Pedestrian Motion Tracking and Crowd Abnormal Behavior Detection Based on Intelligent Video Surveillance

Fan Zhao
School of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou, China
Email: 1424126802@qq.com

Jin Li*
School of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou, China;
Contemporary Business and Trade Research Center of Zhejiang Gongshang University, Hangzhou, China
*Corresponding author: lijinkingdom@163.com

***Abstract*—Pedestrian tracking and detection of crowd abnormal activity under dynamic and complex background using Intelligent Video Surveillance (IVS) system are beneficial for security in public places. This paper presents a pedestrian tracking method combing Histogram of Oriented Gradients (HOG) detection and particle filter. This method regards the particle filter as the tracking framework, identifies the target area according to the result of HOG detection and modifies particle sampling constantly. Our method can track pedestrians in dynamic backgrounds more accurately compared with the traditional particle filter algorithms. Meanwhile, a method to detect crowd abnormal activity is also proposed based on a model of crowd features using Mixture of Gaussian (MOG). This method calculates features of crowd-interest points, then establishes the crowd features model using MOG, conducts self-adaptive updating and detects abnormal activity by matching the input feature with model distribution. Experiments show our algorithm can efficiently detect abnormal velocity and escape panic in crowds with a high detection rate and a relatively low false alarm rate.**

***Index Terms*—Pedestrian Tracking; Behavior Detection; Intelligent Video Surveillance; Abnormal Activity**

I. INTRODUCTION

Intelligent Video Surveillance (IVS) is a synthetic application of a variety of sciences and technologies related to computer vision. It employs image processing, pattern recognition, artificial intelligence and other technologies to process and analyze the video image sequences captured by monitoring system, intelligently understands video content and makes quick response. Generally, what intelligent video surveillance concerns and studies is the moving object in video. By means of detecting, identifying, tracking, comprehending and etc., features such as color, shape, contour and gradient are extracted. Attitude, velocity, trajectory and other movements are recorded to help researchers identify target class, comprehend target behavior and therefore make judgment, early warning, management and other related decisions correctly.

With the widespread applications of IVS technology in numerous fields of human society, it has become one of the core requirements concerning human detection and behavior analysis. After "9•11" attacks, every country in the world has a new insight of international terrorism and domestic security situation, and strengthens the monitoring management of densely crowd places. More and more video surveillance systems aiming at pedestrians emerge in airports, stations, ports, banks, squares and some other important places. Crowd control and pedestrian security problem in public places have been a great challenge. By 2008, there are approximately 2 million cameras used for city surveillance and alarm system in China, and by 2011, "National Urban Alarm and Monitoring System Construction Pilot Project (3111 pilot project)" has covered every prefecture-level city of all the provinces, municipalities and autonomous regions nationwide. As shown in "China Security Industry Twelfth Five-Year (2011-2015) Development Plan", the output value of China's video surveillance system will exceed 100 billion Yuan by 2015, accounting for more than 55% of security electronic products. Yet in these projects and programs, dynamic detection, dynamic early warning, intelligent analysis and processing of video image, biometrics identification and other major content are inseparable from the related researches of pedestrian motion detection and analysis.

Recent several years has witnessed enormous amount of efforts invested to the research of intelligent video surveillance area from academia to industry, and many practical results have already been achieved. In 1997, the U.S. Defense Advanced Research Projects Agency (DARPA) established Video Surveillance and Monitoring (VSAM) which was led by Carnegie Mellon University and jointly participated by Massachusetts Institute of Technology and many other higher education institutions and research institutions, what were mainly studied was the video understanding technology of real-time

doi:10.4304/jnw.9.10.2598-2605

automatic monitoring of military and civilian scenes . Haritaoglu et al. [1] developed a subsystem W^4 (Who are they? When do they act? Where do they act? What are they doing?) of VSAM, which used single camera and the captured grayscale image to locate and segment pedestrians in complex outdoor environment and implemented real-time tracking of several pedestrians. VSAM-based multi-sensor technology was proposed by Collins et al. [2] to develop the monitoring system applicable for campus. Some mature systems have also been widely used in terms of civilian research. Wren et al. [3] developed a Pfinder (person finder) system suitable for single non-occluded pedestrian and fixed camera case, allowing real-time detection and tracking of pedestrian under complex environment. Lipton et al. [4] studied a system using network connected to multiple cameras, which achieved detection and tracking of multiple pedestrians and vehicles in a large scale range and monitoring their activities over a relatively long period of time. In 2005, several European organizations jointly developed the ISCAPS (Integrated Surveillance of Crowded Areas for Public Security) projects, which primarily studied human auto-monitoring technology used for discovering potential security threat in dense region of crowd.

In recent years, researches that can be applied to the analysis of human behavior mainly focused on the detection, tracking, recognition of object and a higher level of behavior analysis. Continuous improvement in detection accuracy, robustness and rapidity has been regarded as major direction, therefore a number of new solutions and algorithms arise. For target tracking, the primary methods included those focusing on local object feature [5] [6], approaches of establishing tracking model [7-9] and those based on active contour [10] [11], the emphasis of these literatures involved multi-objective, camera motion and other complex issues. On the basis of traditional use of Kalman filter, the method adopting particle filter [12] is now in rapid development. In terms of pedestrian identification, the research program concentrated on using pattern recognition method for classification, and research content mainly consist of feature extraction and classifier construction. Methods based on motion characteristics [13] [14] and shape property [15] [16] were applied with regard to feature extraction, motion characteristics primarily referred to the specific rigidity and periodicity of human motion, while shape property can be comprised of region dispersion of image, aspect ratio, gradient and many other characteristics. As for the construction of classifier, methods include SVM and Boost are commonly adopted in research with the intent of shortening training time and improving classification rate during detection process.

Some algorithms for human behavior recognition, crowd behavior analysis and abnormal detection have also gained extensive attention over the past few years [17-22]. Assheton and Hunter [23] presented the mixture of uniform and Gaussian Hough Transform for shape-based object detection and tracking, proposed a variant of the generalized Hough transform. Liu, Chang and Guo [24] proposed a probability-based pedestrian mask pre-filtering to filter out non-pedestrian regions meanwhile retaining most of the real pedestrians.

Particle filtering method has proven to be useful in dealing with non-linear, non-Gaussian systems, therefore, it can be applied to the handling of complex dynamic scenarios during target tracking process. In recent years, tracking algorithm which employs particle filter method continues to develop, however, targeted processing under complex situations such as camera lens moving and pedestrian scale change is still difficult to achieve when dealing with target pedestrian in video, which may lead to tracking errors and even loss of target. In order to improve tracking result of video image-based pedestrian tracking, we adopt the structure of basic particle filtering method to propose a method of pedestrian motion tracking which integrates particle filter with HOG feature detection. HOG provides an explicit description of the shape of local object in the image depending on the statistics of histogram distributions in gradient orientation, which exhibits significant effect on the classification of target pedestrians.

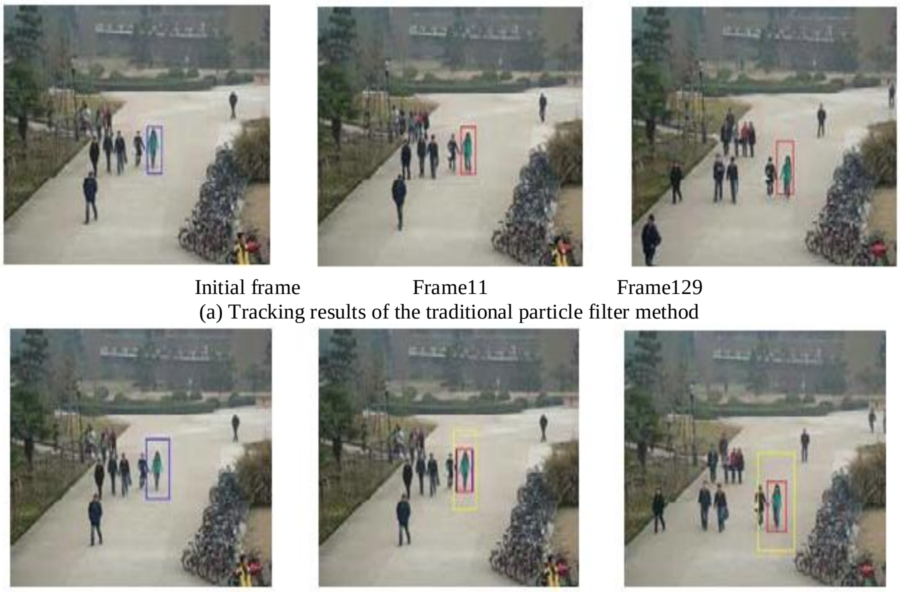
Humans are social animals and urban environment is the distribution center of human. Therefore, capturing crowd in usual city video surveillance systems is a quite common phenomenon. While crowd behavior, especially abnormal behavior usually contains some important information required in monitoring and early warning. As a consequence, detection of abnormal crowd behavior has become the object that intelligent video surveillance system focuses on and studies as well as a research hotspot in pedestrian motion analysis in recent years. Hence, we shall study the detection of abnormal activity in crowds, by extracting crowd feature, estimate motion parameters of crowd-interest points via block matching method, and then employ motion parameters analysis method to detect crowd gathering, dispersing, stranding, running and other group behaviors.

II. PEDESTRIAN TRACKING INTEGRATING HOG DETECTION WITH PARTICLE FILTER

During the process of pedestrian tracking, the proposed method firstly establishes a dynamic model which describes the target location and then adopts particle filter algorithm to conduct pedestrian tracking. After obtaining the region of target pedestrian, we use the method of HOG feature detection to implement classification of pedestrians in the specified region, and modify the position of the tracked object according to the result of pedestrian detection, thus improve the sampling of particle filter to reduce the tracking error.

A. Tracking Model

The position of target pedestrian is defined using a rectangular detection window, the state vector is described as $X = [x, y, w, h]$, where (x, y) is in the center of the detection window, w, h are width and height of window respectively. Based on the second-order autoregressive model, the system is described as follows:



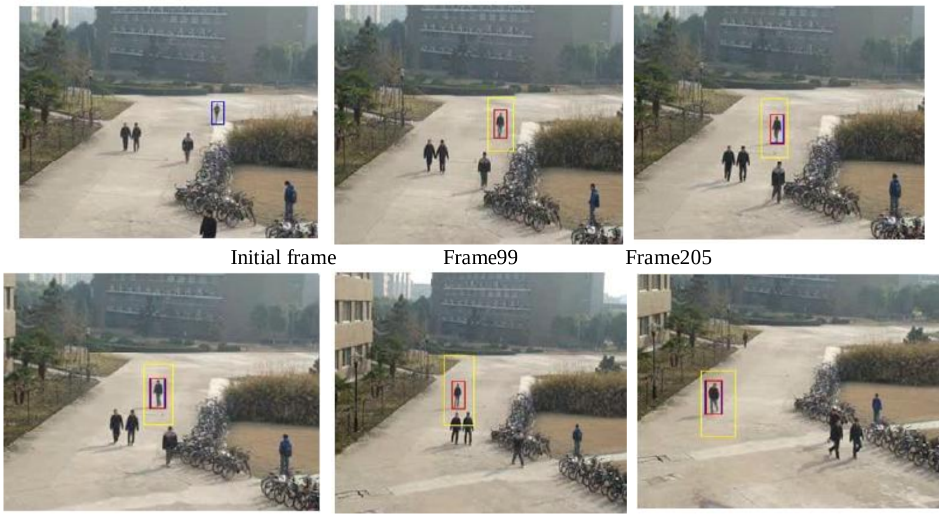
Initial frame Frame11 Frame129

(a) Tracking results of the traditional particle filter method

Initial frame Frame3 Frame117

(b) Tracking results of our algorithm

Figure 1. Comparisons of tracking results under dynamic background



Initial frame Frame99 Frame205

Frame260 Frame300 Frame491

Figure 2. Pedestrian tracking process under camera movement

$$X_k = A_1(X_{k-1} - X_0) + A_2(X_{k-2} - X_0) + \upsilon_k \quad (1)$$

where, υ_k is a zero-mean Gaussian random process vector whose variance vector is adjusted appropriately according to scenarios.

The scale and rotation invariance of the color characteristic of an object makes it suitable to be used in tracking process. Hence, our algorithm adopts color histogram distribution as the observation model and performs in HSV (hue, saturation, value) space. Taking 10 segments quantified on three components H, S, V respectively, the histogram is consequently divided into segments M=1000. Let p be the target model, $q(x, y)$ denotes the color histogram of the (x, y) centered detection region, $l(p, q(x, y))$ and $d(p, q(x, y))$ are similarity and distance between the two respectively. According to Bhattacharyya coefficient, we have:

$$l(p, q(x, y)) = \sum_{i=1}^{M} \sqrt{h_p(i), h_{q(x,y)}(i)} \quad (2)$$

$$d(p, q(x, y)) = \sqrt{1 - l(p, q(x, y))} \quad (3)$$

where $h_p(i)$, $h_{q(x,y)}(i)$ are their respective color histogram component.

When the position of object in time $k-1$ is obtained through observation and estimation, expand a certain range around from this position (the scale of the extended area can be preset according to experiences), HOG pedestrian detection can be conducted in this relatively larger region. The region of pedestrians is decided via HOG classifier, and its credibility is calculated based on the position information and color histogram. Select the region with the maximum credibility and exceeding the threshold η_l as a new template, re-enter particle filter initialization step and continue tracking process.

B. Algorithm Steps

The proposed tracking algorithm in this paper is integrated with HOG detection on the basis of general particle filter algorithm flow. The steps can be described as follows:

Step 1. Initialization. Establish the model when $k = 0$, extract color histogram from the targeted template, then extract N initial particles $\{X_0^i, 1/N, i = 1, 2, \cdots, N\}$ from prior distribution;

Step2. Particle prediction. Based on the particle swarm $\{X_{k-1}^i, 1/N, i = 1, 2, \cdots, N\}$ in time $k-1$, according to the second-order autoregressive dynamic model, we can obtain particle swarm $\{X_k^i, \omega_k^i, i = 1, 2, \cdots, N\}$ in time k;

Step3. Particle observation. Use color histogram to calculate observation likelihood for each particle, thus we have $\{L_k^i, i = 1, 2, \cdots, N\}$;

Step4. Update the particle weight. After having the observation likelihood normalized, update the particle weight to be $\omega_k^i = L_k^i / \sum_{j=1}^{N} L_k^j, i = 1, 2, \cdots, N$;

Step5. HOG pedestrian detection. Calculate the probability after estimation, the obtained position is $X_k^i = \sum_{i=1}^{N} \omega_k^i X_k^i$, expand the region X_k^m by β (default value) times, and then conduct HOG pedestrian detection in this expanded region. If there is no pedestrian target in classification result, continue to the next step; If there are P regions belonging to pedestrians, calculate their credibility $\{\eta_m, m = 1, 2, \cdots, P\}$, select the region with the maximum credibility exceeding the threshold η_l as its new template, then return to the initialization step;

Step6. Resample. To solve particle lacking problem and retain particles with large weight, exclude particles whose weight is less than threshold ω_l and fill the new particle swarm $\{X_{k-1}^i, 1/N, i = 1, 2, \cdots, N\}$, then go to the particle prediction step.

C. Experimentmal Results for Tracking Algorithm

In order to verify the tracking effect of our algorithm, we carry out pedestrian tracking experiment in video sequences. The configuration of the running computer is Pentium Dual Core (2.20GHz CPU) and 2G Mb ROM. Experimental parameters are set: dynamic model parameters $A_1 = 2$, $A_2 = -1$, credibility threshold η_l = 0.5, particle number $N = 100$.

We performed experiment on the videos with dynamic background existing dynamic panning and zooming changes, the tracked object is among a plurality of pedestrians and varies in size attributed to the difference in distance in video sequence. Compared with general particle filter algorithm the experimental results are shown in Figure 1. We can see from the comparison: our algorithm outperforms the general particle filter tracking algorithm in tracking results with less tracking error in dynamic video, it offers a better adaptability to background and target scale changes with the integration of HOG detection which can modify particle sampling during the tracking process

We further carry out tracking experiment for pedestrians in more complex and versatile video image sequences, the results are shown in Figure 2. During the process of video capture, there exists simple movement of lens stretch and directional movement in camera and significant changes in the size of tracked pedestrians. Sequential tracking of pedestrians in this video cannot be achieved using common data-driven algorithm and traditional particle filter methods; however, our algorithm implements comparatively complete tracking which can handle these complex situations more effectively.

III. CROWD ABNORMAL DETECTION

The study of current crowd behavior analysis can be divided into two categories according to the difference of the target concerned: individual-based analysis and entirety-based analysis. Individual-based analysis aims at single pedestrian target, which recognizes individual behavior pattern through analysis and furthermore analyzes abnormal behavior after obtaining crowd motion information. However, in cases of comparatively dense crowd, segmentation and tracking of individuals will become difficult owing to the severe occlusion in pedestrians. Hence, the entirety-based analysis which extracts crowd characteristic parameters to detect abnormal activity has attracted considerable attention. In this paper, we present an algorithm i.e. analysis and modeling of the feature parameter of crowd interest point. First, extract POI (Points of Interest) in crowd within the monitored region, then analyze the statistical eigenvalues such as number of POI, density, velocity and direction of movement, extract these features to build the Gaussian mixture model of crowd eigenvalues to describe crowd behavior and perform updating and maintenance. As to the detection of crowd behavior, we match the POI feature extracted from the input image each time with the Gaussian mixture model established after a period of training, abnormal event is considered to have been detected if there is a mismatch.

A. Gaussian Mixture Background Model

Gaussian mixture background model can well describe the feature distribution of pixel when operating background modeling and object segmentation on video image sequence. In this section, we apply this model to the establishment of the description model of group feature in crowd video image sequence. Set X to be the color value of certain pixel, then it can be approximated by the weighted sum of several Gaussian distributions, suppose that a serials of historical data set of X in time t is expressed as $\{X_1.X_2, \cdots, X_t\}$, thus the probability of current X_t is defined as follows:

$$P(X_t)=\sum_{i=1}^{K}\omega_{i,t}\eta(X_t,\mu_{i,t},\Sigma_{i,t}) \quad (4)$$

where K denotes the number of normal distributions, $\omega_{i,t},\mu_{i,t},\Sigma_{i,t}$ are weight, mean and covariance of the i th Gaussian distribution in time t, respectively. η describes the probability density function of the i th Gaussian distribution, it is given by:

$$\eta(X_t,\mu,\Sigma)=\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}e^{-\frac{1}{2}(X_t-\mu_t)^T\Sigma^{-1}(X_t-\mu_t)} \quad (5)$$

Use online *K*-means approximation method to estimate the parameters of X distribution, the weight is updated as follows:

$$\omega_{k,t}=(1-\alpha)\omega_{k,t-1}+\alpha M_{k,t} \quad (6)$$

where α is learning rate. If current X is located in a tripling standard deviation range of certain normal distribution, we consider that a match occurs, otherwise it is deemed as unmatched. The mean and variance remain unchanged for each unmatched normal distribution, while for those matched they are updated as below:

$$\mu_t=(1-\rho)\mu_{t-1}+\rho X_t \quad (7)$$

$$\sigma_t^2=(1-\rho)\sigma_{t-1}^2+\rho(X_t-\mu_t)^T(X_t-\mu_t) \quad (8)$$

$$\rho=\alpha\eta(X_t\mid\mu_k,\sigma_k) \quad (9)$$

After obtaining the parameters of each current normal distribution, we determine which normal distribution the current variable value belongs to. The self-adaptive background model is ultimately comprised by the mixture of the first B distributions which possess relatively large weight and small volatility, it can be obtained as follows:

$$B=\arg\min_b\left(\sum_{k=1}^{b}\omega_k>T\right) \quad (10)$$

where T is a threshold. Then, we determine whether the current pixel values match the mixture of normal distribution of background or not, those unmatched are regarded as foreground object.

B. Abnormal Detection Process

We characterize the distribution condition of crowd feature via Gaussian mixture model, detect features which are unmatched with the established group feature model by the way of background segmentation, thus detect anomalous event occurred in crowd. In the following section, we will take the analysis of velocity feature anomaly of some ROI (Region of Interest) as an example to discuss the establishment and update method of feature model. Let V be the crowd velocity feature of certain region in the image, describe the probability distribution of V in certain cell region of video image using the mixture of K normal distribution, if V is characterized as the mixture of three kind of movement patterns: high, medium and low, the value of K is taken as 3. Take the following steps to establish and update the model of crowd velocity parameter V:

Step 1. Initialization

Assume that velocity feature is uniformly distributed in [0,1], divide [0,1] into three segments to describe the range of three kinds of movement patterns: low, medium and high. Now we have K =3, thus the expectation μ_k of three patterns falls on the center position of these 3 segments, the variance σ_k of pattern is initialized to 0.066 and weight ω_k is initialized as 0.333.

Step 2. Matching of feature distribution

For a frame of newly input image, use the feature parameters V in ROI to match the 3 distributions in Gaussian mixture model sequentially. If it satisfies $V_{ROI}\in[\mu-2.5\sigma,\mu+2.5\sigma]$ with regard to one of these distributions, then we consider that the input feature is matched with the current distribution, vice versa.

Step 3. Model update

Update the parameter of matched distribution according to the following equations:

$$\mu_{k,t+1}=(1-\rho)\mu_{k,t}+\rho V \quad (11)$$

$$\sigma_{k,t+1}^2=(1-\rho)\sigma_{k,t}^2+\rho(V-\mu_{k,t+1})^T(V-\mu_{k,t+1}) \quad (12)$$

where $\mu_{k,t}$ and $\sigma_{k,t}$ denote mean and variance of the k th distribution, respectively. The learning rate of this model is $\rho=\alpha/\omega_{k,t}$, α is the weight learning rate, and $\omega_{k,t}$ is described as the weight, indicating the probability of the occurrence of the k th distribution. The weight can be updated below:

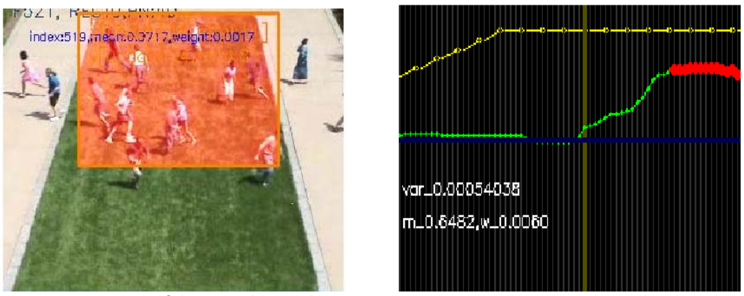
$$\omega_{k,t}=(1-\alpha)\omega_{k,t}+\alpha M_{k,t+1} \quad (13)$$

where $M_{k,t+1}$ takes the value of 1 if a match occurs, otherwise it is 0.

Step 4. Abnormal detection

Sort the k distributions in descending order according to ω/σ. With the proceeding of model training process, distributions with frequent occurrences and relatively small variance will be placed to the front of the queue after several times of sorting. We consider the eigenvalues which frequently occur as a representation of normal event, analogically we can think of it as the background in the image; while anomalous event, which corresponds to the part that does not match the model, is regarded as the foreground analogically. Hence, a method of background segmentation of Gaussian mixture model can be employed to detect anomalous event, that is, if the currently input feature is matched with one of the first B distributions, we consider it is normal condition in current ROI, otherwise, an exception is considered to have been detected. Here, B is defined as follows:

$$B=\arg\min_b\left(\sum_{k-1}^{b}\omega_{k,t}>T_B\right) \quad (14)$$



(a) An instant of dispersed crowd running (b) Detection Schematic

Figure 3. Testing result of sudden dispersed running crowd

TABLE I. COMPARISONS OF THE DETECTION RESULTS FOR THE THREE ABNORMAL DETECTION ALGORITHMS

	Detection rate	False alarm rate
MDT algorithm	75%	25%
SF algorithm	80%	19%
The proposed Algorithm	81%	2%

where T_B is a preset threshold, the bigger T_B is, the smaller the chances that detected abnormal condition occurs in history are.

Similarly, a variety of anomalous events occur in crowd scenarios can be detected. The modeling analysis of POI features is suggested in order to analyze anomalous event in ROI. First, select the POI of the characterized crowd in ROI, next extract quantity, velocity, direction and other statistical characteristics of POI, and then utilize the Gaussian mixture model which established multiple feature of the crowd. When operating background segmentation through model training and updating, the input feature which matches features such as quantity, velocity and movement direction of POI belongs to the background distribution; and the unmatched input feature with small occurring probability thus belongs to the foreground, indicating the occurrence of abnormal condition. Abnormal motion of crowd can be detected after the segmentation of these foregrounds.

C. Experimental Results and Analysis

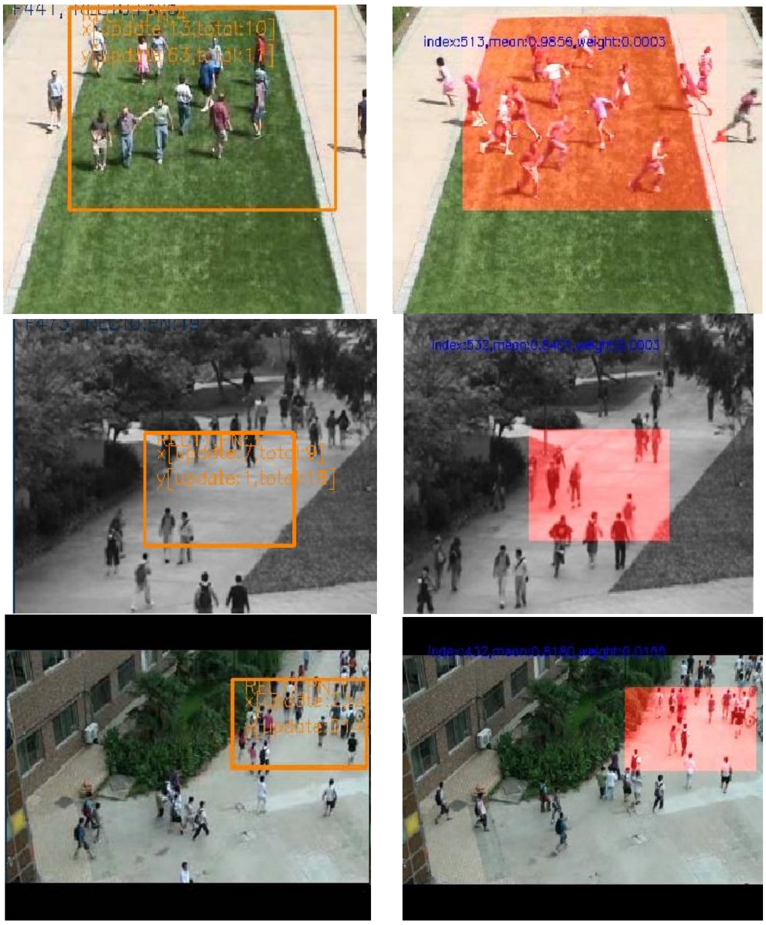
Experiment is conducted on our algorithm using VC and OpenCV2.0 platform and algorithm performance is tested as well. The configuration of the running computer is Pentium Dual Core (2.20GHz CPU) and 2G Mb memories. The video data set used in tests consists of two parts, namely, UCSD Library (Unusual crowd activity dataset of University of Minnesota; UCSD anomaly detection dataset) and video segments collected in campus road circumstances during and after class time period.

In the experiment, we employ our algorithm to conduct anomaly detection to each frame in the video sequence of crowd motion and figure out the number of frames where an anomaly is detected, when the number of abnormal frames in a video sequence exceeds the preset threshold, anomalous event in crowd is considered to have occurred in this video segment. Figure 3 shows the condition of algorithm detection when there exists crowd anomaly in video, the crowd is walking randomly at first, and soon afterwards fleeing phenomenon occurs suddenly. As is shown in chart (a), we intercept a frame of detection result image in the process of dispersed crowd running, the rectangular box is the interest detection region set by users and the inside area of rectangle is lit up (covered with a red mask) attributed to the detection of anomalies; chart (b) describes the state curve plotted through calculation during the detection process of algorithm, where the curve below shows the status in testing process, the normal state point is indicated by a small dot while the abnormal state point is represented by a large dot, therefore, the latter part which is thicker refers to a series of abnormal state points having been detected, it means abnormal events occur when the number of points exceeds the preset threshold.

We have examined several videos which involve anomaly event in crowd. As is shown in Figure 4, images in column (a) describe normal crowd motion before the occurrence of abnormal condition; column (b) corresponds to a frame of abnormal event occurring subsequently in video. The sequence of testing video is from the top down and abnormal conditions include sudden crowd dispersal, swift passing of rider and unexpected curved running. The rectangles in figures refer to the interest detection region that users select to set, when there is an anomaly with crowd detection parameters in detection box, the algorithm will light up the inside area of detection box (filled with a red mask).

We record the results of abnormal detection in these three testing videos and compare the results of our algorithm with SF algorithm [25] and MDT algorithm [26], the results are shown in table 1.

In the accuracy aspect of detection results, our abnormal detection model algorithm reaches 81% in average detection rate with the average fault alarm rate as low as 2%, that is, we achieve relatively high detection rate while keeping the fault alarm rate comparatively low. And in view of the execution speed of algorithm, the proposed method can basically detect 5 frames of images with a scale of 320×240 per second. However, MDT algorithm is low in execution speed with the time required to detect an even smaller image up to 20 seconds.



(a) Normal movement crowd (b) Crowd abnormal detection

Figure 4. Results of crowd abnormal detection

IV. CONCLUSION

In this paper, we study the pedestrian tracking and detection of abnormal crowd behaviors. For the pedestrian tracking, we adopt the model-driven idea, considering the unsuitability of general tracking algorithm in cases of pedestrian target scale change and camera motion, particle filter framework is adopted to establish the observation model via color histogram and implement sampling correction using HOG detection. Our algorithm takes advantage of the insensitivity of color histogram feature to target scale change and partial occlusion and utilizes the substantially high detection accuracy rate of pedestrian under dynamic background in HOG algorithm. The experimental results show that our algorithm achieves pedestrian target tracking under the video condition of dynamic background and camera motion with less tracking error compared to conventional particle filter algorithm, moreover, it exhibits significant tracking effect on complex videos with camera motion and pedestrian scale change which cannot be successfully tracked using traditional method.

For the detection of abnormal events in crowd, we present the following method in this paper: analyze the parameters of feature point from an overall perspective, establish crowd feature Gaussian mixture model and perform self-adaptive updating, and detect abnormal event in crowds by the matching operation between the input feature and model background distribution. Segmentation and tracking of individual object is dispensable in our algorithm and the training of Gaussian mixture model is relatively simple and swift, thus enabling quick and efficient feature extraction and anomaly detection of crowd. As experimental results demonstrate, abnormal phenomena in crowd such as fleeing and speed jump can be detected in a relatively low false alarm rate by the proposed method, moreover, our algorithm outperforms SF and MDT algorithm with higher detection accuracy and faster velocity.

ACKNOWLEDGMENT

The authors wish to thank the reviewers for their valuable comments. This work was supported in part by the Zhejiang Provincial Commonweal Technology Applied Research Projects of China (Grant No. 2013C33030), Humanities and Social Sciences Foundation of Ministry of Education of China (Grant No. 12YJC630091), Zhejiang Provincial Natural Science Foundation of China (Grant No. LQ12G02007, Z14G02 0001, Y14G020006), the National Natural Science Foundation of China (Grant No.71302035), and the Fourth Excellent Yong Talents Foundation of Zhejiang Gongshang University (Grant No. QY13-23).

REFERENCES

[1] I. Haritaoglu, D. Harwood, L. Davis L., "W4: real-time surveillance of people and their activities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, 2000.

[2] R. Collins, A. Lipton, H. Fujiyoshi, et al., "Algorithms for cooperative multisensor surveillance," *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1456-1477, 2001.

[3] C. Wren, A. Azarbayejani, T. Darrell, et al., "Pfinder: real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780-785, 1997.

[4] A. Lipton, H. Fujiyoshi, R. Patil, "Moving target classification and tracking from real-time video," *Proceedings of IEEE Workshop on Applications of Computer Vision*. Princeton, New Jersey, USA: IEEE, pp. 8-14, 1998.

[5] T. Kaneko, O. Hori, "Feature selection for reliable tracking using template matching," *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Madison, Wisconsin, USA: IEEE Press, vol. 1, pp. 796-802, 2003.

[6] P. Tissainayagam, D. Suter, "Object tracking in image sequences using point feature," *Pattern Recognition*, vol. 38, vol. 1, pp. 105-113, 2005.

[7] I. Karaulova, P. Hall, A. Marshall, "A hierarchical model of dynamics for tracking people with a single video camera," *Proceedings of the British Machine Vision Conference*. Bristol, UK, pp. 352-361, 2000.

[8] R. Mazzon, A. Cavallaro, "Multi-camera tracking using a Multi-Goal Social Force Model," *Neurocomputing*, vol. 100, no. 6, pp. 41-50, 2013.

[9] R. Urtasun, D. J. Fleet, P. Fua, "Temporal motion models for monocular and multiview 3D human body tracking," *Computer Vision and Image Understanding*, vol. 104, no. 2–3, pp. 157-177, 2006.

[10] N. Paragios, R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 266-280, 2000.

[11] O. Jokinen, "Tracking of facial deformations in multi-image sequences with elimination of rigid motion of the head," *SPRS Journal of Photogrammetry and Remote Sensing*, vol. 84, pp. 52-68, 2013.

[12] M. Isard, A. Blake, "Contour tracking by stochastic propagation of conditional density," *Proceedings of the European Conference on Computer Vision*. Cambridge, England, pp. 343–356, 1996.

[13] R. Cutler, L. "Davis, Robust real-time periodic motion detection, analysis, and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 781-796, 2000.

[14] G. Goudelis, K. Karpouzis, S. Kollias, "Exploring trace transform for robust human action recognition," *Pattern Recognition*, vol. 46, no. 12, pp. 3238-3248, 2013.

[15] S. F. Zhang, N. J. Li, X. Cheng, Z. Y. Wu. "Adaptive object detection by implicit sub-class sharing features," *Signal Processing*, vol. 93, no. 6, pp. 1458-1470, 2013.

[16] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," *Proceedings of the Conference on Computer Vision and Pattern Recognition*. San Diego, USA, vol. 1, pp. 886-893, 2005.

[17] T. Moeslund, A. Hilton, V. Krüger. "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, pp. 90-126, 2006.

[18] R. Poppe, "Vision-based human motion analysis: an overview," *Computer Vision and Image Understanding*, vol. 108, no. 1-2, pp. 4-18, 2007.

[19] A. Cheriyadat, R. Radke, "Detecting dominant motions in dense crowds," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 4, pp. 568–581, 2008.

[20] A. Albiol, M. Silla, A. Albiol, et al. "Video analysis using corner motion statistics," *Proceedings of the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, pp. 31-38, 2009.

[21] D. Conte, P. Foggia, G. Percannella, et al., "A method for counting moving people in video surveillance videos," *EURASIP Journal on Advances in Signal Processing: Special issue on video analysis for human behavior understanding*, pp. 1-10, 2010.

[22] X. M. Hu, H. Zheng, W. W. Wang, X. Li, "A novel approach for crowd video monitoring of subway platforms," *Optik-International Journal for Light and Electron Optics*, vol. 124, no. 22, pp. 5301-5306, 2013.

[23] P. Assheton, A. Hunter. "A shape-based voting algorithm for pedestrian detection and tracking," *Pattern Recognition*, vol. 44, no. 5, pp. 1106-1120, 2011.

[24] Y. F. Liu, J. M. Guo, C. H. Chang. "Low resolution pedestrian detection using light robust features and hierarchical system," *Pattern Recognition*, vol. 47, no. 4, pp. 1616-1625, 2014.

[25] R. Mehran, A. Oyama, M. Shah. "Abnormal crowd behavior detection using social force model," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 935-942, 2009.

[26] V. Mahadevan V, W. Li, V. Bhalodia, et al. "Anomaly detection in crowded scenes," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1975-1981, 2010.