International Journal of Smart Home
Vol. 9, No. 12, (2015), pp. 149-160
http://dx.doi.org/10.14257/ijsh.2015.9.12.16

Global Anomaly Crowd Behavior Detection Using Crowd Behavior Feature Vector

Yong Yin*, Qiannan Liu and Shibiao Mao

College of Communication Engineering, Chongqing University, Chongqing 400044, China
yongyin@cqu.edu.cn

Abstract

In the area of crowd abnormal detection, the parameter of population density, is seldom used to the global crowd behavior detection. Some of the references simply use the LBP or spatial-temporal LBP features to fulfill the abnormal detection. They don't make full use of the crowd density characteristics and dynamic characteristics. This paper proposes a novel method by increasing the dimension of feature vector to increase the information content so as to improve the recognition accuracy. That is to say the crowd dynamic information and crowd density information will be combined together to form a higher dimension of feature vectors, which is named as the crowd behavior feature vector in this paper to improve the robustness of the algorithm. Finally, Support Vector Machines (SVM) is adopted to detect the abnormal events using the crowd behavior feature vector. This work utilizes the Local Binary Pattern Co-Occurrence Matrix (LBPCM) for crowd density estimation to ensure the excellent accuracy. At the same time, it adopts high accuracy optical flow histograms of the orientation of interaction force to extract the crowd dynamic information (HOIF). After verification, we discovered this algorithm not only can get the good discrimination on the benchmark dataset UMN, but also can achieve the pretty high recognition rate about the web dataset.

Keywords: *anomaly crowd behavior, crowd behavior feature vectors, local binary pattern co-occurrence matrix, histograms of the orientation of interaction force*

1. Introduction

Crowd abnormal detection is an important issue in intelligent video surveillance, which has become an important research approach of computer vision in recent years. Anomaly detection, also named as outlier detection, dictionary defines abnormal as: deviating from the ordinary type, especially in a way that is undesirable or prejudicial; contrary to the normal rule or system: unusual, irregular, aberrant. Based on the research in this field, previous work in abnormal video event detection, can be categorized into two classes, which is shown in Figure 1. They are Local Abnormal Event (LAE) and Global Abnormal Event (GAE). (i) LAE: The behavior of an individual is different from its neighbors. As shown in Fig. 1(a), the motion pattern of the red one is different from its neighbors, which is a spatial abnormal event. (ii) GAE: The group behavior of the global scene is abnormal. Figure 1(b) shows an GAE scene, where the pedestrians suddenly scattered due to an abnormal event, e.g. , an explosion [1].

In this paper, we focus on the detection of the global abnormal events in crowded scenes. For a human observer, it is very difficult to observe each pedestrian's behavior in extremely crowded scenes because conventional methods designed for surveillance applications fail drastically for the following reasons: (1) overlapping between individual subjects; (2) random variations in the density of people over time; (3) low resolution

*Corresponding author.
Email address: yongyin@cqu.edu.cn (Yong Yin).

ISSN: 1975-4094 IJSH

videos with temporal variations of the scene background [2].The understanding of human behaviors in crowded scenes is a far-reached problem which has not been fully resolved.

The main objectives of crowd behavior analysis are of paramount importance since early detection, or even prediction, which may reduce the possible dangerous consequences of a threatening event, and can help alerting a human operator for the sake of inspecting the ongoing situation more carefully.

(a) (b)

Figure 1. (a) Local Abnormal Event (LAE), (b) Global Abnormal Event (GAE)

Throughout the literature, anomaly detection has different classification methods, such as, explicit detection and deviation methods. Saira Saleem Pathan firstly measures the concerning uncertainty of underlying field with social entropy, and uses SVM to detect abnormal directly [3]. Tian Wang and Hua Yang utilize histograms to describe crowd dynamic, and adopt SVM to classify abnormal events [4,5]. Some other methods take energy term, HMM, LDA,GMM, SRC etc., to measure the abnormal events [1,6-9].

From another point of view, the works on this area are reviewed in [9,10]. The research routes of group behaviors can mainly be divided into three types. The first one is the traditional object-based approach which considers the group as a collection of individuals [11]. This approach also ignores the correlation among pedestrians. The second approach focuses on analyzing the entire video frame or extracting specific subject information [13]. The general object can be an image patch, a spatial-temporal block, mixture dynamic texture etc. A classic approach is to use optical flow method to characterize the motion features [2,4,7]. Unfortunately, perfect optical flow cannot be achieved in the extremely crowd scenes or the ones with severe light changes. In the third kind of approach, the above frameworks are combined in research works. They not only analyze the entire video, but also track the pedestrians. This method has a big computation load [12].

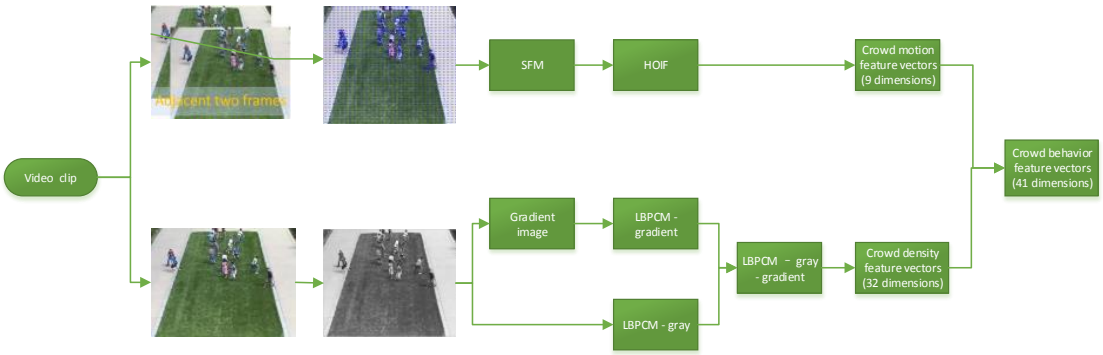


Figure 2. Framework of Getting the Crowd Behavior Feature Vectors

In the most of the articles, firstly a grid of particles is placed over the image by treating the video frame as a whole. In local pressure model [5,7], adopting interaction force as crowd motion dynamic, spatial-temporal local binary pattern can be used to get the local density distribution. However, the optical flow computation and density model in the approach limits the application to more complex scenes. In high-frequency and spatio-temporal features mode [12], it uses high-frequency information after wavelet transformation to characterize the changes in the time direction. LDA is utilized for the global abnormal crowd detection [14]. And in local scene, it builds HMMs to describe the normal behavior types in monitoring area. The probability of this obtained observation is selected as the evaluation standard.

When a video clips input, this paper firstly get the optical flow of each adjacent two frames. After a series of processing, this work gets a 9 dimensions motion feature vector for each particle, which will explains explicitly in section 2.2. Then we conduct gray and gradient co-occurrence matrix processing for each of the frame of the video, and use the parameters of energy, contrast, homogeneity and entropy to fully depict the crowd density feature using the 32 dimensions feature vector, which is described in section 2.3 in details. By combining the 9 dimensions motion feature vector and the 32 dimensions crowd density feature vector this work gets the 41 dimensions feature vector which is name as crowd behavior feature vector in this paper. Finally this work uses the feature vector to form the SVM to train and detect abnormal events [20]. Figure 2. shows the overall pipeline of getting the crowd behavior feature vector.

2. Global Abnormal Crowd Behavior Detection

Because of the complexity of group events, the crowd occlusion, illumination and noise, the crowd features is difficult to be extracted. In order to solve the problem of inaccuracy crowd motion direction and speed caused by the crowd occlusion and illumination, the high accurate optical flow algorithm is adopted which is used to divide the image frames into grid of particles using particle dynamics on behalf of population dynamics. This work introduces the Histogram of Interaction Force (HOIF) which will described in detail in section 2.2. This work uses the social force model (SFM) to extract the interaction force, meanwhile the velocity and directions of particles under complex environment can be extracted effectively. Then, the tender direction HOIF is adopted to improve the accuracy of the crowd motion feature description. It reduces the computational complexity and increases the robustness.

2.1. Crowd Motion Detection

In order to get the high accuracy optical flow, we adopt literature [15] method to generate high accuracy optical flow based on a theory for warping. Using the social force model we can get the interaction force. This model describes the pedestrian motion by combining the personal motivations and environment constrains with the function model as follows which is based on the theories of physics and social psychology.

$$\mathrm{F} = m_i \frac{dv_i}{dt} = F_p + F_{int} = \frac{1}{\tau}\left(v_i^p - v_i\right) + F_{int} \tag{1}$$

In this equation we can get the actual force F on pedestrian i with the mass of m_i. F consists of two parts: (1) F_p, the personal desire force of pedestrian i; (2) F_{int}, the interaction force of the pedestrian i.

(a) (b) (c)

Figure 3. Shows the High Accuracy Optical Flow from the UMN Dataset [16] of Three Scenes.(a) Scene 1 (b) Scene 2 (c) Scene 3

We can get F_p by the expression: $\frac{1}{\tau}(v_i^p - v_i)$, where τ is relaxation parameter. v_i denotes the actual velocity. Considering the effect of panic, v_i^p (denotes the desired velocity) can be replaced by the following equation:

$$v_i^q = (1 - p_i)v_i^p + p_i v_i^a \tag{2}$$

where p_i (0~1) is the panic weight parameter, the higher the herding effect is, the lager the p_i is and the higher the local density is, v_i^a is the average velocity of the neighboring pedestrians. Without loss of generality, we set $m_i = 1$.

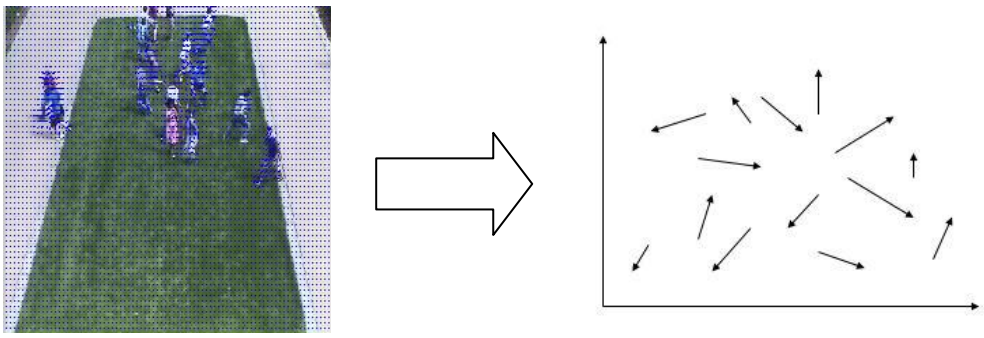
Thus, interaction force can be estimated by the following equation:

$$F_{int} = \frac{dv_i}{dt} - \frac{1}{\tau}(v_i^p - v_i) \tag{3}$$

This paper adopts optical flow to estimate the velocity and v_i^a is the average of optical flow in spatial volume around the particle. In order to get high accuracy optical flow, this work adopts the algorithm in [15] to generate optical flow based on a theory for warping. The particle advection method used in this paper is similar to [7, 13] placing a grid of particles over the image and move them with the underlying flow field. Then we get the certain and excellent interaction force. Figure 3. shows the high accuracy optical flow obtained from the three scenes. Blue arrows represent the motion of optical flow.

2.2. Histograms of the Orientation of Interaction Force

Histograms of the orientation of interaction force HOIF is utilized to handle the interaction force[21]. The interaction forces are two dimensional vectors received from the optical particle for each frame.



(a) High accuracy optical flow of one frame of the video

(b) The corresponding social force

Figure 4. Interaction Force of One Frame of Scene 1

The original intention is to segment the video into normal and abnormal parts according to the distribution of the direction of the interactive force of the frame in the video. The corresponding frame will be judged to be abnormal if the distribution is uniform and normal if concentrating. The angles of these vectors can be any value from 0 degrees to 360 degrees, as shown in Figure 4.

If we treat every angle value as one dimension of the motion feature vector, motion vector will be formed by dozens or hundreds of dimensions. This kind of vector will be the parts of input to the classifier to be trained and tested, which will form a large amount of computation.

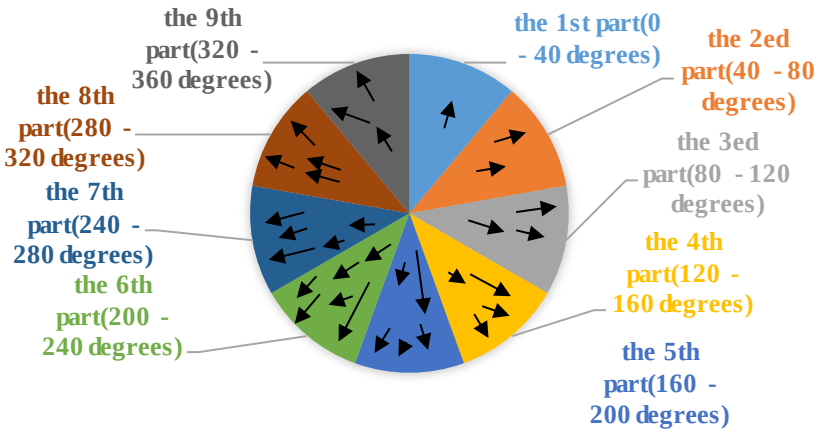


Figure 5. To Evenly Classify the Space of the Interactive Force into 9 Parts

However, we just want to know whether the distribution of the angular of the interaction force is uniform or not. The large computation is in no need. So we propose to classify these angular to reduce the calculation amount. The work evenly classifies the space of the interactive force into 9 parts (from the 1st part to the 9th part as follows:[0 degree - 40 degrees), [40 degrees - 80 degrees), [80 degrees - 120 degrees), [120 degrees - 160 degrees), [160 degrees - 200 degrees), [200 degrees - 240 degrees), [240 degrees - 280 degrees), [280 degrees - 320 degrees), [320 degrees - 360 degrees).) then calculates the sums of the amplitude of the interactive force of the angular falling into the same parts separately as shown in Figure 5.

The values of the nine sums consist the nine dimensions of the motion feature vectors

which is equal to the values of the nine bins of the Histograms of the Orientation of Interaction Force for each two adjacent frames.

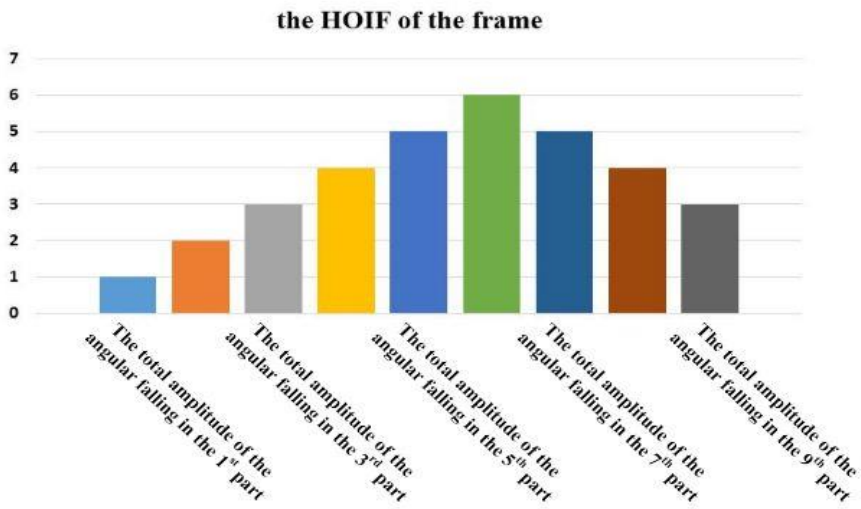
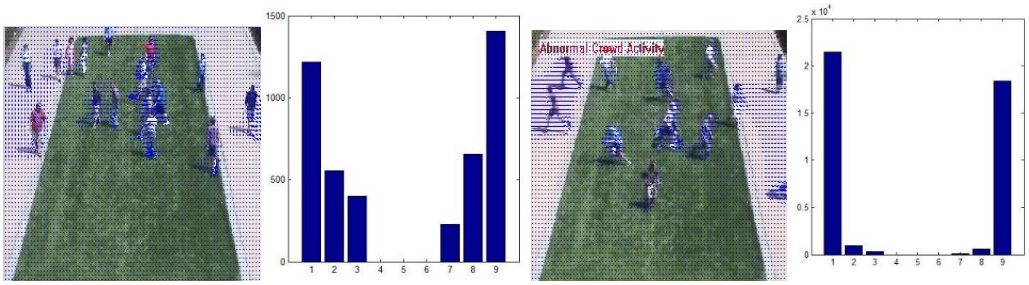


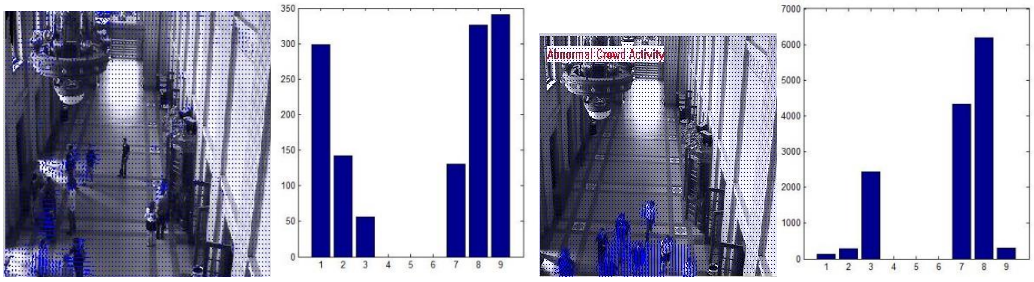
Figure 6. Corresponding Histogram of the Interactive Force

The HOIF are computed and combined to a monolithic vector, as shown in Figure 6. But in UMN database, people suddenly begin to run towards all the directions after walking normally for a while in these three scenes. The corresponding optical flow was obtained and there are speed values in all directions, that is to say, the nine bins will all have their values of the corresponding histogram, no matter the scene is normal or abnormal.

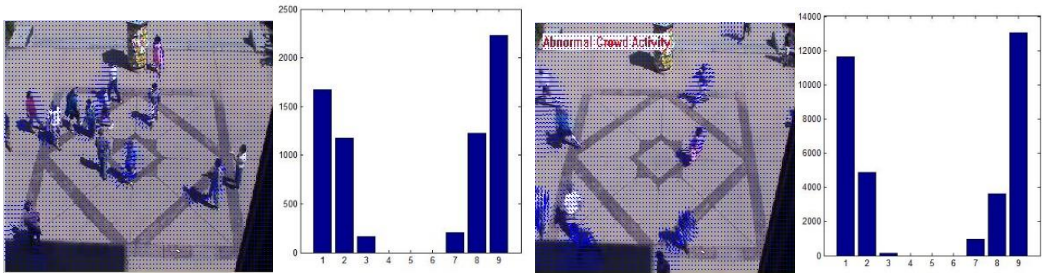
In that case, how could we utilize the information better provided by the optical flow? After analyzing the optical flow, we found that the speed value of optical flow in abnormal scene (as shown in Figure 7.(c),(g),(k)) is much bigger than in the normal scene (as shown in Figure 7.(a),(e,) (i)).



(c) Normal scene of scene1 (a) (b) (b)The corresponding histogram of (a) (c) Abnormal scene (d) The corresponding



(e)Normal scene of scene2 (f) The corresponding histogram of (e) (g) Abnormal scene of scene2 (h) The corresponding histogram of (g)



(i) Normal scene of scene3 (j) The corresponding histogram of (i) (k) Abnormal scene of scene3 (l) The corresponding histogram of (k)

Figure 7. (a),(b),(c),(d) Show the Motion Information of Scene 1; (e),(f),(g),(h) Show the Motion Information of Scene 2; (i),(j),(k),(l) Show the Motion Information of Scene 3

In order to use the founding better, we make each of the value of the nine dimensions of the motion vector equals to the sums of the amplitude of the speed value falling into the same parts to increase the differentiation between the normal and abnormal scene, as shown in Figure 7. The values of the bins of the histogram in the normal scene(as shown in Figure 7.(b),(f),(j)) is much bigger than the values of the bins of the histogram in the abnormal scene(as shown in Figure 7.(d),(h),(l)).

2.3. Local Binary Pattern Co-Occurrence Matrix

Many methods based on texture features have been proposed to solve this problem. Most of the existing algorithms only estimate crowd density on the whole image while ignoring crowd density in local region. When a video clips input, RGB images will be converted to gray scale image and gray-grad image. As Figure 8 shows, for each pixel, it is compared to its neighbors (number of neighbors can be 8, 12 or even more).

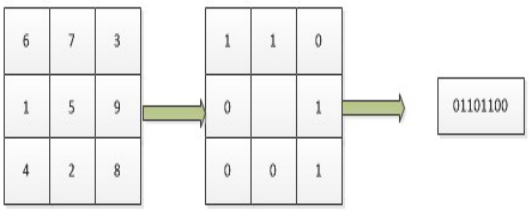


Figure 8. An Example of LBP

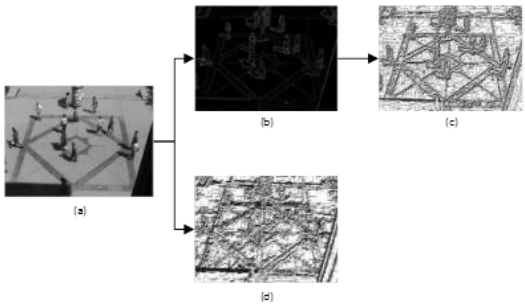


Figure 9. Shows the LBP Image of One Frame of Scene 3

In this work, we calculate LBP for each pixel with distance as 1 pixel and consider all the 8 neighbors. Then each pixel, except the boundary of the input image block, gets a LBP value. The LBP value of a pixel is exactly an integer between 0 and 255, which makes the LBP map of the whole image block processed as a gray scale image, as shown in Figure 9.(c), (d). In Figure 9, (a) shows a gray image of one frame of scene 3 of UMN, (b) is the gradient image of (a); (c) shows the LBP image of (b);(d) shows the LBP image of (a).

This work calculates Co-Occurrence Matrix of the LBP-gradient and LBP-gray image

separately. Co-Occurrence Matrix is proposed by Haralick[17] and Marana[18]. Co-occurrence matrix is a statistical method based on the estimation of second-order joint conditional probability density functions, $f(i,j|d,\theta)$ .Each $f(i,j|d,\theta)$ represents the probability that a pair of gray levels (i,j) occurs at the distance of d along the direction θ in an image. It is difficult to estimate crowd density using the original GLCM. So, we adopt four widely used descriptors for texture measurements: energy, contrast, homogeneity and entropy.

These four volumes can react crowd density information very well, and this paper will compute the four volumes of co-occurrence matrix both in gradient-LBP and gray-LBP separately to generate the feature vector with 8 dimensions. That is to say the values of the four volumes of the gradient-LBP and gray-LBP consists the 8 dimensions of the feature vector. We calculate four sets of co-occurrence features with parameters d as 1 and θ as 0°, 45°, 90° and 135° respectively. Finally the 32 dimensions density feature vectors of the gray-grad-LBPCM can be obtained.

3. Global Abnormal Crowd Behavior Detection

This paper combines the crowd density feature vectors (32 dimensions) with the crowd motion feature vectors (9 dimensions) together to form a new vector (41 dimensions), called crowd behavior vector. Then our work utilized the crowd behavior vector to establish the pattern recognition model Social Force Model (SFM) to distinguish the normal and abnormal events.

3.1. UMN Dataset

The proposed features and their applications are tested on the publicly available datasets containing the normal and abnormal crowd videos from youku.com and United States University of Minnesota UMN dataset (UMN) . The specific details are as follows. The resolution of the UMN dataset is 240×320 . UMN database includes 3 different scenes with 11 different kinds of escape videos, each video segment begins with normal behavior of pedestrians with free movement, and ends with the abnormal behavior fleeing in the same direction of the pedestrian. The dataset properly takes into account the number of people of different scenes, different lighting, various factors such as uncertainty and randomness of crowd movements, which is the most widely applied video database in the detection of abnormal behavior.

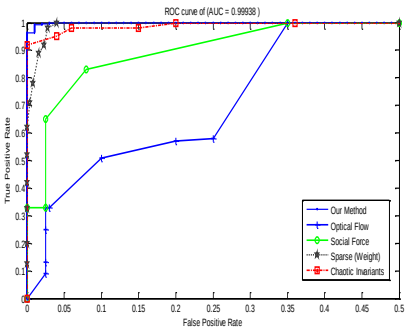


Figure 10. ROCs for the Detection of Abnormal Frames for Scene One in UMN Dataset

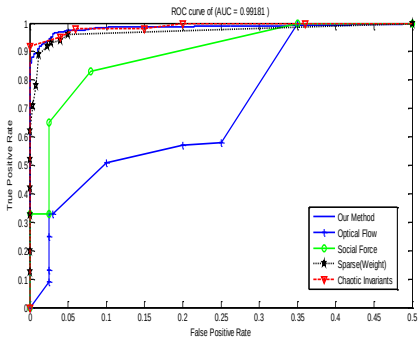


Figure 11. ROCs for the Detection of Abnormal Frames for Scene Two in UMN Dataset

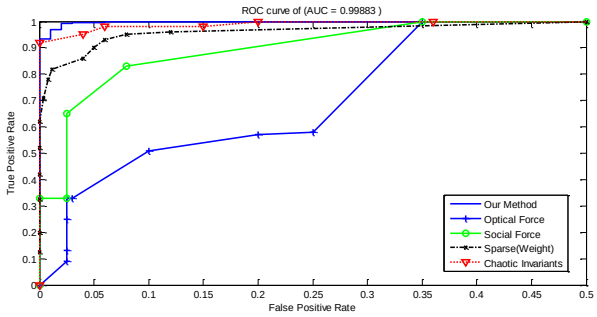


Figure 11. ROCs for the Detection of Abnormal Frames for Scene Three in UMN Dataset

The simulation results ROC curving by frame-level measurement are shown to compare our method to four other measurement (Optical Flow Method, Social Force Model, Sparse Reconstruction, Chaotic Invariants Method) for three scenes of UMN dataset respectively.

Table 1.Comparison of the Proposed Method with the State-of-the-art Methods for Detection of the Abnormal Events in UMN Dataset

Method	AUC
Chaotic Invariants[10]	0.99
Social Force[18]	0.96
Optical Flow[18]	0.84
Sparse Scene1[30]	0.995
Sparse Scene2[30]	0.975
Sparse Scene3[30]	0.964
Our Method-Scene1	0.99938
Our Method-Scene2	0.99181
Our Method-Scene3	0.99883

From the graph, we found that our method outperforms the other methods. However, Figure 11 shows that our results were slightly worse than other two scenes (Figure 10,

Figure 12), mainly because there are the existence of a large number of lighting effects in scenario two. Our method does not perform the light and shadow processing, for we consider to reduce the complexity and improve computing efficiency.

Moreover, Table 1 provides the quantitative comparisons to the state-of-the-art method. And the AUS is 0.99938, 0.99181, 0.99883. The proposed method is a more general solution in a signal scene for the global abnormal detection. The result demonstrates that the LBPCM-gray-grad can well describe the crowd information from all angles. Meanwhile, it can certify the robustness of our HOIF-LBPCM method.

3.2. Web Dataset

The sufficient tests on the abnormal detection in the global scenes have also been carried out on more challenging data collected in public places from youku.com, with the purpose of evaluating our method in practical applications.

Figure 13. The Fight Scene.(a) Shows the Abnormal Fight Scene,(b) Describes the Optical Flow of the Abnormal Corresponding to (a).(c) is the ROC for Detection of Abnormal Crowd Behavior based on the Datasets form Web

The resolution of the web data is 512×288, the frame rate is 15, and this video describes a violence event, occurring in a public place. The experimental results are shown in Figure 13. Through MATLAB simulation, the detection rate reaches 98.75%. Thus we infer that our method has strong generalization and robustness.

4. Conclusions

In this paper, we propose a novel approach to detect the global abnormal crowd behavior by using HOIF and gray-gradient-LBPCM features. We adopt four widely used descriptors for texture measurements to characterize crowd density. Meanwhile, we introduce a histograms of the orientation of the interaction force as the descriptors for reliably capturing crowd motion information of the monolithic video frame, which can effectively be used to simulate abnormal crowd behavior. The experiments on two dataset show favorable results when compared with the state-of-art method.

References

[1] Y. Cong, J. S. Yuan and J. Liu, “Abnormal event detection in crowded scenes using sparse representation”, Pattern Recognition, vol. 46, no. 7, **(2013)**.
[2] R. Raghavendra, A. Del Bue, M. Cristani and V. Murino, “Optimizing interaction force for global anomaly detection in crowded scenes”, ICCV, **(2011)**.
[3] S. S. Pathan, A. Al-Hamadi and B. Michael, “Incorporating social entropy for crowd behavior detection using SVM”, ISCV, **(2010)**.
[4] T.Wang and H. Snoussi, “Histograms of optical flow orientation for visual abnormal events detection”, AVSS, **(2012)**.

[5] H. Yang, Y. H. Cao and S. Wu, "Abnormal crowd behavior detection based on local pressure model", APSIPA ASC, **(2012)**.
[6] S. H. Cho and H. B. Kang, "Integrated multiple behavior models for abnormal crowd behavior detection", SSIAI, **(2012)**.
[7] R. Mehran, A. Oyama and M. Shah, "Abnormal crowd behavior detection using social force model", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, **(2009)**.
[8] S. D. Wu, B. E. Moore and M. Shah, "Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes", CVPR, **(2010)**.
[9] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin, L. Q. Xu and B. B. Zhan, "Crowd analysis: a survey", Machine Vision and Applications, vol. 19, no. 5-6, **(2008)**.
[10] N. Haering, P. Venetianer and A. Lipton, "The evolution of video surveillance: an overview", Machine Vision and Applications, vol. 19, no. 5, **(2008)**.
[11] A. Chan and B. Vasconcelos, "Bayesian poisson regression for crowd counting", Computer Vision of IEEE 12th International Conference of IEEE, **(2009)**.
[12] B. Wang, M. Ye, X. Li, F. Zhao and J. Ding, "Abnormal crowd behavior detection using high-frequency and spatio-temporal features", Machine Vision and Applications, vol. 23, no. 3, **(2012)**.
[13] S. Ali and M. Shah, "A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, **(2007)**.
[14] D. M. Blei, A. Y. Ng and M. I. Jordan, "Latent dirichlet allocation. Journal of Machine Learning Research", vol. 3, **(2003)**.
[15] T. Brox, A. Bruhn and N. Papenberg, "High accuracy optical flow estimation based on a theory for warping", ECCV, **(2004)**.
[16] Unusual crowd activity dataset of University of Minnesota, available from http://mha.cs.umn.edu/proj_events.shtml#crowd.
[17] Haralick and M. Robert, "Statistical and structural approaches to texture", Proceedings of the IEEE, vol. 67, no. 5, **(1979)**.
[18] A. N. Marana, S. A. Velastin, L. F. Costa and R. A. Lotufo, "Estimation of crowd density using image processing", Image Processing for Security Applications (Digest No.: 1997/074), IEE Colloquium on IET, **(1997)**, pp.11-1.
[19] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", Proceedings of the Conference on Computer Vision and Pattern Recognition, **(2005)**; San Diego, California, USA.
[20] C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines", ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 1, **(2011)**.

Authors

Yong Yin, He received the Ph.D degree in control engineering from Chongqing University, China, in 1999. In 1984, he joined College of Communication Engineering, Chongqing University, China, where he was engaged in research on digital image and video processing. He is a professor in Chongqing University. His current research interests are centered around image processing and recognition and image and video.

Qiannan Liu, She sreceived the B. Sc. degree in Zhoukou Normal University, China in 2013.She is currently a master candidate of Chongqing University, China, where her research interests include pattern recognition and digital image processing.

Shibiao Mao, He received his B. Sc. degree in Communication engineering from Hu Bei University of Technology, China in 2011,He is currently a master candidate of Chongqing University, China, where his research interests include pattern recognition and computer vision.

