Computer Life (CPL) ISSN: 1819-4818
DELIVERING QUALITY SCIENCE TO THE WORLD

Public safety prewarning based on machine learning

Meng Liu [1, a], Tianle Dai[2, b] and Ranran Wei[2, c]

[1]National Demonstration Center for Experimental Electrical and electronic technology, Chanchun University of Science and Techonlgy, ChangChun, Jilin, China

[2]National Demonstration Center for Experimental Electrical and electronic technology, Chanchun University of Science and Techonlgy, ChangChun, Jilin, China

[a]869218104@qq.com, [b]438289381@qq.com, [c]1603409002@qq.com

Abstract: Public safety has become a topic of greater concern to the population. At present, the traditional public safety work mainly adopts the security patrol and video monitoring methods, which makes it difficult to realize the blind zone and deal with sudden security problems in a timely and effective manner. In response, The real-time monitoring function of the camera is used by the model to analyze the pedestrian movement behavior. Background modeling method, HOG feature extraction of objects, support vector machine and direction conversion vector data graph are combined by this model. The algorithm is verified by video acquired in real time. Moving target recognition, pedestrian detection and human behavior recognition are applied to real-time monitoring and processing. Potential safety problems are recognized in advance, and security personnel are informed in a timely manner. Security problems are effectively warned and dealt with in a timely manner.
Keywords: Public safety, pedestrian detection, real-time tracking, support vector machine, behavior recognition.

1. INTRODUCTION

In recent years, as more people choose to travel, public safety becomes more important. Public safety accidents are consist of abnormal behaviors such as trailing theft and fighting. At present, the traditional security work mainly adopts the patrol of security personnel and manual monitoring observation and other methods. It has the disadvantages of poor timeliness and high manpower cost, and can not do the early warning and timely processing work before the sudden situation occurs, The use of manual shift to view surveillance caused a great deal of human resources and financial waste, it even caused a lot of missed or misjudged phenomena[1].

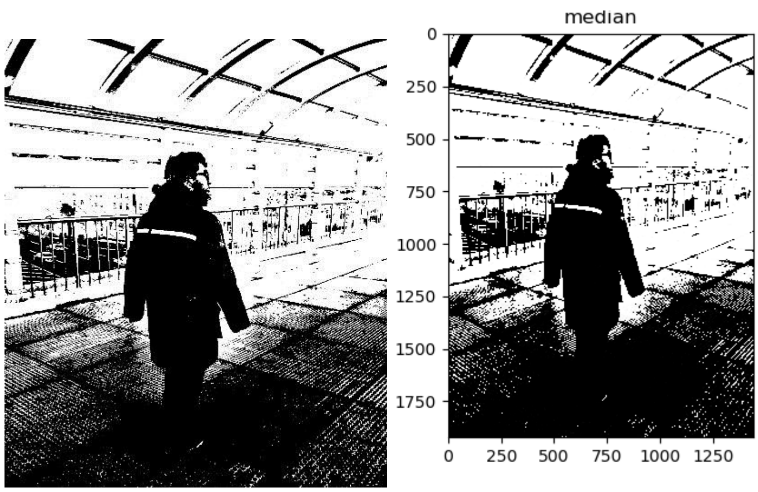
With the development of machine learning, it is beneficial to realize the detection of moving objects, pedestrian classification, pedestrian real-time tracking and human behavior recognition. Through the real-time tracking of the target and the establishment of the directional conversion vector data graph, the real-time analysis of human behavior and walking speed and walking direction, the real-time analysis and alarm processing of possible abnormal behavior is realized, which is beneficial to social public safety and the conservation of related resources.

2. MOTION OBJECT DETECTION

The computer uses the correlation between adjacent frames in the video to detect the moving objects in the video, where the camera is still shooting, and the core is to use the mixed Gaussian background modeling method.

When the background model is initialized, the number of samples added to the sample is counted, the key is to count the neighborhood in a certain location, the number of times a pixel is saved to the background sample, when the number of times more than 3 times, when the pixel is collected again, the sample is discarded and the pixels in that position are saved directly into the template sample. The corresponding classification of pixels at various locations of the current image is implemented, and it is defined as a background image or foreground pixel. For objects moving inside the video, the pixels of all positions in the current frame image are matched to the background template in real time. If the pixel point matches the background template, the pixel point corresponds to the background point, at which point the pixel value is set to 0, which is black, and when the pixel point does not match the background template, the pixel point is the moving point and the pixel value is set to 255, which is white.

In the actual process, there will be noise disturbance, such as the disturbance of leaves, resulting in more small white dots in the detection background, which will cause great interference to the final result. There is a small black hole in the moving area detected by background modeling, and the contour of the moving object appears incomplete. For noise problems, when median filtering is used, the output value of median filtering is the median point in the template, and the noise is not easy to be selected. The effect is more obvious [2].



(a) Effect diagram after binarization (b) The effect diagram after the mean filter

Fig. 1. Effect diagram after different processing

3. PEDESTRIAN DETECTION BASED ON SVM

Support Vector Machine(SVM) can map low-dimensional data features to high-dimensional space through kernel function, divided the data samples by super flat, effectively solve the linear inseparable situation, here is a one-to-one method to establish the classification surface among two different categories, by the decision function, convolution neural network to identify which of the two categories.

According to theoretical analysis, Support Vector Machines can be summed up as two minimization problems.

The detection of the moving objects in the video needs to be further identified in all moving objects of the movement of people, the current traditional methods of identifying the human body are numerous, here the use of HOG features, By calculating the characteristics of the gradient histogram of the local region of the image, all the local features are concatenated, which makes up the unique HOG characteristics of the whole image, The main steps include the following: first standardize the image, and then the corresponding calculation gradient, divide its cells and regions, and finally combine the characteristics of all the above blocks, at this time to form the image corresponding to the HOG characteristics. At this time, the characteristics of hog mainly use gradient information, when the human body slightly deformation also has a good test results and is not affected by the change of light.

When using the HOG feature as the classification feature of moving objects, the first step is to detect the moving objects in the video, and finally according to the different characteristics of the corresponding classification.

The specific extraction process is as follows: Using the Boundingrect () function in open-cv3 to calculate the external matrix of the moving area of the object, the main parameters are the upper left vertex coordinates of the rectangle and the vertex coordinates in the lower right corner, at which point the motion area of the object can be determined, Then the specific parameters of the outbound rectangle, including the vertex coordinates and the corresponding width and height, are calculated, and the ROI region associated with the object is obtained from the current frame. At this time, the HOG characteristics of moving objects can be obtained, and according to the characteristics of HOG, pedestrian classifiers are used to identify pedestrians and complete the initial tracking work of the moving body[3].

4. REAL-TIME TRACKING OF PEDESTRIANS

When using SVM to classify moving objects, only when the classification result is a pedestrian, the tracking will continue. At this time, the difficulty of real-time tracking will be entered. In real life, pedestrians are prone to blockage during exercise, wearing similar colors, motion blur and other interference situations, which will cause great difficulties for pedestrians to track in real time.

When the tracking target has different degrees of occlusion, size and scale changes, and even motion blur, the KCF algorithm can achieve real-time tracking of targets. Moreover, there is no phenomenon such as character drift during the whole tracking process, and the effect of real-time tracking is more significant than that of the Mean-shift algorithm and the ASMS algorithm. Complete the real-time tracking process for pedestrians. Considering comprehensively, the KCF tracking algorithm is used to train the classifier.

The basic process of the KCF tracking algorithm is as follows:

(1) The location area are determined by corresponding to the candidate area in the current frame according to the location of the target of the previous frame.

(2) The classifier trained by the previous frame calculates the response value f corresponding to the candidate region by using Equation 4.1 and Equation 4.2, and detects the target position of the current frame.

K^z represents the kernel matrix of the training sample and the candidate region, and the training sample and the candidate region are the result of the target sample x and the candidate region Z cyclic shift.

$$k^z = c(k^{xz}) \tag{4.1}$$

$$f(z) = (k^2)^T \alpha \tag{4.2}$$

Where f(z) is a vector containing the detection results of all cyclic samples of the candidate region z, where in the position with the largest response value is the final result of the current frame tracking. In practice, in order to simplify the calculation, a corresponding Fourier transform is performed on it.

(3) Calculating the information characteristics of the target and background regions in the current frame, and using Equation 4.3, Equation 4.4 to track the target training classifier for the next frame. After adding the kernel function, the correlation coefficient matrix α is solved in the dual space.

$$\alpha = (k + \lambda I)^{-1} y \tag{4.3}$$

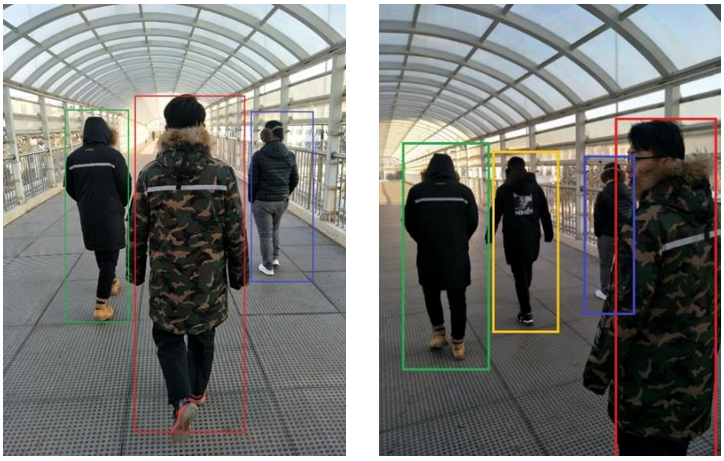
At this time, the Fourier transform of α is $\hat{\alpha}$.

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda} \tag{4.4}$$

Where k^{xx} refers to the first row element of the kernel matrix k=C(k^{xx}), and k^{xx} represents the autocorrelation of x in the frequency domain.

The classifier is similar to a filter, in which case the response value of each candidate region through the filter is calculated, and the target position in the current frame is the maximum position of the response value. In the process of training the classifier, the target information is defined as a positive sample, and the background information is defined as a negative sample. After the classifier training process is completed, the final tracking result of the previous frame is used as the center, and the surrounding areas are defined as candidate targets, and then each candidate target is independently filtered. At this time, the candidate target filtering response value is obtained, and the target are obtained. The area is the maximum response value at this time. In practice, the number of background samples is generally small, and a large number of related samples are obtained by cyclically shifting the samples. The algorithm is further simplified by operations such as Fourier Transform.

The multi-scale detection process is added to the KCF tracking algorithm to process the tracking process of the target deformation. After determining the optimal target position of the current frame, a different size detection is performed on the position, and the size difference is large, and three corresponding Response Peak, using three peaks to finally determine the optimal size of the target in this frame [4].



(a) The 20th frame (b) The 90th frame

Fig. 2. Different frame detection effects

5. REAL-TIME TRACKING OF PEDESTRIANS

In real life, observing a person's behavior mainly through his head movement, lower limb exercise, gesture recognition, etc. In the existing human behavior recognition methods, there are mainly template matching methods, state transition graph model behavior analysis methods and based on A behavioral analysis method for similar metrics.

Keyframes are the key information that best represents an image. The keyframes provide a good representation of the content inside the lens. In different situations, different numbers of choices are made, and the operation of key frames can be reduced by reducing the amount of video operations [5].

Key frame extraction based on shot boundary. A complete video is divided into multiple shots, and each of the divided shots is subjected to a corresponding extraction key frame operation, and the first frame, the middle frame and the last frame are taken as key frames.

Collecting the amplitude-weighted direction angle of the human body as the mainly feature of key frame extraction, due to the number of clusters existing in the clustering algorithm and the problem that the cluster center is difficult to determine, The unsupervised clustering algorithm is improved accordingly, and the automatic threshold selection is added, so that the unsupervised clustering algorithm can adaptively extract different key frames, and calculate the cross entropy between adjacent frame images for the images obtained after clustering. Difference degree sequence, A frame with a larger image entropy is selected as the final key frame. Finally, the direction angle of the moving target in the key frame image obtained by clustering is taken as the input on the direction conversion modulus data graph, and the complexity of the direction conversion modulus data graph is taken as the total characteristics of different behaviors of the human body to realize abnormal behavior detection.

5.1 Direction conversion vector data diagram

When the real-time motion state of an object's motion is studied, it mainly studies its motion speed and direction, and cannot show the degree of change of the motion state of the moving object, that is, the intensity. In the general direction histogram, the main factors are direction and speed, which can not reflect the degree of change of the state. The direction is converted between the two moments to visually represent the intensity of real-time motion change. During the entire monitoring process, the camera position remains unchanged, setting the camera to face direction 0^0, clockwise rotation to the positive direction, and each conversion angle is 30^0 during the acquisition process.

Setting the direction conversion modulus data of the region R be as follows:

$$H(t) = \{h_j(t)\}_{j=1...m} \tag{5.1}$$

Equation (5.1), m represents the number of intervals of the histogram, and t represents the acquisition of the t-th frame video image.

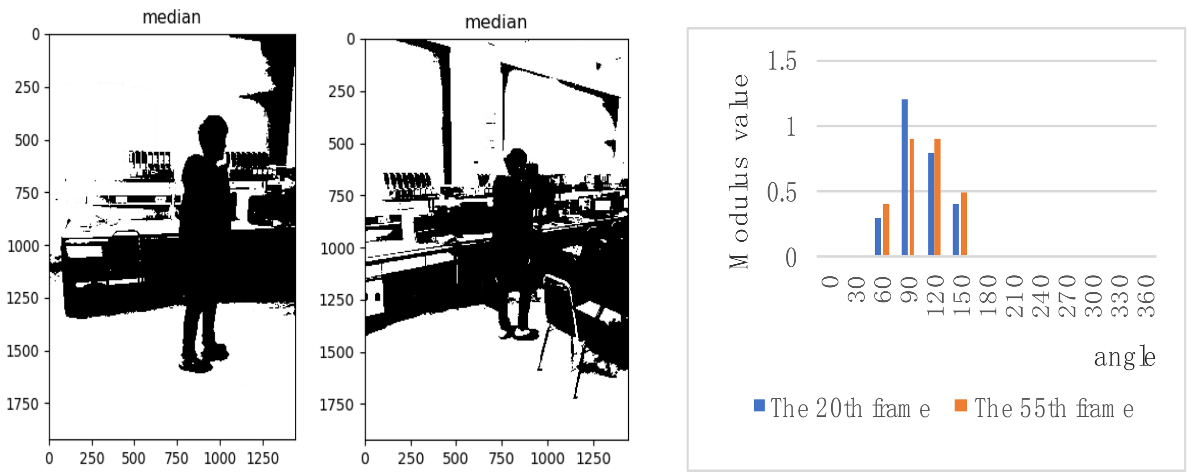
By adding the direction transformation vector, the small-angle directional angle mixing theory caused by small noise can be effectively suppressed, and the relevant features of the moving target can be better reflected at this time. The complexity of a histogram is defined:

$$E_{Ht} = -\Sigma_{j=1}^{m} h_j(t) \log_2(h_j(t)) \tag{5.2}$$

The size of the E_{Ht} value can intuitively reflect whether the behavioral state of the moving target is normal behavior. It is supposed that $E_{\max} = \max(E_{Ht}^1, E_{Ht}^2)$, E^1_{Ht}, E^2_{Ht} represent the complexity of the upper body movement state of the moving target and the complexity of the lower body movement state change. Under normal circumstances, the degree of movement of the upper body is generally small, and when abnormal behavior occurs, the movement state of the lower body changes greatly. It is simple to judge whether the behavior is normal behavior or abnormal behavior by the size of the complexity value. For the movement of pedestrians, the direction conversion modulus data map can be used to intuitively digitize the different states of pedestrian movement.

5.2 Behavioral feature extraction result

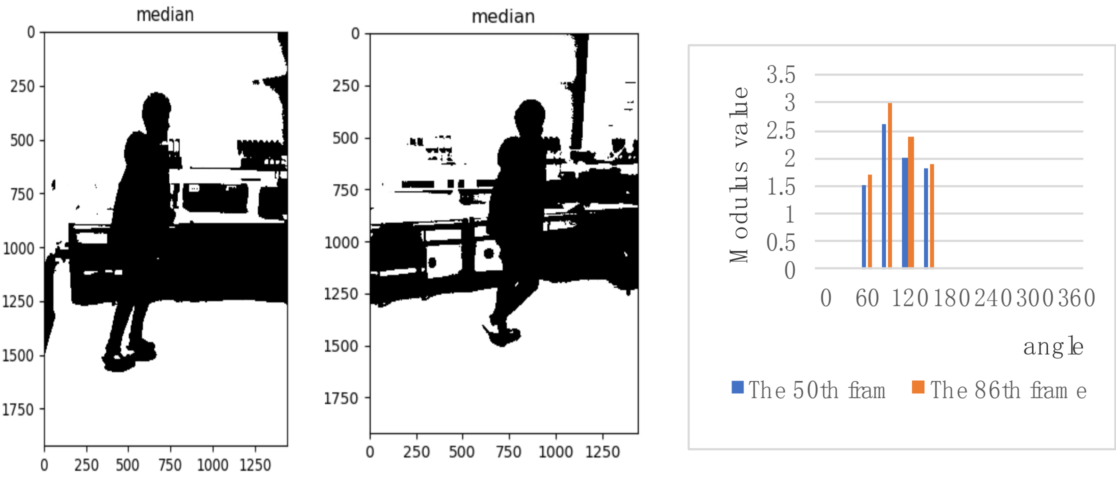
Different direction of motion conversion modulus data is different. Directional transformation modulus data graph can represent different motion behaviors, such as walking, running, trailing and fighting.



(a)The 20th frame (b) The 55th frame (c) Direction conversion vector data diagram

Fig. 3. Walking behavior

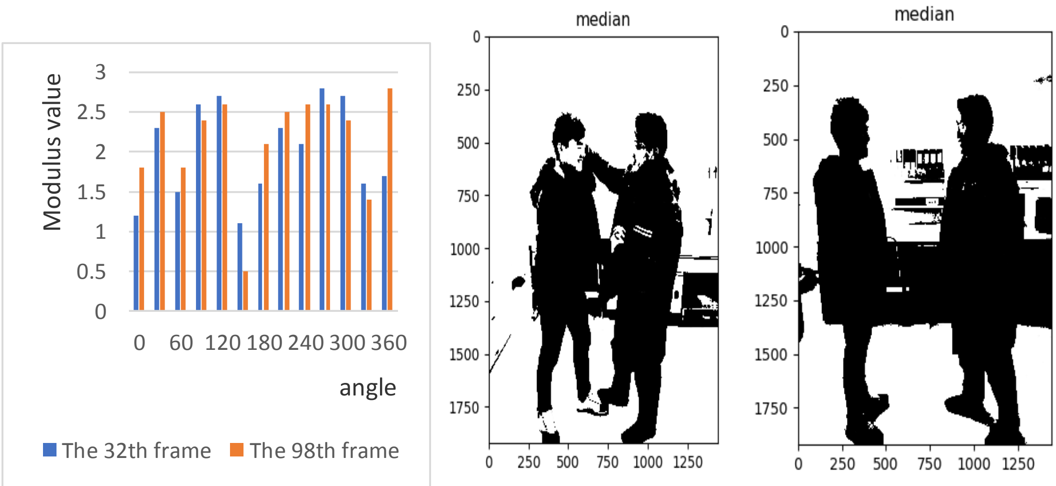
Fig. 3(a) is an effect diagram of the first time the target of the 30th frame is entered into the monitoring area, and (b) is an effect diagram of the normal walking of the target of the 68th frame in the monitoring range. (c) The direction conversion modulus data map corresponding to the above two figures is shown. At this time, the direction of the pedestrian's movement is relatively clear, and the main movement direction is concentrated at 60^0, 90^0, 120^0, 150^0.



(a)The 50th fram (b)The 86th frame (c) Direction conversion vector data diagram

Fig. 4. Running behavior

Fig. 4(a) is an effect diagram of the first acquisition of the pedestrian running of the camera in the 50th frame, (b) is an effect diagram of the above-mentioned target leaving the monitoring area at the 86th frame, and (c) corresponding to the above The direction of the two graphs is converted into a modulus data graph. During the motion, the direction of the target motion is basically the same. At this time, the molded values are larger than normal walking, which conforms to the motion law.



(a)The 32th fram (b)The 98th frame (c) Direction conversion vector data diagram

Fig. 5.Fighting behavior

Fig. 5(a) is an extraction effect diagram in which the distance between two moving targets decreases gradually in the 32th frame, and (b) is an effect diagram extracted when the two moving targets are struck in the 98th frame, (c) Corresponding to the direction conversion modulus data map of the above two figures, the moving direction of the moving target is relatively clear at the 32nd frame, the molding is large, and the directions of the two target objects are basically opposite. When the fighting behavior occurs, the movement of the target object There is no uniform direction and the modulus is large.

5.3 Behavioral feature extraction result

In this surveillance video, each camera has a fixed monitoring range of 24 hours, and sufficient light is guaranteed throughout the day. In the real world, there will be some interference, which will cause a certain degree of interference to the accuracy of the final recognition result. Set a threshold for the collection area, remove the interference item that is less than the threshold, and be a valid target when it is greater than the set threshold[6].

Running behavior mainly has the following characteristics:

The normal walking speed of a person is within the range of 120 steps/min. When a person's walking speed does not fall within this range, another non-walking behavior may be in progress. In physics, velocity is usually used to represent the motion state of an object at a moment, and acceleration represents the degree of motion change. At a certain time, the acceleration can directly reflect the trend of motion of the object. When the absolute value of the acceleration is larger, the speed of the target changes more. Compared with walking, when the speed of the human motion and the acceleration of the same time period are greater than the threshold set in advance, it is determined that the person is running.

Trailing behavior has the following characteristics:

(1) The number of moving targets in the surveillance video is at least two or more.

(2) The targets appearing in the area with the radius R as the center of the follower are marked accordingly.

(3) The relative distance between the marked moving target and the marked follower exhibits different changes at different times, but the magnitude of the value exceeds the threshold set above.

(4) Determine whether the trailing person and the trailing person have the same main moving direction of the modulus-weighted direction angle. In real life, the real-time distance and the relative distance between the trailing person and the trailing person will change in real time. Through these real-time changes, it is determined which behavior it belongs to.

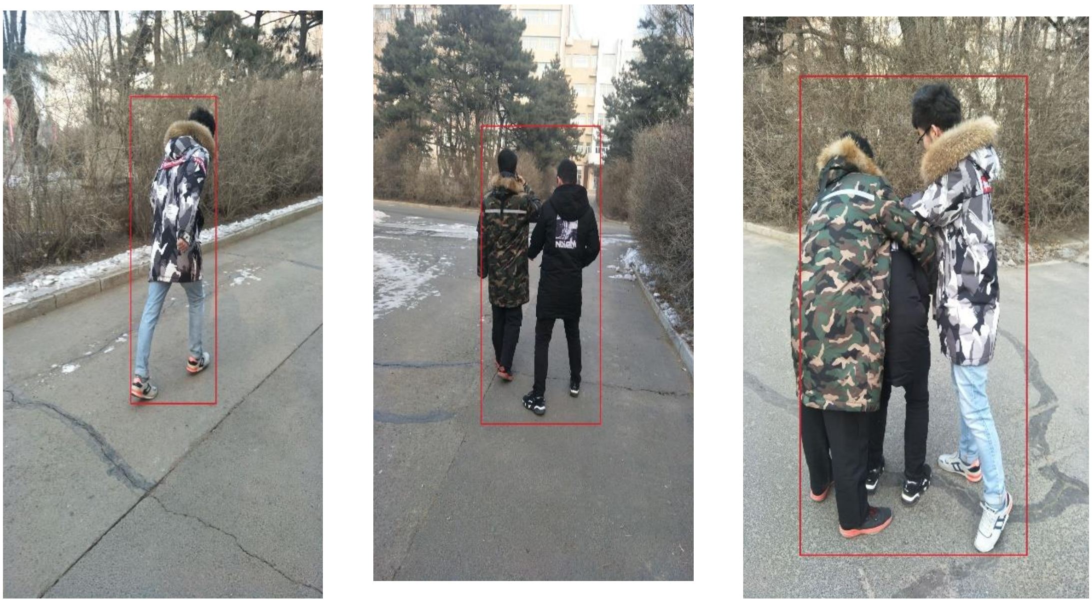
The fight behavior mainly has the following characteristics:

(1) The number of people who have detected the movement in the surveillance video is at least 2 or more.

(2) In this scene, the direction of the modulus-weighted direction angle of all moving targets changes at different moments, and there is no uniform direction in the direction in which the histogram is presented in the modulus direction.

(3) When the fighting behavior occurs, the moving direction of all moving targets is always in a changing state, and the magnitude of the moving direction is also in a state of constant change.

5.4 Detection algorithm verification



(a)running behavior (b) Trailing behavior (c) Fighting behavior

Fig. 6.Abnormal behavior

5.5 Abnormal behavioral model

Determining objects appearing in a video sequence by inputting a video sequence. if it is't the moving target, recorded in the monitoring video and end the process; If it is a moving target, calculate the amplitude-weighted direction angle of all targets and the Euclidean distance between the frames and the family, and adopt adaptive unsupervised clustering to obtain the preliminary key frame and calculate the cross entropy of the frame image. If the cross entropy is not the maximum value, the calculation of the amplitude weighted direction angle is returned. If the cross entropy is the maximum value, the extracted key frame image is calculated, and the entropy value of the direction conversion vector data graph of the moving target is calculated.

If $E^1_{Ht}<E^2_{Ht}$, the alarm has an abnormal behavior. Otherwise, the amplitude weighted direction angle is recalculated, and the cluster obtains a new key frame image. Whether the process completes the determination of the target behavior needs to be detected. If the determination is not completed, the process needs to be returned. Initialization, recalculate the amplitude weighted direction angle, otherwise the process ends.

6. CONCLUSION

Through the detection of moving objects, the pedestrians are classified by the support vector machine and other related classification methods. The KCF tracking algorithm realizes real-time tracking of pedestrians, extracting key frames, and intuitively digitizes different behavioral features by direction conversion vector data map. Through the establishment of the abnormal behavior model, real-time analysis and alarm processing of possible abnormal behaviors can be realized. The model accuracy rate can reach 98.75%.

REFERENCES

[1]Yan Wei. Application and Scene Analysis of Video Surveillance in Colleges and Universities [J]. China New Technology and New Products, 2018(17): 22-24.
[2]Ting Zhang. Detection and tracking of moving objects in surveillance video [J]. Digital Communication World, 2018 (04): 121.
[3] Wei Liu. Chengwei Duan , Bing Yu, et al. Multi-pose pedestrian detection based on posterior HOG features J. Chinese Journal of Electronics, 2015 4302): 217-224.
[4] Juanjuan Tian. Research on pedestrian detection and tracking methods in video [D]. Xi'an University of Technology, 2018.
[5] Yunyun Sun , Zhaohui Jiang , Guipeng Shan, Haiqiu Liu, Yuan Rao. Key frame extraction for optimal distance clustering and feature fusion expression [J/OL]. Journal of Nanjing University of Science and Technology, 2018(04): 416- 423
[6] Dandan He. Research on abnormal behavior detection under surveillance video scene [D]. Jiangnan University, 2018.