

Research Article

An Algorithm of Occlusion Detection for the Surveillance Camera

Peng Shi , Bin Hou , Jing Chen , and Yunxiao Zu 

School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Yunxiao Zu; zuyx@bupt.edu.cn

Received 11 October 2020; Revised 30 January 2021; Accepted 15 February 2021; Published 22 February 2021

Academic Editor: Ting Yang

Copyright © 2021 Peng Shi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As more and more surveillance cameras are deployed in the Internet of Things, it takes more and more work to ensure the cameras are not occluded. An algorithm of detecting whether the surveillance camera is occluded is proposed by comparing the similarity of the images in this paper. Firstly, the background modeling method based on frame difference is improved. The combination method of the background difference and frame difference is proposed, and the experimental results showed that the combination algorithm can extract the background image of the video more quickly and accurately. Secondly, the LBP (Local Binary Patterns) algorithm is used to compare the similarity between the background image and the reference image. By changing the window size of the LBP algorithm and setting an appropriate threshold, the actual demands can be satisfied. So, the algorithms proposed in this paper have high application value and practical significance.

1. Introduction

In the context of the Internet of Things and communication technology being ever-changing, from smart home to smart city, the coverage of the Internet of Things is getting wider and wider. There are more and more surveillance cameras deployed in the Internet of Things. These surveillance cameras are closely related to many fields of our life and work. The surveillance cameras have many functions, such as live watching, video watching, and abnormal warning, which are very important for maintaining personal and social security. However, the camera will be occluded due to various accidents or human factors. For example, some criminals and suspicious people deliberately occlude the cameras in order to avoid being caught [1, 2]. Therefore, it has very important application value and practical significance to ensure the surveillance camera is not occluded.

At present, the methods detecting whether the surveillance camera is occluded are mainly based on the difference between frames [3, 4]. This kind of method is aimed at the monitoring image changing significantly in a short time when the camera is occluded, so it can detect whether the camera is occluded by comparing the difference between frames. But any error will lead to an inaccurate detection

result; the rates of false negatives and false positives are high and the application range is small.

In actual application, the scene monitored by the same camera is unchanged, so the monitoring image can be divided into two parts: the unchanged background image and the changed foreground image. Based on this feature, the methods [5, 6] that determine whether the camera is occluded or not are proposed by many scholars. These methods mainly measure the difference between the current frame image and the reference image by using some appropriate image feature vectors. However, the presence of foreground in the current frame image would have a certain impact on the detection results.

In view of the above methods' shortcoming, the existing background modeling method based on frame difference is improved in this paper by combining the background difference. This improved method can be used to extract the background image of the video more quickly and accurately. Compared with other image features, LBP (Local Binary Patterns) feature is easier to extract, the calculation of the LBP feature is simpler, and the accuracy of the LBP feature is higher. Therefore, the LBP feature is used to construct a feature vector to measure the similarity between the background image and the reference image in this paper. And

whether the surveillance camera is occluded can be determined according to the similarity. The calculation of occlusion detection based on comparison of image similarity is simple, so it is easy to be implemented. Occlusion detection based on comparison of image similarity not only effectively eliminates the influence of foreground but also is robust to illumination.

2. Materials and Methods

2.1. The Principle of the Background Modeling Algorithm Based on Background Difference and Frame Difference. In machine vision, background modeling is the basic technology of video processing. In recent years, a large number of background modeling methods have been proposed by many experts and scholars. Currently, the widely used background modeling methods mainly include median background modeling method [7], mean background modeling method, Gaussian distribution background modeling method [8], and ViBe algorithm [9]. Not only do these algorithms require a large number of video frames for background modeling but also the calculation is complex and requires a long time to extract the video background, which is difficult to meet the real-time requirements when detecting whether the surveillance camera is occluded.

The frame difference method is the most primitive and the simplest background modeling method. The frame difference method can be implemented quickly and it has a wide range of applications. In literature [10], an improved frame difference method is proposed for background modeling of video. Combined with the idea of time series statistics, the background model is established by counting the number of continuous frames. If the gray value of a pixel changes little for several continuous frames, the gray value is considered as the gray value of the background at the point. If the gray value of a pixel changes a lot between two adjacent frames, the gray value is considered as the gray value of the foreground. The gray value of the two adjacent frames will continue to be compared until the gray value of the background at this point can be determined. For the background point whose gray value cannot be determined for a long time, the gray value of the pixel point in the last frame is taken as the gray value of the background point. Because the area is small, it has little influence on the background model.

The above-improved frame difference method still needs a certain number of video frames for background modeling. So the real-time requirements still cannot be satisfied when detecting whether the surveillance camera is occluded. In this paper, the background modeling method is further improved by combining the background difference method. The gray value of the current frame is compared with the gray value of the reference image at the same pixel point. If the difference is small, the gray value of the current frame is considered as the gray value of the background at this point. If the difference is large, the improved frame difference

method in literature [10] is used to determine the gray value of the background at this point.

In the improved frame difference method in literature [10], determining whether the gray value of every pixel point is the gray value of the background at this point needs at least *Thread2* frames (*Thread2* is an adaptive threshold). In the background modeling algorithm based on background difference and frame difference, determining whether the gray value of every pixel point is the gray value of the background at this point only needs at least 1 frame. So, the background modeling algorithm based on background difference and frame difference is easier to meet the real-time requirements.

2.2. The Process of the Background Modeling Algorithm Based on Background Difference and Frame Difference. The specific process of the background modeling algorithm based on background difference and frame difference is shown in Figure 1.

Img1 is the previous frame image, *Img2* is the current frame image, *BImg* is the background image which needs to be built, *RImg* is the reference image extracted from an unoccluded surveillance video, *BimgFlag* is the number of continuous frames whose gray value change little at the same pixel point, *Flag* is the number of pixels whose gray value is not determined in the background image. In the background modeling algorithm based on background difference and frame difference, there are three adaptive thresholds *Thread1*, *Thread2*, and *Thread3*: *Thread1* is the difference threshold of two frames, *Thread2* is the number threshold of continuous frames whose gray value changes little at the same pixel point, and *Thread3* is the number threshold of pixels whose gray value is not determined in the background image. The specific process is as follows:

- (1) Initialize: The first frame image *Img1* and the reference image *RImg* are loaded, and both the values of *BImg* and *BimgFlag* are set as the full-zero matrix. The value of *Flag* is set as the number of pixels in the image *Img1*.
- (2) Judge the number threshold of continuous frames: a new frame image *Img2* is loaded. For every pixel point in *Img2*, if $BimgFlag(x) < Thread2$, let $DiffImg(x) = |Img2(x) - RImg(x)|$ and turn to Step 3. Otherwise, Step 2 will be performed to continue the iteration.
- (3) Judge the difference threshold between the current frame and the reference image: for each pixel that meets the condition in Step 2, if $DiffImg(x) < Thread1$, let $BimgFlag(x) = Thread2$, $BImg(x) = Img2(x)$, $Flag = Flag - 1$, and turn to Step 5. Otherwise, let $DiffImg(x) = |Img2(x) - Img1(x)|$ and turn to Step 4.
- (4) Judge the difference threshold between the current frame and the previous frame: for each pixel that

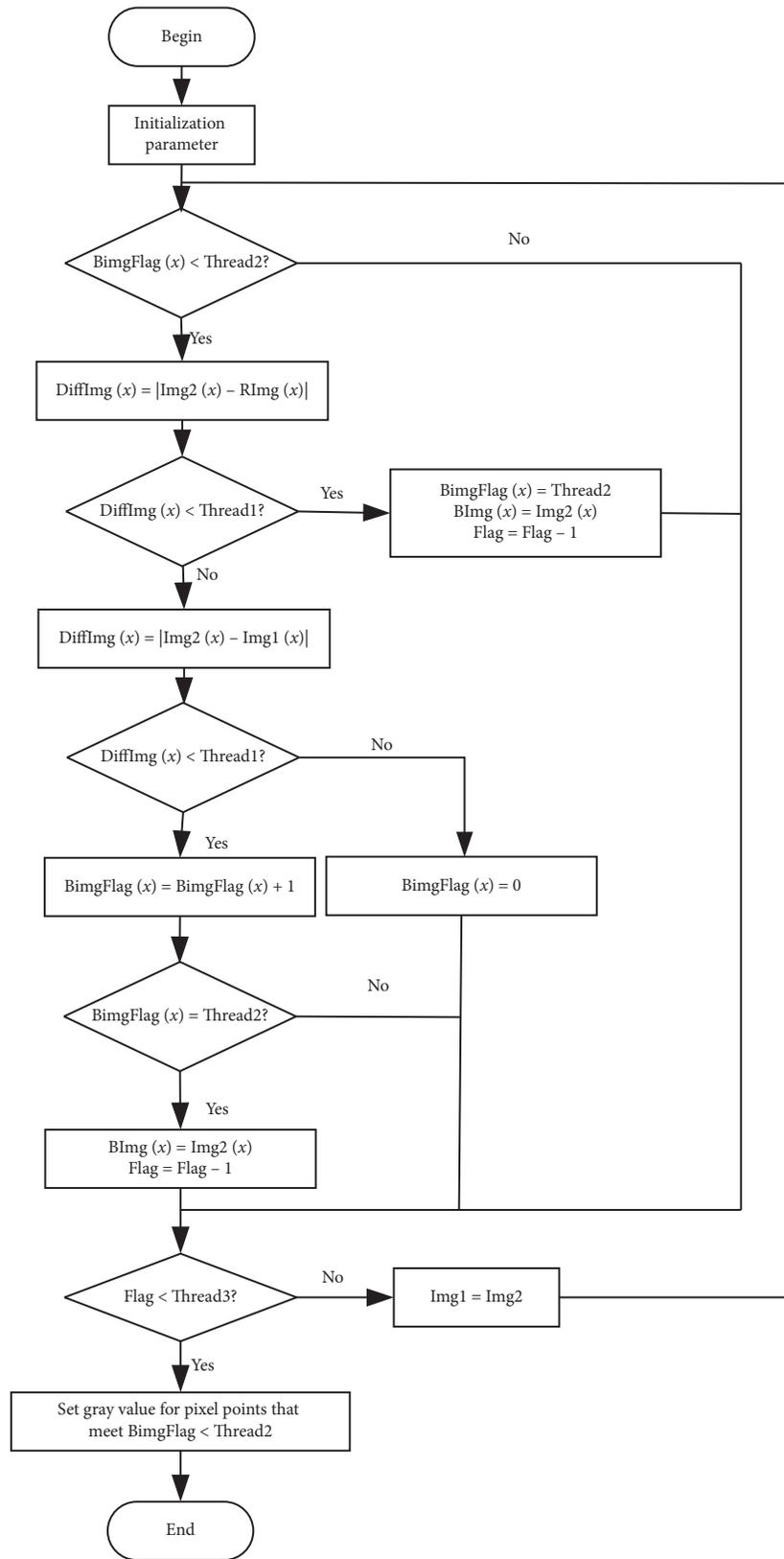


FIGURE 1: The specific process of the background modeling algorithm.

meets the condition in Step 3, if $\text{DiffImg}(x) < \text{Thread1}$, let $\text{BimgFlag}(x) = \text{BimgFlag}(x) + 1$. Otherwise, let $\text{BimgFlag}(x) = 0$. If $\text{BimgFlag}(x) = \text{Thread2}$, let $\text{BImg}(x) = \text{Img2}(x)$, $\text{Flag} = \text{Flag} - 1$.

- (5) Judge the iterative condition: When $\text{Flag} < \text{Thread3}$, it shows that the constructed background image has met the requirements. Then, the gray value of the pixel points which meet $\text{BimgFlag} < \text{Thread2}$ in the background image will be set. Otherwise, let $\text{Img1} = \text{Img2}$ and turn to Step 2, and the iteration is done.

In the background modeling algorithm based on background difference and frame difference, the threshold Thread1 is the key factor to determine the gray value of the background pixel point. Its value depends on the difference in the gray value between the background and the foreground. The threshold Thread2 mainly depends on the speed of the object moving in the video. The slower the object moves, the larger the value should be. The threshold Thread3 is the condition of iteration, its value is related to the micromovements in the video. The value of Thread3 affects the number of iterations and the quality of training. By setting the thresholds, the noise interference can be removed to a large extent and the background image can be obtained more ideal.

2.3. The Principle of LBP Algorithm. There are many features that can be used to measure the differences between different images, including gray histogram, edge histogram, color histogram, corner feature, and scale invariant feature [11]. Through the analysis and study, it is found that there is a certain difference in texture features of the background between the occluded video and the unoccluded video. Therefore, the LBP algorithm (local binary mode) is selected to measure the difference of different background images in this paper. The LBP algorithm is widely used in face recognition [12], facial expression recognition [13], image retrieval [14], image classification [15], and other fields, and it has achieved good results. Compared with other simple methods of feature extraction, the recognition accuracy of the LBP algorithm is higher. Compared with other methods with high recognition accuracy, the calculation of the LBP algorithm is more simple and the LBP feature is easier to extract. Based on the above characteristics, the requirements of real-time and accuracy can be better met by using the LBP algorithm when detecting whether the surveillance camera is occluded.

The LBP algorithm is not only a nonparameter algorithm to describe the difference of the gray value between the center pixel and its neighborhood pixels in the image, but also an efficient algorithm that describes local texture features. The original LBP operator takes the gray value of the central pixel point as the threshold in the window of 3×3 . The gray values of eight neighborhood pixel points are

compared with the threshold. If the gray value of the neighborhood pixel point is greater than the threshold, the coding value of the neighborhood point is 1. Otherwise, the coding value of the neighborhood point is 0. Then, the coding value of each neighborhood pixel point is assigned weight 2^i , $i = 0, 2, \dots, 7$. Through the above coding, the coding values of the eight neighborhood pixel points can form into an eight-bit binary. The decimal value represented by the binary is the LBP value that we want to find. The LBP value can effectively reflect the texture information in the window and it will be used to replace the gray value of the original center pixel points.

As shown in Figure 2, in the window of 3×3 , it is assumed that the gray value of the center pixel is p_c and the gray values of eight neighborhood pixels are $p_i, i \in [0, 7]$. If $p_i > p_c$, the coding value of p_i is 1. Otherwise, the coding value of p_i is 0. The calculation process of LBP value is as follows:

$$\text{LBP}_{P,R} = \sum_{i=0}^{P-1} s(p_i - p_c)2^i. \quad (1)$$

In formula (1), P is the number of neighborhood pixel points, its value is 8. s is the symbolic function:

$$s(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (2)$$

The LBP algorithm only subtracts the gray values of the central pixel points and the neighborhood pixel points in the selected window. It can simply and quickly extract the local texture feature of the image without a complex learning process. So, the calculation of the LBP algorithm is simple and the range of LBP algorithm's applications is wide.

2.4. The Algorithm of Occlusion Detection for the Surveillance Camera. First, the improved frame difference method in literature [10] is used to extract the reference image by using an unoccluded surveillance video. Then, the background modeling method based on background difference and frame difference is used to extract the background image of the video which needs to be detected. The LBP algorithm is used to dispose the reference image and the background image, respectively, and the image disposed by the LBP algorithm is called mapping. In practical application, the blocked histogram of the mapping is used to construct the feature vectors. The feature vectors are compared by using the nonparametric method to measure the difference between images. There are many nonparametric methods that can be used to compare the difference of two histograms, such as Euclidean distance, chi-square statistics, Histogram Intersection, and logarithmic likelihood statistical method. The chi-square statistics is used to measure the difference of two histograms in this paper, and the formula of chi-square statistics is shown as

p_7	p_0	p_1
p_6	p_c	p_2
p_5	p_4	p_3

FIGURE 2: The window of the LBP algorithm.

$$X^2(S, M) = \sqrt{\sum_i \frac{(S_i - M_i)^2}{S_i + M_i}}. \quad (3)$$

In formula (3), S and M are two different feature vectors and S_i and M_i are the values of the same location in the different vectors S and M .

Figure 3 shows the specific process of the algorithm that detects whether the surveillance camera is occluded. The process that the background modeling method based on background difference and frame difference and the LBP algorithm are used to detect whether the surveillance camera is occluded is as follows:

- (1) Firstly, a surveillance video which is not occluded is used to extract the reference image. Then, the LBP algorithm is used to dispose the reference image and obtain the texture mapping. Finally, the blocked histogram of the mapping is calculated to construct the feature vectors. The feature vectors will be stored and it will be called the reference feature vectors.
- (2) Firstly, the video sequence which will be detected is loaded and the background modeling method based on background difference and frame difference is used to extract the background image. Then, the LBP algorithm is used to dispose the background image and obtain the mapping. Next, the blocked histogram of the mapping is calculated to obtain the feature vectors. Finally, the chi-square statistic is used to measure the similarity between the feature vector and the reference feature vector. And whether the video is occluded at this moment will be determined according to the similarity. The higher the similarity is, the less likely the video is occluded; and the lower the similarity is, the larger the area of the video is occluded.
- (3) If the last frame of the video has not been read, Step 2 will continue to be performed. Otherwise, according to the continuous number that the video is detected as occlusion, whether the surveillance camera is occluded will be output. The more the continuous number is, the more serious the surveillance camera is occluded.

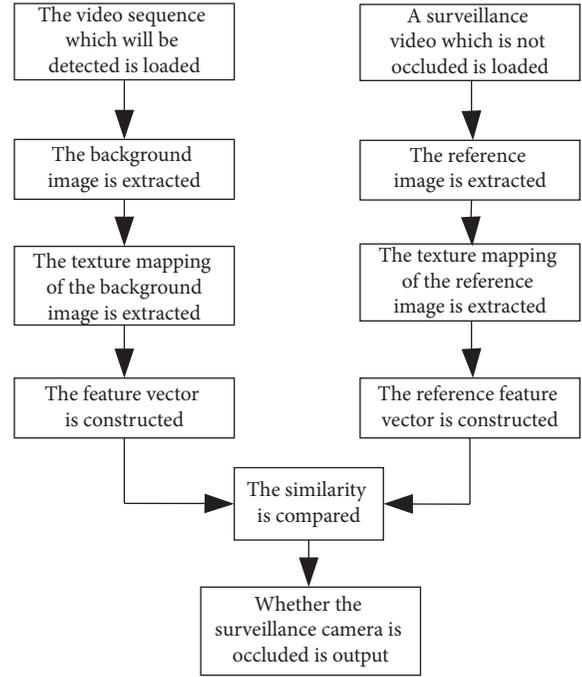


FIGURE 3: The process of occlusion detection for the surveillance camera.

3. Results and Discussion

3.1. The Background Modeling Algorithm Based on Background Difference and Frame Difference. Set Thread1 is 8, Thread2 is 10, and Thread3 is 10. The background image of surveillance video is extracted by using the improved frame difference method in literature [10] and the background modeling method based on background difference and frame difference proposed in this paper. Figure 4 shows the background images extracted from two different videos.

From Figure 4, it can be found that the background modeling algorithm based on background difference and frame difference can extract a better background image with fewer video frames. By comparing the background images extracted from the two videos, it can be found that the background modeling algorithm based on background difference and frame difference has more obvious advantages compared with the original method when the foreground objects whose movement speed is slow exist in the video.

In order to further verify the effectiveness of the background modeling algorithm based on background difference and frame difference, the videos with different occluded areas are shot by using mobile phones. The above two background modeling methods are used to extract the background image of the videos, respectively. Figure 5 shows the results.

From Figure 5 it can be seen that fewer video frames will be used when the background modeling algorithm based on background difference and frame difference is used to

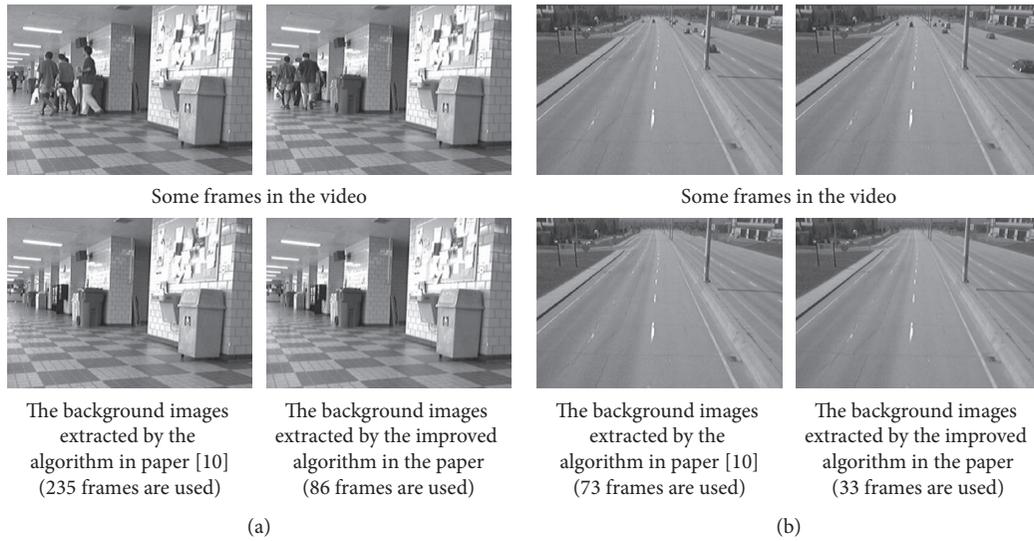


FIGURE 4: The background images extracted by two different algorithms. (a) The background images extracted from the video whose surveillance scene is the hallway. (b) The background images extracted from the video whose surveillance scene is the highway.

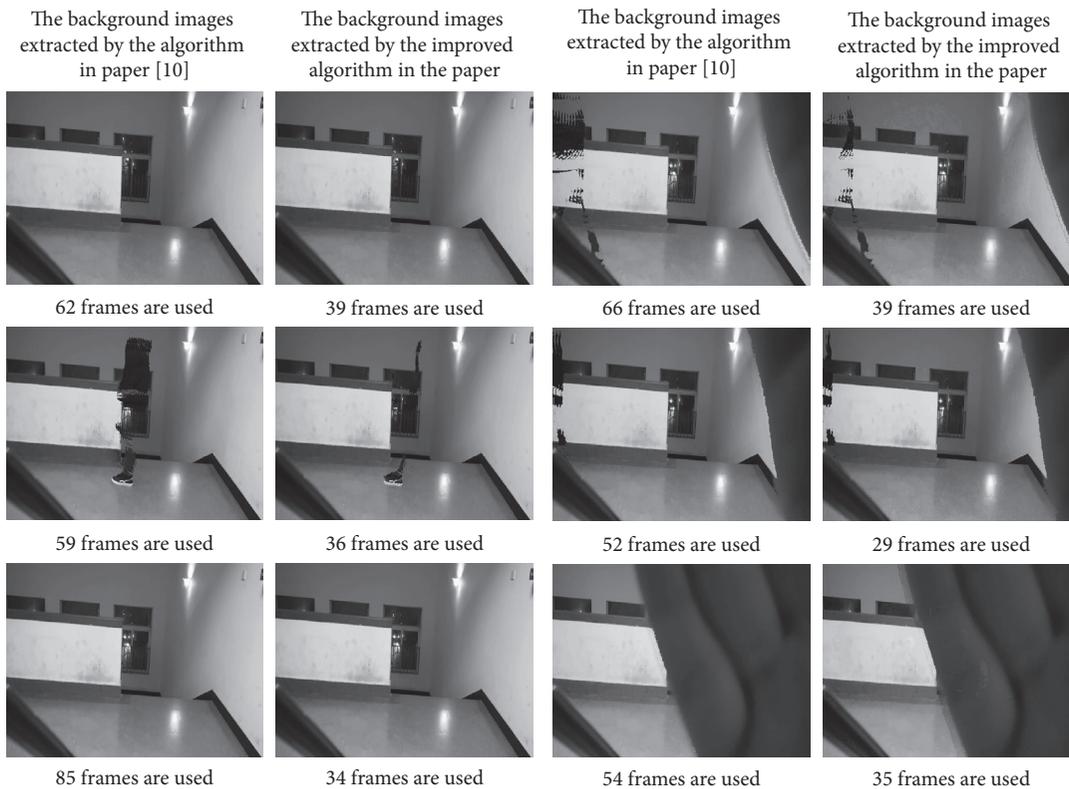


FIGURE 5: The background images extracted from the videos with different occluded areas.

extract the background image for the videos with different occluded areas. By comparing the background images extracted by the two algorithms, it can be found that the background image extracted by the background modeling algorithm based on background difference and frame difference is more effective.

In order to quantitatively compare the improved frame difference method in literature [10] and the background modeling algorithm based on background difference and frame difference, the same videos with different occluded areas are disposed. Each video is 13 seconds long and consists of 400 frames. Using the background modeling



FIGURE 6: When the size of the window is 3×3 , the similarity of the background image with different occluded areas and the reference image. (a) No area occluded 0.14. (b) A small area occluded 0.34. (c) A large area occluded 1.70. (d) A large area occluded 1.83.

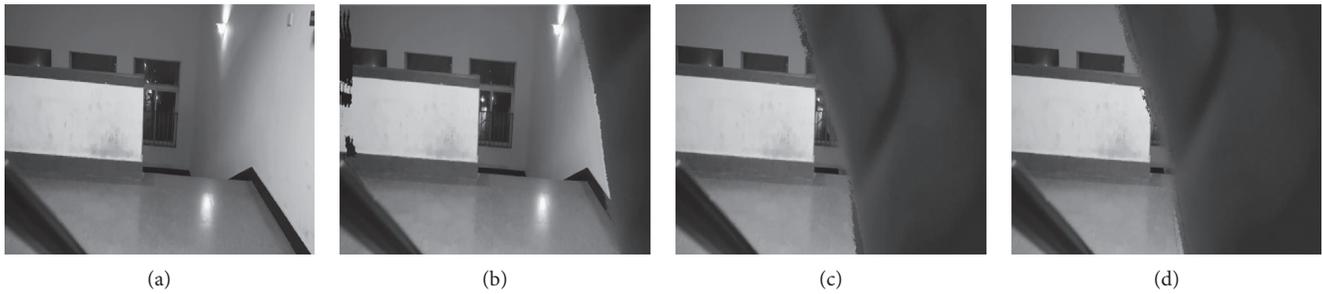


FIGURE 7: When the size of the window is 3×3 , the similarity of the background image with different occluded areas and the reference image. (a) No area occluded 0.16. (b) A small area occluded 1.09. (c) A large area occluded 2.82. (d) A large area occluded 2.89.

algorithm based on background difference and frame difference can extract 87 background images and the success rate of background modeling is 82.76%. Using the improved frame difference method in literature [10] only can extract 48 background images and the success rate of background modeling is 77.08%. So, the background modeling algorithm based on background difference and frame difference is better than the improved frame difference method in literature [10].

3.2. The Algorithm of Occlusion Detection for the Surveillance Camera. The similarity between the background images with different occluded areas and the reference image is calculated by the original LBP algorithm, and the results are shown in Figure 6. Because the occluded areas are a white wall and there is little texture information in it, the texture information in the background image will not change significantly when a small area of the white wall is occluded. For example, the similarity between the background image and the reference image is 0.34 when a small area of the white wall is occluded in the background image. The similarity between the background image which is not occluded and the reference image is 0.14. So, whether the surveillance camera is occluded cannot be easily determined according to the similarity. Only when the similarity gradually increases with the occluded area increase, whether the surveillance camera is occluded can be determined according to the similarity.

After analysis, it is found that this problem can be solved by expanding the window size of the LBP algorithm. Although there is no difference in gray between the white wall

and the white wall, there is a difference in gray between the white wall and the ground. So, the texture features can be extracted according to the difference in gray between the white wall and the ground by adjusting the window size. When the area of the white wall is occluded, the difference in gray between the white wall and the ground will change. So, the texture information extracted will be obviously different. Then, whether the surveillance camera is occluded can be determined according to the similarity between the background image and the reference image. Figure 7 shows the result that the similarity between the background image and the reference image gradually increases with the occluded area increase when the window size is 41×41 in the LBP algorithm. According to the result, the similarity threshold should be preliminarily set at around 1.0. When the similarity between the background image and the reference image is greater than the threshold, it can be determined that the surveillance camera is occluded, and the larger the similarity is, the larger the occluded area will be.

In order to further verify the effectiveness of occlusion detection for the surveillance camera based on a comparison of image similarity, the videos with different occluded areas are disposed. Whether the camera is occluded is determined according to the similarity between the background image and the reference image. Each video is 13 seconds long and consists of 400 frames. When set Thread1 is 8, Thread2 is 10, and Thread3 is 10, about 10 background images can be extracted from each video. It means whether the camera is occluded can be determined every 1.3 seconds. 87 background images are extracted from all videos, and when the similarity threshold is set as 0.8, 0.9, 1.0, 1.1, and 1.2, the

TABLE 1: The result of recognition with different similarity thresholds.

Similarity threshold	Accuracy rate (%)	False positives rate (%)	False negatives rate (%)
0.8	90.80	6.90	2.30
0.9	89.65	6.90	3.45
1.0	90.80	3.45	5.75
1.1	89.65	2.30	8.05
1.2	88.54	1.15	10.34

accuracy rate of recognition, the false positives rate, and the false negatives rate are shown in Table 1.

From Table 1, it can be seen that the accuracy rate of recognition, the false positives rate, and the false negatives rate are better when the similarity threshold is set as 1.0.

4. Conclusions

Considering previous background modeling methods have the disadvantages that the calculation is complex and constructing the background image takes a long time, the background modeling method based on frame difference is improved in this paper. Combining the background difference, a new background modeling method based on background difference and frame difference is proposed. The simulation results show that fewer video frames are used when the background modeling algorithm based on background difference and frame difference is used to extract the background image, and the background image extracted is better. The above advantages are a good foundation that whether the camera is occluded can be determined by comparing the similarity of the background image and the reference image because the real-time requirements will be satisfied. In the algorithm of occlusion detection for the surveillance camera based on comparison of image similarity, the LBP algorithm is used to compare the similarity between the background image and the reference image. By setting an appropriate similarity threshold, the actual demand can be well met and the application value is very high.

Data Availability

The data used to support the findings of this study have not been made available because they involve the authors' privacy.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Beijing Key Laboratory of Work Safety Intelligent Monitoring, Beijing University of Posts and Telecommunications, Beijing, China.

References

- [1] C. Kuang, *Research on Detection and Classification of Camera Tampering Events*, Harbin Institute of Technology, Harbin, China, 2016, in Chinese.
- [2] M. Qiu, *Research and Achievement of the Diagnoses of the Surveillance Video Image Quality*, East China University of Science and Technology, Shanghai, China, 2014, in Chinese.
- [3] E. Ribnick, S. Atev, O. Masoud, N. Papanikolopoulos, and R. Voyles, "Real-time detection of camera tampering," in *Proceedings of the IEEE International Conference on, 2006 Video and Signal Based Surveillance AVSS'06*, Sydney, Australia, November 2006.
- [4] D.-T. Lin and C.-H. Wu, "Real-time active tampering detection of surveillance camera and implementation on digital signal processor," in *Proceedings of the 2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, Piraeus-Athens, Greece, July 2012.
- [5] H. Yin, X. Jiao, X. Luo, and C. Yi, "Sift-based camera tamper detection for video surveillance," in *Proceedings of the 25th Chinese Control & Decision Conference*, Guiyang, China, May 2013.
- [6] P. Gil-Jimenez, R. Lopez-Sastre, P. Siegmann, J. Acevedo-Rodríguez, and S. Maldonado-Bascón, "Automatic control of video surveillance camera sabotage," in *Proceedings of the IEEE International Conference on the Interplay Between Natural and Artificial Computation*, vol. 4528, pp. 222–231, La Manga del Mar Menor, Spain, June 2007.
- [7] H. Liu and Y. Guo, "A vision-based fall detection algorithm of human in indoor environment," *Proceedings of the SPIE*, vol. 10256, Article ID 1025644, 2017.
- [8] Z. Lian and Z. Wang, "Research on vehicle detection method based on background modeling," *International Journal of Advanced Network, Monitoring and Controls*, vol. 3, no. 2, pp. 6–9, 2018.
- [9] T. Yu, J. Yang, and W. Lu, "Background modeling with extracted dynamic pixels for pumping unit surveillance," *Mathematical Problems in Engineering*, vol. 2018, Article ID 8938673, , 2018.
- [10] X. Kong, *Research on Video Anomaly Detection Method Based on Image Analysis*, Beijing University of Posts and Telecommunications, Beijing, China, 2017, in Chinese.
- [11] Y. Yuan, *Research of Identification for Leaf Occlusion in Surveillance Video*, Wuhan University of Science and Technology, Wuhan, China, 2015.
- [12] L. Shi, X. Wang, and Y. Shen, "Research on 3D face recognition method based on LBP and SVM," *Optik*, vol. 220, Article ID 165157, 2020.

- [13] H. Bejaoui, H. Ghazouani, and W. Barhoumi, "Sparse coding-based representation of LBP difference for 3D/4D facial expression recognition," *Multimedia Tools and Applications*, vol. 78, no. 16, pp. 22773–22796, 2019.
- [14] M. Garg and G. Dhiman, "A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants," *Neural Computing & Applications*, vol. 33, pp. 1311–1328, 2020.
- [15] Z. Ye, R. Dong, B. Lin, C. Jin, and Y. Nian, "Hyperspectral image classification based on segmented local binary patterns," *Sensing and Imaging*, vol. 21, no. 2, pp. 556–567, 2020.