Research Article

Crowd Motion Analysis Based on Social Force Graph with Streak Flow Attribute

Shaonian Huang,[1,2] Dongjun Huang,[1] and Mansoor Ahmed Khuhro[1]

[1]*School of Information Science and Engineering, Central South University, Changsha 410083, China*
[2]*School of Computer and Information Engineering, Hunan University of Commerce, Changsha 420005, China*

Correspondence should be addressed to Shaonian Huang; hsn@hunnu.edu.cn

Over the past decades, crowd management has attracted a great deal of attention in the area of video surveillance. Among various tasks of video surveillance analysis, crowd motion analysis is the basis of numerous subsequent applications of surveillance video. In this paper, a novel social force graph with streak flow attribute is proposed to capture the global spatiotemporal changes and the local motion of crowd video. Crowd motion analysis is hereby implemented based on the characteristics of social force graph. First, the streak flow of crowd sequence is extracted to represent the global crowd motion; after that, spatiotemporal analogous patches are obtained based on the crowd visual features. A weighted social force graph is then constructed based on multiple social properties of crowd video. The graph is segmented into particle groups to represent the similar motion patterns of crowd video. A codebook is then constructed by clustering all local particle groups, and consequently crowd abnormal behaviors are detected by using the Latent Dirichlet Allocation model. Extensive experiments on challenging datasets show that the proposed method achieves preferable results in the application of crowd motion segmentation and abnormal behavior detection.

1. Introduction

With the rapid increasing demand of public safety and security, crowd management is becoming one of the most challenging tasks of government department. It has been reported that there were 215 human stampede events worldwide resulting in 7069 deaths and over 14000 injuries between 1980 and 2007. Manually analyzing surveillance video is obviously time-consuming. Moreover, nowadays high-throughput video surveillance systems include a great amount of image data, which makes manual analysis virtually impossible. Therefore, how to model crowd motion plays a significant role in video surveillance system.

Crowd motion analysis has been intensively studied in computer vision fields; the current methods can be roughly divided into three categories from a high level perspective: microscopic approach, macroscopic approach, and hybrid methods. Microscopic approach needs to analyze pedestrian individuals' motivation in the video and then spread individual behavior to crowd motion. Most of these methods need to firstly detect the movement of pedestrian individual or pedestrian groups and then analyze local crowd motion behavior [1–4]. Helbing et al. firstly introduced social force model to simulate pedestrian dynamics. A few methods used social force model to detect and localize unusual events in crowd video [5–7]. Recently, some researchers adopt trajectory-based approach to analyze crowd behavior. Zhou et al. [8] and Bae et al. [9] modeled crowd behavior by extracting individual pedestrian's trajectories which were obtained by KLT tracker [10]. However, object tracking still has some unsolved challenges, such as occlusion or pedestrians entering/existing.

Macroscopic approach reviews crowd as an entity instead of modeling the motion of individual. This kind of algorithms usually takes optical flow to capture the crowd motion dynamic [11–13]. However, optical flow is instantaneous information between consecutive frames, and it fails to capture the motion by itself in a long-range time. A few researchers modeled fluid dynamic to characterize global crowd behaviors. Ali and Shah constructed a Finite Time Lyapunov Exponent (FTLE) field to segment crowd flow [14],
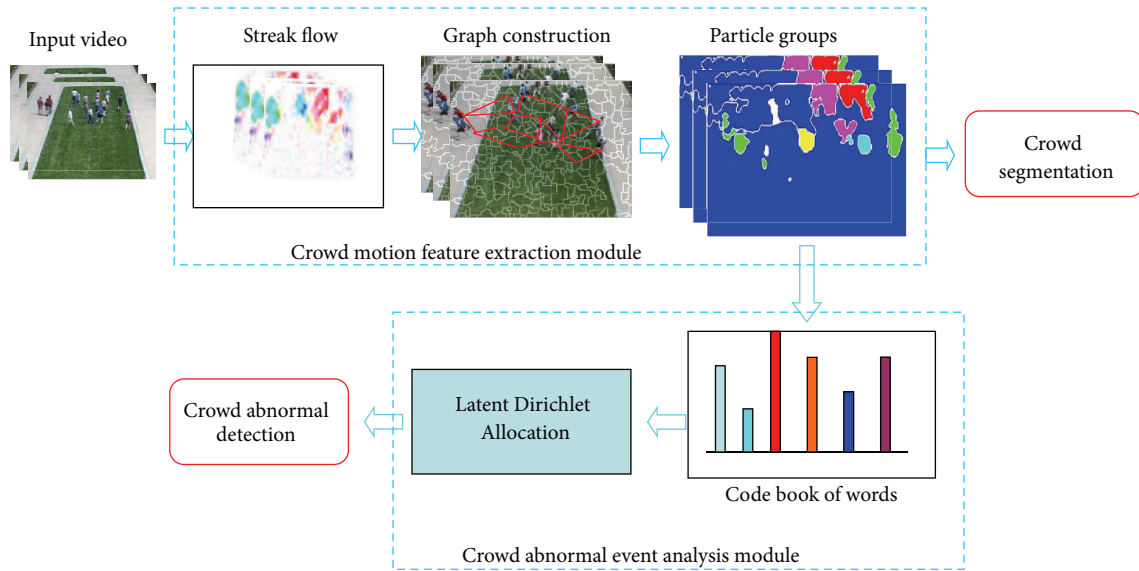
FIGURE 1: Flow chart of the proposed crowd motion analysis method.

but the segmented regions were unstable when the crowd density was not very high or crowd flow varied quickly.

Hybrid methods model crowd motion from macroscopic view as well as microscopic view. Hybrid models need to combine benefits of both macro and micro approaches. Pan et al. firstly studied human and social behavior from the perspectives of human decision-making and social interaction for emergency egress analysis [15]. Borrmann et al. presented a bidirectional coupling technique to predict pedestrian evacuation time [16]. Recently, Bera and Manocha modeled crowd motion based on discrete and continuous flow models; the discrete model is based on microscopic agent formulation and the continuum model analyzes macroscopic behaviors [17]. Inspired by the above hybrid method, we model crowd motion pattern by the combination of a macroscopic and a microscopic model of crowd dynamics. The macroscopic motion model is derived from streak flow representation which can depict motion dynamic over an interval of time [18]. And the microscopic motion model is based on the social force between the particle groups of crowd sequence.

The general flow chart of our proposed method to model crowd motion and detect abnormal behavior is illustrated in Figure 1. We first segment crowd sequence into spatiotemporal patches based on the similarity of streak flow. Then, we construct weighted social force graph by particle adventure scheme [14] and cluster the crowd video into particle groups to represent different motion patterns of crowd. Finally, we use Bag of Words (BoW) to generate codebook and then analyze crowd motion behavior based on Latent Dirichlet Allocation (LDA) [19] model. We evaluate our method on three public datasets and perform comparison with a number of state-of-the-art methods. The experimental results illustrate that the proposed method can achieve preferable results in application of crowd motion segmentation and crowd abnormal behavior detection.

In summary, the main contributions of this paper include the following: (1) a proposed framework to analyze crowd motion behavior which takes the global spatiotemporal motion and local pedestrian interaction into account; (2) construction of weighted social force graph to extract local motion pattern based on multiple social properties; (3) utilization of the LDA model to represent the relationship between the distribution of codebook and the crowd behavior pattern.

The remainder of this paper is organized as follows. Section 2 describes how to extract streak flow which can integrate the optical flow over an interval of time. Section 3 provides details on the construction of weighted social force graph for extracting crowd local motion. Section 4 demonstrates the strength of our method in application of crowd motion segmentation and crowd abnormal behavior detection. The experimental results from different data sets and the comparison results with existing methods are shown in Section 5. Finally, Section 6 includes the conclusion and potential extensions of our paper.

2. Streak Flow Representation of Crowd Motion

Streak flow is defined as an accurate representation of motion field of crowd video [18]. Different from optical flow which only provides motion information in a small time interval, streak flow can provide temporal evolution of motion objects in a period of time. Our approach of computing streak flow is similar to that proposed in [18]. However, for completeness, we will now shortly introduce the steps of this method. Firstly, we initialize regular particles at time t and then update the position of particles according to the optical flow at instant time. Let $(x_{f_t}^p, y_{f_t}^p)$ denote the position of particle p at the instant time t, which is initialized at point p and frame f for
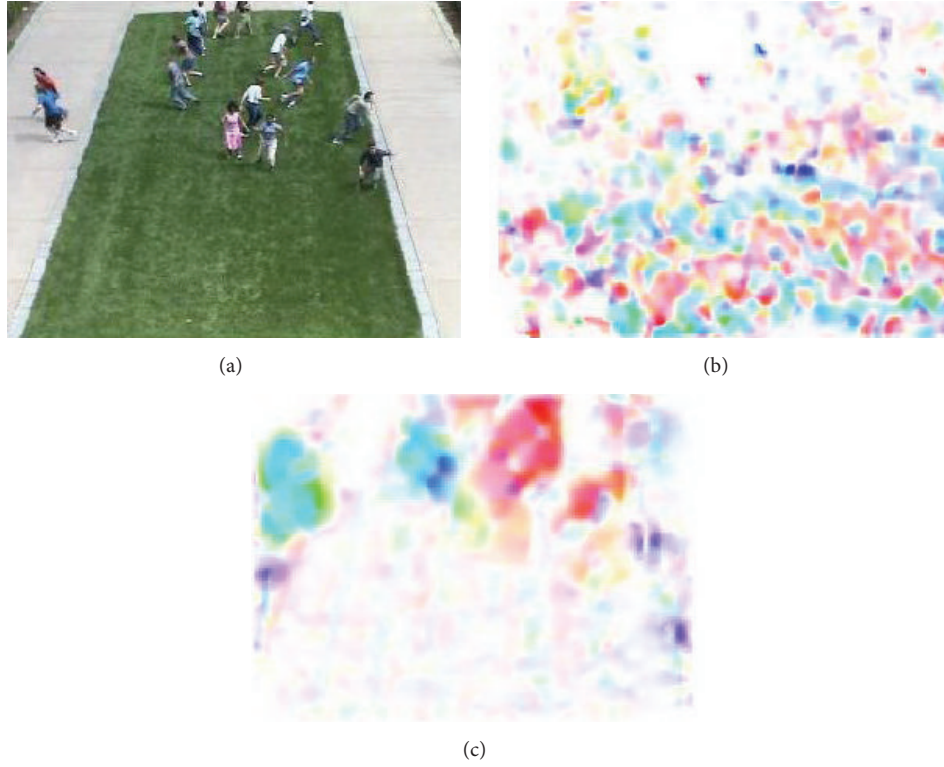
(a) (b)

(c)

FIGURE 2: Comparison of optical flow and streak flow. (a) The original image. (b) Optical flow of the original image. (c) Streak flow of the original image.

$t = 0, 1, 2, \ldots, T$. The position of particle p at any instant time is denoted as [18]

$$
\begin{aligned}
x_{f_t}^p &= x_{f_{t-1}}^p + u\left(x_{f_{t-1}}^p, y_{f_{t-1}}^p, f_{t-1}\right), \\
y_{f_t}^p &= y_{f_{t-1}}^p + v\left(x_{f_{t-1}}^p, y_{f_{t-1}}^p, f_{t-1}\right),
\end{aligned}
\tag{1}
$$

where u and v represent the velocity field of particle p at time t. This will generate a series of curves which start from particle p. For steady motion flow, these curves are the same as the paths of particles, but they will vary dramatically in unsteady flow. In order to fill the gaps in unsteady flow, we propagate the particles backward based on current flow filed using fourth-order Runge-Kutta-Fehlberg algorithm [14].

However, the positions of particles have usually subpixel accuracy. In order to represent streak flow $s = (s_u, s_v)$ of every pixel, we utilize the three nearest neighbor particles to compute the streak flow by linear interpolation. Here, we describe the computation of s_u; the method of estimating s_v is the same. We use $O_u = [O_u^i]$ to represent the streak flow of each particle i in vertical direction. Then, we define

$$
O_u^i = a_1 * s_u^i(n_1) + a_2 * s_u^i(n_2) + a_3 * s_u^i(n_3), \tag{2}
$$

where n_i is the index of the neighbor pixels and a_j is the known triangulation function [18] for the jth neighbor pixel.

Applying formula (2) on all of the data of s_u, we can obtain the following equation:

$$
As_u = O_u. \tag{3}
$$

We resolve the above equation and obtain the value of s_u. Figure 2 illustrates the difference between optical flow and streak flow. We can see that streak flow can better depict the motion changes.

3. Social Force Graph for Crowd Motion Analysis

The aforementioned streak flow represents particle motion in a period of time. However, in the process of crowd behavior analysis, we are concerned more about the motion of particle groups than a particular particle. Most of the existing methods extract group motion by clustering spatiaotemporal feature. However, the interaction between particles plays an important role in crowd motion fluid. The particular pedestrian determines and changes his own motion state according to the motion of his neighbor and the global motion in his perspective. Recently, social force model is adapted to simulate the dynamic of pedestrian's movement [1, 5, 7, 20, 21]. In some crowd applications, social force model has achieved satisfying performance.

Based on the above motivation, we propose the idea of motion particle group representation by social force graph with streak flow attribution. We use streak flow to represent

(a)   (b)

FIGURE 3: An example of spatiotemporal analogous patches. (a) The original image. (b) Partition patches overlap on the original image.

the global motion state in a period of time; meanwhile, we extract the social interaction force between particles to represent the influence of individual movement and then construct a weighted social force graph to extract particle groups revealing the local crowd motion patterns.

3.1. Partition Spatiotemporal Analogous Patches. In order to model crowd motion, we first partition crowd sequence into analogous patches which have similar spatiotemporal features. Specifically, we generate analogous patches by clustering the pixels with similar motion features. For a specific image in crowd sequence, point $p_i(x, y)$ represents any pixel of the image. We define the similarity of different pixels based on the streak flow similarity and distance similarity.

For any two pixels p_i and p_j, the distance similarity is defined as

$$\text{dis}_{ij} = \left\| x_i - x_j \right\|_2^2 + \left\| y_i - y_j \right\|_2^2. \quad (4)$$

In order to describe similarity of motion direction, we divided the whole direction plane into sixteen equal folds which are denoted as $\omega = \{i \mid i \in (1, 2, 3, \ldots, N; \ N = 16)\}$. The motion direction of any pixel is quantized into one of the sixteen bins based on the streak flow direction. Then, the direction similarity of two pixels is defined as

$$\text{dir}_{ij} = \left| \cos\left(\omega_i * 2 * \frac{\pi}{N} \right) - \cos\left(\omega_j * 2 * \frac{\pi}{N} \right) \right|. \quad (5)$$

And motion magnitude similarity of pixels is denoted as

$$\text{mag}_{ij} = \frac{\left(s_u^i - s_u^j \right)^2 + \left(s_v^i - s_v^j \right)^2}{\delta^2}, \quad (6)$$

where δ is the average magnitude of crowd streak flow and s_u^i and s_v^i are the streak flow of pixel p_i in horizontal direction and vertical direction.

In order to partition crowd sequence into spatiotemporal analogous patches $\{P_n\}_{n=1}^N$, we firstly divide the crowd image into N regular patches and then put seeds on each regular patch. The process of partition is to iteratively find the nearest seed for each pixel and then add the pixels into the patch

where the nearest seed belongs. In each iteration process, we update the position of seed as the center of the patch. The determination of nearest seed is based on minimizing spatiotemporal motion similarity and spatial distance as

$$\min \sum_{n=1}^N \sum_{i \in P_n} \left(\alpha \text{dis}_{i\mu^{(n)}} + \beta \text{dir}_{i\mu^{(n)}} + \gamma \text{mag}_{i\mu^{(n)}} \right), \quad (7)$$

where i is any pixel in patch P_n, $\mu^{(n)}$ is the center pixel of P_n, N is the number of spatiotemporal analogous patches specified in our experiment, and the parameters α, β, and γ are balance conditions according to different crowd sequence.

Figure 3 illustrates an example of the analogous patches based on the above method. We overlap the partition patches on the original image. The result shows that our partition method based on spatiotemporal similarity of particles can capture the change of motion objects. We can see that the pixels with similar motion pattern tend to be divided into the same patch, for example, the patches with labels 1, 2, and 3.

3.2. Social Force Graph for Particle Groups. Spatiotemporal analogous patches reflect the global visual similarity of moving pedestrian, but these patches neglect the interaction behavior between pedestrians. Here we describe how to construct a weighted social force graph $G = (V, E)$ to model the group pattern of crowd, where V is a set of vertices and E is a set of edges. In this paper, the vertices set V represents the set of regular particles in particle adventure scheme. The edges set E measures the individuals' motion similarity from the respective interaction social force. By the graph G, we can cluster similar particles together into particle groups to form the basis of crowd motion behavior analysis.

We need firstly to compute the interaction social force of particles. Interaction social force model describes crowd behavior by considering the interaction of individuals and the constraints of environment. The model is described as [5]

$$F_{\text{int}}^i = \frac{1}{\tau} \left(V_i^q - V_i \right) - m_i \frac{dV_i}{dt}, \quad (8)$$

where F_{int}^i is the interaction social force of particle i, V_i^q denotes the personal desired velocity, and V_i indicates the

actual velocity of individual. In this paper, V_i is represented as the streak flow of particle i defined in Section 2. And V_i^q is defined as

$$V_i^q = (1 - \eta)\, s^i + \eta \overline{s^p}, \tag{9}$$

where s^i is the streak flow of the particle i and $\overline{s^p}$ is the average streak flow of all particles in patch p where the particle i belongs.

As shown in [5], any high magnitude of interaction social force implicates that particle motion is different from the collective movement of the crowd. However, here interaction force is just only the motion information of one particle. In real-life crowd analysis, we usually need to use particle groups with similar motion patterns to characterize crowd behavior. Therefore, in this paper, we utilize a weighted social force graph to extract different particle groups by considering multiple factors, for example, distance factor and the initial analogous patches factor. For the weighted graph $G = (V, E)$, the weight w_{ij} of edge from particle i to particle j is used to measure the dissimilarity of different particles. And the value of w_{ij} is defined as

$$w_{ij} = \omega_{ij}^d * \omega_{ij}^g * \left\| F_{\text{int}}^i - F_{\text{int}}^j \right\|_2^2, \tag{10}$$

where ω_{ij}^d and ω_{ij}^g are the weighting factors based on distance factor and analogous patch factor, respectively. The smaller the value of w_{ij} is, the more similar the two particles are.

ω_{ij}^d is used to measure the influence of distance, and we think the motion similarity of different particles is inversely proportional to the distance between particles, which means the influence of particle i on neighbor particles is larger than on far particles.

Based on the above observation, ω_{ij}^d is defined as

$$\omega_{ij}^d = \exp\left(\frac{\sigma}{d_{ij}}\right), \tag{11}$$

where σ is the influential factor and d_{ij} is the Euler distance between particles. ω_{ij}^g is used to judge whether the two particles belong to the same initial analogous patch extracted in Section 3.1. If the two particles are in the same patch, then the value of ω_{ij}^g will be zero; otherwise, the value of ω_{ij}^g is one.

We use the clustering method [22] to extract particle groups to indicate similar motion patterns in a period of time. The details of obtained particle groups will be described in Section 4.

4. Application of Social Force Graph with Streak Flow Attribute

Using the weighted social force graph with steak flow attribute, we demonstrate the strength of our method for crowd motion segmentation and crowd abnormal behavior detection.

4.1. Crowd Segmentation. Based on the social force graph, we can obtain particle groups which have the similar motion patterns. In the graph G, the values of edge weights indicate the dissimilarity of motion pattern between particles. As shown in [22], a greedy method to segment image into regions was presented based on the intensity difference across the region boundary. Motivated by this method, in this paper, we firstly define a predicate value to represent graph region boundary and then segment the graph into different components. Each component of the graph has a smaller motion difference; therefore, different components represent the dissimilar particle groups. In order to measure the dissimilarity of two components, we utilize Dif_{in} to measure the motion difference of all particles in a component, and $\text{Dif}_{\text{in}}(R)$ is defined as

$$\text{Dif}_{\text{in}}(R) = \underset{e \in \text{MST}(R, E)}{\text{Max}} (w(e)), \tag{12}$$

where $\text{MST}(R, E)$ is the minimum generation tree of component R. We define Dif_{bew} to indicate the motion difference between different components:

$$\text{Dif}_{\text{bew}}(R_i, R_j) = \underset{v_i' \in R_i, v_j' \in R_j, (v_i', v_j') \in E}{\text{Min}} \left(w_{v_i' v_j'} \right). \tag{13}$$

When the value of $\text{Dif}_{\text{bew}}(R_i, R_j)$ is smaller than the predicate value Th_{merg}, we will merge the two components and then update the value of Th_{merg} as

$$\begin{aligned} \text{Th}_{\text{merg}} &(R_i, R_j) \\ &= \min\left(\text{Dif}_{\text{in}}(R_i) + \tau(R_i), \text{Dif}_{\text{in}}(R_j) + \tau(R_j) \right), \end{aligned} \tag{14}$$

where τ is a threshold function specified in the experiment based on the size of the components.

The results of crowd segmentation are presented in Section 5.

4.2. Crowd Abnormal Behavior Detection. Abnormal behavior detection is one of the most challenging tasks in the field of crowd analysis because the crowd abnormal behavior is difficult to be defined formally. Generally speaking, crowd abnormal behavior is defined as a sudden change or irregularity in crowd motion, such as the phenomenon of escape panic in public places or the dramatic increase of crowd density. In this section, we will describe how to detect crowd abnormal behavior based on our crowd motion model.

Based on aforementioned crowd segmentation, we have clustered crowd sequence into particle groups with similar motion pattern. We view each group as a visual word $\{w\}$ which is represented by interaction social force feature and the streak flow feature. And then we generate a codebook $D = \{w_1, w_2, \ldots, w_K\}$ by K-means scheme. The process of crowd abnormal behavior detection is regarded as the problem of word-document analysis. We adopt Latent Dirichlet Allocation [19] model to infer the distribution of visual words and then judge the emergence of abnormal event.

LDA is a well-known topic model for document analysis. It views the generation of document as the random selection of document topics and topic words. Particularly, in the crowd analysis, crowd behavior is represented as the result
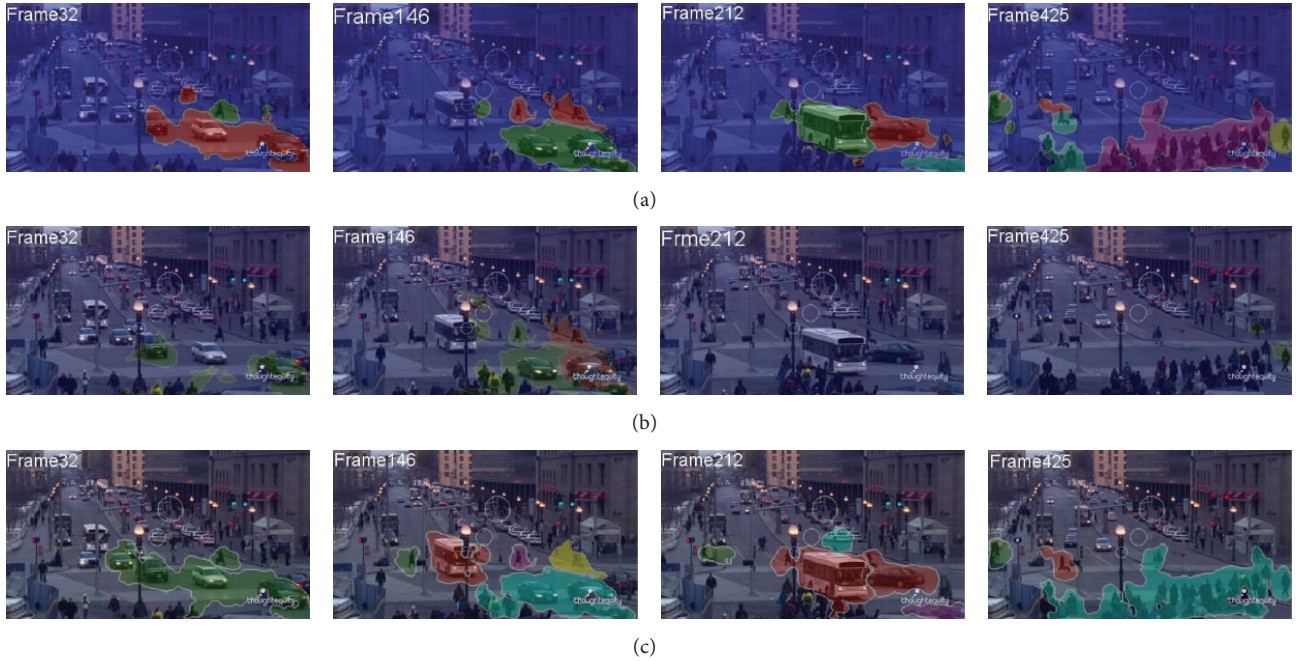
(a)

(b)

(c)

Figure 4: Comparison of segmentation results by using proposed method (1st row) and proposedNC (2nd row) and SRF (3rd row).

of random distribution of latent topics where the topics are denoted as the distribution of visual words. In this paper, we view the whole crowd video as a corpus, each crowd image as a document. Any document is modeled as the mixture of topics $z = \{z_1, z_2, \ldots, z_n\}$ with the joint distribution parameter θ.

LDA model is trained to obtain the value of parameters α and β, which denotes the corresponding latent topics and the conditional probability of each visual word belonging to a given topic, respectively. During the inferential process, the distribution of topic can be inferred by LDA model. According to the value of α and β, the distribution of topic is denoted as

$$p\left(\theta, z \mid v, \alpha, \beta\right) = \frac{p\left(\theta, z, v \mid \alpha, \beta\right)}{p\left(v \mid \alpha, \beta\right)}. \tag{15}$$

Crowd abnormal behavior detection is then achieved by comparing the topic distribution of current crowd frame to the model of training data.

5. Experiments and Discussion

In this section, we will perform different experiments on some publicly available datasets of crowd videos. We apply our method to some common crowd motion analysis problems. The goals of our experiments are (1) to segment the crowd video sequence into different motion pattern and (2) to detect abnormal behavior of crowd video.

5.1. Results of Crowd Segmentation.
Here we provide the results of our crowd segmentation algorithm. We use the same experiment video as that in [18], which is an intersection

video at Boston. Boston video includes 531 frames, and the video has three crowd behavior phases between twice traffic light: (1) traffic is formed, (2) traffic lights change and a new traffic flow emerges, and (3) traffic lights change again and another flow develops. We compare the proposed method with the method in [18] (denoted as SRF) and the proposed method without initially spatiotemporal analogous patches (denoted as proposedNC).

Figure 4 presents the segmentation results of the above three methods. We describe the results of our method in each frame. In frame 32, the north vehicles are moving to south and one person is walking on the sidewalk from east to west. We can see that our proposed method correctly segments the vehicle flow (1st row, red) and the person (1st row, green). But SRF is unable to distinguish the different motion patterns of vehicle and people (3rd row, green), and the proposedNC even neglected the motion of the people (2nd row, green). In frame 146, pedestrians who are walking on the sidewalk are detected as a group (1st row, red); however, SRF views the pedestrian as different groups (3rd row, yellow and purple). In frame 212, our method correctly segments the cars having different motion directions (1st row, green and red), but SRF views them as the same group (3rd row, red). In frame 425, the bottom pedestrian flow actually contains two different motion directions: some people are walking east and some people are going straight. Our method correctly detects the two motion patterns (1st row, purple and aquamarine); SRF, however, views the two motion patterns as a group (3rd row, aquamarine). From the above experimental results, we can see clearly that our proposed method not only is able to capture the temporal changes of crowd sequence (e.g., the car flow from north to south in frame 32) but also depicts the local motion changes (e.g., the cars from different direction

(a) Correctly segmented motion results

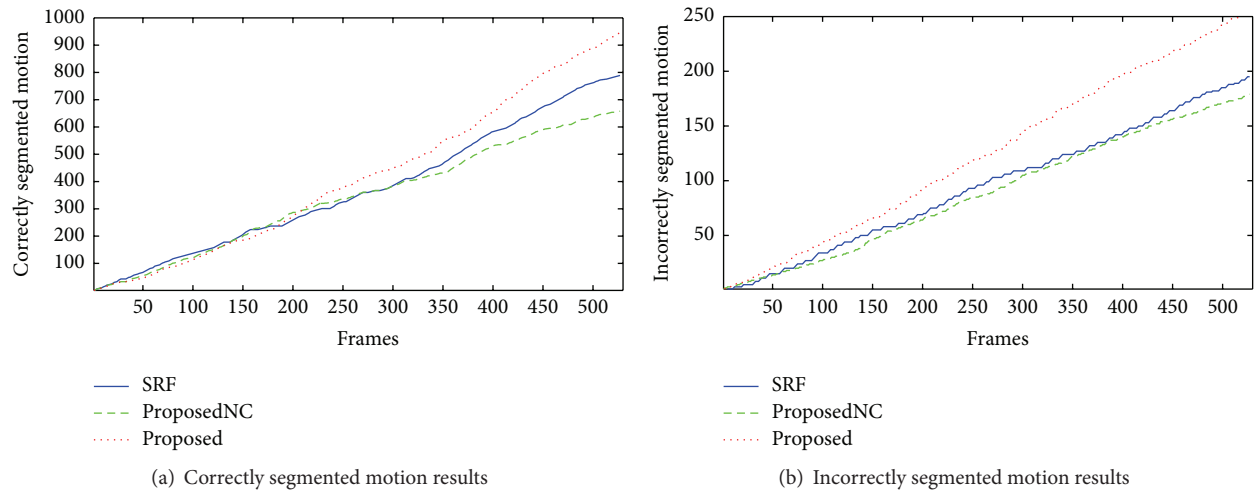(b) Incorrectly segmented motion results

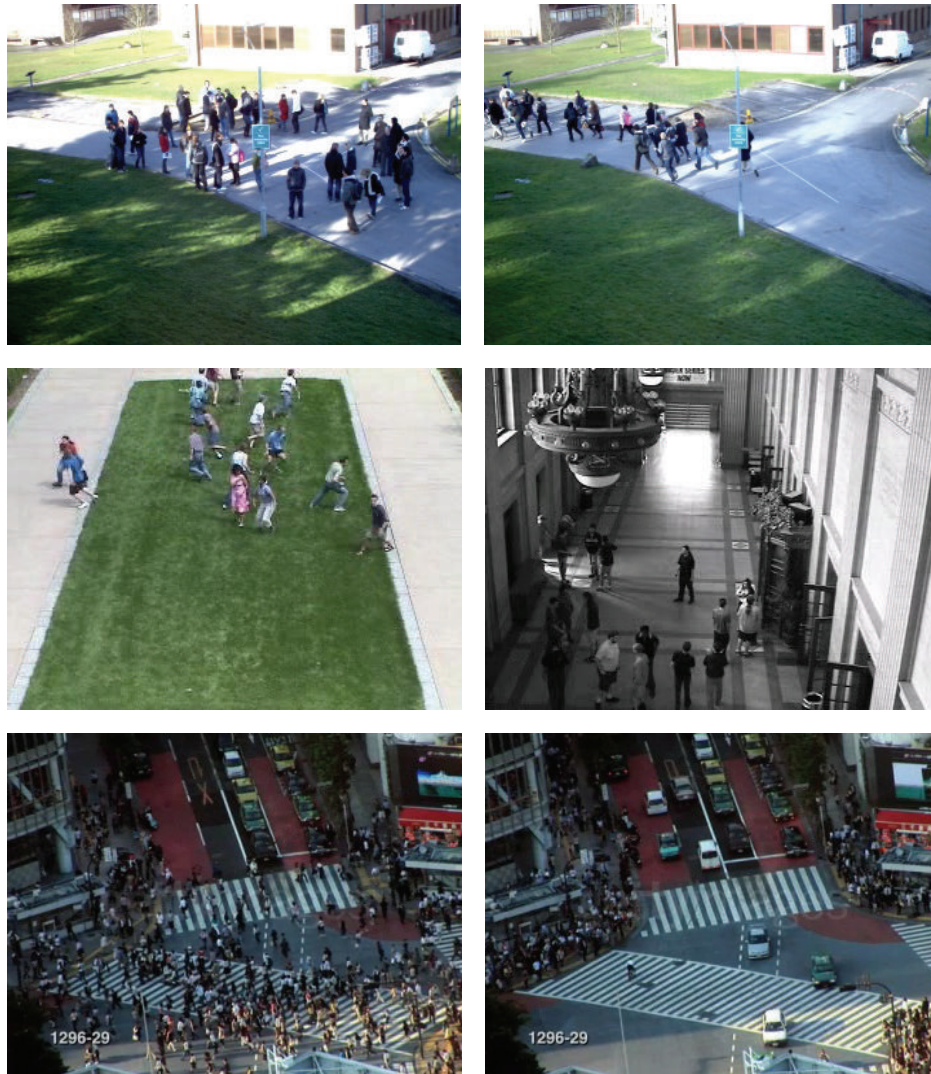FIGURE 5: Comparison of segmentation results by using SRF (blue), proposed method (red), and proposedNC (green).

FIGURE 6: Sample frames from three datasets.

(a) Sample videos from PETS2009

(b) Abnormal detection results of PETS2009

FIGURE 7: Comparison results of abnormal detection for sample videos from PETS2009.

in frame 212 and the two motion patterns of pedestrians in frame 425).

Figure 5 shows the quantitative comparison of the three methods. In our experiment, we manually label the segmented regions frame by frame according to the actual motion scene and then count the correctly segmented regions and incorrectly segmented regions in crowd sequence. Figure 5 demonstrates that our method is inferior to the other two methods in number of incorrect segment motions, which are due to the different label method in [18] and our method. In [18], the segment region is viewed as correct segmentation when difference of direction between the object and the majority of the objects in the region is less than 90 degrees. However, in our experiment we count the segment result according to the actual direction of motion groups, which results in the numbers of incorrect segmentation sharply increased.

5.2. Results of Crowd Abnormal Behavior Detection. In this part, our method is tested on three public datasets from PETS2009 (http://www.cvg.reading.ac.uk/PETS2009/a.html), University of Minnesota (UMN) (http://mha.cs.umn.edu/

proj_events.shtml#crowd), and some video collection from websites [14]. PETS2009 contains different crowd motion patterns, such as walking, running, gathering, and splitting. We select 1066 frames from "S3" sequence as our experiment object, and in this dataset, crowd movement from walking to running is defined as abnormal behavior. UMN includes eleven different kinds of escape events which are shot in three indoor and outdoor scenes. We select 1377 frames from UMN dataset in our experiment, and the abnormal behavior in this dataset is the crowd escape. The web dataset includes crowd videos captured at road intersection, which contains 3071 frames. This crowd video contains two types of motion patterns: waiting and crossing, and pedestrian crossing is considered as abnormal behavior. We also manually label all videos as ground truth. From the three datasets, we randomly select 40% frames for training and the other frames for testing. Figure 6 illustrates the sample frames of the three datasets. The first row is the sampled frames from PETS2009, the second row from UMN, and the last row from web dataset.

In the streak flow phrase, we downsampled the original pixels to fasten the process of experiment. The sample rate

(a) Sample videos from UMN
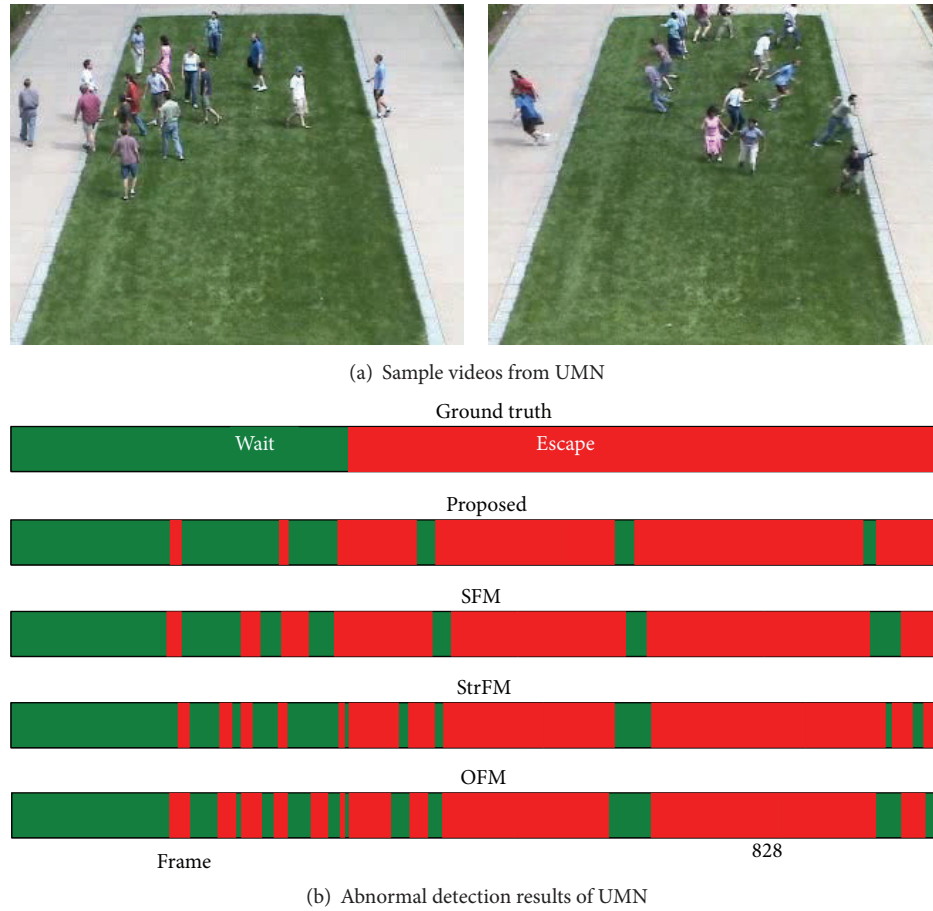

(b) Abnormal detection results of UMN

FIGURE 8: Comparison results of abnormal detection for sample videos from UMN.

of the three datasets is 30%, 50%, and 20%, respectively. In order to utilize the different characteristic of the videos, the parameters in formula (7) are set to $\alpha = 0.2$, $\beta = 0.5$, and $\gamma = 0.3$. After constructing social force graph, we segment the video sequence into different motion groups, and then visual words are generated based on clustering all motion groups. The length of visual words is 32, 64, and 100, respectively. In the LDA phrase, we learned $N_z = 16$ latent topics for abnormal behavior detection. We compare our method with the social force based method [5] (denoted as SFM), the streak flow based method [18] (denoted as StrFM), and the optical flow based method (denoted as OFM). In the process of comparison, we obtain visual words based on corresponding visual features and then train the words by LDA model.

Figures 7–9 show some sample detection results of the three datasets. In each figure, the first picture is first frame and the second picture is abnormal frame of the experimental scene. We then use color bar to represent the detected results of each frame in the crowd sequence. Different color of the bar denotes the status of current frame: green is normal frame and red is the abnormal frame. The first bar indicates the ground truth bar, and the rest of the bars show the results from our proposed method, SFM, StrFM, and OFM, respectively. All in all, these results show that our proposed

method is able to detect the crowd abnormal dynamic, and in most cases, the results of our method outperform the other algorithms.

Figure 7 shows the partial results on running video sequence of PETS2009. Figure 8 illustrates some of the results on escape crowd video of UMN. Figure 9 exhibits part of the results on web video. We can see that the results of the proposed method better conform to the ground truth because we not only consider the temporal characteristic of the motion objects but also exploit the spatial social force relationship between objects. It is observed that, in Figure 7, SFM obtains an even better result compared to our proposed method because the social force is sharply increased in running scene. However, SFM results in more false detection in some crowd scene, for example, crowd walking (Figure 8) and road crossing (Figure 9), because most of the social force in these scenes is nearly zero and thus SFM faces difficulty in depicting this type of crowd motion. OFM is obviously inferior to our method because the optical flow fails to capture the motion in a period of time. In particular, in Figure 9, there is more false detection for road crossing behavior because most of the walking behavior is irregular, but optical flow fails to capture irregular motion pattern. The results of StrFM are better than OFM in the three figures; however, this method still has more false detection compared to our

(a) Sample videos from WEB
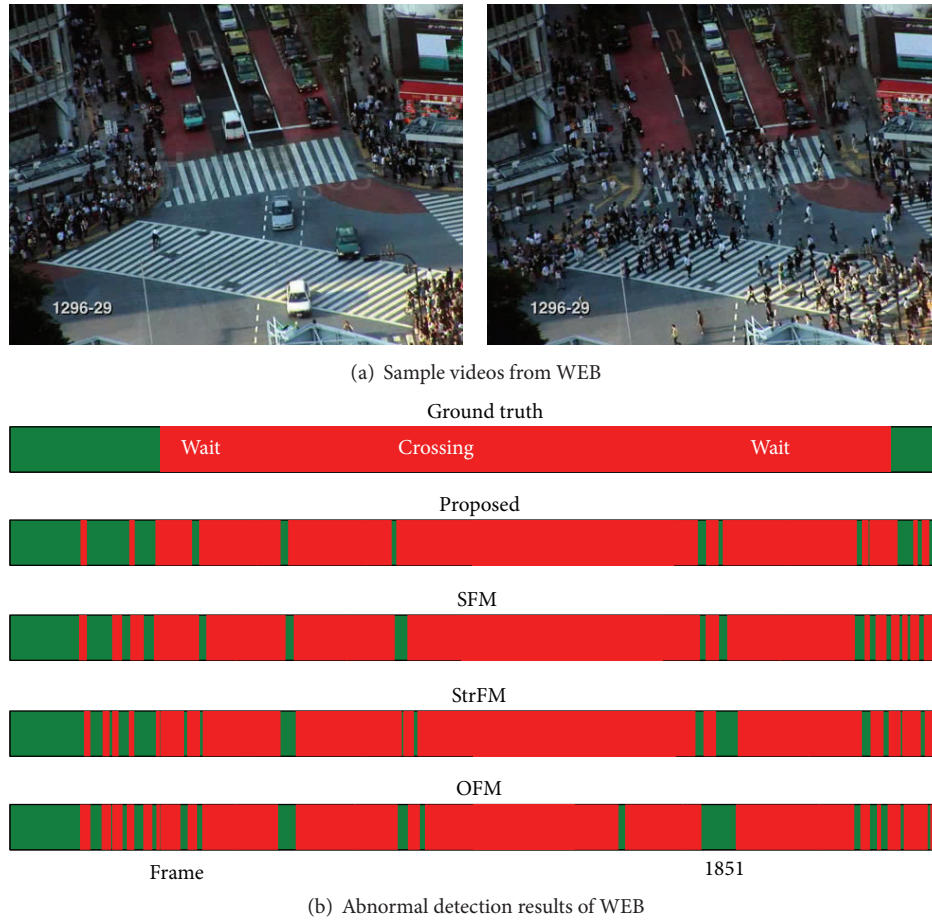

(b) Abnormal detection results of WEB

FIGURE 9: Comparison results of abnormal detection for sample videos from WEB.

TABLE 1: Accuracy comparison of the proposed method with SFM, StrFM, and OFM in crowd abnormal detection for the three datasets.

	Proposed method	SFM	StrFM	OFM
PETS2009	83.33	82.19	73.08	68.52
UMN12	85.10	80.25	79.52	65.31
WEB	96.21	91.23	94.02	95.89
Overall accuracy	88.21	84.79	82.21	76.57

method and SFM. This method attains a comparable result when the motion dynamic is relatively simple (Figure 7), but it undergoes a performance descent when a substantial new flow constantly enters into the scene (Figure 9).

In order to evaluate our method quantitatively, the accuracy results of abnormal detection are shown in Table 1. The results show that our proposed model has improved nearly by 5% accuracy compared with social force model. In particular, in the higher density crowd scene, the social force of particle groups will have a slowly changing ratio; thus, the social force model is not very obvious for normal behaviors. For example, the SFM accuracy is lower than our proposed method and even lower than optical flow method in Table 1.

Figure 10 shows the ROC curves for abnormal behavior detection based on our proposed method against the alternative methods. It is observed that our method provides better results.

6. Conclusion

In this paper, we proposed a novel framework for analyzing motion patterns in crowd video. On the aspect of spatiotemporal feature, we extracted the streak flow to represent the global motion, while for the interaction of crowd objects, we constructed a weighted social force graph to extract particle groups to represent local group motion. Finally, we utilized LDA model to detect crowd abnormal behavior based on the proposed crowd model. The experimental results have shown that our proposed method successfully segments crowd motion and detects crowd abnormal behavior. And the proposed method outperformed the current state-of-the-art methods for crowd motion analysis. As part of our future work, we plan to further study the crowd motion behavior in high density scenes and increase the robustness of our algorithm.

Conflict of Interests

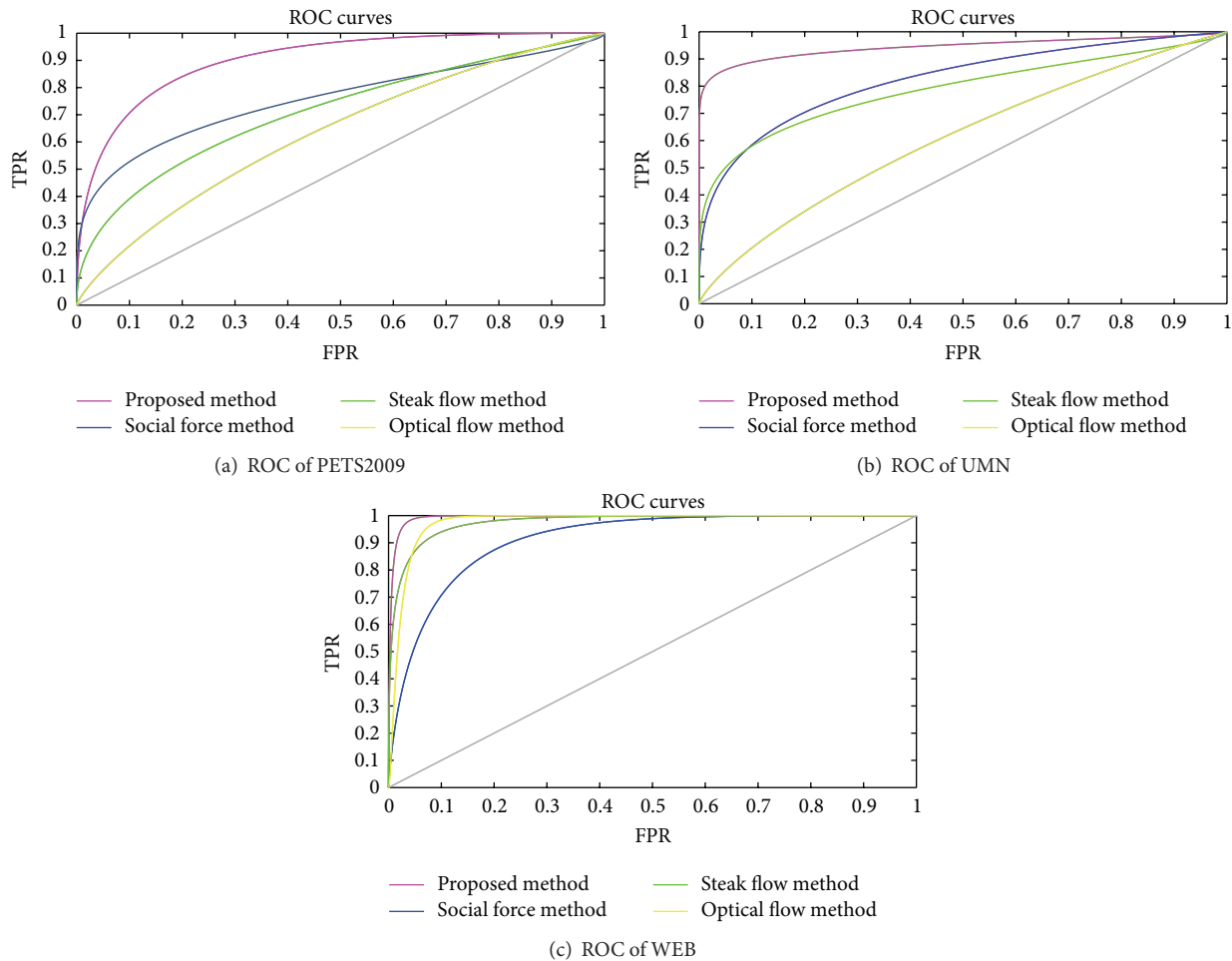The authors declare that there is no conflict of interests regarding the publication of this paper.

(a) ROC of PETS2009

(b) ROC of UMN

(c) ROC of WEB

FIGURE 10: The ROCs for abnormality detection on three datasets.

Acknowledgment

References

[1] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Physical Review E*, vol. 51, no. 5, pp. 4282–4286, 1995.

[2] W. Ge, R. T. Collins, and R. B. Ruback, "Vision-based analysis of small groups in pedestrian crowds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 5, pp. 1003–1016, 2012.

[3] M. Rodriguez, S. Ali, and T. Kanade, "Tracking in unstructured crowded scenes," in *Proceedings of the 12th International Conference on Computer Vision (ICCV '09)*, pp. 1389–1396, IEEE, Kyoto, Japan, October 2009.

[4] L. Kratz and K. Nishino, "Tracking pedestrians using local spatio-temporal motion patterns in extremely crowded scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 5, pp. 987–1002, 2012.

[5] R. Mehran, A. O. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 935–942, Miami, Fla, USA, June 2009.

[6] R. Raghavendra, A. Del Bue, M. Cristani, and V. Murino, "Optimizing interaction force for global anomaly detection in crowded scenes," in *Proceedings of the IEEE 12th International Conference on Computer Vision (ICCV '11)*, pp. 136–143, Barcelona, Spain, November 2011.

[7] R. Li and R. Chellappa, "Group motion segmentation using a spatio-temporal driving force model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 2038–2045, IEEE, San Francisco, Calif, USA, June 2010.

[8] B. Zhou, X. Tang, H. Zhang, and X. Wang, "Measuring crowd collectiveness," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 8, pp. 1586–1599, 2014.

[9] G. T. Bae, S. Y. Kwak, and H. R. Byun, "Motion pattern analysis using partial trajectories for abnormal movement detection in crowded scenes," *Electronics Letters*, vol. 49, no. 3, pp. 186–187, 2013.

[10] C. Tomasi and T. Kanade, "Detection and tracking of point features," Tech. Rep. CMU-CS-91-132, Carnegie Mellon University, Pittsburgh, Pa, USA, 1991.

[11] E. L. Andrade, S. Blunsden, and R. B. Fisher, "Modelling crowd scenes for event detection," in *Proceedings of the 18th*

International Conference on Pattern Recognition (ICPR '06), vol. 1, pp. 175–178, IEEE, Hong Kong, August 2006.

[12] T. Li, H. Chang, M. Wang, B. Ni, R. Hong, and S. Yan, "Crowded scene analysis: a survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 3, pp. 367–386, 2015.

[13] N. Ihaddadene and C. Djeraba, "Real-time crowd motion analysis," in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR '08)*, pp. 1–4, December 2008.

[14] S. Ali and M. Shah, "A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–7, Minneapolis, MN, USA, June 2007.

[15] X. Pan, C. S. Han, K. Dauber, and K. H. Law, "Human and social behavior in computational modeling and analysis of egress," *Automation in Construction*, vol. 15, no. 4, pp. 448–461, 2006.

[16] A. Borrmann, A. Kneidl, G. Köster, S. Ruzika, and M. Thiemann, "Bidirectional coupling of macroscopic and microscopic pedestrian evacuation models," *Safety Science*, vol. 50, no. 8, pp. 1695–1703, 2012.

[17] A. Bera and D. Manocha, "REACH—realtime crowd tracking using a hybrid motion model," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '15)*, pp. 740–747, IEEE, Seattle, Wash, USA, May 2015.

[18] R. Mehran, B. E. Moore, and M. Shah, "A streakline representation of flow in crowded scenes," in *Computer Vision— ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5–11, 2010, Proceedings, Part III*, vol. 6313 of *Lecture Notes in Computer Science*, pp. 439–452, Springer, Berlin, Germany, 2010.

[19] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[20] M. Mancas, N. R. Riche, J. Leroy, and B. Gosselin, "Abnormal motion selection in crowds using bottom-up saliency," in *Proceedings of the 18th IEEE International Conference on Image Processing (ICIP '11)*, pp. 229–232, IEEE, Brussels, Belgium, September 2011.

[21] X. Zhu, J. Liu, J. Wang, W. Fu, and H. Lu, "Weighted interaction force estimation for abnormality detection in crowd scenes," in *Proceedings of the 11th Asian Conference on Computer Vision (ACCV '12)*, pp. 507–518, Daejeon, Republic of Korea, November 2012.

[22] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.