Construction Crane Detection under Transmission Line Based on Improved Vibe Algorithm

Xinwen Li, Li Lu *
School of Electrical Engineering, Shanghai Dian Ji University, Shanghai, China

*Corresponding author e-mail: lul@sdju.edu.cn

Abstract. In this paper, the moving object detection part of crane’s illegal construction detection under transmission line is studied, and two scenes are simulated. The Visual background extraction (Vibe) algorithm is selected as the moving object detection algorithm used in this paper. For the problems of the algorithm, including ghost phenomenon in detection results and the presence of falsely detected foreground points under dynamic background, the Vibe algorithm is improved by the reconstructed background model, dynamic radius threshold and connected domain analysis. Results of the experiment show that ghost area in the detection results of the improved Vibe algorithm can quickly disappear, and the number of falsely detected foreground points is significantly reduced. Compared with the original Vibe algorithm, the average value of the comprehensive evaluation index F1 is increased by 5.8%, and the detection effect of moving object is significantly improved.

1. Introduction

With the continuous improvement of China's economic level, the scale of transmission lines is also expanding. In its 13th five year plan, the State Grid Corporation of China proposed to add 401000 km of 110kV and above voltage level lines, and the total mileage of transmission lines has increased by 45% compared with the 12th five year plan [1]. At the same time, transmission line accidents often occur, and illegal construction of crane is one of the main reasons for transmission line accidents. According to statistics, more than 80% of transmission line accidents in a province in one year are caused by illegal construction of large cranes and floating objects attached to the lines [2]. With the continuous construction of infrastructure, the number of transmission line accidents caused by illegal construction of crane is increasing significantly, which brings great economic losses to power supply enterprises and affects people's production and life.

The traditional methods to prevent illegal construction of crane include assigning special person to patrol the line, building a transmission line video monitoring system for centralized monitoring, etc. These methods mainly rely on manual judgment of crane’s illegal construction behavior, sometimes can't timely alarm, and the consumption of human and material resources is huge. Under the background of intelligent power grid, it has become a new research direction to integrate computer vision algorithm into the crane’s illegal construction detection under transmission line [3].
2. Research scenarios of this paper

This paper mainly studies the moving object detection part of the crane’s illegal construction detection under transmission lines. Two scenes are simulated and two video sequences are obtained, which are referred to as crane stay and crane arm too high, the two video sequences contain the following scenes.

(1) After the crane enters the monitoring screen, the moving object detection algorithm is used to detect the crane foreground. If the crane appears in the monitoring screen for a short time, the monitoring system outputs low-risk alarm information to remind the operator to pay attention to the crane entering the monitoring screen; if the crane appears in the monitoring screen for a long time, it indicates that the crane has a tendency to stay for a long time, the monitoring system outputs high-risk alarm information to remind the operator to pay attention to the follow-up action of the crane and take corresponding measures according to specific situation, as shown in Figure 1, in this video sequence, the phenomenon of the target crane appears in the first frame and the slight shaking of transmission tower, wires and leaves in the background are simulated.

(2) During the lifting operation, due to the negligence of the driver, the crane arm is raised too high when the crane enters the monitoring screen. At this time, the safe distance between the top of the crane arm and the high-voltage line is not satisfied, which is easy to cause high-voltage line discharge, resulting in wire damage and electric shock accidents. Therefore, it is necessary to use the moving object detection algorithm to detect the foreground of the crane. According to whether the safe distance between the top of the crane arm and the transmission line is satisfied and the duration of the crane arm too high state, the alarm is classified, as shown in Figure 2. In this video sequence, the phenomenon of slight shaking of transmission towers, wires and leaves in the background is simulated.

3. Principle of Vibe algorithm

Visual background extraction algorithm, Vibe algorithm for short, is a pixel level moving object detection method combining temporal and spatial information. The algorithm initializes background
model through the first frame, and detects the foreground from the second frame. The calculation process of the Vibe algorithm is simple and the real-time performance is good. The specific implementation process of Vibe algorithm is as follows.

3.1. Initialization of background model
Vibe algorithm initializes the background model of each pixel through the first frame of video sequence. For the pixel \( x \) in the first frame, the algorithm selects the pixels in the 8 neighborhood of pixel \( x \) as the sample set, then the algorithm randomly selects the pixel values of 20 pixels in the sample set as the background sample set of pixel \( x \) to complete the initialization of the background model.

The background sample set of pixel \( x \) is shown in formula (1).

\[
M(x) = \{v_1, v_2, \ldots, v_n\}
\]

In formula (1):
- \( M(x) \) —— The background sample set of pixel \( x \);
- \( v_i \) —— Sample point in the background sample set;
- \( n \) —— Number of sample points, its value is 20.

3.2. Foreground detection
Vibe algorithm starts foreground detection from the second frame. For pixel \( x \), set its pixel value as \( V_t(x) \), and \( t \) means that the current frame number is \( t \). Construct a circle \( S_R(V_t(x)) \) with \( V_t(x) \) as the center and \( R = 20 \) as the radius, as shown in Figure 3.

In Fig. 3, \( V1, V2, \ldots \) are the sample points in the background sample set of pixel \( x \). The distance between the pixel value \( V_t(x) \) of pixel \( x \) in the current frame and the corresponding background sample points is calculated, and the variable \( H \) is defined as the number of background sample points falling within the range of \( S_R(V_t(x)) \), as shown in formula (2).

\[
H = H + 1 \text{ if } \text{dist}(V_t(x), V_i) < R
\]

In formula (2):
- \( \text{dist}(V_t(x), V_i) \) —— Difference of pixel values between \( x \) and background sample point.

If the difference of pixel values is less than the radius threshold \( R \), the variable \( H \) is added with 1. Traverse the background sample set of pixel \( x \), calculate the difference of pixel values respectively, count the value of variable \( H \), and judge whether the pixel \( x \) is the foreground point, as shown in formula (3).
The formula (3) shows that if the number of background sample points falling in the circular range \( S_R(V_i(x)) \) is less than the background matching threshold \( H \min \), then the pixel \( x \) is the foreground point, its pixel value is set to 255; otherwise, the pixel \( x \) is the background point, its pixel value is set to 0, and the value of background matching threshold \( H \min \) is 2.

\[
V_i(x) = \begin{cases} 
255 & H < H \min \\
0 & \text{else}
\end{cases}
\]  

(3)

3.3. Background model update
The Vibe algorithm uses random sampling strategy to update its background model. When the pixel is judged as a background point, there is a \( \frac{1}{16} \) probability to update its own background sample set, the update method is to replace the randomly selected sample value in its background sample set with the value of the pixel. At the same time, there is a \( \frac{1}{16} \) probability to update the background sample set of pixels in the 8 neighborhood of the pixel, the update method is to select a pixel in the neighborhood with \( \frac{1}{16} \) probability, and replace the randomly selected sample value in the sample set of selected neighborhood pixel with the value of the pixel [4, 5].

4. Problems of Vibe algorithm
Vibe algorithm uses the first frame of video sequence to initialize the background model. If there is a moving object in the first frame, when initializing the background model, the algorithm will include the pixels of the moving object area into the background model. In the subsequent detection process, the algorithm will mistakenly detect the moving target area in the first frame as the foreground points, resulting in ghost phenomenon, which affects the accuracy of moving target detection and increases the processing time of the subsequent algorithm, as shown in Figure 4 [6].

![Frame 1 of crane stay video](image1)

(a) Frame 1 of crane stay video

![Processing result of Vibe algorithm at frame 69](image2)

(b) Processing result of Vibe algorithm at frame 69

**Figure 4.** Ghost phenomenon of Vibe algorithm.

In the process of detecting the moving crane, due to the influence of the site environment, there is a slight shaking of the transmission tower, wires and leaves, forming the dynamic background [7]. Vibe algorithm mistakenly detects the dynamic background as foreground points, which not only affects the accuracy of moving target detection, but also affects the subsequent processing process, and increases the processing time of the subsequent algorithm, as shown in Figure 5.
5. **Improved Vibe algorithm**

5.1. **Reconstruct the background model**

This paper improves the initialization method of background model of Vibe algorithm for the ghost phenomenon in the detection results. In this paper, we take the first four frames of video sequence, for pixel \((x, y)\) in each frame, we take 5 pixel values in the 8 neighborhood of \((x, y)\) as sample values, a total of 20 sample values constitute a new background model. The reconstructed background model reduces the proportion of the number of moving target pixels in the first frame to the total number of background model samples, which can speed up the ghost elimination, as shown in Figure 6.

$$\bar{M} = \left\{ V_1, V_2, \ldots, V_n \right\}$$

$$\overline{\bar{M}} = [\sum_{i=1}^{n} V_i] / n$$

5.2. **Dynamic radius threshold**

In order to weaken the influence of dynamic background, this paper uses dynamic radius threshold for foreground detection. The variance of the reconstructed background model is calculated as follows.

$$M = \{ V_1, V_2, \ldots, V_n \}$$

$$\overline{\bar{M}} = [\sum_{i=1}^{n} V_i] / n$$

![Figure 5. Detection results of Vibe algorithm under dynamic background.](image)

![Figure 6. Improved initialization method of background model.](image)
\[
\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (V_i - \bar{M})^2
\]

In the above formulas:
\(M\) ——— The reconstructed background model;
\(n\) ——— The number of sample points in the reconstructed background model, its value is 20;
\(\sigma^2\) ——— The variance of the reconstructed background model.

Calculate the average distance between the pixel \(x\) in the first frame of the detection process and the sample points of the corresponding reconstructed background model, as shown in formula (7).

\[
disavg = \frac{1}{n} \sum_{i=1}^{n} |V_i - V_i'|
\]

In formula (7):
\(V_i\) ——— The value of the pixel \(x\) in the first frame of the detection process;
\(V_i'\) ——— Sample value of the reconstructed background model corresponding to the pixel \(x\).

Define the background complexity of the current frame, as shown in formula (8).

\[
fzd = \alpha \cdot \sigma^2 + (1 - \alpha) \cdot disavg
\]

In formula (8), \(\alpha\) is the weighting coefficient, this paper takes 0.2, and the adjusted radius threshold is shown in formula (9).

\[
R = \begin{cases} 
20 + 0.5 \cdot zd & fzd \geq 20 \\
20 - 0.5 \cdot zd & fzd < 20 
\end{cases}
\]

The upper and lower limits of radius threshold \(R\) are given, as shown in formula (10).

\[
R = \begin{cases} 
10 & R \leq 10 \\
40 & R \geq 40 
\end{cases}
\]

5.3. Connected domain analysis

In order to further suppress the influence of dynamic background, this paper uses the two pass scanning method in connected domain analysis to mark the foreground of the detection result of the above improved methods, delete small foreground areas and retain the target construction crane.

The two pass scanning method scans the image twice. In the first pass, each foreground pixel position is given a label value. In the scanning process, foreground pixels belonging to the same connected area may be given different label values, therefore, it is necessary to record the equal relationship between different label values to merge them. In the second pass, the foreground pixels marked by the label values with equal relation are classified into a connected domain and given the same label value. Usually, the label value is the minimum of the label values with equal relation. Finally, all connected domains are found [8].

5.4. Flow chart of improved Vibe algorithm

The whole flow of the improved Vibe algorithm is shown in Figure 7:

![Flow chart of improved Vibe algorithm](image)

**Figure 7.** Flow chart of improved Vibe algorithm.
6. Experiment and analysis
Select the crane stay video and crane arm too high video, for the two video sequences, Vibe algorithm and the improved Vibe algorithm in this paper are used for moving object detection, and the detection results are shown in Figure 8.

![Frame comparison](image)

**Figure 8.** Comparison of detection results of two Vibe algorithms.

As shown in Fig. 8 (a), Fig. 8 (c) and Fig. 8 (e), in the crane stay video, due to the existence of the target crane in the first frame, the Vibe algorithm initializes the pixel value of the target crane into the background model, resulting in the ghost phenomenon in the detection result of the algorithm. With the update of the Vibe algorithm’s background model, the ghost area does not completely disappear in the 68th detection result of Vibe. Due to the influence of dynamic background, the number of other falsely detected foreground points in the detection result of Vibe algorithm is more than that of the improved Vibe algorithm. In this paper, the improved Vibe algorithm reduces the influence of the target crane in the first frame by reconstructing the background model. From the detection result of the improved Vibe algorithm, it can be seen that the ghost area basically disappears in the 68th frame, and the number of other falsely detected foreground points under the dynamic background is also significantly reduced.

As shown in Fig. 8 (b), Fig. 8 (d) and Fig. 8 (f), in the crane arm too high video, due to the influence of dynamic background such as the slight shaking of transmission tower, wires and leaves, there are
many falsely detected foreground points in the 78th detection result of Vibe algorithm. The improved Vibe algorithm suppresses the influence of dynamic background by setting the dynamic radius threshold and analyzing connected domains, the upper and lower limits of radius threshold are also given. The number of falsely detected foreground points in the 78th detection result of improved Vibe algorithm is significantly reduced. Since there is no target crane in the first frame of the video sequence, there is no ghost phenomenon in the detection results of the two Vibe algorithms.

The comprehensive evaluation index F1 value is introduced to compare the detection effect of Vibe algorithm and the improved Vibe algorithm in this paper. The index reflects the overall situation of moving object detection. The larger the F1 value is, the better the detection effect of the algorithm is. The definition of F1 value is shown in formula (11).

$$F1 = \frac{2PR}{P + R}$$

In formula (11), the variable $P$ is called Precision, it reflects the proportion of correctly detected foreground points in all detected foreground points. The larger the value of $P$ is, the less the number of falsely detected foreground points is.

In formula (11), the variable $R$ is called Recall ratio, it reflects the proportion of correctly detected foreground points in all real foreground points. The larger the value of $R$ is, the more complete the detected target is [9].

Read the detection results of crane stay video and crane arm too high video under Vibe algorithm and the improved Vibe algorithm in this paper, take 10 frames each to calculate the comprehensive evaluation index F1 value, and draw the F1 value change curve of detection results, as shown in Figure 9 and Figure 10.

![Figure 9. F1 value curve of crane stay video.](image1)

![Figure 10. F1 value curve of crane arm too high video.](image2)
Fig. 9 and Fig. 10 show that the improved Vibe algorithm in this paper speeds up the ghost elimination in the detection results of the algorithm and effectively suppresses the influence of the dynamic background by reconstructing the background model, setting the dynamic radius threshold and analyzing the connected domains. However, compared with the original Vibe algorithm, the detection results of the improved Vibe algorithm in two video sequences show a slight decrease in the integrity of the target crane. In the detection results of two video sequences, the number of falsely detected foreground points of the improved Vibe algorithm is less than that of the original Vibe algorithm, and the comprehensive evaluation index F1 values of the detection results of the improved Vibe algorithm in two video sequences are higher than that of the original Vibe algorithm, which shows that the detection effect of the improved Vibe algorithm is better than that of the original Vibe algorithm.

Calculate the average F1 values of the detection results of Vibe algorithm and the improved Vibe algorithm in two video sequences, as shown in Table 1.

| Moving object detection algorithms | Average F1 value of crane stay video | Average F1 value of crane arm too high video | Average F1 value |
|-----------------------------------|-------------------------------------|--------------------------------------------|-----------------|
| Vibe                              | 0.80                                | 0.92                                       | 0.86            |
| Improved Vibe                     | 0.88                                | 0.94                                       | 0.91            |

The data in Table 1 show that after the improved Vibe algorithm introduces the reconstructed background model, dynamic radius threshold and connected domain analysis, the speed of ghost elimination is accelerated, and the number of falsely detected foreground points caused by dynamic background is effectively reduced, but the integrity of the target crane in the detection results of improved Vibe algorithm in two video sequences is decreased. Compared with the original Vibe algorithm, the average F1 values of the improved Vibe algorithm in two video sequences are improved. The average F1 value of the improved Vibe algorithm is 0.91, which is 5.8% higher than the original Vibe algorithm. This shows the effectiveness of the improved Vibe algorithm.

7. Conclusion

This paper mainly studies the moving object detection part of the crane illegal construction detection under the transmission line, simulates two scenes, and obtains two video sequences. Vibe algorithm is selected as the algorithm used in this paper. This paper analyzes the problems of Vibe algorithm, including ghost phenomenon in detection results and the presence of falsely detected foreground points under dynamic background. The Vibe algorithm is improved by the reconstructed background model, dynamic radius threshold and connected domain analysis. In the two video sequences, the Vibe algorithm and the improved Vibe algorithm are used to detect moving objects. Results of the experiment show that compared with the Vibe algorithm, the ghost regions in the detection results of the improved Vibe algorithm can disappear quickly, and the number of falsely detected foreground points is also significantly reduced. The average F1 value of detection results of the improved Vibe algorithm is 0.91, which is 5.8% higher than that of Vibe algorithm, which verifies the effectiveness of the improved Vibe algorithm.

References

[1] ZHAO Zhenbing, QI Hongyu, NIE Liqiang. Research Overview on Visual Detection of Transmission Lines Based on Deep Learning [J]. Guangdong Electric Power, 2019, 32(09): 11-23.

[2] LU Jiazheng, ZHOU Tejun, WU Chuanping, et al. Fault Statistics and Analysis of 220kV and Above Power Transmission Line in Province-level Power Grid [J]. High Voltage Engineering, 2016, 42(01): 200-207.

[3] LIU Yangyu. Development trend of transmission line inspection based on machine vision [J].


Automation and Instrumentation, 2020(04): 183-186+190.

[4] Barnich O, Van Droogenbroeck M. ViBe: A universal background subtraction algorithm for video sequences [J]. IEEE Transactions on Image processing, 2010, 20(6): 1709-1724.

[5] WU Haiping, HOU Tao, NIU Hongxia. Real-time Monitoring of Railway Dynamic Foreign Body Intrusion Based on Improved ViBe Algorithm [J]. Journal of Lanzhou Jiaotong University, 2019, 38(05): 55-60+113.

[6] WANG Shuqing, ZHOU Bowen. Research on Moving Target Recognition Method Based on Improved ViBe Algorithms [J]. Journal of Hubei University of Technology, 2019, 34(04): 58-62.

[7] YANG Dan, DAI Fang. Improved ViBe algorithm for detection of moving objects [J]. Journal of Image and Graphics, 2018, 23(12): 1813-1828.

[8] LIU Chuan. Research on Intelligent monitoring system for transmission lines against external damage [D]. Xihua University, 2015.

[9] LIANG Shuo. Research on Motion Target Detection and Tracking Algorithms Based on Background Subtraction [D]. Xi'an Shiyou University, 2019.