of the isolator and breaker when undertaking a switching operation in a substation. Zhao et al. [4] proposed an image processing-based monitoring method to ascertain the running state of HV power
equipment. In [5], Rahmani et al. proposed an electrical equipment fault detection method using the Zernike moment feature of a thermal image and a SVM. Their simulation experiments were done on databases containing real images of the distribution networks in the North West of Tehran and obtained an acceptable accuracy of 83%. Wang et al. [6] proposed a state recognition method for substation HV line circuit breakers based on a shape-prior active contour model which combines shape information with a coefficient of variance model. They achieve good accuracy and limited applicability, however there remain some shortcomings: e.g., if the initial curve deviates far from the target position, the evolution curve cannot converge to the correct solution when initializing on a small target contour. In [7], Liu and Meng developed an image-based state recognition approach: they propose extraction of texture features using a Gabor transformation, then the state of the isolator is classified by the SVM. Reddy et al. [8] proposed a state monitoring approach for insulators using K-means clustering, a discrete orthogonal S-transform (DOST), and an adaptive neuro-fuzzy inference (ANFI) system. Their method was incorporated into a distribution automated system. Using image and video analysis techniques, Yang et al. [9] established an intelligent monitoring system for electrical equipment. Based on the signboard image recognition method, Cai et al. [10] constructed a novel anti-misoperation system to prevent man-made misoperation. To maintain substation operational safety, they argue that the proposed system can be combined with an active five-prevention system.

3. Overview of the Proposed Method

The images acquired in substation (e.g., Fig. 1) contain some particular prior knowledge: the image background is complicated; different items of electrical equipment have similar colours and weak textural features; the morphological features were more obvious and stable.

As mentioned, a disconnecting switch, and its fixed-contact had a one-to-one relationship: the shape characteristic of the fixed-contact was relatively stable and less noisy (in image clarity terms) than that of the disconnector. Moreover, the fixed-contact affixed to the busbar exhibited bilateral symmetry. Based on these findings, we transformed the state recognition problem of a disconnector into one where we only needed to recognize the status of the fixed-contact part of the disconnector. Then, what was required was to discriminate fixed-contacts from other background image features, and determine their status.

Here, the proposed method can be sub-divided into two stages (see Fig. 2). Specifically, the proposed method is as follows:

1) Detection: the HOG for the fixed-contact and an LDA are used to classify input sub-images, and discover where the disconnector regions are.

2) State recognition: the proposed NGF of a fixed-contact and Gaussian kernel SVM are used to discriminate between disconnector states.

4. Detection Method

4.1 Fixed-Contact Detection with HOG Feature Extraction

The slide window technique was exploited for fixed-contact detection amidst the massive dataset of non-fixed-contact sub-images. Hence, what was needed was a robust feature capable of distinguishing fixed-contact sub-images while being provided with invariance to some given image variations. For the fixed-contact of a disconnector, the shape feature was the most prominent compared to other features (colour, texture, etc.). HOG [11] is a type of robust local region descriptor in which the core idea is that the appearance and complex shapes of one object can be described by its statistical distribution of gradient directions. It can describe object edges and is insensitive to illumination changes and small offsets.

For one image sample region, its gradient image was first built by calculating the image difference. Then the gradient image was divided into several small connected regions (cells): 1-D HOG was computed for each cell. Furthermore, the block-based histograms were normalized to reduce the variance therein. Finally, those normalized block-based histograms were combined sequentially to form the final feature vector, namely the HOG feature. Specifically, the main steps of HOG feature extraction for fixed-contact image \( I \) (measuring \( 64 \times 64 \) pixels) are as follows:

Step 1. Colour space transform: convert input colour image to greyscale image.
**Figure 3.** HOG feature generation flowchart.

**Step 2.** Global image normalization: normalize the colour space using a gamma-correction method.

**Step 3.** Gradient image generation: the grey value of each pixel is $I(x,y)$. The corresponding gradient magnitude and orientation are $M(x,y)$ and $\Theta(x,y)$, respectively. The gradients of the $x$- and $y$-axes are respectively denoted by $I_x(x,y)$ and $I_y(x,y)$.

$$I_x(x,y) = \frac{I(x+1,y) - I(x-1,y)}{2}$$  
(1)

$$I_y(x,y) = \frac{I(x,y+1) - I(x,y-1)}{2}$$  
(2)

$$M(x,y) = \sqrt{I_x(x,y)^2 + I_y(x,y)^2}$$  
(3)

$$\Theta(x,y) = \arctan\left[\frac{I_y(x,y)}{I_x(x,y)}\right]$$  
(4)

**Step 4.** Divide the gradient image into some small non-overlapping connected cells (each cell contains $8 \times 8$ pixels). The orientation range ($0^\circ$ to $180^\circ$) is divided into nine bins and 1-D HOG are constructed $H_{cell}^i$ for each cell, where $H_{cell}^i \in \mathbb{R}^{9 \times 1}$ and $i = 1, 2, \ldots, 64$.

**Step 5.** Generate the histogram $H_{block}^j$ of a block over cells in any overlapping square block consisting of four as:

$$H_{block}^j = \begin{bmatrix} H_{cell}^1 \; H_{cell}^2 \; \cdots \; H_{cell}^{64} \end{bmatrix}^T$$  
(5)

where $H_{block}^j \in \mathbb{R}^{36 \times 1}$ and $j = 1, 2, \ldots, 36$. Then the histogram of each block is normalized by L2-norm, with a hysteresis threshold, to give:

$$\overrightarrow{NH_{block}^j} = \frac{H_{block}^j}{\sqrt{\|H_{block}^j\|_2^2 + \varepsilon^2}}$$  
(6)

where $\varepsilon$ is a small constant used to avoid division by zero.

**Step 6.** The final HOG can be built by integrating all normalized block-based histograms as:

$$\overrightarrow{hog} = \left[ \begin{array}{c} NH_{block}^1 \; NH_{block}^2 \; \cdots \; NH_{block}^{49} \end{array} \right]^T$$  
(7)

where $\overrightarrow{hog}$ is the final HOG feature descriptor and $\overrightarrow{hog} \in \mathbb{R}^{1764 \times 1}$. Figure 3 shows the flowchart through HOG feature generation.

### 4.2 Linear Discriminant Analysis

LDA [12] is an effective method for feature extraction and dimensionality reduction. The core idea of LDA is to extract the most discriminative low-dimensional features from the original high-dimensional feature space and simultaneously obtain the smallest within-class variance and the largest between-class variance. The within-class scatter matrix $S_W$ and the between-class scatter matrix $S_B$ are respectively defined as follows:

$$S_W = \sum_{k=1}^{K} \sum_{x_i \in \Omega_k} (x_i - \mu_k)(x_i - \mu_k)^T$$  
(8)

$$S_B = \sum_{k=1}^{K} n_k (\mu_k - \mu)(\mu_k - \mu)^T$$  
(9)
where $K$ is the number of classes of the dataset and $\Omega_k$ denotes the set of $k$th category, $x_i$ is the $i$th sample within the $k$th category, $\mu_k$ is the centre of the $k$th cluster, $\mu$ is the mean of all data, and $n_k$ is the sample size of the $k$th category.

The optimal projection matrix $W_{opt}$ is found by solving the optimization problem, as follows:

$$W_{opt} = \arg \max_{W} \left[ \frac{W^T S_B W}{W^T S_W W} \right] = \left[ w_1, w_2, \ldots, w_n \right]$$

where $W$ is the projection matrix formed by generalized eigenvector $w_i$ and the optimization problem ($i.e.,$ (10)) can be solved using eigenvalue decomposition:

$$S_B w_i = \lambda S_W w_i$$

In LDA, there are at most $K - 1$ non-zero generalized eigenvalues because of the rank of $S_B$ \( \text{rank}(S_B) \leq K - 1 \). Thus, a training sample $x \in \mathbb{R}^{d \times 1}$ is mapped to $y \in \mathbb{R}^{K - 1}$ using the transformation $y = W^T x$ where $K - 1 \ll d$; more details can be found elsewhere [13].

The detection part of a disconnecting switch can be deemed to be a binary classification problem (disconnector $versus$ non-disconnector, $i.e.,$ $K = 2$), therefore the goal is to find a straight line where the samples in HOG feature space are projected to achieve best separation performance in 1-D space (see Fig. 4): here, LDA is used for its efficiency and accuracy to detect the fixed-contact by using HOG.

Finally, the classification label $c$ of unknown input sample $x$ can be categorized by the following decision rule:

$$c = \arg \min_k \| W^T x - \bar{\mu}_k \|$$

where $k \in \{1, 2\}$, $\bar{\mu}_1 = W^T \mu_1 + \sigma$ and $\bar{\mu}_2 = W^T \mu_2 - \sigma$, and $\sigma$ is a decision threshold.

5. State Recognition Method

The aforementioned detection method is used to locate the fixed-contact of a disconnecting switch in an image. Then, we need to recognize the status thereof. In this section, status (open or closed) recognition is posed as a binary classification problem. To make full use of the prior characteristic of bilateral symmetry, the proposed NGF feature is first extracted from the fixed-contact image dataset with two kinds of state (open and closed). Then the trained classifier is used to determine the final switch state.

5.1 State Recognition with NGF Feature Extraction

The 1-D gradient field (GF) is a graphical spatial distribution computed by the image gradient summing all columns or lines. The spatial distribution and variation of the GF can be given as:

$$GF(y) = \frac{1}{H} \sum_{x=1}^{H} |M(x, y)|$$

where $M(x, y)$ is the image gradient, $(x, y)$ denote the pixel coordinates in a gradient image, $\gamma > 0$ is a normalization constant, $x = 1, 2, \ldots, H$ and $y = 1, 2, \ldots, W$. $H$ and $W$ are the image size.

GF is a feature that is computed by summing all lines. From the GF of each fixed-contact image (the middle line of Fig. 5(a) and 5(b)), it can be seen that there is a “double peak saddle” type distribution curve at both ends of the GF. There is a “single peak uplift” type distribution curve at the middle of the GF when the disconnecting switch is closed, while no “single peak uplift” type distribution curve is seen in the middle of the GF when it is open. However, GF discrimination may not be good enough as is sensitive to noise. $E.g.,$ Fig. 5(a)-(1) and Fig. 5(b)-(1) are pure ones. Fig. 5(a)-(5) and Fig. 5(b)-(5) are the corresponding GF features. Once the fixed-contact encounters background noise, the corresponding GF features are different from those in Fig. 5(a)-(5) or Fig. 5(b)-(5). Also, if an open state fixed-contact encounters background noise (such as that in Fig. 5(a)-(4)), it will have a feature distribution that is similar to that pertaining to its closed state.

The fixed-contact was a rigid structure. To some extent, it was symmetrical regardless of status. It was hoped that the symmetrical characteristic could be described by a type of discriminative feature, which can result in accurate classification when combined with a linear classifier.
Figure 5. Several examples of GF, and NGF, feature generation for open state: (a) and closed state; (b) The first row of (a) denotes an open fixed-contact while the first row of (b) denotes a closed state. The second row and last row respectively denote GF, and NGF, features in both (a) and (b).

(e.g., SVM). To this end, we design the following kernel function:

\[ H(x) = \begin{cases} 
- \left( \frac{x}{h} \right)^p + 1, & -l \leq x \leq l \\
0, & \text{others} 
\end{cases} \quad (14) \]

where \( l > 0 \) affects the range of variable \( x \), \( p > 0 \) is a control factor. The given constant \( h > 0 \) is the bandwidth of kernel function \( H(x) \) (in this work we empirically set \( p = 4 \) and \( h = 6 \)). The NGF can be formulated by

\[ \text{NGF}(x) = \frac{1}{\eta} \int_{-\infty}^{+\infty} GF(\tau) H(x - \tau) d\tau \quad (15) \]

where \( \eta > 0 \) is normalization constant. This paper sets \( l = 64 \). Hence we get a kernel vector \( H(x) \in \mathbb{R}^{129 \times 1} \). Then the convolution operation between \( GF(x) \) and \( H(x) \) will produce a result \( \text{NGF}(x) \) with \( 64 + 129 - 1 = 192 \) dimensions. The result vector \( \text{NGF}(x) \) has many zero elements. We plot a range of \( \text{NGF}(x) \) in the third row of Fig. 5.

Intuitively, Fig. 5(a) shows that there is a “triple peak” type distribution in the NGF when the disconnecting switch is closed state while a “double peak” type distribution arises in the NGF when it is open. Using the triple peak feature to recognize the state of a fixed-contact seems to work. However, the triple peak distribution may be weak when a fixed-contact encounters background noise (Fig. 5(a)-(12) or 5(b)-(12)). In addition, using the triple peak feature to recognize the status of a fixed-contact is a threshold-judging method, and is sensitive to noise. On the contrary, we use a statistical learning theory-based method (i.e., SVM) to learn a hyper-plane to recognize the state of a fixed-contact.

5.2 Support Vector Machine Method

In our state recognition part, SVM [14], [15] was used as the classifier for its accuracy. In an SVM learning strategy, the training data were mapped into a high-dimensional feature space, where data samples with different categories were separated by an optimal hyperplane.

An SVM will construct a binary classifier based on the given labelled training dataset \( \{(x_i, y_i) | i = 1, 2, \ldots, n\} \), where \( x_i \in \mathbb{R}^{m \times 1} \) and \( y_i \in \{-1, 1\} \). Now the purpose is to determine such a hyperplane function \( f : \mathbb{R}^{m \times 1} \rightarrow \{-1, 1\} \) that can maximize the margin to separate samples from two different categories in high-dimensional space. To generate this optimal hyperplane, it is necessary to solve the following quadratic programming problem:

\[ \min_{w,b} \frac{1}{2} ||w||^2 + C \sum_{i=1}^{n} \max \{0, 1 - y_i (w^T \varphi(x_i) + b)\} \quad (16) \]

where \( \max \{0, 1 - y_i (w^T \varphi(x_i) + b)\} \) is the standard hinge loss, \( w \) and \( b \) are respectively the weighted normal vector and bias. \( C \) is a penalty parameter emphasizing the loss caused by outliers and \( \varphi(\cdot) \) denotes a feature mapping which maps the input vectors into a high-dimensional feature space.

The convex quadratic programming problem shown in (16) can be solved through its dual problem as follows:

\[ \max_{\alpha_i} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \kappa(x_i, x_j) \quad (17) \]

\[ \text{s.t.} \sum_{i=1}^{n} \alpha_i y_i = 0, \ 0 \leq \alpha_i \leq C, \ i = 1, 2, \ldots, n \]
where $\kappa(x_1, x_2) = \langle \varphi(x_1), \varphi(x_2) \rangle$ is the kernel function that implicitly identifies the non-linear mapping from input space to a high-dimensional Hilbert feature space. Besides, it can solve the problem of linear inseparability for an SVM.

Finally, the optimal decision hyperplane can be obtained by solving (17) and then the unknown input pattern $x$ can be classified by the final decision function with the following form:

$$f(x) = \text{sign} \left( \sum_{i=1}^{n} \alpha_i y_i \kappa(x, x_i) + b \right) \quad (18)$$

6. Experiments

Experiments were conducted on the image dataset (each image measured $2048 \times 1360$ pixels) to test the performance of the proposed method. The dataset was collected from images recorded at a real substation in China.

6.1 Dataset Description

6.1.1 The Detection Unit

The models of all disconnectors are of the same 220 kV substation. Therefore the quantity of positive samples is small. In other words, the detection of disconnecting switch state belongs to the category of object instance detection. The serious imbalance of sample size between positive and negative samples (e.g., $P/N < 0.01$, where $P$ and $N$ are the numbers of positive and negative samples, respectively) will cause the trained classifier to show bias towards one category. To solve this problem, we can increase the number of positive samples by collecting data from different positions. Taking into account the actual situation of the monitoring scene in a substation, we keep the viewpoint at no more than $60^\circ$ ($\pm 30^\circ$) when acquiring images. We cut out the fixed-contact regions (including open and closed states) from 130 training images. A small translation (3 to 5 pixels) and rotation (less than $3^\circ$) are also carried out to generate more positive samples. Finally, we resize all these sample regions to normalized images measuring $64 \times 64$ pixels to generate a positive sample set (578 samples). The advantage is that not only can it increase the number of positive samples, it can also improve the performance of the classifier in the presence of translation and rotation. The negative sample set is constructed by randomly cutting out rectangular non-disconnector regions from the training images. In addition, the training images are searched for false positives (called hard examples) and then the final LDA detector is produced by re-training the augmented set (initial 578 positive samples, 3023 negative samples, and hard examples).

6.1.2 The State Recognition Unit of Fixed-Contact

The dataset is divided into a positive sample set (open state, 363 samples) and a negative sample set (closed state, 215 samples). Then the final SVM state classifier is produced by training this dataset.

6.2 Experimental Set-Up

The proposed method was implemented in C++ with OpenCV libraries on a PC with a Core i5 3.1 GHz chipset, and 4 GB RAM. The size of image used in the experiment was $2048 \times 1360$ pixels. The positive and negative samples were normalized to $64 \times 64$ pixels and some of the samples are shown in Fig. 6. In fixed-contact detection, the search window size was $64 \times 64$ pixels and the detection step was $8 \times 8$ pixels. We downsampled the test images in 12 steps and the scaling factor was 1.1. In addition, we also trained the SVM classifier with a linear kernel, polynomial kernel, and a Gaussian kernel (radial basis function kernel). Their optimal values were chosen by grid search method [16] in each fold of the cross-validation experiment on the training dataset.
6.3 Quantitative Evaluation Methodology

For a binary classification problem, all tested samples can be divided into four types: true positive (positive samples are correctly classified), false negative (positive samples are misclassified), true negative (negative samples are correctly classified), and false positive (negative samples are misclassified): these are respectively denoted by $TP$, $FN$, $TN$, and $FP$.

Here, the performance of the proposed method is evaluated by two quantitative evaluation metrics, namely the precision-recall (PR) curve and the detection error tradeoff (DET) curve. The PR curve is usually used to evaluate the performance of a classifier. It is difficult to achieve high precision while simultaneously producing a high recall rate. The closer the PR curve is to the upper right corner, the better the performance. Formally, precision and recall are respectively defined as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (19)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (20)$$

The DET curve is widely used in object detection research. Here, the DET curve is plotted on a semi-logarithmic scale. In the DET curve, the vertical axis denotes the miss rate and the horizontal axis denotes the logarithm of the number of false positives per image (FPPI). The lower the DET curve, the better the performance. Formally, miss rate and FPPI are respectively defined as:

$$\text{Miss Rate} = \frac{FN}{FP + FN} \quad (21)$$

$$\text{FPPI} = \frac{FP}{n} \quad (22)$$

where $n$ denotes the number of test image samples.

6.4 Experimental Results

In these experiments, a set of 113 test images measuring 2048 x 1360 pixels were used to test the performance of the proposed method.

6.4.1 Detection Experiments

HOG-LDA was used to detect the fixed-contact component of the disconnecting switch. The parametric study settings for HOG are given in Fig. 7 which plots the PR curves and DET curves with respect to different cell, and block sizes. From Fig. 7, cells measuring 8 x 8 pixels and blocks measuring 2 x 2 cells performed best. The corresponding PR curve was closest to the upper right-hand corner and the DET curve was the lowest (the miss rate was 5.5% at $10^{-1}$ FPPI). Note that the PR curve and DET curve are not plotted at a cell size of 16 x 16 pixels and a block size of 3 x 3 cells due to the zero precision and 100% miss rate thereat.

Here, we compared the proposed LDA detector with some other methods: HOG-SVM, local binary pattern [17] (LBP)-LDA, and LBP-SVM. Figure 8 shows that the HOG-LDA method, compared with other methods, performed best with regard to its PR curve and lowest miss rate at $10^{-1}$ FPPI.

6.4.2 State Recognition Experiments

In these state recognition experiments, the performance of different methods for switch status detection was tested with a different decision threshold $\sigma$ for each LDA target detector. We also compared our method with recognition methods based on GF, and SIFT-based method (in SIFT, templates are predefined for open and closed states). We compared with SIFT-based method because the works [2], [3] which were also for substation automation used SIFT. In this paper, we implemented a version of the SIFT-based method that was similar with the work [3]. To perform
Figure 8. PR and DET curves for different detection methods.

Figure 9. Quantitative results (PR curves) for open: (a) and closed; (b) state recognition.

Table 2
The Recall Rates of State Recognition Methods for a Disconnector at a Precision of 90%

<table>
<thead>
<tr>
<th>State</th>
<th>SIFT (%)</th>
<th>GF-SVM (%)</th>
<th>NGF-LinearSVM (%)</th>
<th>NGF-PolySVM (%)</th>
<th>NGF-GaussianSVM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>0</td>
<td>88.3</td>
<td>92.6</td>
<td>92.4</td>
<td>95.3</td>
</tr>
<tr>
<td>Closed</td>
<td>0</td>
<td>90.3</td>
<td>92.5</td>
<td>96.3</td>
<td>98.2</td>
</tr>
</tbody>
</table>

SIFT feature matching between the template images and the monitoring images, we used KNN to find optimal matching and then the mismatches were eliminated by means of RANSAC. Figure 9 shows the PR curves for the proposed method for open and closed state recognition. Figure 9 shows that the proposed method was superior to GF- and SIFT-based recognition methods. In addition, the method using a Gaussian kernel SVM gave the best results compared with linear, and polynomial, kernel-SVMs. So the Gaussian kernel SVM was used here as the classifier where it achieved the best results, namely 95.4% precision and 94.2% recall rate for open state recognition with a synchronous 97.5% precision and 98.3% recall rate for closed state recognition. In addition, Table 2 shows recall rates of state recognition for the switch at a precision of 90%, while precisions are given in Table 3 for a recall rate of 95%.

The DET curves for open and closed state recognition are shown in Fig. 10: each curve was produced using a different decision threshold $\sigma$ in an LDA detector with different kernel SVMs, a GF-based recognition method, and a SIFT-based recognition method. Figure 10 shows that the method using an NGF-Gaussian SVM had the lowest miss rates which respectively were less than 6% and 3% at an FPPI of $10^{-1}$: Table 4 shows the miss rates at an FPPI of $10^{-1}$. 

<table>
<thead>
<tr>
<th>State</th>
<th>SIFT (%)</th>
<th>GF-SVM (%)</th>
<th>NGF-LinearSVM (%)</th>
<th>NGF-PolySVM (%)</th>
<th>NGF-GaussianSVM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>0</td>
<td>88.3</td>
<td>92.6</td>
<td>92.4</td>
<td>95.3</td>
</tr>
<tr>
<td>Closed</td>
<td>0</td>
<td>90.3</td>
<td>92.5</td>
<td>96.3</td>
<td>98.2</td>
</tr>
</tbody>
</table>
Table 3
The Precisions of State Recognition Methods for a Disconnector at a Recall Rate of 95%

<table>
<thead>
<tr>
<th>State</th>
<th>SIFT (%)</th>
<th>GF-SVM (%)</th>
<th>NGF-LinearSVM (%)</th>
<th>NGF-PolySVM (%)</th>
<th>NGF-GaussianSVM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>0</td>
<td>30.1</td>
<td>87.5</td>
<td>81.2</td>
<td>91.2</td>
</tr>
<tr>
<td>Closed</td>
<td>0</td>
<td>82.3</td>
<td>95.7</td>
<td>80.2</td>
<td>98.7</td>
</tr>
</tbody>
</table>

![Figure 10. Quantitative results (DET curves) for open: (a) and closed: (b) state recognition.](image)

Table 4
The Miss Rates of State Recognition Methods When FPPI Was 10⁻¹

<table>
<thead>
<tr>
<th>State</th>
<th>SIFT (%)</th>
<th>GF-SVM (%)</th>
<th>NGF-LinearSVM (%)</th>
<th>NGF-PolySVM (%)</th>
<th>NGF-GaussianSVM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>&gt;50</td>
<td>17.8</td>
<td>9.8</td>
<td>10.8</td>
<td>5.8</td>
</tr>
<tr>
<td>Closed</td>
<td>&gt;50</td>
<td>16.5</td>
<td>6.1</td>
<td>5.7</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Table 5
The Cost of the Proposed Method for Images of Different Sizes

<table>
<thead>
<tr>
<th>Image Size (Pixels)</th>
<th>LinearSVM (ms)</th>
<th>PolySVM (ms)</th>
<th>GaussianSVM (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2048 × 1360</td>
<td>2090.6</td>
<td>2092.8</td>
<td>2082.1</td>
</tr>
<tr>
<td>1024 × 680</td>
<td>583.1</td>
<td>586.5</td>
<td>585.1</td>
</tr>
<tr>
<td>640 × 425</td>
<td>254.9</td>
<td>252.4</td>
<td>254.4</td>
</tr>
<tr>
<td>512 × 340</td>
<td>171.4</td>
<td>172.8</td>
<td>173.1</td>
</tr>
</tbody>
</table>

Revisiting Figs. 9 and 10, the SIFT-based recognition method cannot achieve an acceptable result. The reason for this was that it was a method based on feature key-point matching. Regarding the fixed-contact, the biggest difference between the two different states was whether there is a moving-contact (see Fig. 1). However, there may be some noise (clutter or background) in the image (see Figs. 5 and 6). In addition, it was also found that the most dominant feature key-points were located on the fixed-contact body. Although the GF-SVM was superior to the SIFT-based method, it still cannot obtain a good result when compared with NGF-SVMs. The main reason was that the NGF was convoluted with a symmetrical kernel $H(x)$ which can extract the symmetrical characteristics with discriminative ability. However, GF is a feature that was only computed along the vertical direction. Therefore it was sensitive to background noise. The open state fixed-contact may suffer from clutter background noise. It maybe had a feature that was similar to that of the closed state fixed-contact, or other image background (e.g., other equipment).

This research also tested the efficiency of the proposed method. Table 5 lists the average recognition cost of the proposed approach using different kernel SVMs. It took approximately 2s to detect and recognize the state of disconnectors within one image measuring 2048 × 1360 pixels. To meet the demands of actual application in a real substation, Table 5 also lists the average recognition cost of the proposed method for images of different sizes: the efficiency of the proposed method with different kernel SVMs was
Figure 11. Examples of the proposed method as used in a substation. The detection and state recognition results are indicated by white and dark rectangles denoting open and closed states, respectively. The full lines are for the actual detected targets, while dashed lines are for the full pair of fixed-contacts.

almost the same and our method was efficient and could be applied in practice.

Finally, Fig. 11 shows some sampled detection and state recognition results. Most fixed-contacts were correctly detected and recognized. The rectangles represent the recognition windows for different states and scales: the dark and white dashed rectangles denote detected disconnectors.

7. Conclusion

An effective computer vision-based automatic detection and state recognition method for disconnecting switches in substation automation was proposed. According to prior knowledge, the HOG feature of the fixed-contact of a disconnector switch was used to design an LDA detector used to generate candidate images of a disconnecting switch. Based on the relatively stable shape characteristic of such fixed-contacts, the proposed discriminative NGF feature was then used to design an SVM classifier with kernels. The experimental results revealed that the proposed method was effective and achieved the best precision and recall rate at the same time using HOG-LDA with a Gaussian kernel SVM. Quantitatively, it performed well with 95.4% precision and a 94.2% recall rate for open state recognition and 97.5% precision and 96.3% recall rate for the closed state. A small miss rate was also found which demonstrated the efficacy of the proposed method and its applicability to practical substation automation operations.

The proposed method was able to handle the usual background clutter in a substation, which would be other horizontal wires, vertical poles, and some other electrical equipment. Our approach may produce misdetection when it encounters heavily unstructured random background which may be behind the substation. Now, it is more suitable to stations in open areas, where the background is expected to be open sky and less to urban stations where there could be random buildings around. On-going, and future, work will focus on a more efficient, robust approach with higher precision, recall and anti-heavy background clutter ability, as well as on its application in situ. Regarding the methodological study, the prior confidence region and robust feature descriptor of the disconnector will be considered; while regarding its wider application, any future approach will be tried in live substations.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under Grants 61271432 and 61273337. This work was also supported in part by the State Grid Science and Technology Project: the Research of Multimedia and Streaming Media Based on Machine Learning.
References


Biographies

**Hongkai Chen** received his B.Sc. degree in Control Science and Engineering from Xiamen University, Xiamen, China, in 2011. He is currently pursuing his Ph.D. degree from the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. His current research interests include robot vision, pattern recognition, target detection, and visual tracking.

**Xiaoguang Zhao** received her B.Sc. degree in Control Engineering from Shenyang University of Technology, Shenyang, China, in 1992, and her M.Sc. and Ph.D. degrees in Control Theory and Control Engineering from the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China, in 1998 and 2001, respectively. She is a professor with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. Her current research interests include advanced robot control, wireless sensor network, and robot vision.

**Min Tan** received his B.Sc. degree in Control Engineering from Tsinghua University, Beijing, China, in 1986, and his Ph.D. degree in Control Theory and Control Engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 1990. He is a professor with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. His research interests include advanced robot control, multitirobot systems, biomimetic robots, and manufacturing systems.

**Shiying Sun** received his B.Sc. degree in Control Science and Engineering from Central South University, Changsha, China, in 2013. He is currently pursuing his Ph.D. degree from the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. His current research interests include robotics and computer vision.