

Overhead ground wire detection by fusion global and local features and supervised learning method for a cable inspection robot

Xuhui Ye

School of Power and Mechanical Engineering, Wuhan University, Wuhan, China

Gongping Wu

School of Power and Mechanical Engineering, Wuhan University, Wuhan, China and
Guangdong Keystar Intelligence Robot Co., Ltd., Guangzhou, China

Fei Fan

School of Power and Mechanical Engineering, Wuhan University, Wuhan, China, and

XiangYang Peng and Ke Wang

EPRI of Guangdong Power Grid Co., Ltd., Guangzhou, China

Abstract

Purpose – An accurate detection of overhead ground wire under open surroundings with varying illumination is the premise of reliable line grasping with the off-line arm when the inspection robot cross obstacle automatically. This paper aims to propose an improved approach which is called adaptive homomorphic filter and supervised learning (AHSL) for overhead ground wire detection.

Design/methodology/approach – First, to decrease the influence of the varying illumination caused by the open work environment of the inspection robot, the adaptive homomorphic filter is introduced to compensation the changing illumination. Second, to represent ground wire more effectively and to extract more powerful and discriminative information for building a binary classifier, the global and local features fusion method followed by supervised learning method support vector machine is proposed.

Findings – Experiment results on two self-built testing data sets A and B which contain relative older ground wires and relative newer ground wire and on the field ground wires show that the use of the adaptive homomorphic filter and global and local feature fusion method can improve the detection accuracy of the ground wire effectively. The result of the proposed method lays a solid foundation for inspection robot grasping the ground wire by visual servo.

Originality/value – This method AHSL has achieved 80.8 per cent detection accuracy on data set A which contains relative older ground wires and 85.3 per cent detection accuracy on data set B which contains relative newer ground wires, and the field experiment shows that the robot can detect the ground wire accurately. The performance achieved by proposed method is the state of the art under open environment with varying illumination.

Keywords Adaptive homomorphic filter, Global and local feature fusion, Overhead ground wire detection, Supervised learn method

Paper type Research paper

1. Introduction

The research on power-line inspection robot which has the ability to cross obstacles automatically has been raised highly attention (Pouliot and Montambault, 2012; Debenest and Guarnieri, 2010; Hongguang *et al.*, 2010; Wang *et al.*, 2014). The key problem of automatically obstacles crossing technology is how to detect the ground wire accurately. After the comprehensive analysis of various detection methods, it can be concluded that there are two broad categories methods: non-vision based and vision based. With a laser sensor installed on the bottom of robot manipulator, one method in literature Cuilian *et al.* (2006) is proposed to detect ground wire by analyzing rising edge signal of

the laser sensor when the laser spot is sheltered by the ground wire. Two laser sensors are adopted directly in literature Xinglong *et al.* (2006) to detect points along two line edges. With the help of the robot dynamics model, the position of ground wire relative to

© Xuhui Ye, Gongping Wu, Fei Fan, XiangYang Peng and Ke Wang. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial & non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

This work is supported by Guangdong Robot Special Project (2015B090922007), Foshan Technical Innovation Team Project (2015IT100143) and South Wisdom Valley Innovative Research Team Program (2015CXTD01). The authors also wish to thank anonymous reviewers, each of whom provided comments that have direct and important improvements in the manuscript.

Received 9 August 2017

Revised 21 November 2017

Accepted 13 December 2017

The current issue and full text archive of this journal is available on Emerald Insight at: www.emeraldinsight.com/0260-2288.htm



Sensor Review
38/3 (2018) 376–386
Emerald Publishing Limited [ISSN 0260-2288]
[DOI 10.1108/SR-08-2017-0154]

the arm is determined. One method is proposed in literature SunSin and JangMyung (2008) to detect the ground wire through the combination of an infrared detector with high measurement accuracy and an ultrasonic sensor which is not affected by light. An electromagnetic sensor is used to inspect power transmission line when energized in literature Zhongwei et al. (2006). Literature by Richard et al. (2014) and Pouliot et al. (2012) introduces a compact and less-expensive LIDARs system, UTM-30LX, to detect ground wire and obstacles by the comprehensive analysis of distance, diameter and signal intensity.

Literature by Yunchu et al. (2007) presents a ground wire detection scheme with a fixed-focused length camera installed on the shoulders of the mechanical arm. By observing the distribution characteristic of binary pixels of the captured image, the horizontal position of the axis center of ground wire can be determined through vertical projection, and the horizontal coordinates of up and down points of ground wire are determined through symmetry constraint. The method is effective for ground wire detection, but it neglects the effect of light on the binary image distribution. Literature by Ludan et al. (2007) models the transmission line as a cylinder with infinitely length and uniform diameter. A stereo vision method is proposed based on monocular camera to determine the location and attitude of ground wire according to the geometrical characteristics and the imaging principle. The mathematical model has been improved in literature Cong and Wei (2011) for detecting ground wire based on the prior literature. Both of the two kinds of solution introduce the stereo technology which is effective for the ground wire detection, but it is lack of camera calibration operation and the light factors affecting imaging that are neglected. Literature by Weibing et al. (2012) defines the transmission line offset distance and angle in image space to characterize the position of ground wire. After extracting the edge of the gray image and calculating length of each detected lines, the longest two parallel ones are selected as the boundary of ground wire by Hough transformation. The method is intuitive and easy to implement, but the experimental results under the cluster background which may interfere the lines detection are not given. One method is proposed in literature Wang et al. (2014) to calculate angle and intercept of ground wire through the following steps: intercepting region of interest (ROI) area, binarization, hole filling, corrosion, connected domain calculation, fitting connected domain and characteristics calculation. For this method, the influences of changing illumination and the cluster background on detection accuracy are considered, but no contrast experiments are conducted. Literature by Wenming (2014) analyses the texture of ground wire and proposes one method to determine position of ground wire based on the texture feature local binary pattern (LBP) and fuzzy c-means method (FCM). Statistical characteristics are used to improve LBP operator and reduce the cost of calculation. The first two-order Shannon entropy of the image intensity histogram is selected as the first two-dimensional characteristic of texture descriptors, and the edge density of canny edge image is selected as the third dimension. One method is proposed in literature Hyunho et al. (2013) which filters image with Gabor filter to characterize the position of the ground wire based on health coefficient. One method is put forward in literature Song et al. (2014), which is mainly implemented in three steps, to identify

the normal line, broken strand, counterweight by using histogram of gradient (HOG) feature and support vector machine (SVM). First, an ROI rectangle is extracted from the captured image. Second, the HOG feature of the ROI is extracted. In the third step, HOG features are put into a hybrid classifier composed of two SVM to train final classifier. This method can get good performance when identifying the normal line. However, when Hough transformation is used to detect the two edges of the ground wire, the influences of illumination and complex background are not considered. Literature by Zhenhui et al. (2015) proposes the pose estimation method for transmission line based on its textual features and integral projection method.

Both non-vision-based and vision-based methods have their own advantages, but vision-based methods are better than non-vision-based methods for three reasons. First, the vision sensor has the advantages of small volume, light quality and convenient installation, while the laser sensor, infrared sensor and radar sensor do not have that advantages. Second, the vision sensor is not affected by temperature and changes of magnetic field, but the infrared image is greatly influenced by temperature. Third, information obtained from vision sensor is comprehensive, while the non-vision-based method acquires local information.

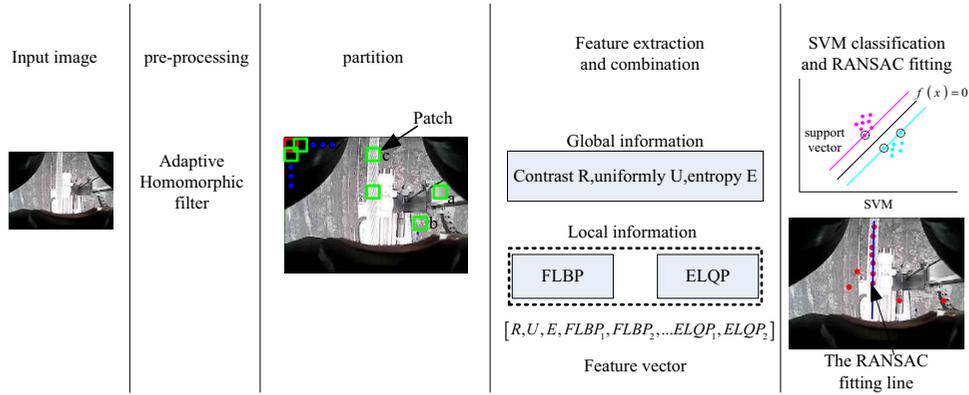
Based on hand-eye-vision system, this paper proposes one new method called adaptive homomorphic filter and supervised learning (AHSL) method, which adopts adaptive homomorphic filter, global and local feature fusion approach followed by image partition as well as random sample consensus (RANSAC) algorithm for ground wire detection. The proposed method AHSL works by pre-processing image with adaptive homomorphic filter and partitioning images into overlapped square patches. Then global and local features are extracted from all of the patches, and all the combined feature vectors are put into the binary classifier trained by supervised learning method SVM. After classifying the patches, the patches which belong to the ground wire are fitted to a straight line to represent the ground wire by using RANSAC method.

Section 2 describes the details of the proposed method which include pre-processing; features extraction and combination; and ground detection and representation. Section 3 describes the experiments and results, and Section 4 finally concludes the work we have done and what we are looking forward to in the future.

2. The proposed method adaptive homomorphic filter and supervised learning

To detect ground wire under the influencing factors mentioned above, this paper proposed one method called AHSL. In general, the method can be divided into four steps. The first step is to blur the image with adaptive homomorphic filter to compensate illumination. The second step is to divide the whole image into small blocks, which is called "patch". For each patch, global and local features are extracted and combined into a feature vector to represent the image. On the basis of the training data set, a binary classifier is generated by using the supervised learning method SVM to determine whether the patch belongs to the ground line. Then through recording all the patches belonging to the ground wire, the RANSAC algorithm is applied to fit them to 2D straight line. The flow chart of the AHSL is shown in Figure 1.

Figure 1 The flow chart of proposed method AHSL



2.1 Pre-processing based on adaptive homomorphic filter

Based on the Lambert reflection model (Gonzalez and Woods, 2010), an overhead ground wire image I is defined by:

$$I(x, y) = R(x, y)L(x, y) \tag{1}$$

where $I(x, y)$, $R(x, y)$ and $L(x, y)$ represent the pixel value positioned in image coordinate (x, y) ; the intensity of illumination reflected by surface of the ground wire and the air; and the intensity of illumination shine to the ground wire, respectively. Seen as high frequency signal, R is corresponding to the high contrast area, such as ground wire, the off-line arm and the pinch roller. L is seen as low frequency signal corresponding to the background of the ground wire image because the image illumination component between adjacent pixels changes slowly (Wang et al., 2011).

To remove the influence of illumination variation, the homomorphic filter $H(u, v)$ is applied to suppress low frequency illumination component and enhance high frequency reflection component at the same time, which achieves the goal of illumination compensation for the ground wire image.

To perform exponential operation on both sides of equation (1), R and L are separated:

$$\ln I(x, y) = \ln R(x, y) + \ln L(x, y) \tag{2}$$

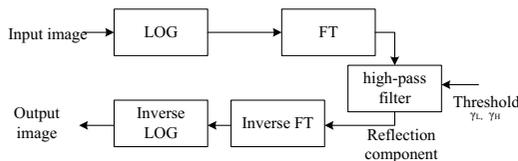
Performing the following operation on image I according to Figure 2, the output image keeps the reflection component.

Applying the same method proposed in literature Jain (2011), $H(u, v)$ is defined by:

$$H(u, v) = (\gamma_H - \gamma_L) \left[1 - \exp^{-c(D^2(u, v)/D_0^2)} \right] + \gamma_L \tag{3}$$

where $D(u, v)$ is the distance from (u, v) to the original of the center Fourier transform, D_0 is the cutoff distance measured

Figure 2 The flow chart of homomorphic filter



from the origin, $\gamma_L < 1$ and $\gamma_H > 1$ are the parameters of the filter, and c is a constant to control the sharpness of the slope of the filter function as transition take place.

Literature by Faraji and Qi (2015) has proven that c is the key parameter of the homomorphic filter. Every ground wire image has different unknown amount of illuminations that require a different filter to compensate illumination. Therefore, we produce the adaptive homomorphic filter based on each input image by adjusting the parameter c .

In the spectrum diagram of ground wire, the origin of the Fourier transform contains the slowest varying frequency component which corresponds to the average gray level intensity of the image. While moving away from the origin, the frequency changes from low to high. Because the illumination component mainly concentrated in the low frequency part, the origin center contains most of the varying of the illumination. Therefore, we consider a square window around the origin of the Fourier transform and use a ratio among the low frequency component to determine parameter c , which represents the changing rate of the low frequency component, that is the change rate of illumination component. Parameter c is defined by:

$$c = \frac{Mag_1}{Mag_2} \tag{4}$$

where Mag_i is the i th largest value in the square window.

In the experiment, the side length of the square is set to the 15 per cent of the smaller value of image width and height. If the image size is 320×240 , the size of the square window is 36×36 . The other parameters are empirically set to be $\gamma_H = 1.1$, $\gamma_L = 0.5$ and $D_0 = 36$. D_0 , the cutoff distance, is set to the side length of the square window.

The effect of adaptive homomorphic filter performed on ground wire image is shown in Figure 3. It indicates that the homomorphic filter attenuates the low frequency component. That is the illumination component of a ground image.

2.2 Feature extraction

2.2.1 Ground wire image partition

For each captured frame, it is divided into some $m \times m$ square window as shown in Figure 4. These square windows are arranged from the top-left corner to the right-bottom corner, from left to right, from top to bottom. The horizontal and

Figure 3 The result of homomorphic filter performing on ground wire image

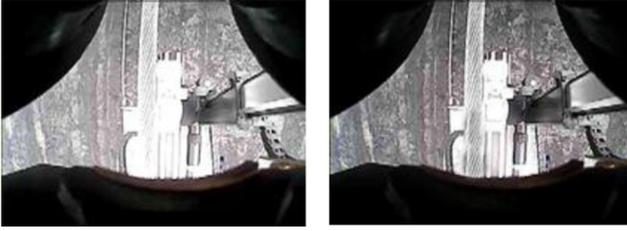
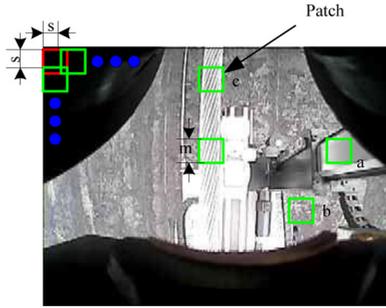


Figure 4 The schematic diagram of image block



vertical moving stride is s which is set to $m \times 0.75$. For every movement, a sub-image is cropped from the image contained in the square window; thus, the ground wire image is divided into many sub-images overlapped each other, and the sub-image is called “patch” in this paper.

So, for a ground wire image whose size is $H \times W$, it can be divided into $N m \times m$ square patches and N is given by:

$$N = \left(\left\lfloor \frac{H-m}{s} \right\rfloor + 1 \right) \times \left(\left\lfloor \frac{W-m}{s} \right\rfloor + 1 \right) \quad (5)$$

Specifically, $H = 240$, $W = 320$, $m = 12$, $s = 9$ and N is 910. So each image is divided into 910 square patches. The value of parameter m is related to the diameter of the ground wire and the distance from the lens of the hand-eye camera to the upper surface of ground wire. The farther the distance is, the smaller m is. In the experiment and field work, the distance is limited to a certain value.

2.2.2 Global features extraction

The characterizations of texture can totally be divided into three categories, namely, smoothness, coarseness and regularity (Gonzalez et al., 2009). Three features are used to represent the global information of the ground wire patch characterized by texture.

For a ground wire patch Z which is shown in Figure 4, let z be a random variable denoting intensity and let $p(z_i)$, $i = 0, 1, 2, \dots, L-1$, be the corresponding histogram, where L is the number of distinct intensity level.

The first global statistical feature R is based on the second moment of z which is particular important in texture description. Because z is a measure of intensity contrast, it can be used to establish relative smoothness descriptors. R is given by:

$$R(z) = 1 - \frac{1}{1 + \sigma^2(z)/(L-1)^2} \quad (6)$$

It is 0 for areas of constant intensity and approaches 1 for large values of $\sigma^2(z)$.

The second feature U is texture measures based on histograms including a measure of “uniformity”, given by:

$$U(z) = \sum_{i=0}^{L-1} p^2(z_i) \quad (7)$$

The $U(z)$ is large for smooth area and relatively small for coarse area, and all values are in the range (0,1). It is a measurement of similarity of image pixels.

The third statistical global feature E is an average entropy measurement, which you will recall from the basic information theory. It is defined as:

$$E(z) = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) / L \quad (8)$$

The value of $E(z)$ is in the opposite order and thus leads to the same conclusions as the uniformity.

The three global features of three different type of texture patch (as show in Figure 4) are shown in Table I.

The ground wire is composed of aluminum strands and the appearance texture is regular, while the background of it shown in Figure 4 is either smooth or coarse. For feature R in Table I, patch a gets the lowest value because it lies in a smooth region where all intensity of pixels are almost the same, while patch b gets the highest value as it has extensive variation in intensity. Given a patch of the captured image, the one with the R value which is higher than a threshold and lower than another threshold has little probability to be a ground wire patch. The feature E does have the same characters with feature R , while the feature U does have the opposite characters. From the Table I, we can know that the three global features are useful tools to distinguish ground wire and the background. So the global feature of ground wire is defined as $G = [R, U, E]$.

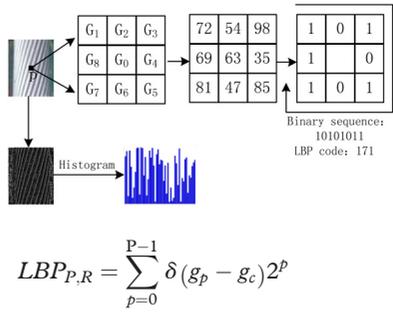
2.2.3 The LBP histogram feature extraction

Derived from a general definition of texture in a local neighborhood, LBP proposed in literature Ojala et al. (2002) is defined as a grayscale invariant texture measurement. It is a useful tool to model texture images. Its effectiveness proves in image classification (Yuan et al., 2015; Nanni et al., 2012), face recognition (Zhou et al., 2013; Shen et al., 2016), vehicle detection (Hu et al., 2016), texture classification (Ojansivu and Heikkilä, 2008) and so on.

As shown in Figure 5, for any point p on the ground wire, the pixel values of 3×3 pixels in the domain are compared with intensity value of point p to obtain a binary sequence. The LBP code at point p is given by:

Table I Texture measures for the sub-image shown in Figure 4

| Standard texture | R | U | E |
|------------------|-------|-------|-------|
| a | 0.003 | 0.103 | 0.052 |
| b | 0.105 | 0.002 | 0.636 |
| c | 0.025 | 0.032 | 0.125 |

Figure 5 The feature extraction of LBP histogram

$$LBP_{P,R} = \sum_{p=0}^{P-1} \delta(g_p - g_c) 2^p \quad (9)$$

where g_c is the intensity value of point p . g_n is the intensity value of neighborhood pixels. R is the radius of neighborhood. P is the number of neighborhood points, and the sign function $\delta(\cdot)$ is given by:

$$\delta(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (10)$$

As shown in Figure 5, the LBP codes of each pixel are calculated through statistical analysis to get the LBP histogram distribution of coding, so the feature vector of the ground image patch is given by $[H_1, H_2, \dots, H_n]$, where H_i is the number of occurrences of intensity value i , and n is the dimension number of feature vector. The feature vector is normalized to get the final vector which is given by $H = [H_1', H_2', \dots, H_n']$, where $H_p' = (H_p - \min(H_n)) / (\max(H_n) - \min(H_n))$.

For the original LBP discussed above, the overall dimension of feature vector is 2^P . An extension of LBP patterns is called circular uniform LBP whose dimension number is $P + 2$. It first links the head and end of the obtained binary sequences to get a circle binary sequence, choose the different node as the starting bit to form P binary sequences, select the smallest LBP code among the P binary sequence as the desired LBP code of point p and finally categorize the LBP codes to the same class which contain two times more than those of transition from 0 to 1 or 1 to 0. Based on the calculation of the original LBP codes of ground wire, the circular uniform LBP histogram are extracted to represent the final feature vector.

2.3 Feature combination

The scheme which fuses different features to boost the description power is popular in computer vision application, and the literature Premebida *et al.* (2009), Yuan *et al.* (2015) and Zhenhui *et al.* (2015) have proven its reliability in pedestrian detection, scene classification and cloud classification.

For each patch shown in Figure 1, the global feature G and local feature H discussed above are extracted, and they are combined to form a description of the patch. The simple way to combine features is to concatenate the two feature vectors and assign the same weights to the two of them, which is given by $V = [G, H]$. The weighting scheme ignores the differences in the discriminative power of global and local features which may lower the performance of combination. A straightforward approach to improvement is to assign weights to features based on their discriminative power. Assuming that the discriminative power of m feature vectors is estimated and given by p_i , $i = 1, \dots, m$, we can simply define the weights of these feature vectors as follow:

$$\omega_i = \frac{p_i}{\sum_{j=1}^m p_j}, \quad i = 1, \dots, m \quad (11)$$

This definition is consistent with our intuitive that feature vectors with a larger discriminative power should be assigned larger weights.

There are two feature vectors in our work, and m is set to 2. The weights of G is set to β , which is called mixing coefficient, and the weight of H is set to $1 - \beta$. The combined feature vector is given by:

$$V = [\beta G, (1 - \beta)H], \quad G = [R, U, E], \quad H = [H_1, \dots, H_n] \quad (12)$$

where G and H are global feature vector and local feature vector, respectively, and n is the dimension of local feature vector which is $P + 2$, and P is the number of considering neighborhood pixels in circular uniform LBP feature extraction.

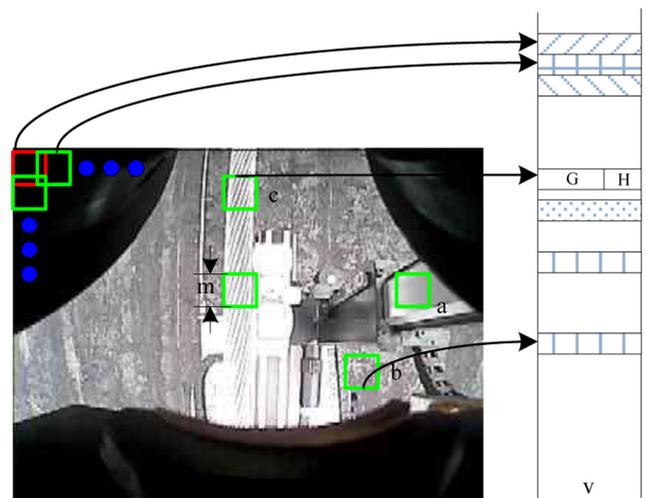
Therefore, for a ground wire image captured by hand-eye camera, the combined features are extracted, and all the feature vectors are sent to the next processing phase as testing data, and the schematic diagram is shown in Figure 6.

2.4 Ground wire detection and representation

2.4.1 Ground wire detection

As the ground wire image is divided into square patches, the purpose of ground wire detection is transformed into the classification of the patches. SVM is a kind of supervised machine learning algorithm based on structural risk minimization. It is commonly used to solve the problem of binary classification and shows particular advantages in the aspect of solving problems of small samples data sets, nonlinear and high-dimensional pattern recognition. Therefore, in this paper, SVM is applied to solve this problem.

Each training samples contained in training samples set D can be expressed as $D_i = (x_i, y_i)$, $i = 1, 2, \dots, n$, where n is the total number of training samples, and x_i is the feature vector discussed in previous stage, y_i is the corresponding label of x_i and $y_i \in \{-1, 1\}$. The label -1 indicates the patch, as patch a in Figure 6, which belongs to background and 1 means the patch, as patch c in Figure

Figure 6 The feature vector

6, which belongs to the ground wire area. For testing data, the label is unknown which is ready to be assigned by SVM.

For a linear separable problem, an optimal separating hyperplane, given by $w^T x + b = 0$ is desired to be found to separate positive sample data and negative sample data, and the classifier function is given by $f(x) = w^T x + b$, where w is the normal vector of the hyperplane. b is the intercept of the hyperplane equation. For any input feature vector x , if $f(x) > 0$, the label of the corresponding patch is set to 1. The hyperplane tries to maximize the gap which is given by $2/\|w\|$ between positive samples and negative samples, and the propose is equal to minimize the object function $\|w\|^2$. For a linear inseparable problem, the slack variable ζ is introduced to allow a certain degree of mistaken classification. The constraint condition is $y_i[w^T x + b] \geq 1 - \xi_i$. The penalty factor C is proposed at the same time to value the importance of mistaken classification points, and large value of C means that there is intolerance with mistaken classification. Accordingly, the problem is converted to minimize the object function under constraint conditions, which is given by:

$$\min : \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \quad (13)$$

where the constraint condition is given by $y_i[w^T x + b] \geq 1 - \xi_i$, $\xi_i \geq 0$, $i = 1, 2, \dots, n$.

The appropriate mapping transformation Φ is selected to map the feature vector of the sample to high-dimensional feature space, so that the inseparable problem can be transformed into separable problem. Because the mapping function ϕ is hard to find out, the appropriate kernel function K is proposed to receive two low-dimensional feature vectors to obtain the result which is equal to the inner produce of two feature vectors in the high-dimensional space, and the equation is given by:

$$K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle \quad (14)$$

where $\langle x_i, x_j \rangle$ is the inner product of vectors x_i and x_j . The parameter w is related to the feature vector and label of the training sample, and we assume that:

$$w = \sum_{i=1}^l (a_i y_i x_i) \quad (15)$$

where a is the Lagrange multiplier. So the classifier function is given by:

$$g(x) = \text{sgn} \left[\sum_{i=1}^l a_i y_i K(x_i, x) + b \right] \quad (16)$$

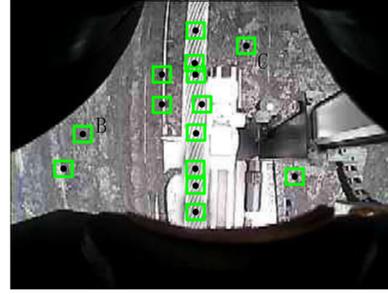
In the experiment, the RBF kernel function is applied and the core parameters are C and γ .

For each test sample, the feature vector x_i is calculated and put into the equation (14), we get 1 or -1 which represents positive sample or negative sample, respectively. Figure 7 shows the classification result of a testing frame. The patches which are classified as positive samples are shown in the image with a black center point. It indicates the position of the sample under the image coordinate.

2.4.2 Ground wire representation

The ground wire in the image may be seen as a 2D straight line overlap with the center line of the ground wire. The result of the

Figure 7 The SVM classification of ground wire patches



detection phase is a collection of candidate patches. But as shown in Figure 7, not all points represent candidate patches which are in the ground wire area, such as point B and Point C which are the mistaken detection samples. The black points in Figure 7 can be divided into inner points and outlier points. The inner points are the ones which lie in the area of ground wire and the outlier points are outside the area of the ground wire. It is difficult to find the accurate line through all the inner points by conventional linear fitting method, i.e. least square method because it is tried to consider all the candidate points which include outlier points.

RANSAC algorithm is a kind of robust parameter estimation method. A model estimation is obtained through the random subspace of all the points to test the remaining points to get a score. Repeating the action certain times, the model with the biggest score is selected as the final model of all the points. The premise of the correct erection for the ground wire by using RANSAC is based on two assumptions. One is that the majority of the detected candidate patches using SVM are in the area of ground wire, and this assumption is verified by the following experiments based on the proposed method AHSL. The other one is that the iteration times k should be large enough to ensure that all the selected points are all the inner points at least one time, and the k is given by:

$$k \geq \frac{\log(1 - \eta_0)}{\log(1 - \varepsilon^m)} \quad 17$$

where η_0 is the confidence coefficient and equal to 0.95 in the following experiment. $1 - \eta_0$ is the probability that an inner point is classified as outlier point. m is the numbers of random selected points each iteration and is set to 2. ε is the ratio of the number of inner points and the number of all points, and it is set to 0.5. According to the equation (17), k should be larger than 10. So k is set to 15 in the experiment.

The fitting results of the detected points in Figure 7 by least square method and the RANSAC method are shown in Figure 8. It can be indicated that the RANSAC method can eliminate outlier points and fit out the ground wire accurately.

After detecting the ground wire, it is represented by two parameters θ and λ , given by Figure 9. θ means the angle between the vertical middle line and the center line of the ground wire. Starting from the middle vertical line, θ is positive in clockwise and negative in anti-clockwise. λ donates the distance between center point of the image and the intersection point of center line of ground wire and center line of the image height orientation. When the ground wire is on

Figure 8 The fitting result (a) least square method and (b) the RANSAC method

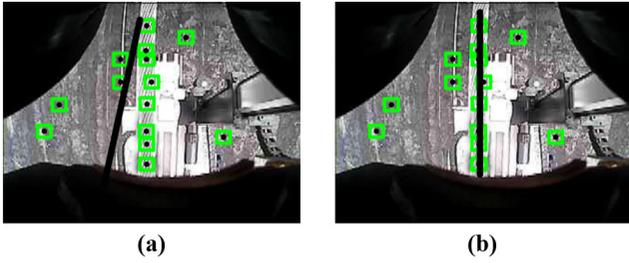
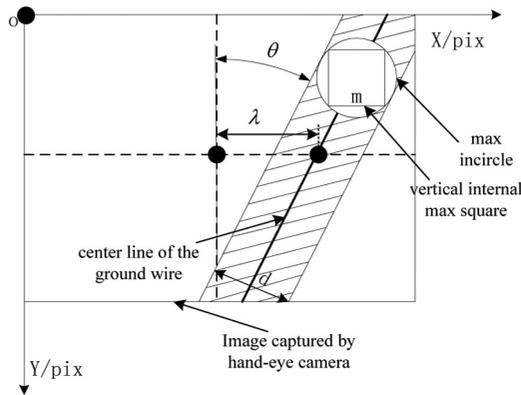


Figure 9 The parameters of detected ground wire



the left side of the central vertical line, γ is positive. Conversely, it is negative.

3. Experiment and result

3.1 Data sets setup

The positive samples of the training set are cropped manually from the image patches which belong to ground wire, and the negative samples of the training set are cropped randomly from the image patches which belong to the background of the image. The images under different illuminations are used to create testing data set.

3.1.1 The establishment of the training set

Considering two kinds of ground wire with different surface that one is a new ground wire and the other is a relative old ground wire, the images are taken under different illumination conditions camera, and the patches which belong to the ground wire are cropped to be used as positive samples. The negative samples are cropped randomly from the images which screen out the ground wire. The size of the cropped patches is 20×20 and part of the training samples is shown in Figure 10.

Figure 10 The part training samples of ground wire



3.1.2 The establishment of the testing set

Considering two-ground wire with different surface, two testing sets called A and B are created based on the two-ground wire. A is relatively older ground wire and B is relatively newer ground wire. For each ground wire, images taken by hand-eye camera are different due to the various illumination in the different period of the day. Five hundred images are selected on an average to be the testing set. For each captured image, the parameters θ and γ are assigned to it, and γ is set to 1000 if there's no ground wire in the image. Some of the testing images in set A and B are shown in Figure 11. The first row of Figure 11 shows the examples of set A and the second row shows the examples of set B.

3.2 Experiment setup

For the two testing data sets, A and B, we perform all processing in grayscale. To evaluate the effects of the adaptive homomorphic filter on the final detection accuracy, the mixing coefficient β is set to 0.5, and a comparison experiment between the with and without adaptive homomorphic filter is performed. To find out the optimal mixing coefficient β , we enumerate the β from 0.1 to 0.9. To prove the effectiveness of the proposed method, the comparison experiment between previous methods and the proposed method is performed.

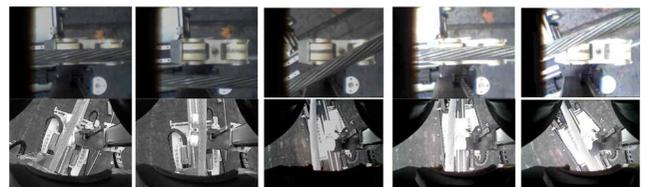
For each experiment, the size m of the square patch shown in Figure 4 is given by $m = d\sqrt{2}$, where d is the diameter of the ground wire shown in Figure 9, and $d = 16$ mm for data A and B. In the ground wire contour, the size of the vertical internal max square from the max incircle is chosen as the one of the patch. This is a rational method because the patch is full of the ground wire with the largest probability. Larger m will lead to higher false negative ratio because the patch which belongs to the ground wire can be classified as background. Smaller m will lead to the increase of the total number of the patches which will consume more calculation resources.

The parameter R is the radius of neighborhood, and P is the number of neighborhood points. They are set to (2, 16) empirically. The core parameters of the SVM (C, γ) used in this paper are optimized by the grid search method, which is provided by the LibSVM toolbox (Chang and Lin, 2011), and the final parameters is (2.0, 45.2) for $\beta = 0.2$.

For a testing sample, if $\Delta\theta \leq 3^\circ$ and $\Delta\lambda \leq 5$ pixels, we treat it as an accurate detection, where $\Delta\theta$ and $\Delta\lambda$ denote the absolute differential value of the θ and λ of the fitting line by RANSAC method and the sample itself. The detection accuracy is given by:

$$\text{accuracy} = \frac{TP}{N} \quad (18)$$

Figure 11 The part testing samples



where TP denotes the number of accurate detection and N is the total number of testing sample. The definition is used to evaluate the performance of all the following experiments.

3.3 Result

3.3.1 The effect of the adaptive homomorphic filter

The illumination changes of the testing images are caused by the changing environment where the robot works. To evaluate the performance that the adaptive homomorphic filter is applied to compensate the illumination variation, a set of controlled experiments were performed on data sets A and B, which are relatively older and newer ground wire, respectively. The first experiment was performed under the condition that there is no pre-processing filter, and the following procedures are the same as the proposed method AHSL. The second experiment was performed under the condition that the core parameter c of the adaptive homomorphic filter is set to fixed value 0.9 empirically, and the following procedures are the same as the proposed method AHSL. The third experiment was under the proposed method with the adaptive homomorphic filter as the pre-processing procedure.

Figure 12 shows the results of the experiments, and Table II gives the details of the performance obtained from our method. It can be concluded that the adaptive homomorphic filter applied in the proposed method improves the detection accuracy of the ground wire tremendously.

From the Table II, the row means the performance of the three experiments on the same data set. It can be seen from the first row in Table II that the method with the adaptive homomorphic filter achieves the highest detection accuracy among the three experiments in the data set A, and the performances of the first two controlled experiments are almost the same. The results shown in the second row reveal the same

Figure 12 The results of adaptive homomorphic filter on data set A and B

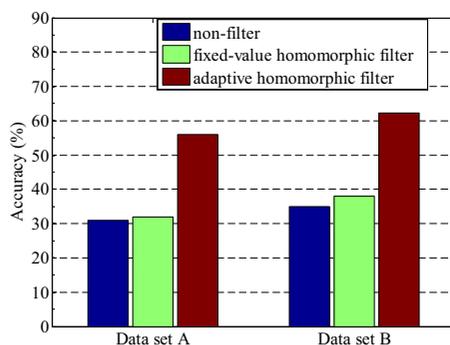


Table II The detection accuracy of the three different methods on data set A and B

| Accuracy (%) | Non-filter | Fixed-value homomorphic filter | Adaptive homomorphic filter |
|--------------|------------|--------------------------------|-----------------------------|
| Data set A | 31.1 | 32.2 | <i>56.0</i> |
| Data set B | 35.2 | 38.1 | <i>62.3</i> |

Note: The highest results are shown in italic

discipline as the first row. The reason is that the robot working environment is characterized by an open space and complex illumination changes. The varying working environment leads to the illumination changes in the captured image, and the interference of illumination differs from image to image. Therefore, we need to filter the image with different amount of compensation for each image. If a fixed illumination compensation parameter is applied on all testing images, the performance of the method is unacceptable. Accordingly, for each input testing image, the parameter of the amount of the illumination compensation should be calculated, namely, adaptive parameters, to achieve improvement in detection accuracy of the ground wire.

From Figure 12 and Table II, we can see that the adaptive homomorphic filter improves the detection accuracy to a certain extent, but to be exact, the improvement is limited. The highest detection accuracy is 62.3 per cent for data set B and 56.0 per cent for data set A which is relatively low for detecting ground wire. The reason is that the features mixing coefficient β is set to a constant value 0.5. That is to say, we assign the same weights to the global and local features. However, the different importance of the two feature vectors are not taken into account. This kind of average weighting method may inhibit testing potential performance of the proposed method AHSL.

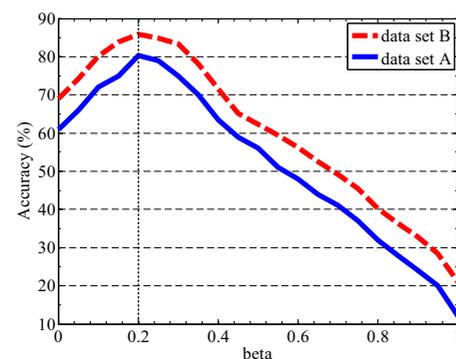
3.3.2 Relationship between mixing coefficient and final performance

The detection accuracy is related with the features mixing coefficient β because the global and local features are combined to form the final feature vectors. β and $1-\beta$ denote the proportion of the global feature vector G and the local feature vector H , respectively. To find the optimal β to get the best performance, we assign 21 different values to it between 0 and 1 in an interval of 0.05, and Figure 13 shows the performance under different value of β .

In Figure 13, it can be concluded from the two curves that when β is approaching 0.2, the highest performance is achieved. That is to say, when the proportion of the local feature vectors is 80 per cent, the optimal performance is achieved. The results show that the local features possess the powerful description ability and discriminative power for the texture of the ground wire.

When β is approaching 1, only the global feature vector is used as final feature vector. When the detection accuracy is

Figure 13 The performance under varying mixing coefficient β



merely 20 per cent, the performance is undesirable. When β is approaching 0, only the local feature vector is used as the final feature vector, and the detection accuracy is about 69 per cent. The reason is that the global features are statistical characteristics and describe the texture of the image roughly which is smooth, coarse and regular. It may be similar about the global features of the patches belonging to the ground wire and background. Under this condition, the ground wire patches and the background patches can be mistakenly identified which may produce more mistaken classification, and the final detection accuracy is relatively low. The local feature vector H is the LBP histogram features which can resist partly changing illumination as well as 2D rotation. It is capable of describing the texture of the ground wire shown in the image with arbitrary angle and under varying illumination. But when the distance between the surface of the ground wire and the lens of the camera varies slightly which may lead to texture density changes. The LBP histogram features of the patches in the same position may be different which will lead to mistaken classification. But another fact is that the global statistical features remain constant under this circumstance. Therefore, the global features can be used as a complementary information for the local feature vectors so as to achieve the effect of the two complementary advantages

Meanwhile, it can be concluded from Figure 13 that the performance on the data set B is better than the performance on the data set A with varying β . The reason is that the surface of the image containing relatively older ground wire in data set A may be corroded by some factors, which will lead to the breakage of the texture.

3.3.3 Comparison with other methods

We compare the performance of our approach with other methods mentioned in the introduction part of this paper on testing data set A and B. The performances of our approach and other methods are summarized in Figure 14 and Table III.

In the data set B, the detection accuracy of the method based on Hough transformation (Cong and Wei, 2011) is merely 35 per cent. All lines in the image are detected, and the longest two lines are selected to determine the parameter for the ground wire. The similar method proposed in literature Song et al. (2014) is applied to extract HOG features and perform classification procedure using SVM, and the detection accuracy is 65.1 per cent. The method proposed in literature Wenming (2014) extracts LBP features of the image and perform image segmentation using FCM, and the detection accuracy is 72.2

Figure 14 The performance comparison with other methods

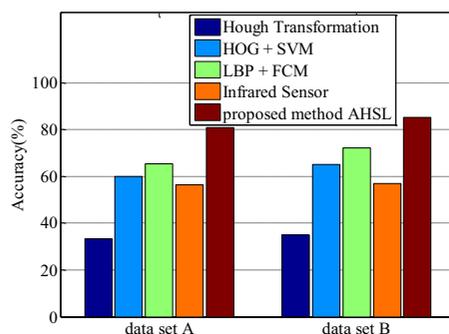


Table III The performance comparison with other methods

| Methods | A(%) | B(%) |
|---|-------------|-------------|
| Hough transformation (Cong and Wei 2011) | 33.2 | 35.0 |
| HOG + SVM (Song et al., 2014) | 60.1 | 65.1 |
| LBP + FCM (Wenming, 2014) | 65.2 | 72.2 |
| Infrared sensors (SunSin and JangMyung, 2008) | 56.3 | 57.0 |
| Our method AHSL | <i>80.8</i> | <i>85.3</i> |

Note: The best performance of each data set is shown in italic

per cent. The method proposed in literature SunSin and JangMyung (2008) with infrared sensors capture the infrared images of the ground wire, and the detection accuracy is 57.0 per cent. This method is implemented on another robot of our laboratory. The proposed method AHSL achieves the best performance which is 85.3 per cent, and the mixing coefficient β is set to 0.2 which is the optimal parameter discussed in previous experiment. The results on testing data A reveal the same discipline as the data set B.

The improvement of our approach on testing data A and B mainly depends on two effectual procedures. One is that the adaptive homomorphic filter is introduced in the proposed method AHSL with invariant illumination. For each input testing image, the optimal illumination compensation coefficient is calculated to remove the effect of the varying illumination. The other is that the combination of the global and local feature is introduced in the proposed method AHSL. The optimal mixing coefficient is obtained through the enumeration method. The combined features possess abundant information about the texture of the ground wire which is more powerful in describing the texture and discriminating the ground wire and the background.

Meanwhile, compared all the results on data set A and B, the performance on B is better than that of A. The reason is that the image surface of testing data A containing relatively older ground wire may be corroded by some factors, which will lead to the breakage of the texture.

3.3.4 The field experiment

To verify the validity of the proposed method AHSL, we conduct experiments on a simulation ground wire and an actual ground wire both under the open and complex changing illumination condition. Figure 15(a) shows the software window when the robot is finding ground wire, and Figure 15 (b) shows the experiment performed on simulation ground wires. Figure 15(c) shows the experiment performed on real projection. In the process of the two experiments, the robot can detect the ground wire accurately under the changing illumination, so it can seize the ground wire and cross the obstacle rapidly. The result shows that the method has certain advantages in practical application to some extent.

4. Conclusion

This paper presents a novel method AHSL by using adaptive homomorphic filter and global and local feature fusion approach based on image partition. This method is followed by RANSAC approach for ground wire detection. The proposed method AHSL works by pre-processing

Figure 15 Field experiment



(a)



(b)



(c)

Notes: (a) The software implementation; (b) the simulation ground wires experiment; (c) the field experiment

image with adaptive homomorphic filter which can decrease the influence of the varying illumination. In the feature extraction step, global and local features are fused with different weights to form the final feature vectors which contain more abundant local and global information than a single global or local feature.

This method AHSL has achieved 80.8 per cent detection accuracy on data set A which contains relatively older ground wires and 85.3 per cent detection accuracy on data set B which contains relatively newer ground wires, and the field experiment shows that the robot can detect the ground wire accurately. The performance which is achieved by proposed method is the state of the art under open environment with varying illumination. In our future work, we will focus on the ground wire detection method aiming at different kinds of ground wire with different aging degree and improving the efficiency of the algorithm as the real time processing is important for the practical application.

References

- Chang, C.C. and Lin, C.J. (2011), "IBSVM: a library for support vector machines", available at: www.csie.ntu.edu.tw/~cjlin/libsvm/
- Cong, W. and Wei, S. (2011), "The movement control of grasping high-voltage transmission line for deicing robot", *Journal of Mechanical Engineering*, Vol. 47 No. 9, pp. 8-15.
- Cuilian, S., Hongguang, W., Ludan, W. and Mingyang, Z. (2006), "An improved obstacle-navigation method based on extra-high voltage power transmission line for inspection robot robotics", *Robot*, Vol. 28 No. 6, pp. 379-384.
- Debenest, P. and Guarnieri, M. (2010), "Expliner-From prototype towards a practical robot for inspection of high-voltage lines", *1st International Conference on Applied Robotics for the Power Industry (CARPI)*, IEEE, 2010, pp. 1-6.
- Faraji, M.R. and Qi, X. (2015), "Face recognition under varying illumination based on adaptive homomorphic eight local directional patterns", *IET Computer Vision*, Vol. 9 No. 3, pp. 390-399.
- Gonzalez, R.C. and Woods, R.E. (2010), *Digital Image Processing*, Publishing House of Electronics Industry.
- Gonzalez, R.C., Woods, R.E. and Masters, B.R. (2009), "Digital image processing, third edition", *Journal of Biomedical Optics*, Vol. 14 No. 2, pp. 331-333.
- Hongguang, W., Yong, J., Aihua, L., Lijin, F. and Lie, L. (2010), "Research of power transmission line maintenance robots in SIACAS", *1st International Conference on Applied Robotics for the Power Industry (CARPI)*, Montreal, pp. 1-7.
- Hu, Q., Paisitkriangkrai, S., Shen, C., Hengel, A.V.D. and Porikli, F. (2016), "Fast detection of multiple objects in traffic scenes with a common detection framework", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17 No. 4, pp. 1002-1014.
- Hyunho, L., Changhwan, K., Sunggi, P., Seok, L., Jaehun, K., Sunho, K. and Taikjin, L. (2013), "Vision-based automatic real time inspection of power transmission line", *44th International Symposium on Robotics*, Seoul, pp. 1-3.

- Jain, A.K. (2011), *Fundamentals of Digital Image Processing*, Prentice Hall, NJ.
- Ludan, W., Hongguang, W., Lijing, F. and Mingyang, Z. (2007), "A control method for grasping the power line transmission based on visual servoing for line inspection robot robot", *Robot*, Vol. 29 No. 5, pp. 451-455.
- Nanni, L., Lumini, A. and Brahmam, S. (2012), "Survey on LBP based texture descriptors for image classification", *Expert Systems with Applications*, Vol. 39 No. 3, pp. 3634-3641.
- Ojala, T., PIETIKÄINEN, M. and MÄENPÄÄ, T. (2002), "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24 No. 7, pp. 971-987.
- Ojansivu, V. and Heikkilä, J. (2008), "Blur insensitive texture classification using local phase quantization", *Image and Signal Processing*, Springer, pp. 236-243.
- Pouliot, N. and Montambault, S. (2012), "Field-oriented developments for LineScout technology and its deployment on large water crossing transmission lines", *Journal of Field Robotics*, Vol. 29 No. 1, pp. 25-46.
- Pouliot, N., Richard, P. and Montambault, S. (2012), "LineScout power line robot: characterization of a UTM-30LX LIDAR system for obstacle detection", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, Vilamoura, pp. 4327-4334.
- Premebida, C., Ludwig, O. and Nunes, U. (2009), "LIDAR and vision-based pedestrian detection system", *Journal of Field Robotics*, Vol. 26 No. 9, pp. 696-711.
- Richard, P.L., Pouliot, N. and Montambault, S. (2014), "Introduction of a LIDAR-based obstacle detection system on the LineScout power line robot", *Journal of Endourology*, Vol. 28 No. 3, pp. 330-334.
- Shen, F.M., Shen, C.H., Zhou, X., Yang, Y. and Shen, H.T. (2016), "Face image classification by pooling raw features", *Pattern Recognition*, Vol. 54, pp. 94-103.
- Song, Y., Wang, H. and Zhang, J. (2014), "A vision-based broken strand detection method for a power-line maintenance robot", *IEEE Transactions on Power Delivery*, Vol. 29 No. 5, pp. 2154-2161.
- SunSin, H. and JangMyung, L. (2008), "Path-selection control of a power line inspection robot using sensor fusion", *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Seoul, pp. 8-13.

- Wang, B., Li, W., Yang, W. and Liao, Q. (2011), "Illumination normalization based on weber's law with application to face recognition", *IEEE Signal Processing Letters*, Vol. 18 No. 8, pp. 462-465."
- Wang, W., Wu, G.P., Bai, Y.C., Xiao, H., Yang, Z.-Y., Yan, Y., He, Y., Xu, X.-J. and Su, F. (2014), "Hand-eye-vision based control for an inspection robot's autonomous line grasping", *Journal of Central South University*, Vol. 21 No. 6, pp. 2216-2227.
- Weibing, G., Hongguang, W., Yong, J. and Peng, S. (2012), "A transmission line inspection robot's automatic line grasping visual servo control", *Robot*, Vol. 34 No. 9, pp. 620-627.
- Wenming, C. (2014), *Research on Visual Control Methods for High Voltage Transmission Line Deicing Robot*, HuNan University.
- Xinglong, Z., Hongguang, W. and Fang, L. (2006), "Algorithm research of inspection robot for searching for poseof overhead ground wires", *6th World Congress on Intelligent Control and Automation*, Dalian, pp. 7513-7517.
- Yuan, L., Chen, F., Zhou, L. and Hu, D. (2015), "Improve scene classification by using feature and Kernel combination", *Neurocomputing*, Vol. 170, pp. 213-220.
- Yunchu, Z., Zize, L., Ming, T. and Wenbo, Y. (2007), "An obstacle-navigation visual servo control method for overhead transmission line patrol robot robot", *Robot*, Vol. 29 No. 2, pp. 111-116.
- Zhenhui, L., Honggunag, W., Chaoyue, W., Yong, J. and Xiang, Y. (2015), "Line-grasping control for a power transmission line inspection robot", *Journal of Jilin University (Engineering and Technology Edition)*, Vol. 45 No. 5, pp. 1519-1526.
- Zhongwei, C., Hua, X. and Gongping, W. (2006), "A navigation methods based on electromagnetic sensors for high pressure patrol robot", *Transducer and Microsystem Technologies*, Vol. 25 No. 9, pp. 33-35.
- Zhou, S.-R., Yin, J.-P. and Zhang, J.-M. (2013), "Local binary pattern (LBP) and local phase quantization (LBQ) based on Gabor filter for face representation", *Neurocomputing*, Vol. 116, pp. 260-264.

Corresponding authors

Xuhui Ye can be contacted at: xhye@whu.edu.cn and **Gongping Wu** can be contacted at: gpwu@whu.edu.cn