Abnormal crowd behavior detection based on local pressure model

Hua Yang, Yihua Cao, Shuang Wu, Weiyao Lin, Shibao Zheng, Zhenghua Yu
Institute of Image Communication and Network Engineering,
Shanghai Key Laboratory of Digital Media Processing and Transmissions,
Shanghai Jiao Tong University, Shanghai, China
E-mail: hyang@sjtu.edu.cn,5061509181@sjtu.edu.cn

Abstract — **Abnormal crowd behavior detection is an important issue in crowd surveillance. In this paper, a novel local pressure model is proposed to detect the abnormality in large-scale crowd scene based on local crowd characteristics. These characteristics include the local density and velocity which are very significant parameters for measuring the dynamic of crowd. A grid of particles is placed over the image to reduce the computation of the local crowd parameters. Local pressure is generated by applying these local characteristics in pressure model. Histogram is utilized to extract the statistical property of the magnitude and direction of the pressure. The crowd feature vector of the whole frame is obtained through the analysis of Histogram of Oriented Pressure (HOP). SVM and median filter are then adopted to detect the anomaly. The performance of the proposed method is evaluated on publicly available datasets from UMN. The experimental results show that the proposed method can achieve a higher accuracy than that of the previous methods on detecting abnormal crowd behavior.**

I. INTRODUCTION

Over the past decades, a wide attention has been paid to crowd control and management in intelligent video surveillance area. Especially, abnormal crowd behavior detection is becoming a challenging issue, and there're many related research work have been done in recent years. One of the most popular models for formulating the crowd movement is the social force model [1], which is based on the theory of physics and social psychology. After that, in [2] Herbing considers the panic effect and constructs a generalized social force model. This model is demonstrated to be effective in detecting the abnormalities of the crowd in [3]. Afterwards, some methods based on social force model to detect abnormal crowd behavior are proposed [4-6]. However, these algorithms using social force model only utilize the temporal characteristic, that is local velocity, and ignore the spatial information such as local density.

In this paper, we construct a novel local pressure model to formulate the dynamics of the crowd. The proposed model considers the local density and local velocity simultaneously, while the generalized social force model and its variants only utilize the local velocity. In fact, it's obvious that the abnormalities of the crowd depend on the changes of local crowd parameters, local density and local velocity. Generally, when emergency happens, drastic changes of local velocity and density will be represented. For example, when sudden gathering or scattering appears, local density and velocity usually changes greatly along with the movement of people. The variation of local velocity indicates the dynamic of the crowd to some degree, but it's not complete. The local density plays the same important role as the local velocity. The difference is that they reflect the state of the crowd in two different dimensions. The variation of the local velocity represents the crowd motion in temporal dimension, while the local density indicates the crowd distribution in spatial dimension. Therefore, in order to model the crowd motion more accurately we need consider the local density and velocity simultaneously. That's the motivation of our local pressure model. When the local density doesn't change with the time, our model degenerates into the generalized social force model.

The rest of this paper is organized as follows. In the next section we introduce the whole system structure and elaborate the local pressure model. In Section III, Histogram of oriented pressure (HOP), SVM and Median Filtering are stated in detail to detect the anomaly. Section IV gives experimental results to demonstrate the superiority of the local pressure model. Conclusions are drawn in the last section.

II. LOCAL PRESSURE MODEL

The system structure of anomaly detection is shown in Fig.1. First of all, local pressure is extracted by combining the local interaction force using social force model and the local density using spatial-temporal local binary pattern. Then the Histogram of Oriented Pressure is established to draw the crowd feature vector. Afterwards, Support Vector Machine (SVM) is used for its simplicity and efficiency to classify the normal and abnormal behaviors. In order to remove the unexpected transient anomaly, median filtering is adopted after the classification.

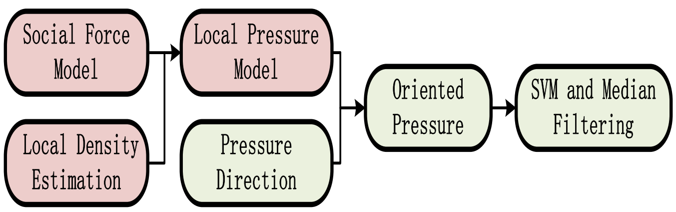


Fig.1 System Structure of Anomaly Detection

The change of the local density can directly indicate the crowd dynamic in spacial domain. And the larger the density change, the higher the pressure between pedestrians is. Based on this point, we propose local pressure model, which considering the local density and interaction forces simultaneously.

In the model, the pressure on a pedestrian i at location $\vec{r}_i$ is defined as

$$P = e^{\left|\frac{d\rho(\vec{r}_i,t)}{dt}\right|^{\alpha}} \cdot F_{\text{int}} \qquad (1)$$

where $\rho(\vec{r}_i, t)$ denotes the local density of location $\vec{r}_i$ at time t, F_{int} denotes the interaction force from other pedestrians and objects around pedestrian i and α is an empirical parameter.

In conventional object-based methods, people in crowd need to be segmented and tracked accurately. But it has the limitation of low accuracy in medium and high density levels due to the occlusions and clusters. Therefore, object-based method is not feasible in dense crowd which has the higher possibility to bring out anomalies. Inspired by [3], we put a grid of particles over the image and treat them as the basic entity instead of tracking each pedestrian. After the particles are divided, interaction forces and local density need to be calculated which are elaborated in the following sections.

A. Social Force Model[3]

The interaction force F_{int} can be obtained via social force model, which is a very important model for crowd behavior analysis in video surveillance and crowd behavior synthesis in computer graphics. Social force model describes the crowd motion dynamics based on the theories of physics and social psychology. Personal desire and environment constraints are mainly considered in this model. The actual force F on pedestrian i with mass of m_i is

$$F = m_i \frac{dv_i}{dt} = F_p + F_{\text{int}} = \frac{1}{\tau}\left(v_i^p - v_i\right) + F_{\text{int}} \quad (2)$$

where F_p describes the personal desire force and is defined as $\frac{1}{\tau}\left(v_i^p - v_i\right)$ in social force model, v_i^p denotes the desired velocity while v_i denotes the actual motion. Considering the effect of panic, v_i^p can be replaced with

$$v_i^q = \left(1 - p_i\right) v_i^p + p_i v_i^a \qquad (3)$$

where p_i (0~1) is the panic weight parameter, the higher the herding effect, the higher the p_i is and the higher the local density is [3], so this parameter can be expressed by the local density. v_i^a is the average velocity of the neighboring pedestrians.

Without loss of generality, we set $m_i = 1$. Thus, F_{int} can be estimated by

$$F_{\text{int}} = \frac{dv_i}{dt} - \frac{1}{\tau}\left(v_i^q - v_i\right) \qquad (4)$$

In this paper, optical flow is utilized to estimate the velocity. v_i^a is described as $O_{ave}\left(x_i, y_i\right)$ which is the average of optical flow in spatial-temporal volume around the particle $i(x_i, y_i)$. Then the desired velocity of particle i is

$$v_i^q = \left(1 - p_i\right) O(x_i, y_i) + p_i O_{ave}(x_i, y_i) \qquad (5)$$

Then the interaction force F_{int} can be calculated out.

B. Local Density Estimation based on spatial-temporal LBP

Compared to other density estimation algorithms which focus more on the density level of the whole scene, spatial-temporal local binary pattern (LBP) can be used to extract the local density distribution. Besides, LBP also has the advantage of simple computation and high performance [7, 8].

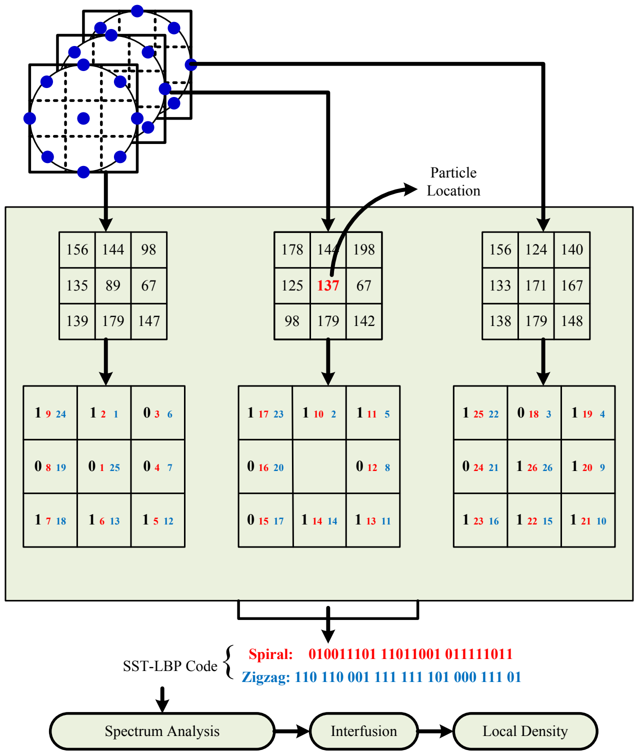


Fig.2 Schematic for Local Density Estimation using Local Binary Pattern

The overall computing procedure for the local density calculation using local binary pattern is given in Fig.2. Firstly, the neighboring pixels within a sphere with the radius R in the spatial-temporal domain are sampled. Afterwards, the binary value will be obtained with the gray value of the particle pixel as threshold. Then the two kinds of LBP code are obtained by arranging the binary value in spiral order and zigzag order respectively. Spiral order means that the values within the frame are firstly loaded by the circle and then the values of the three frames are connected, just like arranging values along the spiral curve, shown as the red number from 1 to 26. Zigzag order means that the corresponding values in the three frames are successively loaded, shown as the blue number. In essence, spiral order is spatial-temporal order while zigzag order is temporal-spatial order. The LBP code of the consistent part such as the background changes seldom while the code of the occultation changes more frequently, so the LBP code extracted from the occultation volume contains more high-frequency components than that from low-density part. Considering that, spectrum analysis is applied on the LBP code. The higher the spectrum analysis result is, the higher the local density is [8]. After the interfusion of the two spectrum analysis, the relative local density can be achieved. Here, we take the average of the two analysis results as the interfusion result. The color map is then adopted to transform the local density where red values represent higher density while blue values represent lower density. In Fig.3, the LBP analysis result has been shown, demonstrating that the proposed method using LBP to calculate the local density is reasonable.

(a)Origin Image (b)Density Distribution

Fig.3 Local Density Distribution using Local Binary Pattern

III. ANOMALY DETECTION

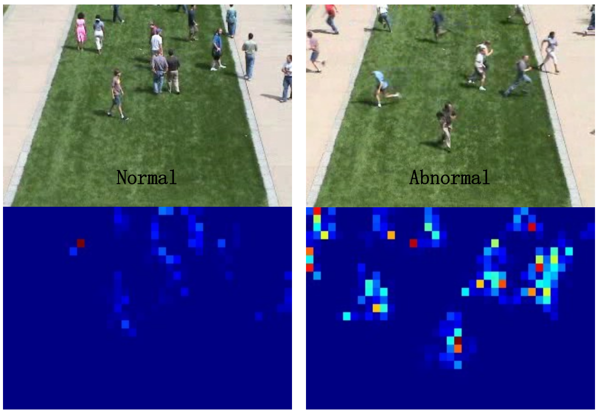
In this section, the follow-up process after extracting the local pressure is introduced. In order to improve the detection accuracy and reduce the computation, histogram is utilized to extract the statistical property of the magnitude and direction of the pressure. The crowd feature vector of the whole frame is obtained through the analysis of histogram of oriented pressure. Each element in the crowd feature vector is the number of the particles whose spectrum analysis result falls into the corresponding bin. Cross statistics of magnitude and direction are adopted instead of parallel statistics so that more information is contained in the histogram. We can divide 2π into m bins and divide the magnitude of the pressure into n bins. In cross statistics, pressures are divided into m categories by the direction first, and then in each category, pressures are divided into n classes, so the overall bins are $m \times n$. While in parallel statistics, the overall bins are $m+n$.

Finally, after the crowd feature is obtained by the HOP, Support Vector Machine (SVM) [9], which has the advantage of small sample, global optimum and strong generalization ability, is used to classify the behavior in the scene. Besides, since the SVM classify the scene by the frame, sometimes, it will generate single-frame or double-frames anomaly. In order to avoid this situation, we apply median filter on the classification result. In this paper, we set the window size of the filter as 5 since larger size may reduce the accuracy of beginning and ending detection of the anomaly and smaller size cannot remove the double-frames anomaly.

IV. EXPERIMENTAL RESULTS

In this paper, experiments are performed on three different scenes of the UMN dataset [10]. Each scene contains some video segments. The resolution of the video is 320×214 . People suddenly begin to run towards all the directions after walking normally for a while in these three scenes. The particle size to calculate pressure is set to be 8×8 in order to achieve a balance between the calculation and the accuracy. The result of the pressure can be seen in Fig.4.

For the histogram, we set the number of direction bins and magnitude bins to be 9 and 7 respectively. Then the overall bins are $9 \times 7 = 63$. Fig.5 shows the histogram of oriented pressure.



(a) Normal Pressure (b) Abnormal Pressure

Fig.4 Comparison of Pressure in Normal and Abnormal Situations

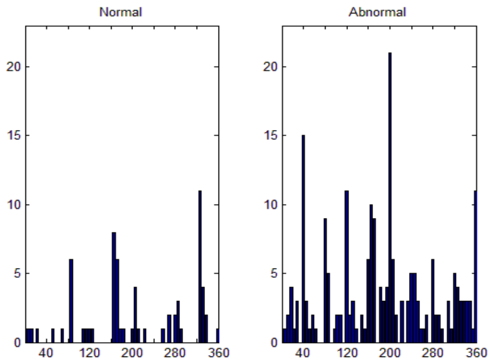


Fig.5 Histogram of Oriented Pressure

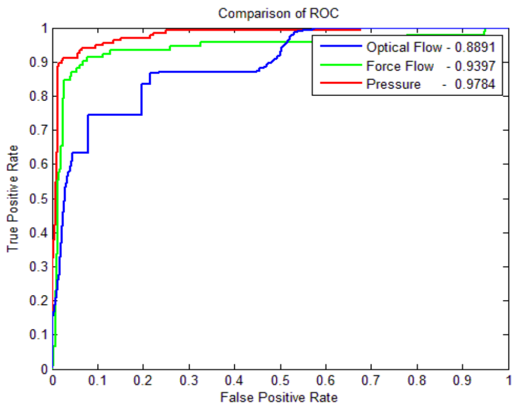


Fig.6 ROC Curve for Anomaly Detection

The ROC curve for the temporal anomaly detection shown in Fig.6 indicates that our method performs better than the optical flow and social force. It proves that combining different local characteristics such as spatial-temporal density variation, local density distribution and direction is effective.

Usually, it's more important to detect the anomaly at the right point so that people can give an early warning to avoid serious accidents. Besides, giving the accuracy time when the anomaly appears can help people to retrieval from large amounts of videos. In Fig.7, the detection results are listed compared with ground truth. The training number of frames for these three segments is 614, 540 and 692 respectively. Compared to the detection result in [3] which only uses social force, the result shows that the proposed method has a large superiority in detecting the begin and the ending of the anomaly. It is almost completely consistent with the ground truth except for the result of the first scene. The reason for the undetection in first scene is that in the calibration, we regard the frames as abnormal from the point when people begin to run to the point when the last person runs out of the scene. In the first scene, it lasts a little too long for the last person to run out of the scene. But in our detection, if the density is too low, it's very probably detected as normal, which is reasonable in the real world. From the result, we can see that the system is suitable and practical for crowd surveillance application.

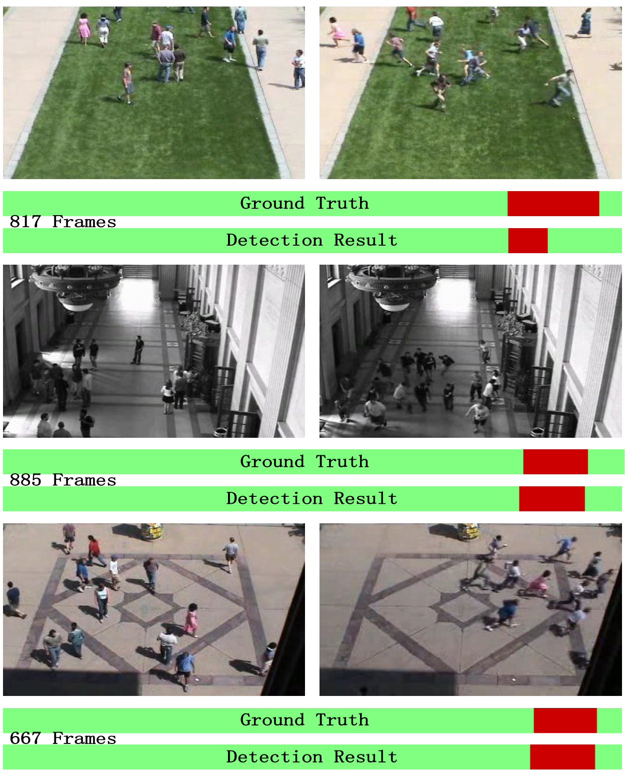


Fig.7 Spatial-temporal anomaly detection results

V. CONCLUSIONS

In this paper, we propose an efficient method of **Histogram of Oriented Pressure (HOP)** to detect crowd anomaly. Social Force Model and Local Binary Pattern are adopted to calculate the pressure. Besides, the direction of the pressure is introduced into the statistics to improve the accuracy. Cross histogram is utilized to produce the feature vector instead of parallel merging the magnitude histogram and direction histogram. Afterwards, SVM is adopted due to its high efficiency and effectiveness. Finally, unexpected transient anomaly is removed by the median filtering. The experimental results demonstrate that the beginning and ending of the anomaly can be detected rightly.

ACKNOWLEDGMENT

This research is supported by NSFC (No.61102099, No.61171172), STCSM (No.10231204002, No.11231203102) and National Key Technologies R&D Program (2011BAK14B02).

REFERENCES

[1] D. Helbing and P. Molnar. Social force model for pedestrian dynamics. *Physical Review E*, 51:4282, 1995.

[2] I. F. D. Helbing and T. Vicsek. Simulating dynamical features of escape panic. *Nature*, pages 487–490, 2000.

[3] R. Mehran, a. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," *2009 IEEE Conference on Computer Vision and Pattern Recognition*, no. 2, pp. 935-942, Jun. 2009.

[4] J. Zhao, Y. Xu, X. Yang, and Q. Yan, "Crowd instability analysis using velocity-field based social force model," *2011 Visual Communications and Image Processing (VCIP)*, pp. 1-4, Nov. 2011.

[5] R. Raghavendra, A. D. Bue, M. Cristani, and V. Murino, "Abnormal Crowd Behavior Detection by Social Force Optimization," Lecture Notes in Computer Science, vol. 7065, pp. 134-145, 2011.

[6] B. Wang, M. Ye, X. Li, and F. Zhao, "Abnormal crowd behavior detection using size-adapted spatio-temporal features," *International Journal of Control, Automation and Systems*, vol. 9, no. 5, pp. 905-912, Oct. 2011.

[7] Guoying Zhao, Pietikainen M., "Dynamic texture recognition using local binary patterns with an application to facial expressions", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 915-928, 2007.

[8] Hua Yang, Hang Su, Shibao Zheng, "The Large-scale Crowd Density Estimation Based on Sparse Spatio-temporal Local Binary Pattern", IEEE International Conference on Multimedia and Expo (ICME),2011, pp. 1-6.

[9] Chapelle O., Haffner P., Vapnik V.N., "Support vector machines for histogram-based image classification", IEEE Transactions on Neural Networks, vol. 10, no. 5, pp. 1055-1064,1999.

[10] Unusual crowd activity dataset of University of Minnesota, available from http://mha.cs.umn.edu/movies/crowd-activity-all.avi.