


Article

# Detection of Salient Crowd Motion Based on Repulsive Force Network and Direction Entropy

Xuguang Zhang <sup>1,\*</sup>, Dujun Lin <sup>1</sup>, Juan Zheng <sup>2</sup>, Xianghong Tang <sup>1</sup>, Yinfeng Fang <sup>1</sup> and Hui Yu <sup>3,\*</sup>

<sup>1</sup> School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China; 172080035@hdu.edu.cn (D.L.); tangxh@hdu.edu.cn (X.T.); yinfeng.fang@hdu.edu.cn (Y.F.)

<sup>2</sup> School of Electrical Engineering, Shandong Huayu University of Technology, Dezhou 253034, China; zjj@sdhyxy.com

<sup>3</sup> School of Creative Technologies, University of Portsmouth, Portsmouth PO1 2DJ, UK

\* Correspondence: zhangxg@hdu.edu.cn (X.Z.); hui.yu@port.ac.uk (H.Y.); Tel.: +44(0)23-9284-5470 (H.Y.)

Received: 8 May 2019; Accepted: 18 June 2019; Published: 20 June 2019



**Abstract:** This paper proposes a method for salient crowd motion detection based on direction entropy and a repulsive force network. This work focuses on how to effectively detect salient regions in crowd movement through calculating the crowd vector field and constructing the weighted network using the repulsive force. The interaction force between two particles calculated by the repulsive force formula is used to determine the relationship between these two particles. The network node strength is used as a feature parameter to construct a two-dimensional feature matrix. Furthermore, the entropy of the velocity vector direction is calculated to describe the instability of the crowd movement. Finally, the feature matrix of the repulsive force network and direction entropy are integrated together to detect the salient crowd motion. Experimental results and comparison show that the proposed method can efficiently detect the salient crowd motion.

**Keywords:** crowd behavior analysis; salient crowd motion detection; repulsive force; direction entropy; node strength

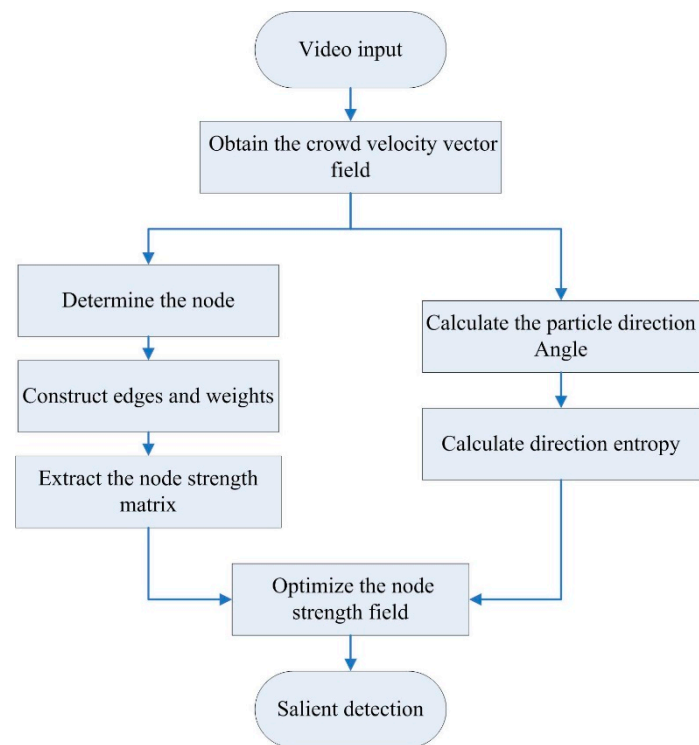
## 1. Introduction

Video surveillance plays an important role in monitoring crowd safety, which is one of the key concerns in our daily life. Since the traditional human-computer interaction between video surveillance and crowd safety is time-consuming and labor-intensive, intelligent video surveillance issues such as target tracking, target detection and crowd analysis have become popular research topics. Crowd motion detection and analysis are essential for crowd behavior understanding [1,2]. It is thus very important to detect the salient motion in the crowd to monitor any potential threats or even damage to social safety. Salient motion has been defined as motion that is likely to result from a typical surveillance target as opposed to other distracting motions [3]. According to this definition, salient crowd motion usually indicates areas that are inconsistent with the mainstream pedestrians' movement. For video surveillance, these areas deserve more attention.

In recent years, due to the rapid development of computer vision technologies, progress has been made in detection of crowd saliency. For example, Lim et al. [4,5] proposed a method for automatically detecting a salient region using time variation of a crowd scene flow field by detecting the fluid activity in a given scene and detecting saliency with a minimum amount of observation region. Some methods for detecting globally salient motion regions for spectral singularity analysis of motion regions in video [6,7] have been also presented. Zhou et al. [8] studied the invariance of coherent neighbors as coherent motion priors, and proposed an effective clustering technique to detect crowd saliency.

Solmaz et al. [9] overlaid the scene from the particle grid of the dynamic system defined by the optical flow, and proposed a method to identify the behavior of five people in the visual scene through time integration. Zhang et al. [10] surveyed physics-based methods for crowd video analysis and sorted out the existing public database of crowd video analysis. Although many methods have shown good performance in crowd salient motion detection, the internal mechanism of crowd movement still needs to be explored. The pattern of crowd movement depends on both individual movement and interaction between individuals. It is of great value to explore a method to describe individual interaction and apply it to crowd salient motion detection.

In this paper, we propose a salient crowd motion detection method based on a direction entropy and a repulsive force network. The optical flow is first obtained using the pyramid-based Lucas-Kanade optical flow algorithm. Then, the weighted network is constructed by the repulsive force and the node strength matrix is obtained by using the node degree as the characteristic parameter. Finally, the particle motion direction entropy is used to optimize the node strength matrix and to detect salient movements of the crowds. The framework of the proposed method is shown in Figure 1. A motion vector field is established by giving each pixel a velocity vector in each image through the Pyramid Lucas-Kanade optical flow algorithm. Each vector in the crowd vector field is treated as a moving micro-particle. In order to build a complex network model, we regard each particle and the relationship between two particles as node and edge in the network, respectively. In order to show whether there is a connection between two particle nodes, we use the interaction force to construct the network. After calculating by optical flow method, the position and velocity parameters of each particle can be determined. Whether there is an edge between the nodes depends on the value of repulsive force between these nodes. The repulsive force can be described by the inertial centrifugal force. The value of the inertial centrifugal force is the weight of the edge and a velocity vector node can be selected accordingly. In the neighborhood of the node, the relevancy between the two velocity vectors is taken as a condition to determine the relevancy between the corresponding nodes.



**Figure 1.** Framework of salient crowd motion detection based on repulsive force network and direction entropy.

A weighted crowd network model is constructed to obtain the adjacency matrix representing the crowd motion information. In order to obtain a complete boundary of salient motion region, the velocity field is reversed and the repulsive force between particles is calculated repeatedly to construct the repulsive force network model. Then, the edge and weight are constructed by the repulsive force model, and the results of the superposition are taken as a construction step. Once all nodes are traversed, the strength of each crowd-weighted network node is extracted as a characteristic parameter to construct the strength matrix of the nodes. By calculating the direction of the velocity entropy of each node in the neighborhood, we can obtain the direction entropy matrix of the node. Then, the normalized direction entropy matrix and the strength matrix of the node are used to further optimize the strength matrix of the node. Once the node strength matrix is obtained, the salient region in crowd movement can be detected.

## 2. Calculation of Crowd Velocity Vector Field

To calculate the velocity vector field, the crowd video is decomposed into image sequences. Then, each pixel of the image is given a velocity vector calculated using an optical flow algorithm. A motion vector field is thus established. In this paper, considering the spatio-temporal information in motion detection [11], we adopt an improved algorithm based on Lucas-Kanade optical flow algorithm [12] for this task, namely pyramid optical flow algorithm [13].

Lucas-Kanade optical flow, in the process of moving the picture, assumes that a pixel  $(x, y)$  on the image has a brightness of  $I(x, y, t)$  at time  $t$ . After a small time interval of  $\Delta t$ , the brightness of the point becomes  $I(x + \Delta x, y + \Delta y, t + \Delta t)$ . The Taylor formula is used to expand and when  $\Delta t$  is small enough to approach zero:

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t \quad (1)$$

The optical flow constraint equation can be obtained from the brightness constant:

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = \frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} = I_x u + I_y v + I_t = 0 \quad (2)$$

According to the uniformity of optical flow, we can establish the optical flow equations:

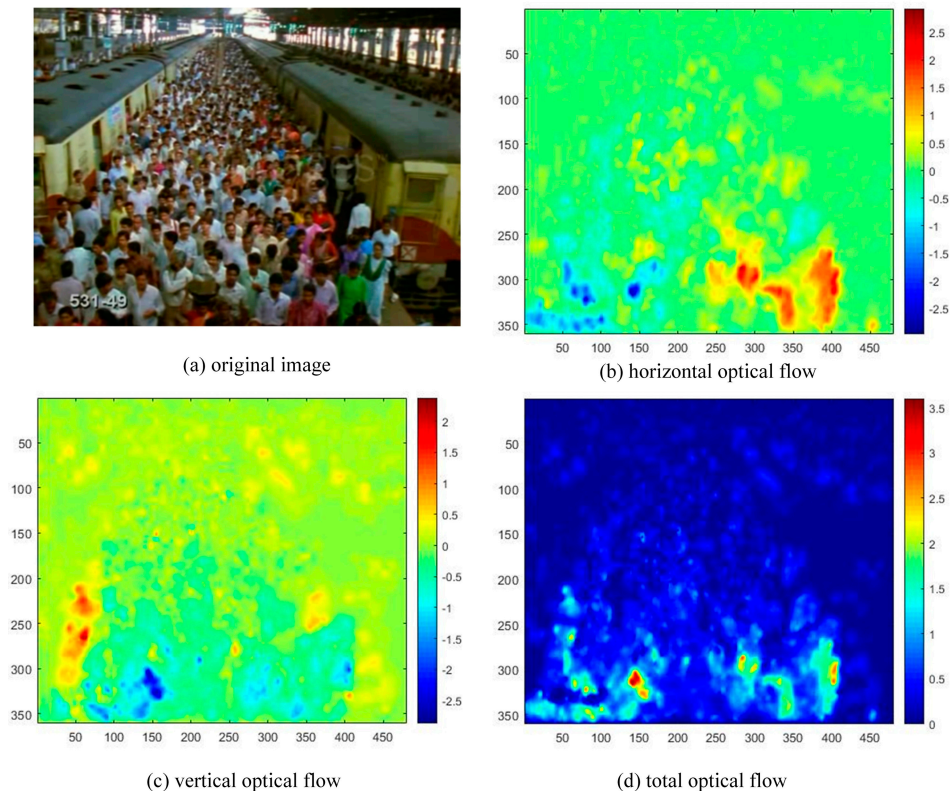
$$\begin{aligned} I_{x1}u + I_{y1}v + I_{t1} &= 0 \\ I_{x2}u + I_{y2}v + I_{t2} &= 0 \\ &\vdots \\ I_{xnu} + I_{ynv} + I_{tn} &= 0 \end{aligned} \quad (3)$$

Then use the least square method to gain the Lucas-Kanade optical flow, where  $u$  is the horizontal velocity and  $v$  is the vertical velocity:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n I_{ix}^2 & \sum_{i=1}^n I_{ix}I_{iy} \\ \sum_{i=1}^n I_{ix}I_{iy} & \sum_{i=1}^n I_{iy}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_{i=1}^n I_{ix}I_t \\ -\sum_{i=1}^n I_{iy}I_t \end{bmatrix} \quad (4)$$

The basic ideas of Lucas Kanade optical flow algorithm are mainly based on three assumptions: (1) constant brightness; (2) time continuous or movement is "small movement"; (3) spatial consistency. If an object is moving fast, the second assumption is not fully satisfied. The value calculated by traditional Lucas-Kanade optical flow will have a larger deviation. Pyramid optical flow algorithm reduces the offset of the target motion by reducing the image layer by layer, which satisfies the hypothesis of optical flow calculation better and weakens the influence of fast target motion. In this

paper, the crowd velocity field  $Q$  is obtained by the pyramid optical flow algorithm. All the velocity values in the horizontal direction and vertical direction are rounded up. The velocity vector field calculated using pyramid Lucas-Kanade optical flow algorithm for a crowd scene is shown in Figure 2.



**Figure 2.** Crowd optical flow field of the sampled frame.

### 3. Construction of Repulsive Force Network

#### 3.1. Establishment of a Network Node

Complex network is a useful tool for describing a complex system. Each element in the system unit is regarded as a node, and the relationship between elements is regarded as a connection. A complex system can be represented as a network [14]. The crowd velocity vector field can be described as a complex network, in which each velocity vector is a node, and the relationship between the velocity vectors is connected. If the properties of the velocity vectors are measured separately, information stored in the velocity vector cross-correlation cannot be obtained, because the correlation of velocity vectors carries more information than the nature of each velocity [15].

We use the interaction force between particles to construct the network [16,17]. After applying the optical flow method, each vector in the obtained crowd vector field is regarded as a moving microscopic particle. The position and velocity parameters of each particle can be then determined. In our crowd complex network, each particle is treated as a node, and the interaction force between two particles is treated as an edge in the network. Whether there is an edge between the nodes depends on the repulsive force between the nodes, the repulsive force can be described by the inertial centrifugal force, and the value of the inertial centrifugal force is the weight of the edge. A weighted undirected network  $G^w$  node set  $Q = \{q_1, q_2, \dots, q_n\}$  can be generated, where  $n$  is the total number of nodes. The number of network nodes is equal to the number of particles in the crowd velocity field.

### 3.2. Establishing the Network Edges Using Repulsive Force Model

In the whole particle field, the size and direction of particle velocity are instantaneous, and the motion of the next times is random. If the moving particle is assumed as an agent, there is a possibility of interaction and collision between particles in motion. Imagine that each particle is an agent. In order to avoid collision between agents, an agent adds a repulsive force element to prevent them from colliding with each other. This repulsive force can be described by inertial centrifugal force [18]. For a given crowds particle field  $Q (M, N)$  in the column  $N$  and row  $M$ , selecting a particle  $q_{x_0y_0}$  as the node, constructing a two-dimensional neighborhood  $\delta$ , the size is  $(x_0 \pm \varepsilon, y_0 \pm \varepsilon)$ . In this region, the connection between  $q_{x_0y_0}$  and other nodes  $q_{xy}(x \neq x_0, y \neq y_0)$  can be described as  $e(q_{x_0y_0}, q_{xy})$ . Whether this connection exists is determined by the following formula:

$$e(q_{x_0y_0}, q_{xy}) \begin{cases} \exists, & \vec{F}_{ij} \neq 0, q_{xy} \in \delta \\ \nexists, & \text{otherwise} \end{cases} \quad (5)$$

The formula for calculating the inertial centrifugal force is as follow:

$$\vec{F}_{ij} = -m_i k_{ij} \frac{v_{ij}^2}{dist_{ij}} \vec{e}_{ij} \quad (6)$$

where  $\vec{e}_{ij}$  is the direction vector and  $m_i$  is the mass of particle  $q_i$ . In this paper, the mass of all particles is set as unit 1.  $v_{ij}$  is the relative velocity of two particles,  $dist_{ij}$  is the distance between two particles,  $k_{ij}$  is a coefficient, the calculation of  $v_{ij}$  and  $k_{ij}$  is determined by the following formula:

$$v_{ij} = \begin{cases} (\vec{v}_i - \vec{v}_j) \cdot \vec{e}_{ij}, & (\vec{v}_i - \vec{v}_j) \cdot \vec{e}_{ij} > 0 \\ 0, & \text{others} \end{cases} \quad (7)$$

$$k_{ij} = \begin{cases} (\vec{v}_i \cdot \vec{e}_{ij}) / v_i, & \vec{v}_i \cdot \vec{e}_{ij} > 0, v_i \neq 0 \\ 0, & \text{others} \end{cases} \quad (8)$$

Then, we can obtain the joint weight, which can be expressed by the magnitude of the repulsive force:

$$We = \left| \vec{F}_{ij} \right| \quad (9)$$

According to the repulsive force formula, if the particle moves away from the affected particle, the repulsive force will be very low. As shown in Figure 3, the arrow represents the moving optical flow, and the blue line represents the repulsive force generated. Figure 3a is the schematic diagram of the repulsive force in the original direction, and Figure 3b is the schematic diagram of the repulsive force in the opposite direction. Thus, for some application of salient region detection, only half of the boundary can be detected.

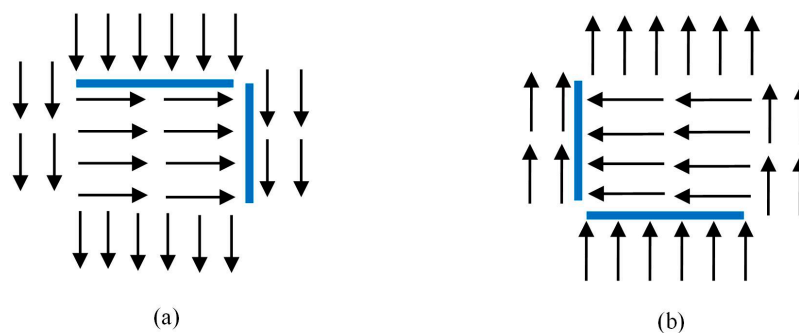
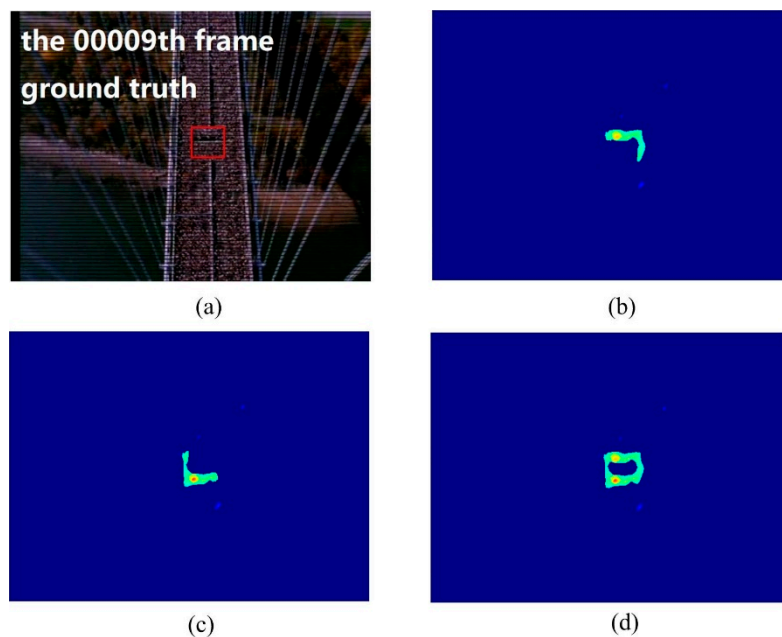


Figure 3. Schematic diagram of repulsive force: (a) original motion flow; (b) motion flow reversed.

In order to get a complete boundary, the velocity field is reversed and the repulsive force between particles is calculated repeatedly. Thus, the repulsive force model can be used to construct the edge and weight of the repulsive force network. The results of the superposition of the two are taken as a construction step. Figure 4 shows an example. If we construct the repulsive force network for the optical flow field of the original video sequence, only half of the boundary can be obtained. If we construct the repulsive force network again after reversing the optical flow field, the other half of the boundary can be obtained.



**Figure 4.** Expressing the effect of optical flow reversal: (a) original sample frame; (b) detection result using original optical flow; (c) detection result using reversed optical flow; (d) detection result after the integration of optical flow.

The two-dimensional crowd velocity field is transformed into weighted undirected network model  $G^w(Q, E, We)$  by repeating the above steps for each node. The corresponding weighted undirected network node is set as  $Q = \{q_1, q_2, \dots, q_n\}$  and the network edge is set as  $E = \{e_1, e_2, \dots, e_m\}$ . In the crowds weighted network model, the connection between nodes and the degree of connection between nodes can be expressed by the following adjacency matrix:

$$A = \begin{bmatrix} \left| \vec{F}_{11} \right| & \left| \vec{F}_{12} \right| & \dots & \left| \vec{F}_{1n} \right| \\ \vdots & \vdots & \ddots & \vdots \\ \left| \vec{F}_{n1} \right| & \left| \vec{F}_{n2} \right| & \dots & \left| \vec{F}_{nn} \right| \end{bmatrix}, \quad (10)$$

### 3.3. Calculation of Node Strength

Statistical characteristic parameters of a network can be used to represent the characteristics of a network, such as node degree, average path length, clustering coefficient. In this paper, node strength is chosen to describe the characteristics of the crowd complex network. In the complex network model, node strength is the generalization of node degree, which integrates the strength between edges and nodes [19,20]. From the adjacency matrix, the node strength  $s(q_i)$  of node  $q_i$  can be expressed as follows:

$$s(q_i) = \sum_{n=1}^j \left| \vec{F}_{ij} \right|, \quad (11)$$

After calculating each point in the crowd velocity field, we can get the node strength of all nodes. The node strength field  $S(M, N)$  is also a two-dimensional matrix containing  $M$  rows and  $N$  columns. There is also a one-to-one correspondence between the node strength field and the crowd speed field:

$$S = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ S_{M1} & S_{M2} & \dots & S_{MN} \end{bmatrix}, \quad (12)$$

In order to facilitate the node strength field optimization operation in later stage, the node strength field is normalized as follows:

$$S' = \frac{S - S_{min}}{S_{max} - S_{min}} \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ S_{M1} & S_{M2} & \dots & S_{MN} \end{bmatrix}, \quad (13)$$

where  $S_{max}$  and  $S_{min}$  are the maximum and minimum values of the nodes in all node strengths.

#### 4. Optimizing Node Strength Field Using Direction Entropy

##### 4.1. Establishment of Vector Direction Entropy Matrix

For a crowd motion field  $Q(M, N)$  of the  $M$  row and  $N$  column, one particle  $q_{x_0y_0}$  is selected, and thus, the direction angle of particle motion is divided into eight directions at 45 degrees interval. The calculation of velocity direction angle and direction grade is determined by the following formula:

$$\theta = \arctan \frac{q_{y_0}}{q_{x_0}}, \quad (14)$$

$$d = \begin{cases} 1 & 0 \leq \theta < \frac{\pi}{4} \\ \vdots & \vdots \\ 8 & \frac{7\pi}{4} \leq \theta < 2\pi \end{cases}, \quad (15)$$

Choose a two-dimensional neighborhood  $\delta$  with the same edge and weight as the repulsive force model with the size of  $(x_0 \pm \varepsilon, y_0 \pm \varepsilon)$ . For a sub-image region, because of the different motion forms of particles, the direction of particle motion is uncertain at eight angles. Shannon entropy is a classical method to measure the uncertainty of information, and is the basis of communication science [21–23]. In this paper, Shannon entropy is used to measure the uncertainty of particle motion direction. In this paper, we employ Shannon entropy to describe the chaotic degree of crowd motion. In a neighborhood  $\delta$ , each particle can be calculated by direction rank formula to get a direction rank  $d$ . Each direction rank occupies a certain probability  $p_i$  in all direction ranks. According to the definition of Shannon entropy [21] and [23], we can assign the velocity direction entropy between the central particle  $q_{x_0y_0}$  and other particles  $q_{xy}(x \neq x_0, y \neq y_0)$  neighboring the central particle. The calculation is determined by the following formula:

$$H_{x_0y_0} = -\sum_{i=1}^n p_i \log p_i, \quad n = \varepsilon^2, \quad (16)$$

For each position, in the crowd particle field  $Q(M, N)$ , the entropy can be calculated by repeating the steps mentioned above. Therefore, the direction entropy of each particle in the crowd particle field can be obtained. The two-dimensional crowd velocity vector field can be transformed into a particle direction entropy matrix:

$$H = \begin{bmatrix} H_{11} & H_{12} & \dots & H_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{M1} & H_{M2} & \dots & H_{MN} \end{bmatrix}, \quad (17)$$

where,  $H_{11}, H_{12} \dots \dots H_{MN}$  is the entropy at the corresponding position of the crowd particle field. In order to facilitate the node strength field optimization operation in later stage, the direction entropy matrix is normalized as follows:

$$H' = \frac{H - H_{min}}{H_{max} - H_{min}} \begin{bmatrix} H_{11} & H_{12} & \dots & H_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{M1} & H_{M2} & \dots & H_{MN} \end{bmatrix}, \quad (18)$$

$H_{max}$  and  $H_{min}$  are the maximum and minimum values in the entropy matrix for all directions.

#### 4.2. Optimizing the Node Strength Field

The direction entropy matrix of crowd movement can describe the degree of changes in the direction of movement of the nodes. Furthermore, the strength field of the repulsive force node describes the degree of repulsion of each node and the surrounding nodes. In order to reduce the noise caused by other interference motion, this paper combines these two kinds of model to optimize the node strength field. It is very important to choose an effective way to integrate these two features, e.g., node strength and entropy. There are many ways to integrate features, such as multiplication and addition. For the application of salient crowd motion detection, the way of feature fusion requires significant expression of specific crowd motion regions and adaptation to the changes of scene. We analyzed the feature of node strength and entropy. The saliency region can be detected by combining the two features by multiplying or add. However, the saliency region obtained by addition is more effective. Because the range of the two features is quite different and there are great changes in different scenarios, it is difficult to determine the combined weights. Therefore, this paper applies a normalized processing of the two features before adding the two features together. Although there are differences in dimension between them, as a normalized feature, it works well when integrating them at the application level.

The direction entropy matrix of crowd motion is in one-to-one correspondence with the strength field of nodes; thus, we have made a comparison according to the following formulas:

$$P_{ij} = \begin{cases} S_{ij}' + H_{ij}' & S_{ij}' \neq 0, H_{ij}' \neq 0 \\ 0 & \text{others} \end{cases}, \quad (19)$$

The optimized node strength field is:

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ P_{M1} & P_{M2} & \dots & P_{MN} \end{bmatrix}, \quad (20)$$

Then, for normalizing the optimized node strength field, the specific calculation formula is as follows:

$$P' = \frac{P - P_{min}}{P_{max} - P_{min}} \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ P_{M1} & P_{M2} & \dots & P_{MN} \end{bmatrix}, \quad (21)$$

After normalizing the strength field of the nodes, we smoothed the node strength field with a  $3 \times 3$  mean filter template. It can eliminate the negative effects of the node strength caused by too high or too low values on the experimental results. In order to intuitively describe and observe the value of



node strength, we use a pseudo-color image display method to visualize node strength. Pseudo-color image shows the pixel value corresponding to the node strength value. In a crowd scene, it is obvious that the node pixel values in salient regions are higher than those in other regions.

## 5. Experimental Results and Analysis

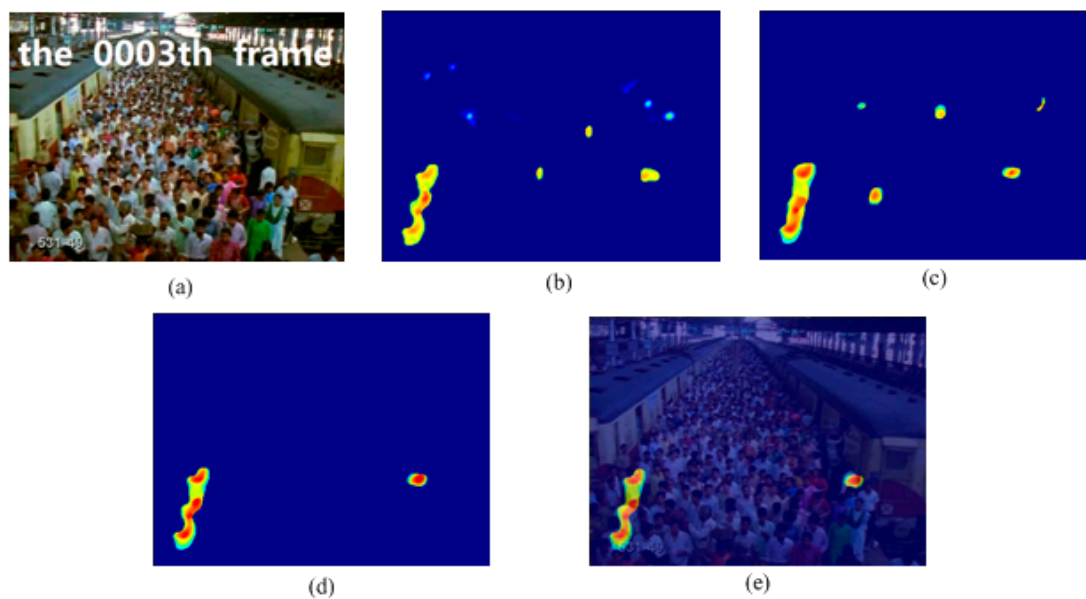
In our experiments, we tested three crowded scene video sequences from Crowd Saliency dataset [5] and a video sequence in [24] to show the performance of the proposed method. Retrograde and instability regions of a crowd were detected in the experiment. For different crowded scenes, the scale of the velocity field  $Q(M,N)$  and the parameters  $\varepsilon$  (the size of neighborhood) in the experiment are shown in Table 1. The proposed method is effective for images used in this experiment, which do not have a high resolution. If it is used to deal with high resolution images, there are two ways to processing the data. One is to reduce the high-resolution image using interval sampling and local mean, and the other is to process optical flow data by interval sampling.

**Table 1.** Different scenes and parameter values.

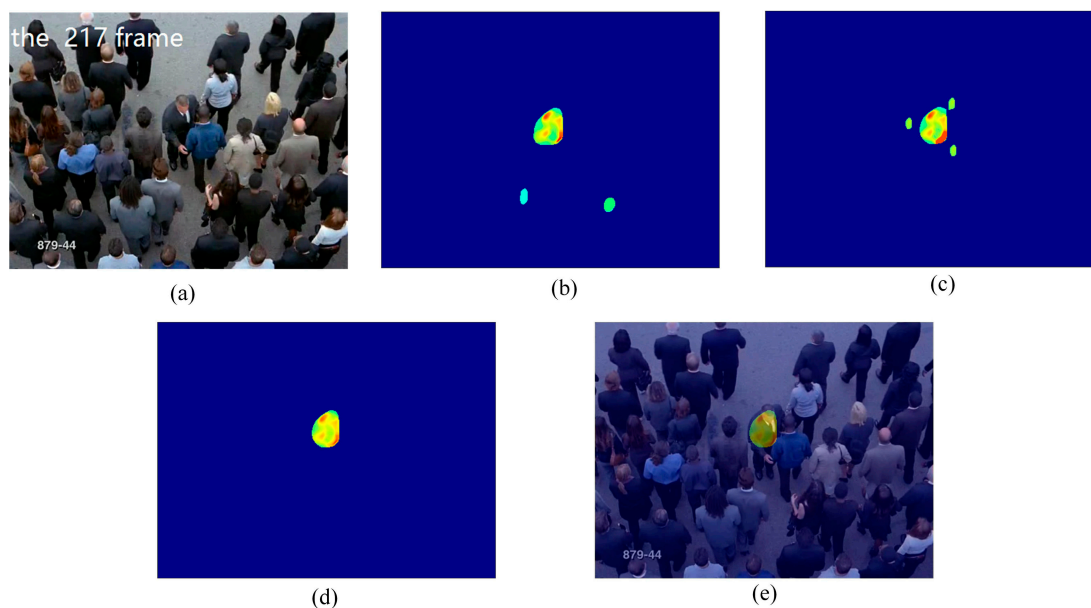
Crowded Scenes	Symbol of Parameter	The Value
Train station scene in Figure 5	$\varepsilon$	13
	$M \times N$	$480 \times 360$
Single retrograde scene in Figure 6	$\varepsilon$	15
	$M \times N$	$480 \times 360$
Marathon scene in Figure 7	$\varepsilon$	11
	$M \times N$	$640 \times 480$
Pilgrimage scene in Figure 8	$\varepsilon$	15
	$M \times N$	$640 \times 480$

### 5.1. Crowd Retrograde Behavior Detection

In this experiment, we used the train station scene and the single retrograde scene to show the salient detection for retrograde behavior. As shown in Figures 5 and 6, some pedestrians do not conform to the flow of the mainstream crowd, hence, a retrograde motion was formed instead. The particles will thus have a larger repulsion force and direction entropy in this region. This proposed method can effectively detect human retrograde movement. Figures 5a and 6a shows the original video frame. The node strength field calculated from the repulsive force network is shown in Figures 5b and 6b. It can be clearly seen that the regions with high node strength represents the retrograde motion. However, there are still some disturbances. As shown in Figures 5c and 6c, though the entropy value of the retrograde region is large, there are still some noise regions. Fortunately, the disturbance regions detected by node strength and direction entropy are different. Therefore, we can optimize the saliency detection results by integrating node strength and direction entropy, as is shown in Figures 5d and 6d. In order to illustrate the detection performance, we overlap the saliency detection results with the original video frames in Figures 5e and 6e. Experiments show that our method can detect pedestrians who even move oppositely to the flow of mainstream crowd.



**Figure 5.** Retrograde motion detection in train station scene: (a) input frame; (b) node strength field of repulsive force network; (c) detection result using direction entropy; (d) salient region detection after optimized; (e) overlap the salient region with input frame.



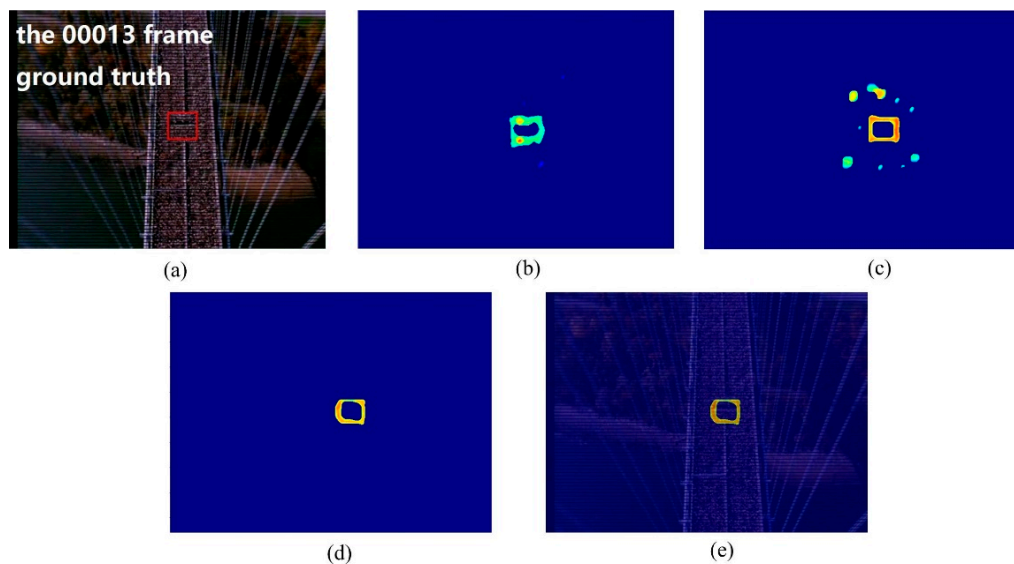
**Figure 6.** Retrograde motion detection in single retrograde scene: (a) input frame; (b) node strength field of repulsive force network; (c) detection result using direction entropy; (d) salient region detection after optimized; (e) overlap the salient region with input frame.

## 5.2. Crowd Motion Instability Region Detection

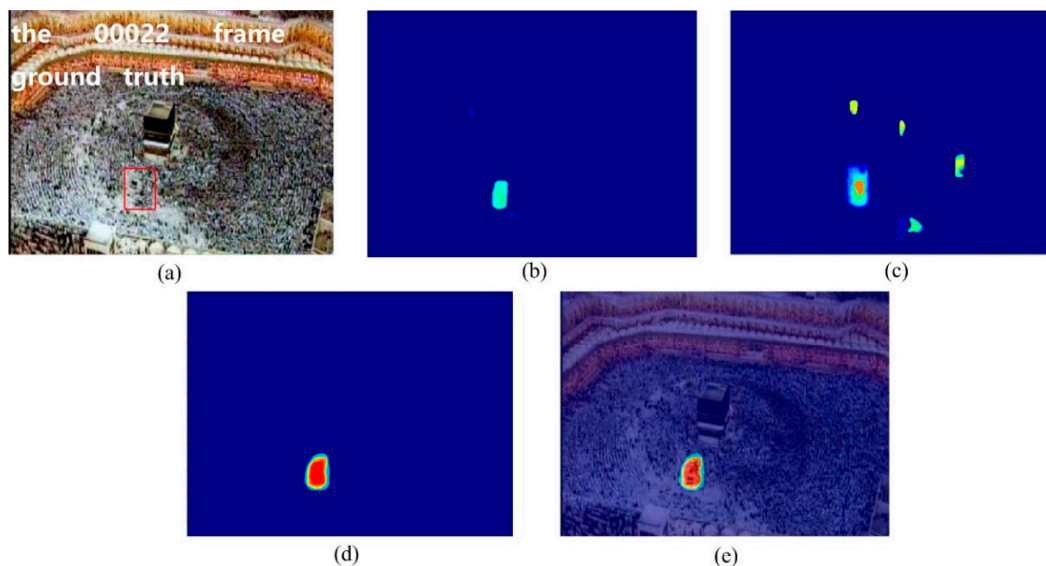
In the crowd surveillance system, the instability area of crowd movement often deserves attention. In this experiment, we used two scenes, including the marathon scene (Figure 7) and the pilgrimage scene (Figure 8) to show the performance of the proposed method for detecting the instability crowd motion.

The sample frames for the two scenarios are shown in Figures 7a and 8a. There are instability motion regions (some pedestrians are different from the mainstream crowd) in these two crowds. The results of node strength fields of two scenes are shown in Figures 7b and 8b, respectively. We can see that the node strength of the repulsive force model is larger in instability motion regions.

The direction entropy fields of two scenarios are shown in Figures 7c and 8c, respectively. The entropy values of the instability region are clearly large. However, there is some noise in the unstable region detected by any single method. After integrating these two methods of node strength and direction entropy, the saliency detection results are optimized and the interference areas are effectively removed, which can be seen in Figures 7d and 8d. Figures 7e and 8e show saliency detection results after overlapping with the original video frame. Experimental results show that the proposed method can detect the salient crowd instability motion in large-scale crowded scenes.



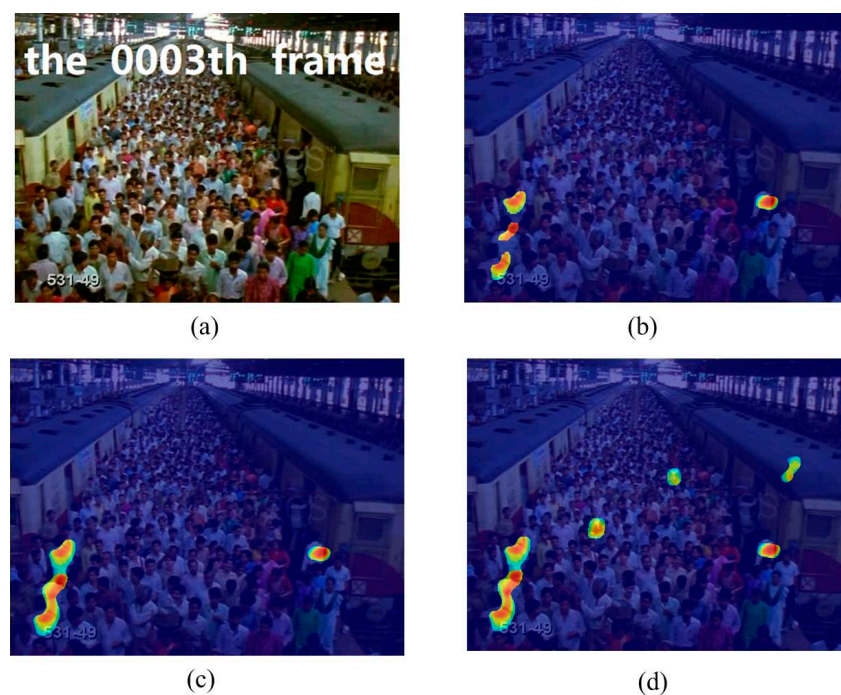
**Figure 7.** Salient crowd instability motion detection in marathon scene: (a) original video frame and ground true (red box); (b) node strength field of repulsive force network; (c) detection result using direction entropy; (d) salient region detection after optimized; (e) overlap the salient region with original video frame.



**Figure 8.** Salient crowd instability motion detection in pilgrimage scene: (a) original video frame and ground true (red box); (b) node strength field of repulsive force network; (c) detection result using direction entropy; (d) salient region detection after optimized; (e) overlap the salient region with original video frame.

### 5.3. Detection Results Using Different Neighborhood Size

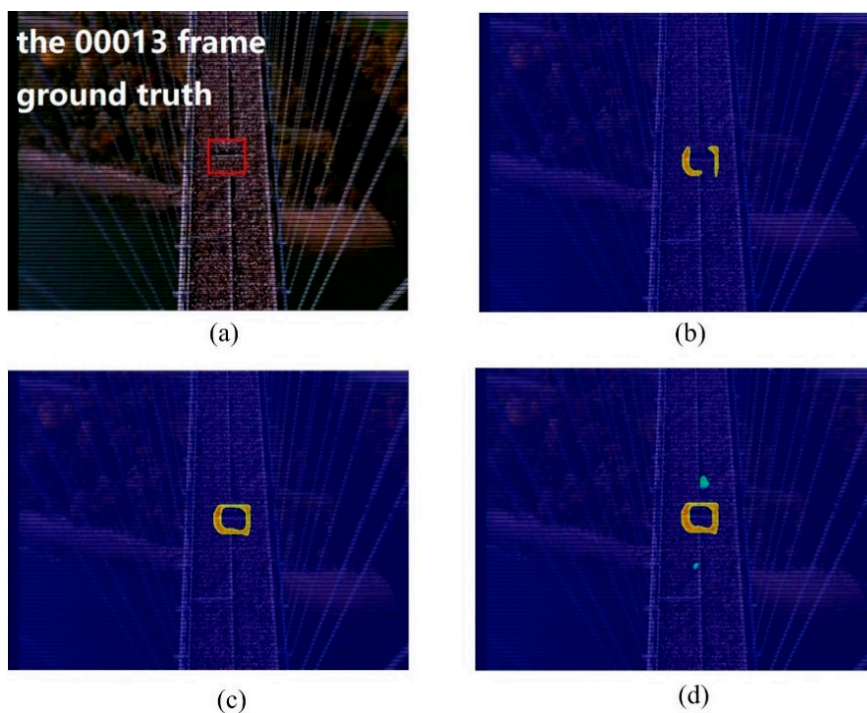
It is very important to select a suitable neighborhood size  $\varepsilon$  to construct complex network. A neighborhood that is too small will not be bias to salient motion, while a neighborhood with too large a scale will introduce more noise. In this section, the salient crowd motion will be detected using different neighborhood sizes  $\varepsilon$ . For retrograde motion detection, the train station scene and single retrograde scene is used to show the performance of the proposed method. From Figure 9, we can see that the salient motion region detected by applying the size of  $5 \times 5$  neighborhood is slightly scattering, while the area detected by the size of  $13 \times 13$  is more complete. A larger the neighborhood  $23 \times 23$ , can cause more noise in the detection results. As for the Figure 10, the salient motion region size detected by using the  $5 \times 5$  neighborhood is small, while the area detected by using the  $15 \times 15$  neighborhood size is more complete. When choosing a larger neighborhood of  $23 \times 23$ , the result includes more noise. For instability motion detection, two scenes were used in this experiment. For the marathon scene, the detection result was usually not closed if the neighborhood size was too small ( $5 \times 5$  neighborhood). Applying a neighborhood size of  $23 \times 23$ , noise interference will be introduced, although closed salient motion regions can still be obtained. After the experiment, the closed salient motion region can be obtained using a neighborhood size of  $11 \times 11$  (Figure 11). For another pilgrimage scene, although the saliency region can also be detected with a  $5 \times 5$  size neighborhood, the result obtained with a  $15 \times 15$  size neighborhood is closer to the ground truth (Figure 12).



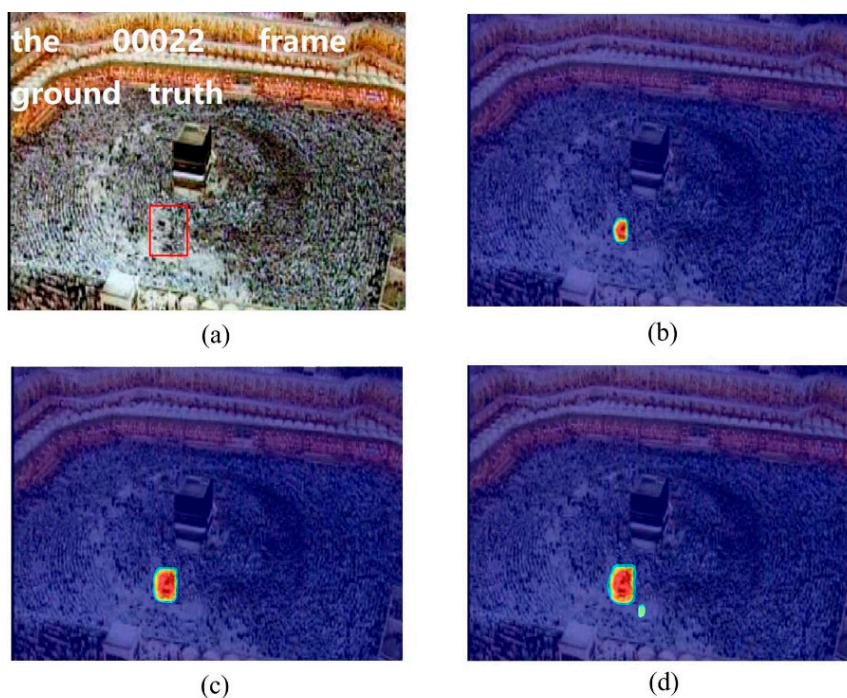
**Figure 9.** Retrograde motion detection in train station scene using different neighborhood size: (a) original video frame; (b) detection result using  $5 \times 5$  neighborhood; (c) detection result using  $13 \times 13$  neighborhood; (d) detection result using  $23 \times 23$  neighborhood.



**Figure 10.** Retrograde motion detection in single retrograde scene using different neighborhood size: (a) original video frame; (b) detection result using  $5 \times 5$  neighborhood; (c) detection result using  $15 \times 15$  neighborhood; (d) detection result using  $23 \times 23$  neighborhood.



**Figure 11.** Instability motion detection in marathon scene using different neighborhood size: (a) original video frame and ground truth region; (b) detection result using  $5 \times 5$  neighborhood; (c) detection result using  $11 \times 11$  neighborhood; (d) detection result using  $23 \times 23$  neighborhood.



**Figure 12.** Instability motion detection in pilgrimage scene using different neighborhood size: (a) original video frame and ground truth region; (b) detection result using  $5 \times 5$  neighborhood; (c) detection result using  $15 \times 15$  neighborhood; (d) detection result using  $25 \times 25$  neighborhood.

#### 5.4. Performance Evaluation and Comparison

The ground truth of crowd salient detection for pilgrimage and marathon scene has been given in the Crowd Saliency dataset [5]. The ground truth is given using a rectangular area. In order to evaluate the performance of the proposed method, we calculate the minimum enclosing rectangle of the detected salient motion region. To quantitatively evaluate the performance of the method, two indicators (precision and recall) are calculated in our experiments. In this paper, precision is the ratio of the number of pixels in the detected region that belong to the ground truth to the number of pixels in the detected area, indicating whether the number of pixels in the detected local motion instability area is accurate, expressed by  $Pr$ . Recall is the ratio of the number of pixels in the detected results region that belonging to the instability motion region to the number of all pixels of the ground truth, represented by  $R$  [25]. The precision and recall can be calculated as:

$$Pr = \frac{TP}{TP + FP} \quad (22)$$

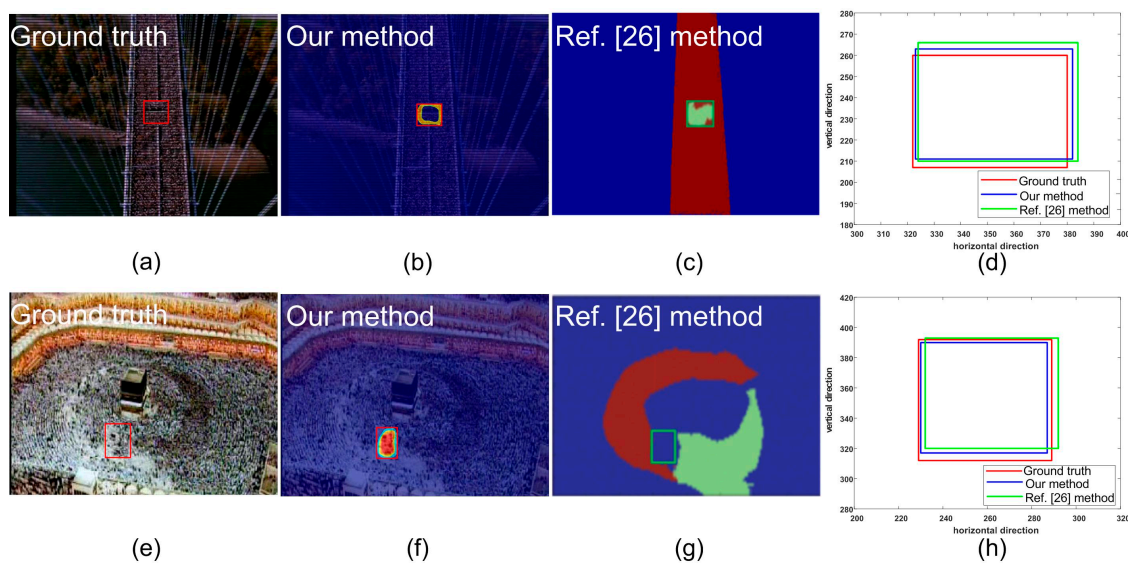
$$R = \frac{TP}{TP + FN} \quad (23)$$

where  $TP$  indicates that both the detection result and the ground truth are positive.  $FP$  indicates that the detection is positive and the actual is negative.  $TN$  indicates that both the prediction and the ground truth are negative.  $FN$  indicates that the prediction is negative but the actual is positive.

The precision and recall calculated from the pilgrimage and marathon scene using different parameters (the size of neighborhood) are given in Table 2. Obviously, according to the parameters selected in this paper, satisfactory detection accuracy can be obtained. If the neighborhood size is too large or too small, the detection accuracy will be seriously affected. Figure 13 shows the detection results of the pilgrimage and marathon scene using different methods. From Figure 13 we can see that both the proposed method and the methods mentioned in [26] can detect the salient region correctly. However, the rectangular region obtained by the proposed method is closer to the ground truth.

**Table 2.** The measurement of the accuracy of the detection results using different parameters.

Crowded Scenes	Statistics	Size of Neighborhood	Results
marathon	$Pr$	$5 \times 5$	0.862
		$11 \times 11$	0.910
		$23 \times 23$	0.531
	$R$	$5 \times 5$	0.841
		$11 \times 11$	0.909
		$23 \times 23$	0.877
pilgrimage	$Pr$	$5 \times 5$	1
		$15 \times 15$	1
	$R$	$25 \times 25$	0.684
		$5 \times 5$	0.244
		$15 \times 15$	0.867
		$25 \times 25$	0.656

**Figure 13.** Comparison of the method in this paper with the article [26]: (a,e) are the ground truth of marathon and pilgrimage scene; (b,f) are the results gained by our method; (c,g) are courtesy of reference [26]; (d,h) are the local enlarged displays of the results.

## 6. Conclusions

In this paper, we proposed a method for crowd salient motion detection based on a direction entropy and a repulsive force network. This paper focused on how to detect saliency regions in crowd movement effectively. Firstly, the crowd video sequence frames are processed by the optical flow algorithm followed by the crowd velocity vector field calculation. Secondly, according to the repulsive force model, the interaction force between two particles is determined as a certain condition. The repulsive force network is obtained and the strength of the crowd weighted network node is extracted as the characteristic parameter to construct a two-dimensional feature matrix. Finally, the velocity vector direction entropy is combined with the repulsive force network characteristic matrix to detect the salient crowd motion structure. The experimental results of four crowd video sequences show that the proposed method can not only detect the region of retrograde behavior of crowd movement but also the region of unstable crowd movement in large-scale crowd scenes. For future work, we will focus on the development of a method for an adaptive threshold and neighborhood calculation.

**Author Contributions:** X.Z. proposed the idea in this paper; X.Z., D.L., J.Z. and X.T. conceived and designed the experiments; D.L., J.Z. and Y.F. performed the experiments; X.Z., D.L., and H.Y. analyzed the data; X.Z. and

D.L. wrote the paper; X.Z., Y.F., X.T. and H.Y. edited and reviewed the paper; All authors read and approved the final manuscript.

**Funding:** This research was supported by National Natural Science Foundation of China (no. 61771418).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Challenger, R.; Clegg, C.W.; Robinson, M.A. Understanding crowd behaviours: Supporting evidence. In *Understanding Crowd Behaviours*; Leigh, M., Ed.; Emergency Planning College: Easingwold, UK, 2009; pp. 1–326.
2. Zhang, X.; Shu, X.; He, Z. Crowd panic state detection using entropy of the distribution of enthalpy. *Phys. A Stat. Mech. Appl.* **2019**, *525*, 935–945. [[CrossRef](#)]
3. Wixson, L. Detecting salient motion by accumulating directionally-consistent flow. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *22*, 774–780. [[CrossRef](#)]
4. Lim, M.K.; Chan, C.S.; Monekosso, D.; Remagnino, P. Detection of salient regions in crowded scenes. *Electron. Lett.* **2014**, *50*, 363–365. [[CrossRef](#)]
5. Lim, M.K.; Kok, V.J.; Loy, C.C.; Chan, C.S. Crowd Saliency Detection via Global Similarity Structure. In Proceedings of the 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; pp. 3957–3962.
6. Loy, C.; Xiang, T.; Gong, S. Salient motion detection in crowded scenes. In Proceedings of the 5th International Symposium on Communications, Control and Signal Processing (ISCCSP), Rome, Italy, 2–4 May 2012; pp. 1–4.
7. Jian, M.; Zhag, W.; Yu, H.; Cui, C.; Nie, X.; Zhang, H.; Yin, Y. Saliency detection based on directional patches extraction and principal local color contrast. *J. Vis. Commun. Image Represent.* **2018**, *57*, 1–11. [[CrossRef](#)]
8. Zhou, B.; Tang, X.; Wang, X. Coherent Filtering: Detecting Coherent Motions from crowd clutters. In Proceedings of the 12th European Conference on Computer Vision (ECCV), Firenze, Italy, 7–13 October 2012; Volume 7573, pp. 857–871.
9. Solmaz, B.; Moore, B.E.; Shah, M. Identifying behaviors in crowd scenes using stability analysis for dynamical systems. *PAMI* **2012**, *34*, 2064–2070. [[CrossRef](#)] [[PubMed](#)]
10. Zhang, X.; Yu, Q.; Yu, H. Physics Inspired Methods for Crowd Video Surveillance and Analysis: A Survey. *IEEE Access* **2018**, *6*, 66816–66830. [[CrossRef](#)]
11. Hao, Y.; Xu, Z.J.; Liu, Y.; Wang, J.; Fan, J.-L. Effective Crowd Anomaly Detection Through Spatio-temporal Texture Analysis. *Int. J. Autom. Comput.* **2019**, *16*, 27–39. [[CrossRef](#)]
12. Kanade, L. An Iterative Image Registration Technique with an Application to Stereo Vision. In Proceedings of the 7th International Joint Conference on Artificial Intelligence, Vancouver, BC, Canada, 24–28 August 1981; pp. 674–679.
13. Bouguet, J.Y. Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm. *OpenCV Doc.* **1999**, *22*, 363–381.
14. Chen, G.; Wang, X.; Li, X. *Introduction to Complex Networks: Models. Structures and Dynamics*; Higher Education Press: Beijing, China, 2012.
15. Joldos, M.; Technical, C.C. A parallel evolutionary approach to community detection in complex networks. In Proceedings of the IEEE International Conference on Intelligent Computer Communication and Processing IEEE, Cluj-Napoca, Romania, 7–9 September 2017; pp. 247–254.
16. Boccaletti, S.; Latora, V.; Moreno, Y. Complex networks: Structure and dynamics. *Phys. Rep.* **2006**, *424*, 175–308. [[CrossRef](#)]
17. Dzwiniel, W.; Yuen, D.A.; Boryczko, K.B. Diverse physical scales with the discrete-particle paradigm in modeling colloidal dynamics with mesoscopic features. *Chem. Eng. Sci.* **2006**, *61*, 2169–2185. [[CrossRef](#)]
18. Mohcine, C.; Seyfried, A.; Schadschneider, A. Generalized Centrifugal Force Model for Pedestrian Dynamics. *Phys. Rev.* **2010**, *82*, 046111. [[CrossRef](#)]
19. Zou, Q.; Sun, W.; Xing, L. Interest Point Detection in Images Based on Topology Structure Features of Directed Complex Network IEEE. In Proceedings of the 36th China Control Conference, Dalian, China, 26–28 July 2017; pp. 1507–1511.
20. Wu, Z.; Lu, X.; Deng, Y. Image edge detection based on local dimension: A complex networks approach. *Phys. A Stat. Mech. Appl.* **2015**, *440*, 9–18. [[CrossRef](#)]



21. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech.* **1948**, *27*, 379–423. [[CrossRef](#)]
22. Shannon, C.E.; Wyner, S.A. (Eds.) *Collected Papers*; IEEE Press: New York, NY, USA, 1993.
23. Tribus, M.; McIrvine, E.C. Energy and information. *Sci. Am.* **1971**, *225*, 179–188. [[CrossRef](#)]
24. Mehran, R.; Oyama, A.; Shah, M. Abnormal crowd behavior detection using social force model. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 935–942.
25. Goutte, C.; Gaussier, E. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.* **2005**, *51*, 952.
26. Ali, S.; Shah, M. A Lagrangian Particle Dynamics Approach for Crowd Flow Segmentation and Stability Analysis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–6.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).