Research Article

Detection Algorithm of Wind Power Equipment Video Image Sequence Based on Artificial Intelligence

Qun Li\(^1\) and M.M. Kamruzzaman\(^2\)

\(^1\)School of Electronic Information Engineering, Ningbo Polytechnic, Ningbo 315800, Zhejiang, China
\(^2\)Department of Computer Science, College of Computer and Information Science, Jouf University, Sakaka, Al Jouf, Saudi Arabia

Correspondence should be addressed to M.M. Kamruzzaman; mmkamruzzaman@ju.edu.sa

Received 24 January 2022; Revised 21 February 2022; Accepted 5 March 2022; Published 25 March 2022

Academic Editor: Muhammad Arif

Copyright © 2022 Qun Li and M.M. Kamruzzaman. 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.

How to detect the data in the video image sequence and analyze it efficiently by using the artificial intelligence method is a frontline topic in the field of computer vision. Global wind energy resources are abundant, widely distributed, clean, and pollution-free and meet the requirements of sustainable economic and social development, so relying on the large-scale development and utilization of wind energy has been the common choice of many countries. A major goal of artificial intelligence research is to enable machines to perform complex tasks that normally require human intelligence to perform. The purpose of this study is to detect the video image sequence of wind power equipment based on artificial intelligence and analyze the effectiveness of the algorithm. In this study, augmented reality is used as an auxiliary means. According to the characteristics of the video image inside the fan, a moving target region detection algorithm based on the background difference method is proposed. The algorithm uses the difference between the current image and the background image, uses the first-order Kalman filter to update the dynamic background image, and then uses the adaptive threshold method to segment the moving region. After filtering, the moving target area can be obtained. The results show that the positive detection rate of 980 test samples is 99.6%, and the training time is only 3.785. It is concluded that the accuracy and the number of support vectors of this algorithm are better than other algorithms in the case of the same value of C. It shows that this study has a good detection effect. It provides an effective method for video image sequence detection of wind power equipment.

1. Introduction

With the development of digital video technology, a large number of video data has been brought in. In the face of huge information, it is impossible to complete the management work only by human beings. Intelligent video detection technology can solve this problem well. The development of intelligent video detection technology is bound to become a trend and a new research direction of the security industry [1, 2]. Due to the popularity of networked digital video detection devices and the combination of powerful processing functions of a computer, the detection system has powerful intelligent analysis and image processing functions, providing users with convenient analysis and early warning and other video information management functions. In addition, compression can greatly improve the efficiency of video information management and utilization and effectively achieve the purpose of these information resources.

With the increasing attention to clean energy and renewable energy generation in various countries, China’s wind power generation industry has been developing rapidly in recent years. A series of centrifugal fans cannot adapt to the technical level of previous special applications. In practical engineering applications, such as electrolytic aluminum exhaust centrifugal fan, the demand for fan consumption is increasing; this greatly deviates from the design system and greatly reduces the efficiency. During the design process, the fan also rotates faster than planned. According to the design limitation, on-site installation, and cost, the
centrifugal fan is transformed into a high-speed axial-flow fan. It is inconsistent with the concept of energy-saving and emission reduction. Therefore, it is more economical to develop a relatively economical centrifugal wing with the help of a video depiction of the internal equipment of the fan. At present, the development of the high-efficiency and high-speed centrifugal propeller fan is industrial energy-saving and emission reduction.

In the research of video image sequence detection algorithm, Malagi proposes a new air tracking learning detection algorithm, which can effectively track single or multiple targets in aerial images. The algorithm refers to the decomposition of the persistent tracking task into three parts: tracking, learning, and detection. The tracker is used to track the target between frames, and the detector detects the target based on the previous appearance model and initializes the tracker in due time. The learner evaluates the errors of the detector and updates it to avoid the same errors in the future. Tracking learning detection (TLD) considers both appearance and motion features to track. It can deal with the occlusion problem to a certain extent and can deal with long-time video sequences well. His algorithm includes compensation for camera motion, an improved algorithm combining appearance and motion cues in multiple target detection and tracking, and enhancement of distance measurement between targets to improve the performance of the tracker when there are many identical targets. His method is not comprehensive enough [3]. Tian proposed an improved method for salient region detection. His method combines binocular normalized gradient for coarse location and low-rank decomposition model for selection. First, he uses the histogram threshold (HT) method to adaptively calculate the optimal threshold to reduce the boundary detection error; second, on the basis of coarse localization of candidate salient regions detected by gradient box, he uses robust principal component analysis to extract low-order components, establishes background model, and uses background difference method to eliminate background regions. His method is not accurate enough [4]. Ray and Chakraborty proposed a new method for target detection and tracking in moving camera video without any additional sensors. His algorithm uses a three-dimensional Gabor filter to analyze the image sequence in time and space, detects the moving object as a spatiotemporal cluster, and uses the minimum spanning tree for fusion. In the tracking process, the data association problem is solved by linear assignment problem, and the occlusion problem is processed by Kalman filter. His method is not practical enough [5]. Elharrough et al. proposed a moving object detection algorithm based on background subtraction. The basic idea of background subtraction is to extract the target region by using the difference operation of different images, specifically by subtracting the current frame image from a continuously updated background model and extracting the motion target in the difference image. His method combines quadtree decomposition with entropy theory to generate a background model. In general, many background subtraction methods are very sensitive to the sudden change of illumination in the scene and cannot update the background image in the scene. The background modeling method proposed by him analyzes the problem of illumination change. After background subtraction based on the background model, the moving target can be detected accurately in every frame of the image sequence. The most ideal method for background subtraction is to acquire a frame as a background in a state without a moving target. In practical situations, this is difficult to achieve due to many factors such as rain and snow, target motion, and so on. His method is not stable [6]. The above studies provide a detailed analysis of video image sequence detection algorithms and Kalman filtering applications. It is undeniable that these studies have greatly contributed to the development of the corresponding fields. We can learn a lot of lessons from the methods and data analysis. However, there are relatively few studies on the detection of video image sequences of wind turbines in the field of artificial intelligence, and there is a need to fully apply these methods to the research in this field.

This paper focuses on the detection algorithm of wind power plant video image sequences based on artificial intelligence, with the following innovations: (1) first, the augmented reality technology and its principle are introduced in this paper. The background subtraction method is also described in detail, including single Gaussian background modeling, model initialization, target detection, and model updating. In this study, the Kalman filtering process is used to update the background; (2) then the threshold adaptive method of moving region segmentation is described to detect the moving image; and (3) the experimental part of this study describes the main components of the fan, as well as its principle parameters, and the design of the test bench. Analyzing the results, detect algorithm accuracy and background differences before and after image extraction by training the classifier. The results show that this algorithm has high accuracy and less training time compared with other algorithms. It shows that the algorithm has a good effect.

2. Fan Internal Results and Detection Algorithm

2.1. Augmented Reality Technology and Its Principle

2.1.1. Augmented Reality. This technology uses computer vision technology to compare virtual information with the real world to generate the corresponding relationship. The 3D model or 2D image information generated by the computer vision system will be attached to the marks or objects in the real world and displayed on the same screen at the same time [7, 8].

In this implementation method, an extended reality tag is prepared to match the computer, and then the tag card is placed anywhere in the actual scene so that the plane is in the real world, and then cameras and camera devices must be used. You can identify the marker and measure the azimuth to determine its position. Mark the center as the origin to generate a consistent 3D model or image. If you want to create images, and links, you can directly override the corresponding tag content. When the 3D model needs to be generated, the environment around the marker needs to be measured to
generate a 2D image centered on the marker. Then the actual space is consistent with the 3D model on the plane to generate two fused 3D models. Such a virtual reality scene can be displayed on the display [9–11]. Augmented reality technology not only shows the real-world information but also displays the virtual information at the same time, and the two types of information complement and superimpose each other. In visualized augmented reality, the user uses a headset display to multiply the real world and computer graphics together and can then see the real world surrounding it.

2.1.2. Principle. First, a series of algorithms are used to extract the features of an object or image, and then they are recorded in the computer system and stored in the computer system. When the camera device scans and collects real-world models or images, the system can extract elements of the model or image, in order to adapt to the characteristics previously stored on the computer [12–15]. Scanning allows you to obtain special point values that will be stored when you arrive at the computer. The system regards the model as the scanning object and displays the sequential content in the model to improve the actual effect [16].

Augmented reality systems have three characteristics: (1) the integration of information from the real world and the virtual world, (2) real-time interactivity, and (3) locating virtual objects in three-dimensional scale space.

2.2. Technologies Used in Augmented Reality. This technology includes object recognition technology, detection technology, display technology, and human-computer interaction technology. First of all, the purpose of object recognition and detection technology is to find and identify the real-world signs, which is a very important step. This is usually used to segment groups with the same characteristics, such as vehicle detection and recognition and biometric detection [17].

The use of tracking registration technology can be divided into two categories: device-based and vision-based. Through the use of a glasses enhancement device, the target can be kept in the vision during continuous use to determine, track, and locate the target. Display technology is divided into monitor display and executive mirror display. Monitors can be divided into display-based and head-mounted displays. It is the basic form of using a common display to display the actual effect of expansion. In this mode, the user uses the camera to obtain the real-world image processed by the computer graphics system, adds or superimposes the corresponding computer processing virtual image or model according to the actual world image, and displays it on the monitor [18, 19].

2.3. Background Difference Method. The basic idea of background subtraction is to select one or more image frames to be modeled and use the model to compare subsequent frames. The part of the image similar to the background is regarded as the background, and various fields are regarded as the foreground. Finally, the background model is updated according to the image information. This method is only applicable to the static monitoring device, and the implementation flow is shown in Figure 1.

Background extraction is the extraction of the background in a sequence of video images; the background is the scene in which the scene is stationary. Because the camera is not moving, each pixel point in the image has a corresponding background value, which is relatively fixed over a period of time. The goal of background extraction is to find the background value of each point in the image based on the video image sequence. The current background extraction algorithms can be divided into two categories. One is based on statistical averaging. The gray level of pixels is selected, and the initial background is extracted by comparing the continuous sequence of frame images (such as average or median method). The other is based on statistical modeling, Gaussian mixture distribution, and other methods [19, 20].

2.3.1. Single Gaussian Background Modeling. The Gaussian distribution function is used to establish the gray model of pixels in the image background area. Among them, the basic idea of the single Gaussian model is to establish the background image of each pixel of the video sequence. If all pixels in the image are assumed to be independent of each other according to Gaussian distribution, then each pixel included in the image can be represented by Gaussian distribution, and the gray value of pixels is distributed according to Gaussian distribution within a certain period of time and then compared the current new image sequence. Finally, the Gaussian distribution parameters corresponding to each pixel are updated [21, 22].

The Gaussian distribution function model is expressed as follows:

$$p(I_t(x,y), \mu_t, \Sigma_t) = \frac{1}{(2\pi)^{1/2} |\Sigma_t|^{1/2}} e^{-1/2 (I_t(x,y) - \mu_t)^T \Sigma_t^{-1} (I_t(x,y) - \mu_t)}.$$

(1)

Suppose any pixel at time t, the value of a specific pixel in the input video sequence image is $I_t(x,y)$. In addition, $\mu_t$ is regarded as the average value of Gaussian distribution corresponding to the same pixel in the image frame, and $\Sigma_t$ is expressed as the covariance matrix of the image frame [23].

2.3.2. Model Initialization. A single Gaussian distribution background model is used to simulate the background of video sequence images. First, the model needs to be initialized. Usually, the video sequence is first collected, and then the average value of each pixel of the video sequence frame $(x, y, t)$ and variance $\sigma(x, y, t)$ are calculated at a
specific time $t$ and then initialized as two parameters of the model to calculate the values of $\mu (x, y, t)$ and variance $\sigma (x, y, t)$. The initialization mathematical expression is as follows:

\[
\mu (x, y, T) = \frac{1}{T} \sum_{k=1}^{T} I_k (x, y, T),
\]

\[
\sigma^2 (x, y, T) = \frac{1}{T} \sum_{k=1}^{T} [I_k (x, y, T) - \mu (x, y, T)]^2.
\] (2)

2.3.3. Target Detection. When the initialization of a single Gaussian model is completed, the moving objects in the video sequence can be detected. According to the output result $D (x, y, T)$, the moving target in the image can be determined, and the calculation result is $d (x, y, T) = 1$, where the pixel is the moving target area of the video sequence image. In addition, the calculated result is $d (x, y, T) = 0$, which belongs to the background region. $\lambda$ is the threshold constant, which is the empirical value based on the experimental results [24, 25].

\[
D (x, y, t) = \begin{cases} 
1, & |I (x, y, t) - \mu (x, y, t - 1)| > \lambda \times (x, y, t - 1) \\
0, & \text{otherwise} 
\end{cases}.
\] (3)

2.3.4. Model Update. The background region of the video sequence image remains unchanged. The change of simple scene may be weak or slow, and the background may change suddenly. At this time, the parameters determined by the Gaussian model must be updated, and the background update must be completed. The following formula updates the parameters of the model:

\[
\mu (x, y, t) = (1 - \alpha)\mu (x, y, t - 1) + \alpha \mu (x, y, t),
\]

\[
\sigma^2 (x, y, t) = (1 - \alpha)\sigma^2 (x, y, t - 1) + \alpha [I (x, y, t) - \mu (x, y, t)]^2.
\] (4)

At this time, $\alpha$ is set as the background update rate of the model, and the update speed of the current video sequence frame is reflected in the background. If we focus on the range of $\alpha$ value $[0, 1]$, the larger the $\alpha$ value, the faster the background update rate of moving target detection.

The moving target detection algorithm based on the Gaussian model can meet the real-time requirements of practical application due to its low time complexity. However, in practice, if the complex external environment, chaotic indoor scene, large change, and sudden change scene are encountered, the single Gaussian model will fail, be unable to correctly describe the background, and be unable to achieve the desired effect [26, 27].

### 2.4. Kalman Filtering Process

Kalman filtering is an algorithm that uses the state equation of a linear system to optimally estimate the system state from the system input and output observations. Since the observed data includes the effects of noise and disturbances in the system, the optimal estimation can also be regarded as a filtering process. Since Kalman filtering is easy to implement by computer programming and can update and process the data collected in the field in real time.

The system state equation can be expressed as follows:

\[
Y (k) = F (k - 1) + BW (k).
\] (5)

The state variable $Y (k)$ is as follows:

\[
Y (k) = [y_k, y_{k-1}, \beta_k]^T.
\] (6)

When the output of the inertial sensor is $Z (k)$, the system observation equation of the Kalman filter can be expressed as follows:
\[ Z(k) = H(k)Y(k) + V(k), \]  

(7)

where \( H \) corresponds to [100] and \( V(k) \) represents measurement error. Suppose that the process noise \( \{W(k)\} \) and the observation noise \( \{V(k)\} \) represent the Gaussian white noise sequence with zero mean value, and their cross-correlation function is zero [28].

Through the system state equation and system observation equation, the above discrete Kalman filter processing process can be obtained. Among them, the estimated error covariance matrix is as follows:

\[ P(k, k) = [I - G(k)H(k)]P(k, k - 1). \]  

(8)

At the beginning of filtering, the error covariance matrix \( P(0) \) is obtained as follows:

\[
P(0) = \begin{bmatrix} Q_2, 0, 0 \\ 0, Q_2, 0 \\ 0, 0, 0 \end{bmatrix},
\]

(9)

where \( Q_2 \) is the dispersion of observation noise extracted by a high pass filter [29].

2.5. Adaptive Threshold Method for Moving Region Segmentation. Based on the brightness distribution of different regions of the image, its local threshold is calculated, so it is able to calculate different thresholds adaptively for different regions of the image, and is therefore called the adaptive thresholding method. And according to the test results, there will be more noise. This algorithm sets its own threshold matching radius \( R(P_x) \) and probability \( T(P_x) \) of updating the background model for each pixel, which can adapt to the complexity of background changes.

\[ L_r(P_x) = \sum_{i=1}^{N} \left( I(P_x) - B_j(P_x) \right)^2 \]  

(10)

where \( I(P_x) \) is the pixel value of the new pixel and \( B_j(P_x) \) is the sample pixel. The variance \( L_r(P_x) \) between the new pixel and the sample set is calculated to measure the complexity of background change.

When the background change is more complex, \( R(P_x) \) should be adaptive. In order to avoid the interference of wrong pixel values in the foreground of the background model, on the contrary, in the scene with low complexity and stable background change, \( R(P_x) \) should be self-adaptively reduced, and the sensitivity of the algorithm to subtle changes in the scene needs to be improved.

\[ R(P_x) = R(P_{x-1}) \times \left| 1 - L_r(P_x) \right|. \]  

(11)

When the background changes more complex, the background model update probability \( T(P_x) \) should be reduced. Although it can maintain the consistency between the background model and the reality, on the contrary, in stable scenes with fewer background changes, \( T(P_x) \) should be adaptive to improve, in order to accurately detect target objects.

2.6. Main Parts of Centrifugal Fan. The function of the inlet part is to introduce gas into the impeller. In order to make the gas flow smoothly into the filter, the inlet pipe must meet the requirements of a specific length. The shape of the inlet section also needs to be carefully designed and fabricated to ensure good air circulation within the embedded clip.

The impeller is to convert the mechanical energy of a motor into gas energy, so as to increase the pressure of the gas. The impeller is composed of blades fixed on the front plate, the back plate, and the center. The geometry, size, and speed of the impeller are determined, and the gas flow characteristics in the impeller are determined. The flow rate and pressure rise of the fan are controlled. The main difference in impeller structure is the difference between the impeller blade and the front disc [30, 31].

The function of the volute is to collect the gas emitted from the propeller, guide it to the outlet of the volute, and then release the gas into the atmosphere. Therefore, generally speaking, the convolution model simplifies the manufacturing, and the convolution of the fan is usually designed in a certain width apart. Because convolution is also one of the main components of the fan, the change of convolution distribution directly affects the flow loss in the convolution and also affects the aerodynamic performance of the fan propeller, thus affecting the efficiency and pressure of the fan. Performance parameters such as lifting and vibration [32–34].

In addition to the above main components, some centrifugal fans are also equipped with a front guide vane and outlet diffuser. The main function of the front guide vane is to change the blade angle and obtain various fan performance curves. In this way, the operating range of the fan can be expanded. The outlet diffuser is set at the convolution outlet, and its function is to convert part of the kinetic energy at the outlet into static pressure energy, so as to reduce the dynamic pressure loss at the outlet.

3. Fan Internal Video Image Sequence Detection Experiment

3.1. Test Hardware. CPU Celeron 1.7 GHz, memory 256M, and hard disk 40 g are used. The operating system is Microsoft Windows 2010 and MATLAB 2016b.

3.2. Experimental Data Set. This system is realized by the mixed programming of MATLAB and C++. Among them, Kitti automatic operation task data is a series of machine vision task data aiming at automatic operation, including spatial modeling, visual process, visual ranging, 3D object detection, 3D object tracking, and so on. The data are from two color cameras and black-and-white cameras, 360-degree riders, and cars with a GPS positioning system. They are running and recording the data in the carnival of medium-sized cities. Audacity data set is the automatic driving vehicle data set of automatic driving routes set up by audacity.

3.3. Working Principle of Centrifugal Fan. The gas of the test device of the centrifugal fan is inhaled from the atmosphere
and flows into the impeller through the inlet pipe, the internal rectification grid, and the inlet collector. The engine drives the impeller to rotate, and the impeller produces centrifugal force in the process of rotation to eliminate the air in the impeller. The air ejected from the propeller is curled and discharged by convection. If the air in the propeller overflows, a negative pressure is formed in the center of the propeller. Due to the influence of atmospheric pressure, the inlet air will be pressed into the propeller. Therefore, if the fan rotates regularly, the fan will continue to draw in air, continue to flow in the pipeline, and release the circulation. As shown in Table 1, the main performance parameters of the fan are compared.

3.4. Design of Test Bench. Centrifugal fan test is carried out on the centrifugal fan test platform of the machine company. The performance test method and data processing refer to the requirements of national standards. The test platform is mainly composed of valve, intake pipe, centrifugal fan, impeller, and control parts. The computer and the data collection system are processed by computer.

The main test equipment is selected and installed as follows:

Inlet and outlet: the centrifugal fan adopts a type C test device to simulate the working conditions as much as possible: pipe inlet and free outlet. When the bench is not working, there is no more than 1 m/s airflow around the entrance and exit of the bench and the test pipe fittings, and there are no other major obstacles near the entrance and exit. Air intake pipeline: all test pipelines are guaranteed to be horizontal; all sections are circular; and there are few indirect gaps in the cross-section. There is a good joint between the various parts of the pipeline; there is no internal expansion; and there is correct positioning. The total length of the pipeline is designed according to the pipe diameter and standard requirements.

Free entry condition for measuring inlet flow. That configuration file is axisymmetric. There are no protruding and protruding sharp edges between the cone and the end face and the hose. Mesh screen: the mesh screen is an antivortex device used to prevent the increase of eddy current in normal axial airflow and is used together with the conical inlet of the pipe.

3.5. Construction and Training of Classifier. The difference between the design of weak classifiers and the common machine learning algorithms is that the linear separation impossibility problem does not need to be considered. The design of weak classifiers is an important part of integrated machine learning, and it is also an important reason why integrated machine learning attracts many researchers’ attention. The concept of the weak classifier is first mentioned in the proposed learning theory. In the learning mode, there is a set of polynomial learning algorithms to identify concepts. If the recognition accuracy is high, the set of concepts can be learned strongly.

The calculation model of centrifugal fan consists of three parts: input, worm, and impeller. These areas are treated with grids of six profiles, and the digital simulation design is to ensure that the thickness of feeder and collector, the gap between collector and air, and the grid density at all wall boundary layers meet the bidding requirements.

### Table 1: Comparison of main performance parameters of fans.

| Parameter (unit)                  | High specific speed centrifugal fan |
|-----------------------------------|-------------------------------------|
| Design flow (m³/h)                | 3807.64                             |
| Specific speed (m/s)              | 120                                 |
| Number of blades                  | 15                                  |
| Blade outlet width, b₂ (mm)       | 206                                 |
| Impeller inlet diameter, D₀ (mm)  | 748                                 |
| Impeller outer diameter, D₂ (mm)  | 900                                 |
| Blade inlet angle, β₁ (°)         | 21.8                                |
| Blade outlet angle, β₂ (°)        | 41.3                                |

4. Fan Video Image Sequence Detection Algorithm Analysis

4.1. Training Classifier Analysis of Detection Algorithm. The improved background difference method is used to train positive and negative fan samples. The parameter is set to 5,000 positive and negative samples, and the sample size is 24 × 24. There are 10 levels of target cascade classifiers. The minimum detection rate of each classifier is 0.996 for positive samples and 0.4 for negative samples. Table 2 shows the examples of the training process parameters (take phase 3 as an example).

In Table 2, N represents the number of classifiers in this layer, ST.THR represents the threshold value of weak classifiers of classifiers in this layer, and HR represents the detection rate of weak classifiers in positive samples of classifiers in this layer. FA shows the error detection rate of weak classifiers in the negative samples of this taxon layer. EXP.ERR was a mistake in anticipation. After training, 10 cascaded classifiers (stage0–stage9) were obtained. The parameters of all layer classifiers in this classifier will be saved in different documents. These parameters can be used for face detection experiments.

4.2. Training Sample Number Analysis of Video Image Sequence Detection Algorithm. When training the face classifier, the maximum detection rate is determined by searching up and down 1,000 times, and the corresponding parameters are used to determine the parameters ε and σ. In all 6,980 training samples, 6,000 training samples were randomly selected, and 980 training samples were used as a test set. In order to reduce the amount of calculation, first determine the target value (the default value given by LIBSVM software according to experience) and search in the range of $C = 2^1, 2^2, \ldots, 2^{10}$ to reduce the range of C value. The standard SVM classifier and PCA-SVM classifier are trained, as shown in Table 3.

It can be seen from Table 3 that SVM is the standard SVM algorithm; PCA-SVM is the principal component
4.3. Accuracy Analysis of Video Image Detection Algorithm. As shown in Figure 2, the classification accuracy of SVM and PCA-SVM is compared. It can be seen from Figure 2 that 980 test samples are classified by using three classifier models with \( C = 2^7 \) and \( \sigma = 2^{-8} \), and \( 2^{-7} \) and \( 2^{-8} \), and the positive detection rate reaches 99.6% (973/997). According to the requirements of parameter selection, the parameters of PCA-SVM are determined: \( C = 2^7 \) and \( \sigma = 2^{-8} \).

The training method of the PCA-SVM classifier is introduced below. Using the projection coefficient as a new sample, the characteristics were selected from 6,000 training samples and 980 validated samples. Training personnel to use new teaching mode and parameter classification personnel. The test samples were classified by the PCA-SVM classifier model and added to the sample set of the new training module. Repeat the above steps until there are more classification errors. As shown in Figure 3, the number of support vectors between the SVM classifier and the PCA-SVM classifier is compared.

It can be concluded from Figure 3 that the classification accuracy and the number of support vectors of PCA-SVM are better than those of standard SVM, and the training time is shorter than SVM.

4.4. Image Extraction and Analysis before and after Background Difference Method. The gray level histogram of the background frame and the gray level distribution of the updated background frame are obviously changed as shown in Figures 4 and 5, respectively. Figures 4 and 5 fully show that external conditions such as illumination and shadow have a great impact on the background image intensity. Meanwhile, it is proved that the Surendra algorithm can effectively update the background model and reduce the noise caused by the changes of illumination, shadow, and other external factors.

The update coefficient \( a \) of the Surendra algorithm represents the background update speed. The larger the value of \( a \), the faster the background update speed, the slower the object will be integrated into the new background, which will affect the accuracy of target detection. The smaller the value, the slower the background update speed. If the background changes too fast, it is difficult to reflect the current background information in the updated background. Therefore, it is very important to choose the

| N   | %SMP (%) | F     | ST.THR | HR   | FA   | EXP.ERR |
|-----|----------|-------|--------|------|------|---------|
| 1   | 100      | —     | -0.7292 | 1.0  | 1.0  | 0.1985  |
| 2   | 100      | —     | -1.2931 | 1.0  | 1.0  | 0.2844  |
| 3   | 100      | —     | -1.8228 | 1.0  | 1.0  | 0.2960  |
| 4   | 100      | —     | -2.4265 | 1.0  | 1.0  | 0.2427  |
| 5   | 85       | —     | -2.8872 | 1.0  | 1.0  | 0.2755  |
| 6   | 86       | —     | -3.3409 | 0.9860| 0.9866| 0.2590  |
| 7   | 85       | —     | -2.9866 | 0.9866| 0.9835| 0.2377  |
| 8   | 79       | —     | -3.0617 | 0.9967| 0.9824| 0.2383  |
| 9   | 80       | —     | -3.0594 | 0.9935| 0.9787| 0.2197  |
| 10  | 81       | —     | -2.9751 | 0.9954| 0.9859| 0.2179  |

| Log2C | SVM Accuracy (%) | PCA-SVM Accuracy (%) | SVM Vector (face/non-face) | PCA-SVM Vector (face/non-face) | SVM Train time (s) | PCA-SVM Train time (s) |
|-------|------------------|-----------------------|---------------------------|-------------------------------|-------------------|-----------------------|
| 1     | 98.18            | 98.65                 | 1,095 (579/518)           | 758 (395/362)                | 30.368            | 6.966                 |
| 3     | 98.73            | 99.60                 | 552 (296/254)             | 428 (239/188)               | 13.489            | 4.302                 |
| 5     | 99.20            | 99.62                 | 448 (190/258)             | 393 (181/211)               | 12.688            | 4.336                 |
| 7     | 99.40            | 99.73                 | 362 (158/206)             | 362 (158/202)               | 11.023            | 3.992                 |
| 8     | 99.47            | 99.76                 | 344 (148/195)             | 362 (158/203)               | 10.630            | 3.785                 |
| 9     | 99.52            | 99.77                 | 344 (136/208)             | 362 (158/203)               | 10.475            | 3.866                 |
| 10    | 99.51            | 99.80                 | 344 (136/208)             | 362 (158/203)               | 10.432            | 3.891                 |
updating coefficient $a$. The best coefficient $a$ of the system is determined by several experiments, and the background model is optimized.

If $a = 1$, it is a staff difference model; if $a = 0$, it becomes background difference. In addition, the smaller $A$ is, the smaller the difference effect is and the greater the noise is.
Multiple values $a = 0.8$ can be used to determine the best effect 8. The advanced real-time update background algorithm is used to integrate the background information of the current frame into the context effectively; Environmental changes affect the accuracy of the algorithm.

5. Conclusion

In this study, the background difference method is used to fuse the improved Gaussian mixture model together, and the one-time Kalman filter algorithm is used to cluster the results. In the complex scene, it can detect the interested moving objects and noise under the influence of light and wave, which is less than the commonly used algorithm, so it can get better results. This paper studies the related theory of moving target detection algorithm in video sequence image, describes in detail the basic idea of several commonly used moving object detection algorithms, analyzes their advantages and disadvantages, carries out several experimental tests, and analyzes the results. First, the adaptive threshold method is used to obtain more accurate candidate regions of video targets, and then the feature maps obtained by the feature extraction network are combined to classify and recognize video targets accurately. Next, the final intuitive results are obtained by adapting the test results using several effective statistical strategies. Experimental results show that the detection algorithm has made great progress in detection accuracy and real-time performance. However, due to the lack of effective shadow detection, the detected moving target has some shadows, and the size is slightly larger than the actual object size. The use of color degree features tracking has little impact on recognition and classification accuracy. The detection algorithm needs to be further optimized, and the shadow recognition algorithm is added to make the size of the detected object close to the actual size. The tracking algorithm can be added to the target trajectory and velocity analysis. Not only can the tracking be realized simply, but also the warning can be replaced by the action of crossing the border manually by computer. The recognition algorithm also has great shortcomings; the amount of data used for training is still small; the recognition rate of known parts of the video has not reached the expected requirements, so it can increase the function of classifier training. Add false identification tags to strengthen the system.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The author(s) declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

This work was funded by the Deanship of Scientific Research at Jouf University under grant no. DSR-2021-02-0339.

References

[1] N. O. Garcia, M. F. D. Velasquez, C. A. T. Romero, and J. H. O. Monedero, “Remote academic platforms in times of a pandemic,” International Journal of Emerging Technologies in Learning, vol. 16, no. 21, pp. 121–131, 2021.
[2] O. I. Khalaf, G. M. Abdulsahib, H. D. Kasmaei, and K. A. Ogudo, “A new algorithm on application of blockchain technology in live stream video transmissions and telecommunications,” International Journal of E-Collaboration, vol. 16, no. 1, pp. 16–32, 2020.
[3] V. P. Malagi, R. B. D.R., and K. Rangarajan, “Multi-object tracking in aerial image sequences using aerial tracking learning and detection algorithm,” Defence Science Journal, vol. 66, no. 2, p. 122, 2016.
[4] X. Tian, H. Li, and H. Deng, “An improved two-steps saliency detection algorithm based on binarized normed gradients and nuclear norm model in video sequences,” Journal of Information Hiding and Multimedia Signal Processing, vol. 9, no. 4, pp. 841–852, 2018.
[5] K. S. Ray and S. Chakraborty, “Object detection by spatio-temporal analysis and tracking of the detected objects in a video with variable background,” Journal of Visual Communication and Image Representation, vol. 58, no. JAN, pp. 662–674, 2019.
[6] O. Elharrouss, D. Moujahid, S. Elkah, and H. Tairi, “Moving object detection using a background modeling based on entropy theory and quad-tree decomposition,” Journal of Electronic Imaging, vol. 25, no. 6, 2016.

[7] S. Ding, S. Qu, Y. Xi, and S. Wan, “Stimulus-driven and concept-driven analysis for image caption generation,” Neurocomputing, vol. 398, 2019.

[8] S. Zhou, M. Ke, and P. Luo, “Multi-camera transfer GAN for person re-identification,” Journal of Visual Communication and Image Representation, vol. 59, pp. 393–400, 2019.

[9] A. Palanci and Z. Turan, “How does the use of the augmented reality technology in mathematics education affect learning processes?: a systematic review,” Uluslararası Eğitim Programları ve Öğretim Çalışmaları Dergisi, vol. 11, no. 1, pp. 89–110, 2021.

[10] J. Rajeshwari, K. Karbasappa, and M. T. Gopalakrishna, “Adaboost modular tensor locality preservative projection: face detection in video using Adaboost modular-based tensor locality preservative projections,” IET Computer Vision, vol. 10, no. 7, pp. 670–678, 2017.

[11] J. Zeng, J. Huang, Z. Yu et al., “Research on human motion detection and tracking algorithm based on adaptive dynamic video image scaling technology,” Revista de la Facultad de Ingenieria, vol. 32, no. 4, pp. 455–463, 2017.

[12] O. I. Khalaf, C. A. T. Romero, A. A. J. Pazhani, and G. Vinuja, “VLSI implementation of a high-performance nonlinear image scaling algorithm,” Journal of Healthcare Engineering, vol. 2021, Article ID 6297856, 10 pages, 2021.

[13] M. Abdolmaleky, M. Naseri, J. Batle, A. Farouk, and L.-H. Gong, “Red-Green-Blue multi-channel quantum representation of digital images,” Optik, vol. 128, pp. 121–132, 2017.

[14] B. Wang, B. F. Zhang, X. W. Liu, and F. C. Zou, “Novel infrared image enhancement optimization algorithm combined with DFOCS,” Optik, vol. 224, pp. 165476, 2020.

[15] B. Wang, B. F. Zhang, and X. W. Liu, “An image encryption approach on the basis of a time delay chaotic system, Optik,” An image encryption approach on the basis of a time delay chaotic system, vol. 225, p. 165737, 2021.

[16] D. Tursun, L. Kai, G. Halelbek, and A. Hamdulla, “A joint approach of Harris corners detection and baseline searching for localization of Uyghur text lines in image sequences,” Journal of Information Hiding and Multimedia Signal Processing, vol. 7, no. 2, pp. 352–361, 2016.

[17] Z. Kun and Z. Jianyan, “A moving object detection and recognition method based on improved background subtraction method,” Boletin Tecnico/technical Bulletin, vol. 55, no. 7, pp. 144–151, 2017.

[18] G. P. Wang, L. W. Ma, L. Guo, and H. F. Wang, “Image detection algorithm for incident of discarded things in highway. Chang’an Daxue Xuebao (Ziran Kexue Ban),” Journal of Chang’an University (Natural Science Edition), vol. 37, no. 5, pp. 81–88, 2017.

[19] K. S. Raviya and D. V Vyas, “Real time depth data refurbishment in frequency domain and 3D modeling map using Microsoft kinect sensor,” International Journal of Image, Graphics and Signal Processing, vol. 8, no. 11, pp. 49–58, 2016.

[20] M. Tiwari and R. Singhai, “A review of detection and tracking of object from image and video sequences,” International Journal of Computational Intelligence Research, vol. 13, no. 1, pp. 745–765, 2017.

[21] M. Cao, Z. Deng, L. Rai et al., “Generating panoramic unfolded image from borehole video acquired through APBT,” Multimedia Tools and Applications, vol. 77, no. 19, pp. 1–31, 2018.

[22] S. Chen, S. Tan, B. Li, and J. Huang, “Automatic detection of object-based forgery in advanced video,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 11, pp. 2138–2151, 2016.

[23] Z. Wang, J. Yu, and Y. He, “Soccer video event annotation by synchronization of attack-defense clips and match reports with coarse-grained time information,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 5, pp. 1104–1117, 2017.

[24] V. Golikov, O. Lebedeva, and M. A. Rodriguez, “Detection of barely discernible subpixel objects using a modified orthogonal subspace projection method,” IEEE Latin America Transactions, vol. 14, no. 2, pp. 899–906, 2016.

[25] P. Lu, B. Gao, W. L. Woo, X. Li, and G. Y. Tian, “Automatic relevance determination of adaptive variational bayes sparse decomposition for micro-cracks detection in thermal sensing,” IEEE Sensors Journal, vol. 17, no. 16, pp. 5220–5230, 2017.

[26] L. Zhang, D. Tjondronegoro, V. Chandran, and J. Eggink, “Towards robust automatic affective classification of images using facial expressions for practical applications,” Multimedia Tools and Applications, vol. 75, no. 8, pp. 4669–4695, 2016.

[27] H. X. Cu, Y. Q. Wang, F. Zhang, and T. Li, “A motion detection approach based on UAV image sequence,” KSii Transactions on Internet & Information Systems, vol. 12, no. 3, pp. 1224–1242, 2018.

[28] L. Bampis, A. Amanatidis, and A. Gasteratos, “Fast loop-closure detection using visual-word-vectors from image sequences,” The International Journal of Robotics Research, vol. 37, no. 1, pp. 62–82, 2018.

[29] H. K. Wong, J. Epps, and S. Chen, “Automatic pupillary light reflex detection in eyewear computing,” IEEE Transactions on Cognitive and Developmental Systems, vol. 11, no. 4, pp. 560–572, 2019.

[30] Z. Zhang, Y. Cao, M. Ding et al., “Candidate regions extraction of intruder airplane under complex background for vision-based search and avoid system,” IET Science, Measurement & Technology, vol. 11, no. 5, pp. 571–580, 2017.

[31] G. Ramirez-Alonso and M. I. Chacon-Murguia, “Object detection in video sequences by a temporal modular self-adaptive SOM,” Neural Computing & Applications, vol. 27, no. 2, pp. 411–430, 2016.

[32] W. Song, M. Wen, Y. Xi et al., “A collaborative client participant fusion system for realistic remote conferences,” The Journal of Supercomputing, vol. 72, no. 7, pp. 2720–2733, 2016.

[33] P. Mohan, N. Subramani, Y. Alotaibi, S. Alghamdi, O. I. Khalaf, and S. Ulanaganathan, “Improved metaphueistics-based clustering with multihop routing protocol for underwater wireless sensor networks,” Sensors, vol. 22, no. 4, p. 1618, 2022.

[34] O. I. Khalaf, C. A. T. Romero, S. Hassan, and M. T. Iqbal, “Mitigating hotspot issues in heterogeneous wireless sensor networks,” Journal of Sensors, vol. 2022, Article ID 7909472, 14 pages, 2022.