WILEY | Hindawi

Retraction

Retracted: Detection and Adaptive Video Processing of Hyperopia Scene in Sports Video

Complexity

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Q. Chen and M. Dong, "Detection and Adaptive Video Processing of Hyperopia Scene in Sports Video," *Complexity*, vol. 2021, Article ID 6610760, 13 pages, 2021.

WILEY | Hindawi

Research Article

Detection and Adaptive Video Processing of Hyperopia Scene in Sports Video

Qingjie Chen[1,2] and Minkai Dong [iD][3]

[1]School of Sports and Health, Linyi University, Linyi 276000, Shandong, China
[2]School of Political Science and Public Administration, Shandong University, Jinan 250100, Shandong, China
[3]P.E. Department, Shanghai University of Finance and Economics, Yangpu, Shanghai 200433, China

Correspondence should be addressed to Minkai Dong; dong.minkai@mail.shufe.edu.cn

In the research of motion video, the existing target detection methods are susceptible to changes in the motion video scene and cannot accurately detect the motion state of the target. Moving target detection technology is an important branch of computer vision technology. Its function is to implement real-time monitoring, real-time video capture, and detection of objects in the target area and store information that users are interested in as an important basis for exercise. This article focuses on how to efficiently perform motion detection on real-time video. By introducing the mathematical model of image processing, the traditional motion detection algorithm is improved and the improved motion detection algorithm is implemented in the system. This article combines the advantages of the widely used frame difference method, target detection algorithm, and background difference method and introduces the moving object detection method combining these two algorithms. When using Gaussian mixture model for modeling, improve the parts with differences, and keep the unmatched Gaussian distribution so that the modeling effect is similar to the actual background; the binary image is obtained through the difference between frames and the threshold, and the motion change domain is extracted through mathematical morphological filtering, and finally, the moving target is detected. The experiment proved the following: when there are more motion states, the recall rate is slightly better than that of the VIBE algorithm. It decreased about 0.05 or so, but the relative accuracy rate increased by about 0.12, and the increase ratio is significantly higher than the decrease ratio. Departments need to adopt effective target extraction methods. In order to improve the accuracy of moving target detection, this paper studies the method of background model establishment and target extraction and proposes its own improvement.

1. Introduction

With the maturity of computer technology, especially multimedia technology, and the processing and analysis theory of digital images, video images, as more direct and richer information carriers, are becoming more and more important research objects. In recent years, with the introduction of highway and digital earth concepts and the widespread application of the Internet, video image information has become an important source and means for humans to obtain and use information. The detection and tracking of target images is based on dynamic image analysis combined with image recognition and image tracking methods to detect targets in image sequences. The process of recognition and follow-up tracking is very important in the field of image processing. After inspection, it was found that there was a sentence segmentation phenomenon, so the incomplete sentence was deleted directly. Inspection technology is the intersection of image processing technology, machine vision technology, and artificial intelligence technology. Therefore, detection technology has broad prospects in scientific theoretical research and also has broad prospects in practical engineering applications.

The research on video processing technology in foreign countries started early, almost accompanied by the birth of black-and-white TV sets, but limited by the technical level, the development is slow. The combination of embedded system and computer vision and image processing technology forms an object-oriented embedded video processing technology. Shukla and Sharma proposed that moving target

detection and moving target tracking algorithms are the most basic components, especially target detection algorithms, which have a direct impact on the overall performance of video surveillance systems and have been a hot spot in the field of image processing and machine vision. After inspection, it was found that there was a sentence segmentation phenomenon, so the incomplete sentence was deleted directly [1]. Yang et al. proposed an RFID security monitoring system that combines motion sensors and RFID modules. The system application relies on motion detection sensors to detect moving targets [2]. Chebi et al. used frame difference sharpness pattern matching to detect vehicles in motion and incorporated the background removal function into a small portable camera, which can better segment the foreground in the background area [3].

The domestic video surveillance industry initially developed from closed-circuit television surveillance. From simple analog video surveillance to the current purely digital network video surveillance, surveillance systems have been widely used in China recently, and their performance is comparable to that of large foreign companies. Quite a gap. Wang et al. uses the interframe difference method in the detection of moving objects in the video. Compared with the optical flow method, the advantage of the frame difference method is that the calculation speed is very fast, but there is still a big disadvantage, that is, in the detection process, the target object is easily detected [4]. Huang et al. proposed to use Wiener filtering to model and used it to predict the pixel value of the model and regarded the pixels that deviate from the estimated value as the former scenic spot [5]. Huang et al. assumed that the pixel value will change linearly with Gaussian changes over time, so a single Gaussian distribution is adopted to simulate the background model. This method mainly uses a Gaussian distribution to represent the characteristic value of a pixel and detects the pixel of the image. The eigenvalue size of will match the Gaussian distribution to classify the pixels [6].

Moving target detection and tracking are at the bottom of the vision system and are the basis for various subsequent advanced processing such as target classification and behavior understanding. This paper uses an adaptive background subtraction algorithm on a fixed single node to better complete the extraction of the moving area under a complex dynamic background and uses the static and dynamic characteristics of the moving target for multitarget tracking. The model of the tracking system under the condition that a single node can be moved is discussed, and a system model is proposed. In this paper, the contour of the target is detected, and the smallest bounding rectangle of the largest contour is used as the tracking frame of MeanShift, combined with the detection results to track the selected target, and the algorithm is applied to the UAV video to detect and track the target.

2. Detection and Adaptive Video Processing of Hyperopia Scene in Sports Video

2.1. Background Modeling Based on Gaussian Mixture Model.
The K-Gaussian model is used to model every pixel in RGB color space to represent every pixel in the image. According to the priority $\omega_I, t(\omega_I, t = 1, I = 2k)$, the K-Gaussian distribution used to describe the color distribution of each point has different weights from high to high ω_I/σ_I. Let us say you have a low order [7, 8]. Given an appropriate background weight and threshold T, only the previous distribution within this threshold is considered as background distribution, and the other distribution is foreground distribution. At time t, X_t is the pixel value, and the probability density function can be written as a linear combination of K-Gaussian distribution.

$$P(x_t) = \sum_{i=1}^{k} \frac{\omega_{,it}}{(2\pi)^{d/2}|\text{Cov}_{t,i}|^{1/2}}, \quad (1)$$

where ω_I, t, μ_I, t, Cov_i, and t are the weight, mean, and covariance matrices of the i Gaussian distribution at time t, respectively. The K-Gaussian distribution is sorted in descending order of priority (ω_I/σ_I), while the previous b Gaussian distribution is used to represent the background distribution, for example,

$$B = \arg\min\left(\sum_{i=1}^{b} w_{i,t} > T\right), \quad (2)$$

where t is the set background threshold. For the convenience of calculation, it is assumed that the red, green, and blue components of pixels are independent of each other.

$$x_t = \left(x_h^t x \le x_b^a\right),$$
$$\mu_{j,t} = \left(\mu_{j,t}^y, \mu_{j,t}^g, \mu_{j,t}^b\right), \quad (3)$$
$$\text{Cov}_{i,t} = \left(\text{Cov}_{i,t}^r, \text{Cov}_{i,t}^g, \text{Cov}_{i,t}^b\right).$$

2.2. Automatic Extraction of Stadium Main Area Based on Histogram Statistics.
The document used statistical information on the differences of color elements to distinguish legal areas. For the main characteristics of the green soccer field and the lawn tennis court, the extraction algorithm is used based on the difference in component color information. In the lawn area, the green element of the pixel is larger than the red and blue elements. The differences between the green component and the other two components were calculated and then the color characteristics of the lawn were obtained by threshold treatment. The grass area is extracted from the image, and the binary image is achieved by threshold processing [9, 10].

$$I(x, y) = \begin{cases} 1, & G(x, y) - R(x, y) > T \text{ and } (G(x, y) - B(x, y)) > T_b, \\ 0, & \text{other.} \end{cases}$$
$$(4)$$

Among them, $R(x, y)$, $G(x, y)$, and $B(x, y)$ are pixel values of color image and T_R and T_B are two thresholds which can be adjusted according to the position. Choosing reasonable parameters in the green football field, the extraction accuracy of the main areas of the stadium is relatively accurate [11, 12]. However, this algorithm can only be

used in lawns and other places with prior knowledge and cannot be extracted adaptively.

After inspection, it was found that there was a sentence segmentation phenomenon, so the incomplete sentence was deleted directly. Due to the influence of noise and other factors in the imaging process, there are many small noises or small areas in the main area of the segmented stadium, so it is impossible to remove them directly by removing small areas. According to the nature of noise, the method of eliminating block motion noise can solve this problem well [13, 14]. The blocking operation used in this paper is noise algorithm which is as follows: the image is divided into $n \times n$ blocks, and the statistical value of stadium color in each block is greater than a certain percentage; the block is regarded as stadium; otherwise, the stadium block is considered. The formula is

$$B(i, j) \begin{cases} 1, & \text{if } \sum_{y=N*j}^{y=N(j+1)} \sum_{x=N*i}^{x=N(i+1)} G(x, y) > T \times (N^2), \\ 0, & \text{other.} \end{cases} \quad (5)$$

Among them, $B(I, J)$ is a binary result graph composed of $n \times n$ blocks, $G(x, y)$ is the primary extracted main area of the course, and t is the proportion of the area. In most cases, t is 0.6 and n is 36.

It is assumed that the gray value of the pixel at the same point does not change significantly in two consecutive frames, so it can be considered to be approximately equal as shown in the following formula:

$$f(x + dx, y + dy, t + dt) = f(x, y, t). \quad (6)$$

Among them, $f(x, y, t)$ is the gray scale of pixel (x, y) at time t and $f(x + dx, y + dy, t + dt)$ is the gray scale of (x, y) moving to pixel $(x + dx, y + dy)$. Since the foundation of the algorithm is based on gray level consistency, it is also necessary to assume that formula (7) holds, namely:

$$\frac{\partial f(x + dx, y + dy, t + dt)}{dt} = 0. \quad (7)$$

After the second-order Taylor formula expansion and simplification of formula (7), we can obtain the optical flow field calculation formula:

$$\frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial t} = 0. \quad (8)$$

Among them, $(\partial f / \partial x)$ can be understood as the gradient of the brightness of the image in the horizontal direction, (dx/dt) can be understood as the gradient in the vertical direction, $(\partial f / \partial t)$ can be understood as the gradient in time, and (dx/dt) and $(\partial f / \partial t)$ are the optical flow in the horizontal direction and optical flow in the vertical direction, respectively.

2.3. Eigen Image Filtering.
Among various image processing methods, adjusting pixels is the most basic image processing method. It can adjust the changes of the image under different illuminations through the logarithmic conversion of the pixel value. In addition, rendering is also a kind of image processing, and reasonable rendering can make the image

clearer [15, 16]. Recently, Finlayson proposed a new intrinsic image algorithm, and by calibrating the camera's sensitive equipment, the color image is converted into one-dimensional gray-scale image, thus eliminating the influence of shadow. We use this method in this article because there are few restrictions on this method, and it is not very strict. At the same time, this paper provides a method to improve the calibration technology and expand the scope of the algorithm.

Assuming that the Lambert model can be used in the imaging process of the camera photoreceptor, an integral formula can be used to describe the response of each color channel of the photoreceptor [17, 18].

$$\rho_k(x, y) = \sigma(x, y) \int E(\lambda, x, y) S(\lambda, x, y) Q_k(\lambda) d\lambda, \quad (9)$$

where $k = B_{GR}$ represents the index of the color channel, Q_k is the spectral response function of the kth color sensing device, $\sigma(x, y)$ is the Lambert light dark coefficient, which is the dot product of the normal vector of the reflection plane and the direction vector of the incident light, $E(\lambda, x, y)$ is the spectral energy distribution function of the incident light, and $S(\lambda, x, y)$ is the reflection function of the object surface. If the following two conditions are met at each point of the photosensitive device: the spectral response function of each color channel is a Dirac function and the illuminance function of the incident light can use Planck model, then $Q_k(\lambda)$ and $E(\lambda, x, y)$ can be expressed as follows:

$$Q_k(\lambda) = q_k \delta(\lambda - \lambda_k), \quad (10)$$

$$E(\lambda, T) \approx I_c \lambda^{-5} e^{-(c_2/T\lambda)}, \quad (11)$$

where Q_k is the response strength of the inductor. Because every pixel needs to be considered, the pixel position parameter in the function is removed. I refers to the absolute parameter of light control c_1 and c_2 refer to the intensity of light, and, t refers to the absolute temperature. Substituting formulas (7) and (8) into formula (6) yields

$$P_k = \sigma I_{c_1} \lambda_k^{-5} e^{-(c_2/T\lambda)} S(\lambda_k) q_k. \quad (12)$$

Now establish the ratio of the two channels:

$$r_k = \frac{p_k}{p_p}. \quad (13)$$

If $p = G$, then $k = R, B$; take the natural logarithm for formula (10) and use two values to form a two-dimensional vector (R/G):

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \log\left(\dfrac{S_R}{S_G}\right) \\ \log\left(\dfrac{S_B}{S_G}\right) \end{bmatrix} + \frac{1}{T} \begin{bmatrix} e_R - e_G \\ e_B - e_G \end{bmatrix}, \quad (14)$$

where $\lambda_1 = R/G$ and $\lambda_2 = B_r k$; from (11), it can be seen that the first term on the right side of the formula is a constant, whose value is determined jointly by the reflecting surface

and the camera, and the second term is determined by the incident light temperature and the camera. With the change of the temperature of the incident light, the image points of the same medium are distributed along a straight line in the two-dimensional plane, and this straight line is only related to the camera, not to the incident light and the reflection plane [19, 20].

The basic idea of the interframe difference method is to obtain the shape of the moving target by performing the difference operation on two consecutive frames of the video image sequence.

The change between the images $f_k(x, y)$ and $f_{k+1}(x, y)$ of frame k and frame $k + 1$ is represented by a two-finger differential image $D(x, y)$, as shown in the following formula:

$$D(x, y) = \begin{cases} 1, & \text{if } |f_{k+1}(x, y) - f_k(x, y)| > T, \\ 0, & \text{others.} \end{cases} \quad (15)$$

In formula (15), 0 corresponds to the area where the two frames of images have not changed and 1 corresponds to the area that has changed. The flowchart of the interframe difference method is shown in Figure 1.

2.4. Background Recognition Modeling.

The common methods of background model building and updating are the statistical average method, coefficient updating method, Gauss model method, etc [21, 22]. The complexity of various methods of background model building and updating is different, and the effect of background model obtained is also different.

The statistical average method is to average the continuous video sequence and take the average value as the pixel value in the current background model. Assuming that the current video sequence is the nth frame, then the current time background model is common, and the formula is as follows [23, 24]:

$$
\begin{aligned}
B_{n-1}(i, j) &= \frac{\sum_{k=1}^{n-1} f_k(i, j)}{n-1}, \\
B_{n-1}(i, j) &= (1 - \alpha)B_{n-2}(i, j) + \alpha f_n(i, j), \\
\mu_{n-1}(i, j) &= \frac{1}{n-1}\sum_{k=1}^{n-1} f_k(i, j), \\
\sigma_{n-1}^2(i, j) &= \frac{1}{n-1}\sum_{k=1}^{n-1}\left[f_k(i, j) - \mu(i, j)\right]^2,
\end{aligned}
\quad (16)
$$

where $f_k(i, j)$ the pixel gray value at the k image coordinate (i, j). When the H-th image is acquired [25, 26], the system will get the gray value. According to the above formula transformation process, the improved Gaussian function formula is as follows:

$$
\begin{aligned}
\mu_{n-1}(i, j) &= (1 - \alpha)\mu_{n-1}(i, j) + \alpha f_n(i, j), \\
\sigma_n^2(i, j) &= (1 - \alpha)\sigma_{n-1}^2(i, j) + \alpha\left[f_n(i, j) - \mu_n(i, j)\right]^2.
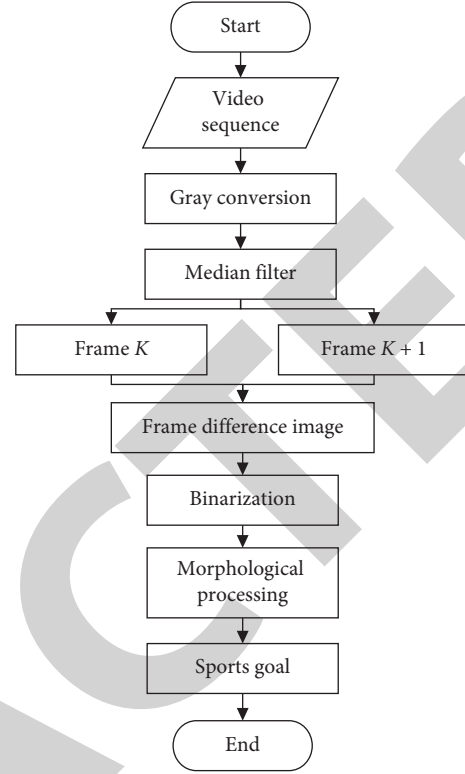\end{aligned}
\quad (17)
$$

Figure 1: Interframe difference method flowchart.

Firstly, the first image of video sequence is used as the initial background, and the current tilt image f_n at n time in the post w sequence matrix is used with the previous moment background image F_0 [27, 28]. A background model updating template is obtained by differential operation, and the common CB_{n-1} is obtained by binarization of the updated template with decision threshold BT_n:

$$
BT_n(i, j) = \begin{cases} 0 \\ 1 |F_n(i, j) - CB_{n-1}(i, j)| \end{cases} > D. \quad (18)
$$

Background image CB_n is updated as follows:

$$
CB_n(i, j) = \begin{cases} CB_{n-1}(i, j), \\ \omega_1 CB_{n-1}(i, j) + \omega_2 F_n(i, j). \end{cases} \quad (19)
$$

3. Hyperopia Scene Detection and Adaptive Video Processing Experiment Design

3.1. System Design.

The image acquisition subsystem divides the video information collected by the USB camera into two parts. The first part is transmitted to the real-time monitoring client through the network; the other part enters the moving target detection subsystem. When a moving target passes in the surveillance area, the system will quickly detect the target. The subsystem calls the alarm subsystem to perform alarm work and at the same time compresses and saves the video pictures and several continuous screenshots when moving objects appear in the monitoring area through the image compression subsystem and the storage subsystem

and saves them locally. When a user wants to find a moving target, he can view the video or picture through the WEB subsystem of the network to quickly find the target object. The process is shown in Figure 2.

3.2. Test Subject. In order to compare the effect and processing speed of the remote scene detection system in this article, the system in this article is compared with several common systems that use other algorithms. The main comparison algorithms are GMG algorithm, GMM detection algorithm, IMBS detection algorithm, KDE detection algorithm, and VIBE detection algorithm; the algorithm in this article is an improved VIBE detection algorithm. The video images of the experimental database are used for qualitative comparison, and these experimental data are collected for data analysis to draw conclusions. In this paper, the precision rate and recall rate are used as indicators to evaluate the detection algorithm, and the processing speed is also used as the evaluation indicator. After inspection, it was found that there was a sentence segmentation phenomenon, so the incomplete sentence was deleted directly. The larger the value, the better the detection of the algorithm; the smaller the value of processing speed, the better the detection of the algorithm.

3.3. Experimental Method. First, it is necessary to fill the image with holes to remove small target impurities, that is, when there are disturbances in the background, such as the shaking of leaves, or the misdetection caused by changes in illumination, this will cause some false small targets in the foreground image. And there will be some defects and holes in the detected moving target, which will have an impact on the subsequent contour detection and tracking. In intelligent monitoring, some false targets will cause false alarms and incomplete targets will cause missed detection. Secondly, the circumscribed rectangle of the detected target after improvement needs to be used as the initial target of the MeanShift tracking algorithm. The contour of the final target image will be detected. After the moving target is obtained, the moving target to be tracked is selected, that is, select tracking target area, tracking the target.

3.4. Statistical Data Processing Method. SPSS23.0 software was used for data processing, and the count data were expressed in percentage (%); k is the number of data in this experiment, σ^2 is the variance of all survey results, and $P < 0.05$ indicates that the difference is statistically significant. The formula for calculating reliability is as follows:

$$a = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma^2}\right). \tag{20}$$

4. Experimental Hyperopia Scene Detection and Adaptive Video Processing

4.1. Evaluation Index System Based on Index Reliability Testing. Reliability refers to the stability and reliability of the questionnaire. This article adopts the α coefficient method created by L.J. Cronbach. The α coefficient can be obtained by reliability analysis in SPSS software. It is generally believed that the α coefficient above 0.8 indicates that the effect of the index setting is very good, and above 0.7 is also acceptable. Here we analyze the reliability of each type of object, and the reliability index we choose for each type of object is slightly different. The results are shown in Table 1.

It can be seen from Table 1 that there are certain differences in the processing results of different scenes, but these processing results can be optimized technically, so the difference between the processing effects of different scenes has an acceptable impact on this experiment ($\alpha > 0.7$). In order to better illustrate the detection effect, corresponding indicators are used to evaluate the detection effect. According to the segmented real foreground image, many indicators are used to evaluate the background modeling algorithm.

4.2. Comparison and Analysis of Detection Algorithms

4.2.1. Effect Analysis of Baseline Video Library Scene. First, we analyze the scenes of the baseline video library and compare the effect analysis of the algorithm used in the system with that of other algorithms. Here, five scenes are selected for comparative analysis: highway scene, indirect motion scene, scene with similar color, small target scene, and hyperopia scene. The results of the target scene and the hyperopia scene are shown in Table 2. We make a bar graph based on this result, as shown in Figure 3.

It can be seen from Figure 3 that although the GMG detection algorithm can detect a relatively complete target, the algorithm is more sensitive to illumination changes and background disturbance; the GMM detection algorithm is more obvious in the detection of the target wheel, but it is not easy to find it inside the target and when the target is hidden. When the background is more complex, the detection effect is not very good; the IMBS detection algorithm has a good detection effect, but it is very sensitive to light changes and background disturbance; the KDE detection algorithm has a bad detection effect, and the algorithm is more sensitive to the contour of the object. The detection rate is very high; the VIBE detection algorithm is more robust to illumination and background disturbance; the detection algorithm in this paper maintains the advantages of the VIBE algorithm and further improves the detection accuracy.

4.2.2. Effect Analysis of Dynamic Background Video Library Scene. We analyze the scenes of the dynamic background video library and compare the effect analysis of the algorithm used in the system with that of other algorithms. Here, five scenes are selected for comparative analysis: the oar shaking scene, water wave shaking scene, leaf shaking scene, fountain scene, and hyperopia scene. The results of the fountain scene and the hyperopia scene are shown in Table 3. We make a histogram based on this result, as shown in Figure 4.

It can be seen from Figure 4 that the GMG detection algorithm is very sensitive to background disturbances, and background disturbances have a great influence on the detection effect; the detection effect of GMM detection algorithm is not ideal for background disturbances; the IMBS
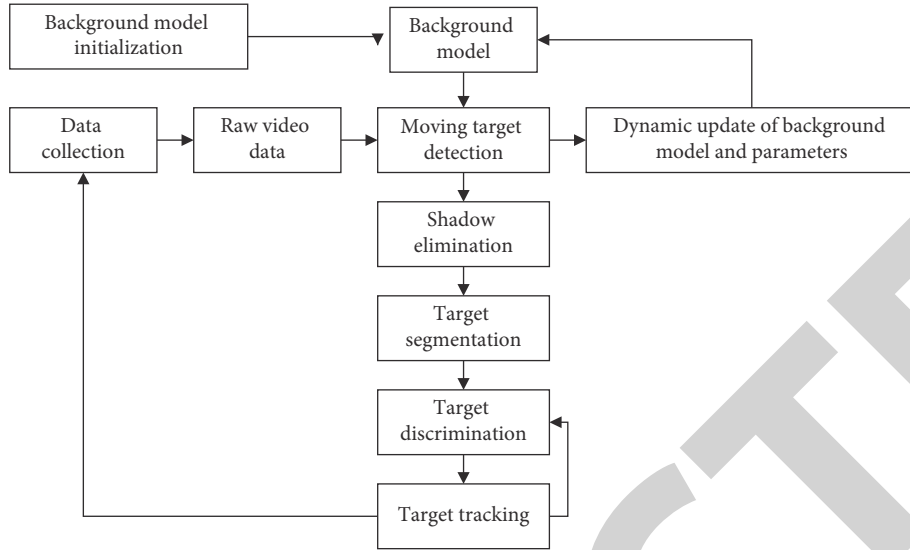
FIGURE 2: Moving target detection and tracking program structure diagram.

TABLE 1: Summary of reliability test results.

Category	Index combination	Alpha coefficient (α)
Highway scene	Heuristic rule Spatial correlation	0.8632
Indirect motion scene	Heuristic rule Spatial correlation	0.8447
Scenes with similar colors	Heuristic rule Spatial correlation	0.7365
Small target scene	Heuristic rule Spatial correlation	0.7746
Hyperopia scene	Heuristic rule Spatial correlation	0.8124

detection algorithm is very sensitive to the interference of leaf shaking, and there is a lot of false detection rate; VIBE detection algorithm still has a little false detection; the detection algorithm in this paper reduces the target of false detection for the rippling of water waves and the shaking of leaves, and it can also detect moving targets.

4.2.3. Effect Analysis of Camera Shake Video Library Scene. We analyze the scenes of the camera shake video library and compare the effect analysis of the algorithm used in the system with that of other algorithms. Here, five scenes are selected for comparative analysis, namely: badminton court scene, highway scene, zebra crossing scene, road scene, and hyperopia scene. The results are shown in Table 4, and we make a line chart based on this result, as shown in Figure 5.

It can be seen from Figure 5 that the GMG detection algorithm has a relatively high false detection rate for camera shake. Although the complexity of the GMM detection algorithm is reduced, its detection accuracy is not very high; although the KDE detection algorithm has a high detection rate, there are still errors, while the IMBS detection algorithm has large errors, and its detection effect needs to be further improved; the false detection rate of the VIBE

detection algorithm is relatively reduced, but there will still be some false warnings. The detection algorithm in this article guarantees the detection accuracy and reduces the false detection rate.

4.3. Evaluation Index Analysis

4.3.1. Analysis of the Accuracy Index of Dynamic Background Video Library Scene Detection. We analyze the detection accuracy index of dynamic background video library scenes and compare the algorithm used in this paper with other algorithms to analyze the motion state of people in the video and determine its accuracy. Here, five scenes are selected for comparative analysis. They are badminton court scene, basketball court scene, street scene, park scene, and hyperopia scene. The results are shown in Table 5. We make a combined picture based on this result, as shown in Figure 6.

It can be seen from Figure 6 that for the five scenes with dynamic background, there are fast-moving people or indirect-moving people, there are changes in illumination, and the leaves are slightly shaking. The improved algorithm in this paper is relatively stable in the detection accuracy index value. Compared with the original VIBE algorithm, it has increased to a certain percentage, so it verifies the effectiveness of the improved method in this paper.

4.3.2. Analysis of the Recall Rate Index of Dynamic Background Video Library Scene Detection. We analyze the detection recall rate index for the scene of the dynamic background video library and compare the algorithm used in this paper with other algorithms to analyze the motion state of people in the video and determine the recall rate. The results are shown in Table 6. This result makes a bar graph, as shown in Figure 7.

The improved algorithm in this paper has different recall rates for different scenarios. As can be seen from Figure 7, the recall rate is higher than that of the VIBE algorithm when

TABLE 2: Baseline video library scene effect analysis data.

| Algorithm	Highway scene	Indirect motion scene	Scenes with similar colors	Small target scene	Hyperopia scene
GMG detection algorithm	1.42	1.6	1.83	2.39	2.56
GMM detection algorithm	1.62	1.85	2.33	2.45	3.11
IMBS detection algorithm	1.89	2.07	2.74	2.81	3.4
KDE detection algorithm	2.17	2.73	3.09	3.22	3.58
VIBE detection algorithm	2.76	2.8	3.16	3.52	3.71
Detection algorithm of this article	2.84	3.35	3.66	3.89	4.99
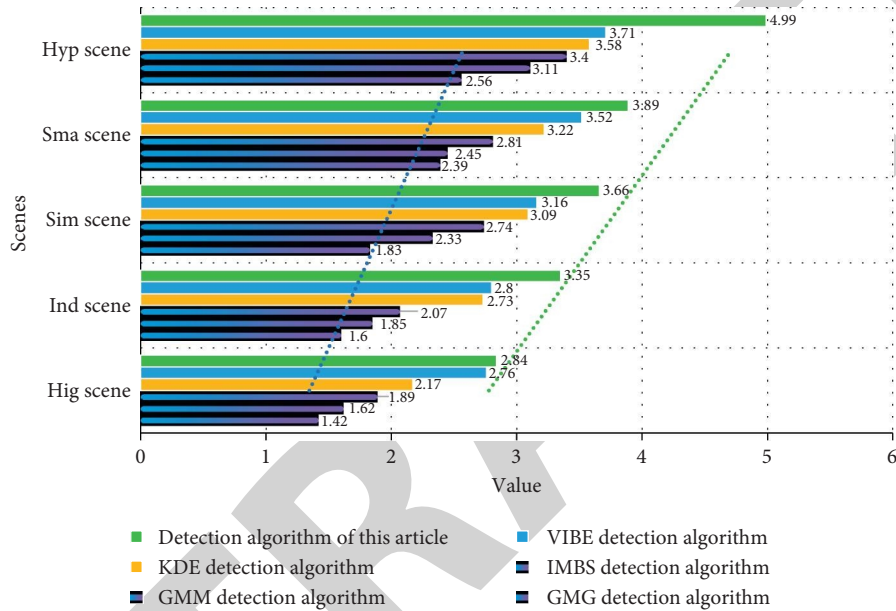
FIGURE 3: Effect analysis diagram of baseline video library scene.

TABLE 3: Effect analysis data of dynamic background video library scene.

Algorithm	Oar shaking scene	Water wave shaking scene	Leaf shaking scene	Fountain scene	Hyperopia scene
GMG detection algorithm	1.58	1.35	1.68	2.32	2.71
GMM detection algorithm	2.11	1.57	2.08	2.33	3.00
IMBS detection algorithm	2.21	2.04	2.13	2.81	3.31
KDE detection algorithm	2.39	2.17	3.11	3.01	3.52
VIBE detection algorithm	3.18	2.42	3.04	3.77	3.79
Detection algorithm of this article	3.23	3.12	3.63	3.93	4.16

the motion state is relatively uniform, and when there are more motion states, the recall rate is slightly better than that of the VIBE algorithm. It decreased about 0.05 or so, but the relative accuracy rate increased by about 0.12, and the increase ratio is significantly higher than the decrease ratio.

4.3.3. Analysis of the Speed Index of the Dynamic Background Video Library Scene Detection. We analyze the detection speed index of the scene of the dynamic background video library and compare the algorithm used in this paper with other algorithms to analyze the motion state of people in the video and determine the detection speed. The results are shown in Table 7. The result is a histogram, as shown in Figure 8.

The improved algorithm in this paper has different detection speed changes for different scenarios. From Figure 8, it can be seen that when the motion state is relatively uniform, the increase in detection speed is not too large. When there are more motion states, the detection speed increases. It can be clearly observed.

4.3.4. Analysis of the Accuracy Index of Camera Shake Video Library Scene Detection. We analyze the detection accuracy index of the camera shake video library scene and compare the algorithm used in this paper with other algorithms to analyze the motion state of people in the video and determine its accuracy. Here, five scenes are selected for
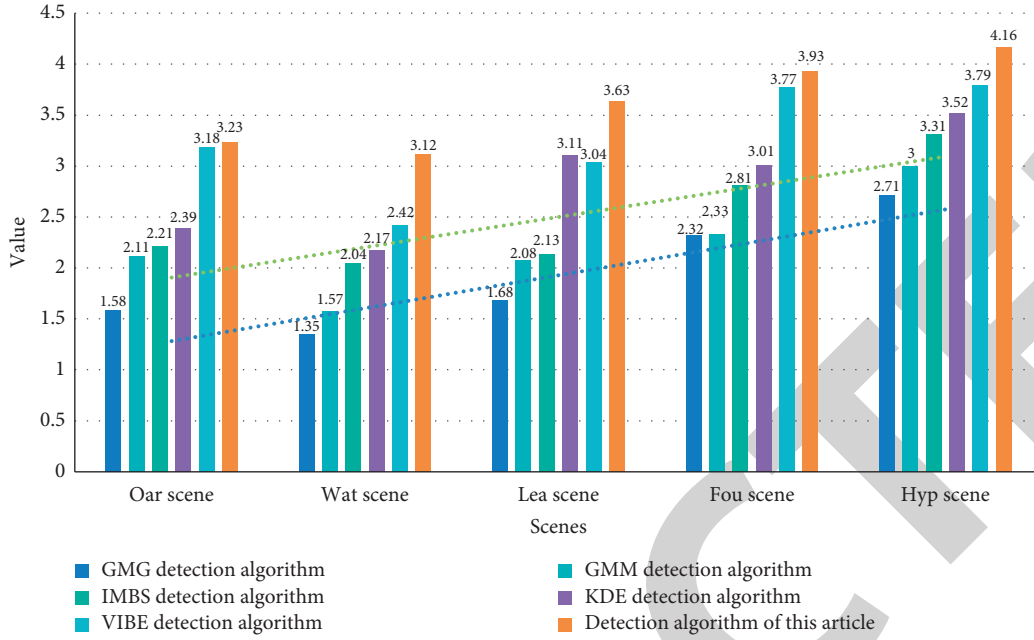
FIGURE 4: Effect analysis diagram of dynamic background video library scene.

TABLE 4: Effect analysis data of camera shake video library scene.

| Algorithm	Badminton court scene	Highway scene	Zebra crossing scene	Road scene	Hyperopia scene
GMG detection algorithm	5.01	4.12	3.84	3.45	4.87
GMM detection algorithm	1.95	2.19	3.02	5.52	4.32
IMBS detection algorithm	2.50	1.32	1.88	2.59	3.70
KDE detection algorithm	2.81	4.15	6.20	5.65	6.53
VIBE detection algorithm	3.41	3.12	4.83	5.19	7.03
Detection algorithm of this article	5.12	4.91	4.80	4.82	7.79
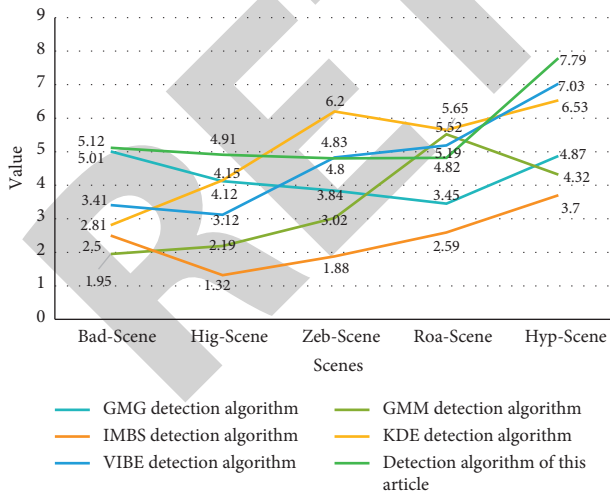
FIGURE 5: Effect analysis diagram of camera shake video library scene.

comparison and analysis. These are subway station scene, highway scene, zebra crossing scene, road scene, and far vision scene. The results are shown in Table 8. We make a histogram based on this result, as shown in Figure 9.

It can be seen from Figure 9 and from the experimental data that for the five complex scenes of camera shake, the effect of the improved algorithm in this paper is relatively stable in the detection accuracy index value, and the false detection rate is lower than that of other algorithms. The VIBE algorithm has been improved, so for camera shake, the detection effect is better than that of other algorithms.

4.3.5. Analysis of the Recall Rate Index of Camera Shake Video Library Scene Detection. We analyze the detection recall rate index of the camera shake video library scene and compare the algorithm used in this paper with other algorithms to analyze the motion state of people in the video and judge the recall rate. The results are shown in Table 9. This result makes a histogram, as shown in Figure 10.

The improved algorithm in this article has different recall rates for different scenarios. As shown in Figure 10, when the motion state is relatively uniform, the recall rate is higher than that of the VIBE algorithm, and when the motion state increases, it is higher than that of the VIBE algorithm. There is only an increase of about 0.02, but the relative accuracy rate has increased by about 0.08, and the improvement ratio is significantly higher.

Table 5: Data of accuracy index for dynamic background video library scene detection.

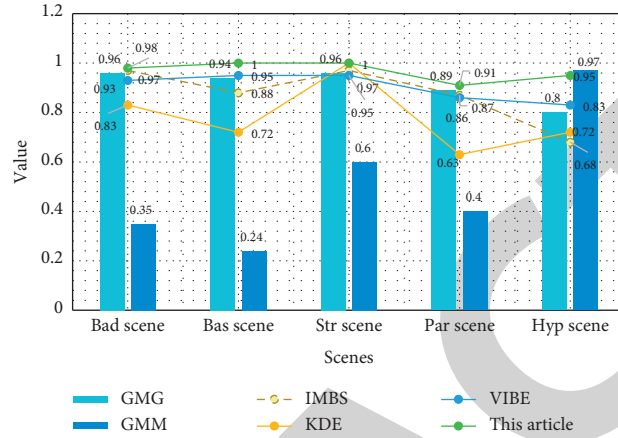| Detection accuracy	GMG	GMM	IMBS	KDE	VIBE	This article
Badminton court scene	0.96	0.35	0.97	0.83	0.93	0.98
Basketball court scene	0.94	0.24	0.88	0.72	0.95	1.00
Street scene	0.96	0.60	0.97	1.00	0.95	1.00
Park scene	0.89	0.40	0.87	0.63	0.86	0.91
Hyperopia scene	0.80	0.97	0.68	0.72	0.83	0.95

Figure 6: The accuracy index map of scene detection for dynamic background video library.

Table 6: Data of recall rate index for dynamic background video library scene detection.

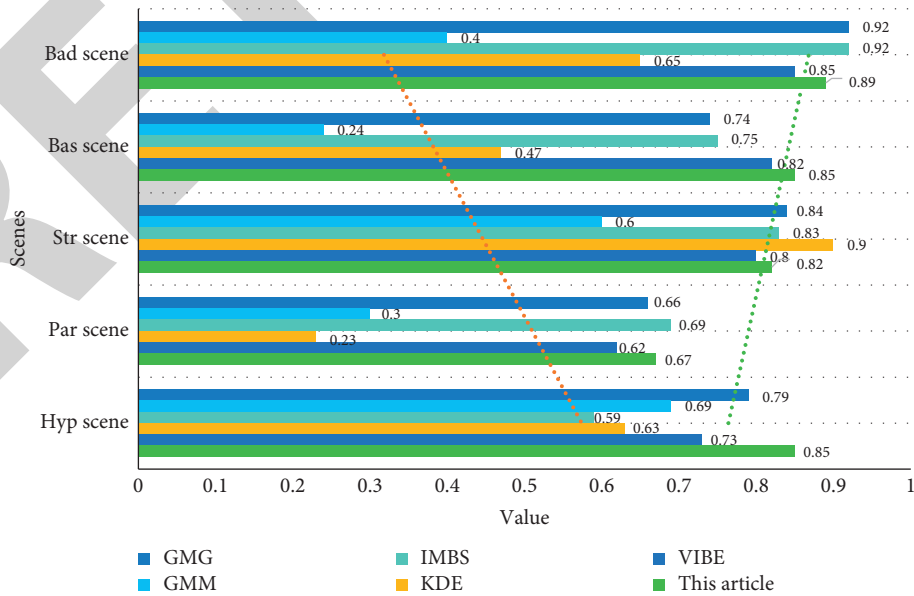| Test recall rate	GMG	GMM	IMBS	KDE	VIBE	This article
Badminton court scene	0.92	0.40	0.92	0.65	0.85	0.89
Basketball court scene	0.74	0.24	0.75	0.47	0.82	0.85
Street scene	0.84	0.60	0.83	0.90	0.80	0.82
Park scene	0.66	0.30	0.69	0.23	0.62	0.67
Hyperopia scene	0.79	0.69	0.59	0.63	0.73	0.85

Figure 7: Recall rate indicator diagram for scene detection of dynamic background video library.

TABLE 7: Data of speed indicators for scene detection of dynamic background video library.

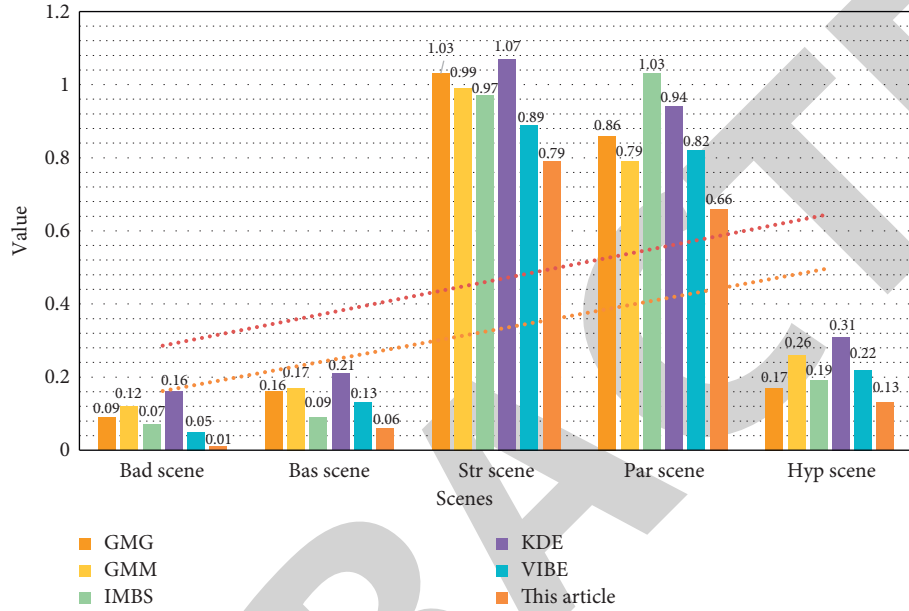| Detection speed	GMG	GMM	IMBS	KDE	VIBE	This article
Badminton court scene	0.09	0.12	0.07	0.16	0.05	0.01
Basketball court scene	0.16	0.17	0.09	0.21	0.13	0.06
Street scene	1.03	0.99	0.97	1.07	0.89	0.79
Park scene	0.86	0.79	1.03	0.94	0.82	0.66
Hyperopia scene	0.17	0.26	0.19	0.31	0.22	0.13

FIGURE 8: Speed detection speed indicator diagram for scene of dynamic background video library.

TABLE 8: Data of accuracy index for camera shake video library scene detection.

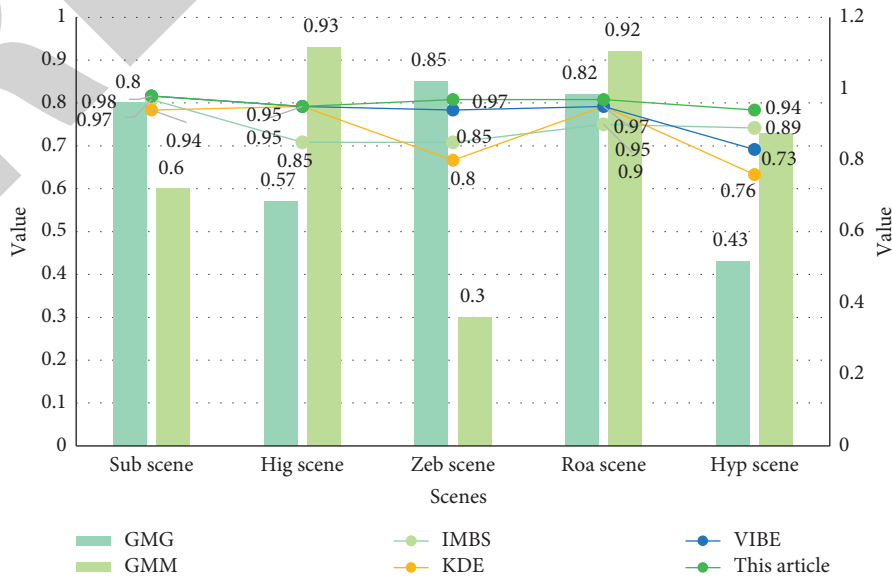| Detection accuracy	GMG	GMM	IMBS	KDE	VIBE	This article
Subway station scene	0.80	0.60	0.97	0.94	0.98	0.98
Highway scene	0.57	0.93	0.85	0.95	0.95	0.95
Zebra crossing scene	0.85	0.30	0.85	0.80	0.94	0.97
Road scene	0.82	0.92	0.90	0.95	0.95	0.97
Hyperopia scene	0.43	0.73	0.89	0.76	0.83	0.94

FIGURE 9: The accuracy index chart of camera shake video library scene detection.

TABLE 9: Data of recall rate index for camera shake video library scene detection.

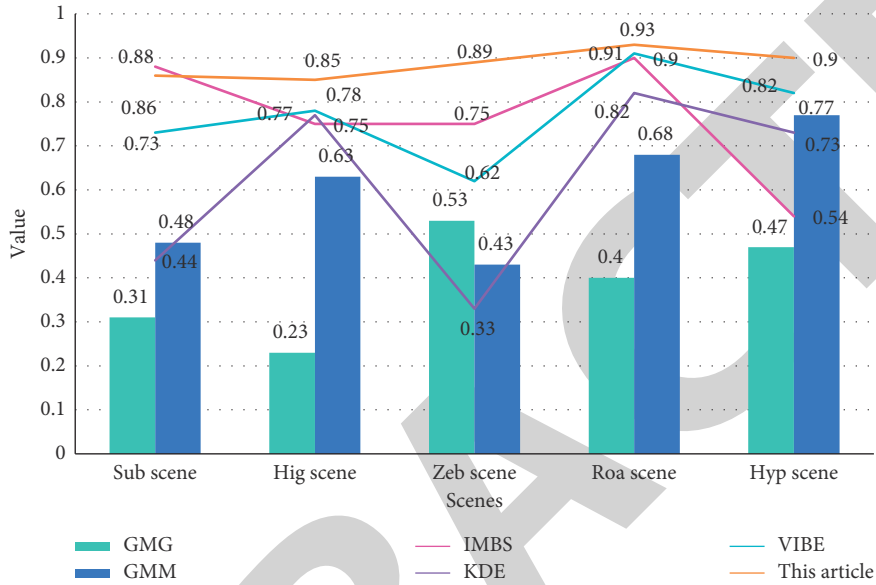| Detection accuracy	GMG	GMM	IMBS	KDE	VIBE	This article
Subway station scene	0.31	0.48	0.88	0.44	0.73	0.86
Highway scene	0.23	0.63	0.75	0.77	0.78	0.85
Zebra crossing scene	0.53	0.43	0.75	0.33	0.62	0.89
Road scene	0.40	0.68	0.90	0.82	0.91	0.93
Hyperopia scene	0.47	0.77	0.54	0.73	0.82	0.90

FIGURE 10: Recall rate index graph for camera shake video library scene detection.

TABLE 10: Data of speed indicators for camera shake video library scene detection.

Detection accuracy	GMG	GMM	IMBS	KDE	VIBE	This article
Subway station scene	0.89	0.76	1.03	0.88	0.64	0.37
Highway scene	0.36	0.13	0.19	0.28	0.15	0.01
Zebra crossing scene	0.35	0.23	0.27	0.65	0.33	0.06
Road scene	0.16	0.34	0.26	0.22	0.19	0.01
Hyperopia scene	0.17	0.26	0.19	0.31	0.22	0.13

FIGURE 11: Speed detection indicators for camera shake video library scene.

4.3.6. Analysis of the Speed Index of Camera Shake Video Library Scene Detection. We analyze the detection speed index of the camera shake video library scene and compare the algorithm used in this paper with other algorithms to analyze the motion state of people in the video and determine the detection speed. The results are shown in Table 10. The result is a histogram, as shown in Figure 11.

The improved algorithm in this paper has different detection speeds for different scenarios. From Figure 11, it can be seen that when the motion state is relatively uniform, the detection speed will increase slightly due to the increase in the number of people to be detected. When the motion state increases, it will not increase because of the increase in the number of people tested.
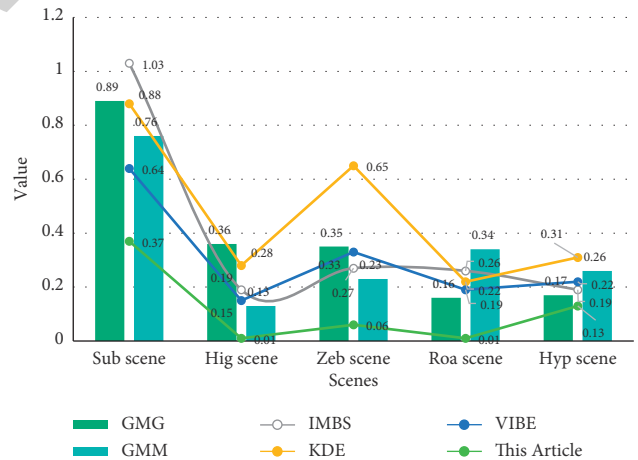
5. Conclusions

This topic is derived from a target image search project of a certain research. It detects intrusive moving targets in the monitored warning area. The background of the video is more complicated. Since the characteristics of the moving target are unknown, the background subtraction method is used. The key step of this method is the establishment of the background

model, the target extraction, and the update of the background model method; only the generated background model has good adaptability to the changes of the complex background and can reflect the real background image, and the target needs to be extracted after the background model is established. Departments need to adopt effective target extraction methods. In order to improve the accuracy of moving target detection, this paper studies the method of background model establishment and target extraction and proposes its own improvement.

Afterwards, an improved detection algorithm is used to clean and delete the shadow area of the moving target in order to better track the target. Combining the improved VIBE algorithm and the shadow removal of multifeature fusion, it has been verified by experiments to effectively remove the shadow of the detected target, making the detected moving target more accurate. Finally, we detect the maximum contour of the target from which the shadow is removed, use the minimum bounding rectangle of the target as the initial window of MeanShift tracking, and combine the detection results to track the target. Through comparison and verification, the MeanShift tracking algorithm combined with the improved detection algorithm in this paper has better tracking effect than the MeanShift algorithm without combined improvement. In this paper, the improved detection and tracking methods are applied to the video images captured by the drone, and the moving targets in the video are detected and the targets are tracked. It can be seen from the experimental results that the moving targets in the video taken by the drone are clearer. The target detection and tracking effect is relatively ideal, which further proves the effectiveness of the improved method in this paper.

By applying the detection and tracking system of moving objects to sports video, the system will process the collected motion information data and obtain the motion parameters of human body. It is very important to send some data to improve the quality of athletes and coaches and expand the development direction of athletes. This paper is based on the background of the research and development of sports video object detection and tracking system, through the Gauss model to separate the scene, through the eigen extraction to extract the characteristics of the players in the stadium, and finally through the difference method to identify the video content, which opens a new idea for the detection and adaptive processing of the athletes in sports video.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] D. Shukla and M. Sharma, "A novel video scene change detection using successive estimation of statistical measure and HiBiSLI method," *International Journal of Information Technology*, vol. 11, no. 1, pp. 47–54, 2019.

[2] X. Yang, D. Wang, W. Zhang, P. Peng, and L. Feng, "Detection of ventricular tachycardia and fibrillation using adaptive variational mode decomposition and boosted-CART classifier," *Biomedical Signal Processing and Control*, vol. 39, pp. 219–229, 2018.

[3] H. Chebi, M. D. Acheli, and M. Kesraoui, "Dynamic detection of abnormalities in video analysis of crowd behavior with DBSCAN and neural networks," *Advances in Science, Technology and Engineering Systems Journal*, vol. 1, no. 5, pp. 56–63, 2016.

[4] Y. Wang, A. W.-C. Y. Hu, and C.-T. Li, "ENF based video forgery detection algorithm," *International Journal of Digital Crime and Forensics*, vol. 12, no. 1, pp. 131–156, 2020.

[5] Z. Huang, J. B. Sui, and G. Jiang, "An intelligent ship image/video detection and classification method with improved regressive deep convolutional neural network," *Complexity*, vol. 2020, Article ID 1520872, 11 pages, 2020.

[6] C. C. Huang, C. E. Lee, and V. L. L. Thing, "A novel video forgery detection model based on triangular polarity feature classification," *International Journal of Digital Crime and Forensics*, vol. 12, no. 1, pp. 14–34, 2020.

[7] Q. Abbas, M. A. Ibrahim, and M. A. Jaffar, "Video scene analysis: an overview and challenges on deep learning algorithms," *Multimedia Tools and Applications*, vol. 77, no. 16, pp. 20415–20453, 2018.

[8] Z. Shao, Z. L. Wang, and W. W. Wu Du, "Saliency-aware convolution neural network for ship detection in surveillance video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 3, pp. 781–794, 2020.

[9] L. Pan, T. Pezzuti, and C. W. Pechmann, "Hyperopia and frugality: different motivational drivers and yet similar effects on consumer spending," *Journal of Business Research*, vol. 95, pp. 347–356, 2019.

[10] S. Feldman, M. M. W. Peterseim, R. H. Trivedi et al., "Detecting high hyperopia: the plus lens test and the spot vision screener," *Journal of Pediatric Ophthalmology & Strabismus*, vol. 54, no. 3, pp. 1–5, 2016.

[11] G. S. Ying and G. E. Quinn, "Uncorrected hyperopia and preschool early literacy: results of the Vision in Preschoolers-Hyperopia in Preschoolers (VIP-HIP) Study," *Ophthalmology*, vol. 123, no. 4, pp. 681–689, 2016.

[12] M. Garcia-Gonzalez, M. L. Iglesias-Iglesias, P. D. Rodriguez-Casanova et al., "Femtosecond laser–assisted LASIK with and without the adjuvant use of mitomycin C to correct hyperopia," *Journal of Refractive Surgery (Thorofare, N.J.: 1995)*, vol. 34, no. 1, pp. 23–28, 2018.

[13] C. Roesler and T. Kohnen, "Changes of functional optical zone after LASIK for hyperopia and hyperopic astigmatism," *Journal of Refractive Surgery (Thorofare, N.J.: 1995)*, vol. 34, no. 7, pp. 476–481, 2018.

[14] M. D. Basar, A. D. Duru, and A. Akan, "Emotional state detection based on common spatial patterns of EEG," *Signal, Image and Video Processing*, vol. 14, no. 3, pp. 473–481, 2020.

[15] W. Sun, J. D. Yan, and C. Sun, "Small-scale moving target detection in aerial image by deep inverse reinforcement learning," *Soft Computing*, vol. 24, no. 8, pp. 5897–5908, 2020.

[16] Naoki, Tanimura, Natsuko et al., "Visual function and functional decline in patients with waterclefts," *Investigative*

Ophthalmology & Visual Science, vol. 60, no. 10, pp. 3652–3658, 2019.

[17] Y. Lu and S. An, "Research on sports video detection technology motion 3D reconstruction based on hidden Markov model," *Cluster Computing*, vol. 23, no. 3, pp. 1899–1909, 2020.

[18] M. Alizadeh and M. Sharifkhani, "Compressed domain moving object detection based on CRF," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 3, pp. 674–684, 2020.

[19] T. Rice, "Hudl video software developer: agile sports technologies," *International Sport Coaching Journal*, vol. 3, no. 3, pp. 364-365, 2016.

[20] Y. Yuan, J. Chu, L. Leng et al., "A scale-adaptive object-tracking algorithm with occlusion detection," *Eurasip Journal on Image & Video Processing*, vol. 2020, no. 1, pp. 1–15, 2020.

[21] M. Qi, A. Y. Wang, and J. Luo, "Sports video captioning via attentive motion representation and group relationship modeling," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 8, pp. 2617–2633, 2020.

[22] A. Tejero-De-Pablos, T. Y. Nakashima, and M. N. LinnaYokoya, "Summarization of user-generated sports video by using deep action recognition features," *IEEE Transactions on Multimedia*, vol. 20, no. 8, pp. 2000–2011, 2018.

[23] K. C. Rahtu and A. Bhandari, "A euclidean distance based super resolution method for sub pixel target detection in hyper spectral images," vol. 15, no. 2, pp. 104–124, 2020.

[24] H. Wei and G. Bu, "The Fusion algorithm based on frame difference and optical flow for sports video detection," *C e Ca*, vol. 42, no. 4, pp. 1498–1502, 2017.

[25] A. Khan, D. Windridge, and J. Kittler, "Multilevel Chinese takeaway process and label-based processes for rule induction in the context of automated sports video annotation," *IEEE Transactions on Cybernetics*, vol. 44, no. 10, pp. 1910–1923, 2017.

[26] Z. Yu, F. Shuo, S. Xiaohua et al., "Research on tracking algorithm for fast-moving target in sport video," *Journal of Computational & Theoretical Nanoence*, vol. 14, no. 1, pp. 230–236, 2017.

[27] M. Zhuang, "Sports video structure analysis and feature extraction in long jump video," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 9, no. 9, pp. 83–92, 2016.

[28] W. . Xinghua, "Application research on sports video analysis using object segmentation algorithms," *International Journal of Simulation: Systems, Ence & Technology*, vol. 17, no. 36, pp. 8.1–8.9, 2016.