Fast Detection of Hidden Dangers in Transmission Line Corridors Based on Region Partitioning and Feature Extraction

Wenjie Zheng¹, Chengqi Li¹, Demeng Bai¹ and Jiafeng Qin¹

¹State Grid Shandong Electric Power Research Institute, Jinan, 250002, China

zhwjchww@163.com

Abstract. Hidden dangers like large-scale construction machinery are the main causes of line trips. In this paper, we propose a fast detection algorithm for hidden dangers in transmission line corridors based on region partitioning and feature extraction. Since the scenes in the sky are simple and stable compared with the scenes on the ground, we detect hidden dangers in the sky and those on the ground separately. For hidden dangers in the sky, we firstly designed an algorithm to calculate the mask image for sky area. After dividing the sky area, we extracted features including colors and shapes of the difference areas in order to eliminate the interference factors. The targets left are the hidden dangers we truly cared about. For hidden dangers on the ground, firstly we design a pre-processing algorithm to eliminate the influence of uneven illumination on subsequent matching. Secondly, we propose a multi-scale gray-weighted average method to fuse multiple channels of multiple color spaces, by which we can effectively suppress noise caused by camera shake and maximize the area that cannot be matched. Thirdly, we use the Haar feature density map to filter the photo to remove discrete pixel points caused by small disturbances. Finally, we extract multiple features and fuse those features on decision level to obtain the final matching result. When deployed on intelligent monitoring devices and tested with massive scene photos, our algorithm has achieved the expected running efficiency and detection accuracy.

1. Introduction

Destruction of external forces such as the illegal operation of tower cranes is the main cause for line trips. To improve the capability to protect overhead transmission lines, the transmission line channel visualization technologies are raised in recent years. Monitoring devices deployed on power transmission towers use solar energy to independently supply power, which can not only monitor the overhead transmission lines, but also monitor the hidden dangers in the protection area. After one photo is captured, it is transmitted to the server through the LTE network. Then, the server performs hidden danger identification and SMS alert. Transferring high-definition photos consumes a lot of power and data traffic, which reduces device availability. Therefore, we study lightweight hidden danger detection algorithm that can be deployed on monitoring devices. The photo is transferred to the server only after hidden dangers are detected, by which we can not only save traffic and power, but also improve the frequency of monitoring.

Aiming at detecting hidden dangers for overhead transmission lines, researchers have proposed many related algorithms. Zhang Wei proposed a transmission line tree detection algorithm based on image processing [1]. The texture photo segmentation was realized by the combination of texture analysis and
threshold segmentation. The Sobel operator was used to detect the edge of the tree area. However, due to the influence of smog or rainy weather, the texture information on the ground area near the skyline is covered. Thus, this method has a large error. Peixiang Yu proposed a power image online monitoring algorithm [2]. The stable period photo is used as the background to calculate the difference between the real-time photo and the background photo. The maximum probability criterion is used to determine the abnormal condition. However, the method is only applicable to the substation with a single scene. The transmission line environment is complex, the target variety is various, and the size and shape of the area are uncertain, which brings great difficulty to the area selection. Chengqi Li proposed a binary photo filtering for object detection based on Haar feature density map [3], which can filter the noise caused by small disturbances in two photos, but cannot determine whether the unmatched connected domains of large blocks are caused by human factors or natural factors. In addition, there are many neural network-based photo recognition technologies [4][5][6][7], but those algorithms are not suitable for solar-powered monitoring devices limited by power consumption.

By comprehensively analyzing hidden dangers for overhead transmission lines, we find that the scenes in the sky area are simple and stable compared with those on the ground. Therefore, hidden dangers in the sky, such as tower cranes and crane arms, are distinguished from other hidden dangers on the ground. We propose a fast detection algorithm to detect hidden dangers based on region partitioning and feature extraction. Firstly, we use the Gaussian background model to obtain the final sky area mask photo, and combine the color and shape information to detect and determine the hidden dangers in the sky area. Secondly, for the hidden dangers on the ground, we use the multi-feature fusion method to realize the fast and intelligent detection of channel hidden dangers under complex background. Experimental results show that the proposed algorithm has both satisfactory running efficiency and high detection accuracy.

2. Detection of hidden dangers in the sky based on background modeling and feature extraction
Tower cranes and crane arms in the sky are the most common hidden dangers for overhead transmission lines, due to their high probability of touching these lines. Considering the detection of tower cranes and crane arms in the sky may be affected by clouds, wires, trees, buildings and mountains, we propose the following detection method based on background modeling and feature extraction.

2.1. Calculation of the mask image for sky area
The calculation of the mask image for sky area mainly has the following 4 steps, i.e.,

**Step 1:** We convert color photos to grayscale photos in order to reduce the amount of subsequent calculations. After that, median filtering is conducted on the above photos with a window size of 5×5.

**Step 2:** We use the Sobel [8] operator to calculate the gradient of the photo and set threshold for binary segmentation on the above photo:

\[
B(x, y) = \begin{cases} 
1, & \text{if } G(x, y) < T \\
0, & \text{otherwise} 
\end{cases}
\]

where \(T\) is the empirical threshold, which is set to be 15 in this study. \(G(x, y)\) is the gradient of the photo at the coordinate \((x, y)\). \(B\) represents the photo after segmentation.

**Step 3:** We choose the connected domain with the largest area and the highest centre of gravity as current sky area.

**Step 4:** We fuse the currently extracted sky area map with that of the historical photo, and then calculate the final sky area mask with the fused photo.

In fact, we can get the sky area mask of the current photo after the first three steps. However, the texture information on the ground area near the skyline can be covered when effected by haze or rain, which leads to large errors. Therefore, we fuse the sky mask of the current photo with the sky mask of the historical photo to get more accurate results.
As shown in Figure 1, compared with the distant mountain in photo (a) taken in sunny weather, the distant mountain in photo (b) taken in smog weather is indistinguishable due to the influence of smog. As a result, the sky mask (c) calculated from the photo (b) obviously counts the distant mountain as a part of the sky area. However, the sky mask (d) obtained after the fourth step can accurately identify the outline of the sky area.

2.2. Feature extraction of hidden dangers in the sky
After dividing the sky area, we analyze the difference areas in the sky and find that there are mainly four kinds of differences, i.e., tower cranes and crane arms, cloud, wire sloshing and shaking trees above the skyline. We consider tower cranes and crane arms as hidden dangers and others as interferences. Considering it is very difficult to find a unified classifier for these complex and diverse objectives, we exclude the above interferences based on their features. The final target we keep is the hidden danger we really care about. Feature extraction of hidden dangers in the sky includes 2 steps, i.e.,

Step 1: Interferences such as wires and trees, are inherent objects that exist for a long time. Due to the fixed position and angle of the camera, the positions of these objects in the photo are relatively fixed. We combine the sky mask of current photo with that of historical photos, and then calculate the final sky mask in the fused photo, in order to eliminate the influence of wire sloshing and shaking trees above the skyline.

Step 2: Clouds appear randomly and cannot be eliminated through background modeling. However, we find there are distinct features between clouds and cranes. Firstly, the color of clouds is white and the brightness is high, while the color of tower cranes and crane arms is usually gray or yellow. Furthermore, the outlines of clouds are curved, while these of tower cranes and crane arms are relatively straight. Based on these features, we can separate them by their colors and outlines.

3. Detection of hidden dangers on the ground based on multi-feature fusion
The scene on the ground is much more complicated than that in the sky, so it is difficult to detect hidden dangers using traditional matching algorithms. We propose the following detection methods based on multi-feature fusion. The main operations are summarized as follows.

3.1. Reference photo selection based on gamma correction and Euclidean distance
To reduce the impact of uneven illumination on subsequent processing, we use gamma correction [9] to improve the contrast of the photo:
where $Y$ is the pixel value in the corrected photo. $X$ is the original pixel value in the original photo. $e$ is the compensation coefficient. $\gamma$ is the gamma value.

Since our algorithm is based on matching two photos, large-scale illumination changes have a huge impact on photo matching. After analyzing massive photos taken by monitoring devices on transmission towers, we find there are three main features, i.e., the long photo interval, low probability of hidden dangers, and slow development of hidden dangers. So we propose a new method for illumination processing. Firstly, we select 13 photos including 6 recent photos, the photo of the same moment from the most recent day of the same weather, and the latest 6 photos before it. Secondly, we calculate the feature vectors of the photo $v_0$ and the 13 photos above $\{v_i\}$, $i = 1$–$13$. Thirdly, we calculate the Euclidean distance between the feature vector of current photo and that of each photo in the 13 photos above. The historical photo that has the highest similarity and closest time to the current photo is selected as a reference photo.

$$D_{\text{min}} = \min_{i=1,...,13} \left[ \sqrt{v_i - v_0} \right]$$

3.2. Photo Fusion based on image difference in multiple color spaces

3.2.1. Image difference in multiple color spaces. Single color component can only present very little information, for example, the R channel in the RGB color space is only sensitive to red. The difference of pixels with large difference in red component in the photo can be more obviously expressed in the R channel, while the difference of pixels with large difference in green or blue is not good. Similar problems exist in other color spaces.

Image difference is a method of matching photos at the pixel level using the size of pixel values. Since different color components of different color spaces are not identical to the same color, the difference in the different color components is also different.

3.2.2. Fusion of difference images. Fusion of differential images can preserve as much information as possible about the matching of pixels in different color channels. Based on which we propose a multi-scale gray-weighted average method to fuse photos:

$$F(x, y) = \frac{1}{N} \left( \frac{\alpha_1}{(2k_1 + 1)^2} \sum_{i=-k_1}^{k_1} \sum_{j=-k_1}^{k_1} I(x+i, y+j) \right) + \frac{1}{N} \left( \frac{\alpha_2}{(2k_2 + 1)^2} \sum_{i=-k_2}^{k_2} \sum_{j=-k_2}^{k_2} I(x+i, y+j) \right) + \frac{1}{N} \left( \frac{\alpha_3}{(2k_3 + 1)^2} \sum_{i=-k_3}^{k_3} \sum_{j=-k_3}^{k_3} I(x+i, y+j) \right),$$

where $I(x, y)$ is the pixel value of the original photo $I$ at the coordinate $(x, y)$. $F(x, y)$ is the pixel value of the photo $F$ after the fusion at the coordinate $(x, y)$. $N$ is the number of fused photos. $k_i$ is the size of the scale window. $\alpha_i$ represents the weight of the scale. In this study, we set $N=11$, $k = \{1,2,3\}$. $\alpha_i = \{0.1,0.2,0.7\}$.

3.3. Filtering the fused photo based on Haar feature density map

Image difference is a method of matching photos at the pixel level without taking into account the neighborhood relationship, so it is very sensitive to noise. We make use of Haar feature density map proposed in our another article [3] to filter the fused photo, by which we can not only remove discrete pixel points caused by small disturbances effectively, but also preserve the information in the connected region and its neighbors well.
3.4. Detection of hidden dangers using multi-feature fusion

After filtering the fused photo based on Haar feature density map, noise caused by small disturbances has been removed. However, further matching is required for large unconnected areas, defined as regions of interest (ROI), so as to determine whether ROI is caused by human factors. To improve matching accuracy, we extract multiple features of ROI and then fuse those features on decision level.

3.4.1. Segmentation of partial regions of photos. We mark and segment the partial region with the help of connected regions, and then segment the original photo portion within the circumscribed rectangle corresponding to the connected region.

3.4.2. Multi-feature extraction. After segmenting the ROI area, we extract and fuse four kinds of features, including geometric features, grayscale features, color features and texture features.

1) Geometric features. The area of the ROI is the total number of white pixels in the mask photo, and the connected domain with too small an area is usually caused by noise. In addition, the eccentricity is used to describe the extent of current ROI, with which we can filter out fake targets such as the high-voltage line in the sky and the edges of the buildings. Finally, we extract area and eccentricity as geometric features.

2) Grayscale features. Grayscale features extracted here include average gray value and gray variance. In general, two photos with large differences have larger differences in their gray mean values, and the average gray value can be used as a feature of photo matching, which can be calculated as follows:

\[
\mu = \frac{\sum_{i=0}^{m} \sum_{j=0}^{n} I(i, j) \cdot II(M(i, j))}{\sum_{i=0}^{m} \sum_{j=0}^{n} II(M(i, j))},
\]

where \( I \) is the grayscale photo. \( M \) is the mask photo of \( I \). \( II(\cdot) \) is the indication function, which equals 0 when the parameter value is 0 and equals 1 when the parameter value is bigger than 0.

After that we calculate the difference between the average gray value of the photo to be matched and the average gray value of the reference photo:

\[
\text{Diff}_\mu = \mu_m - \mu_r
\]

The gray-scale variance describes the degree of change in gray value of the photo. We calculate the grayscale variance under the definition of the mask photo, so as to obtain more accurate results, which can be expressed as follows,

\[
\sigma^2 = \frac{\sum_{i=0}^{m} \sum_{j=0}^{n} (I(i, j) - \mu)^2 \cdot II(M(i, j))}{\sum_{i=0}^{m} \sum_{j=0}^{n} II(M(i, j))}
\]

After that we calculate the difference between the variance of the photo to be matched and the average gray value of the reference photo:

\[
\text{Diff}_\sigma = \sigma_m^2 - \sigma_r^2
\]

3) Color features. We map the reference photo to multiple color spaces mentioned in chapter 3.2 and extract features separately in each color channel. After that, we extract the average color value and histograms of colors as the feature of colors.

4) Texture features. The above features can only reflect the statistical features of the pixels of photos, missing the relative positions of pixels. Textures describe the uneven grooves on the surface of the object, and can effectively describe the spatial position and intensity information of photo pixels. Texture
features used in this project include gray correlation coefficient and Laws texture energy at multi-scale windows.

The correlation coefficient is used to calculate the degree of matching between the two photos, which can be calculated as follows,

\[
C(x, y) = \frac{\sum_{i=-k}^{k} \sum_{j=-k}^{k} (I_1(x+i, y+j) - \mu_1) \cdot (I_2(x+i, y+j) - \mu_2)}{\left(\sum_{i=-k}^{k} \sum_{j=-k}^{k} (I_1(x+i, y+j) - \mu_1)^2\right) \cdot \left(\sum_{i=-k}^{k} \sum_{j=-k}^{k} (I_2(x+i, y+j) - \mu_2)^2\right)^{1/2}},\tag{9}
\]

where \( I_1 \) and \( I_2 \) are gray scale photos of the photo to be matched and the reference photo. \( C \) represents the correlation coefficient. When the corresponding pixel points of the two photos and their surrounding pixel values match well, the correlation coefficient is higher. In order to remove the influence of the matched region on the correlation coefficient, the correlation image can be masked after calculating the correlation coefficient map:

\[
D(x, y) = C(x, y) \cdot H(M(i, j))\tag{11}
\]

After that, we calculate the mean of the correlation coefficient map as a matching feature when matching the matching image with the reference image, defined as MeanCorr-Gray.

The distribution map of texture energy can reflect the distribution of energy in the photo. Smaller scale of the calculated texture energy takes more accurate results. As the scale increases, the energy distribution will spread near the peak. The diffusion of energy enhances the difference between the photo to be matched and the reference photo. Inspired by Laws texture energy analysis [10], we propose a new multi-scale texture energy analysis method to improve the matching accuracy. This method enables multi-scale analysis by changing the size of the template rather than scaling the image, enabling a more efficient description of the texture at multiple scales. Detailed steps are illustrated as follows.

Firstly, we calculate the energy distribution maps of the photo to be matched and the reference photo at four different scales \{1, 2, 3, 4\}.

Secondly, we calculate the average value of each energy distribution map of the two photos, i.e., Mean1 = \{mean1, mean2, mean3, mean4\}, Mean2 = \{mean1, mean2, mean3, mean4\}, as well as the energy of each energy distribution map, i.e., Energy1 = \{energy1, energy2, energy3, energy4\}, Energy2 = \{energy1, energy2, energy3, energy4\}.

Thirdly, we calculate the correlation coefficient between Mean1 and Mean2, as well as the correlation coefficient between Energy1 and Energy2, which are expressed as Corr-Mean and Corr-Energy, respectively.

Fourthly, we calculate the correlation coefficient of the energy distribution maps of the photo to be matched and the reference photo at four scales, i.e., Corr-MultiEnergy = \{correlation1, correlation2, correlation3, correlation4\}.

Finally, we choose the following texture features for the final photo matching: Texture = \{MeanCorr-Gray, Corr-Mean, Corr-Energy, Corr-MultiEnergy\}.

3.4.3. Fusion of multiple features. We train and classify the photos using different features separately, and combine the matching values to obtain the final match result, which is shown in figure 2.
We select 1000 photos and classify them using 4 types of features and tolerance features. The results are demonstrated in Table 1.

**Table 1.** Classification results using various features.

| Feature Type       | Geometric | Grayscale | Color | Texture | Features fusion |
|--------------------|-----------|-----------|-------|---------|-----------------|
| Classification     | 64.35%    | 60%       | 76.1% | 71.7%   | 87.5%           |

**4. Experiments**

In order to verify the effectiveness of the above algorithm, we tested 20,000 photos taken by 200 cameras installed on various power transmission towers. Due to the wide distribution of power transmission towers, differences between those photos are very large. The scenes are divided into 4 categories according to the frequency of human activities:

- **Scene 1:** Scenes almost without human activities, including wild, mountainous and large farmland.
- **Scene 2:** Scenes with moderate level of human activity, which are usually in the outer suburbs containing part of roads.
- **Scene 3:** Scenes with frequent human activities in local areas such as suburban factories.
- **Scene 4:** Scenes with very frequent human activities such as resident settlements.

We categorized the 200 scenes according to the above 4 types. The number and ratio of various scenes are shown as follows,

**Table 2.** Number and proportion of various types of scenes

| Scene   | Number | Ratio   |
|---------|--------|---------|
| Scene 1 | 151    | 75.5%   |
| Scene 2 | 22     | 11.0%   |
| Scene 3 | 19     | 9.5%    |
| Scene 4 | 8      | 4.0%    |

From Table 2, we can see that field scenes account for more than three quarters of the total number of scenes, while Scene 4 with the most frequent human activities account for only about 4% of the total.

We define the number of detected photos containing unmatched ROI areas as the number of detections, and the number of photos with security risks but not detected by the algorithm as the number of missed detections, and the number of photos with security risks detected by the algorithm and confirmed by people as the correct number. Table 3 shows the detection results of 20000 photos for these 4 types of scenarios.

**Table 3.** Test results of five types of scenarios.

| Scene  | Total | Quantity detected | Quantity missed | Quantity correct |
|--------|-------|-------------------|-----------------|------------------|
| Scene 1| 15515 | 305               | 37              | 152              |
| Scene 2| 1685  | 1214              | 25              | 230              |
| Scene 3| 1700  | 1383              | 19              | 310              |
| Scene 4| 1100  | 875               | 13              | 158              |
From the table above we can see that for scene 1, our algorithm can exclude 97% of photos, which means reducing the workload of workers by 97%. The detection rate and missed rate are shown in figure 3 and figure 4 as follows:

**Figure 3.** Detection rate in each scene

**Figure 4.** Missed rate in each scene

From the figures above we can see that for all four types of scenes, the algorithm can eliminate 82.1% of the photos, which can save about 81% of the labor cost. The monitoring device takes about 400 milliseconds on detecting each picture, which can meet the requirements of field application on transmission towers.

5. Conclusion

In this paper, we propose a fast detection algorithm for hidden dangers in transmission line corridors based on region partitioning and feature extraction. We firstly deal with hidden dangers in the sky with background modeling and feature extraction, and then we detect hidden dangers on the ground by fusing multiple features. When deployed on intelligent monitoring devices and tested with massive scene photos, the results proved the efficiency and detection accuracy of our algorithm. Next, we will design an efficient scheduling algorithm to rationally allocate computing resources at the front and back ends so as to improve the availability and effectiveness of our transmission line channel visualization remote inspection system.

Acknowledgment

This paper is sponsored by State Grid Science and Technology Project (520626180067).

References

[1] Zhang Y, Huang X B and Zhou K H 2013 Study on transmission line tree detection algorithm based on image processing Guangdong Electric Power 26(09) pp 26-31
[2] Yu P X, Wu Z, Lin J, et al 2013 An efficient power image online monitoring algorithm East China Electric Power 41(09) pp 1985-6
[3] Li C Q, Ren Z G and Yang B 2017 Binary image filtering for object detection based on Haar feature density map Proc. SPIE 10613 1061303
[4] Tang C, Feng Y S, Yang X, et al 2017 The Object Detection Based on Deep Learning Proc. Int. Conf. on Information Science & Control Engineering pp 723-8
[5] Rahman M, Tan Y H, Xue J, et al 2017 RGB-D object recognition with multimodal deep convolutional neural networks Proc. IEEE Int. Conf. on Multimedia & Expo vol 0 pp 991-6
[6] Yan L, Wang Y X, Song T Z, et al 2017 An incremental intelligent object recognition system based on deep learning Proc. 2017 Chinese Automation Congress vol 2017-January pp 7135-8
[7] Wang J H, Lu J, Chen W, et al 2015 Convolutional neural network for 3D object recognition based on RGB-D dataset Proc. IEEE Conf. on Industrial Electronics and Applications pp 34-39
[8] Deng C X, Ma W F and Yin Y 2011 An edge detection approach of image fusion based on improved Sobel operator Proc. Int. Congr. on Image & Signal Processing vol 3 pp 1189-93
[9] Lee D U, Cheung R C and Villasenor J D 2007 A flexible architecture for precise gamma correction IEEE Transactions on Very Large Scale Integration Systems 15(4) pp 474-8
[10] Govindaraj P and Sudhakar M S 2017 Shape characterization using laws of texture energy measures facilitating retrieval The Imaging Science Journal 66(2) pp 1-8