Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Crowd modeling based on purposiveness and a destination-driven analysis method*

Ning DING[1,3], Weimin QI[2,3], Huihuan QIAN[†‡2,3]

[1]*Institute of Robotics and Intelligence Manufacturing, the Chinese University of Hong Kong, Shenzhen 518172, China*

[2]*Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen 518172, China*

[3]*School of Science and Engineering, the Chinese University of Hong Kong, Shenzhen 518172, China*

[†]E-mail: hhqian@cuhk.edu.cn

Received July 2, 2020; Revision accepted Nov. 4, 2020; Crosschecked Aug. 12, 2021; Published online Aug. 25, 2021

Abstract: This study focuses on the multiphase flow properties of crowd motions. Stability is a crucial forewarning factor for the crowd. To evaluate the behaviors of newly arriving pedestrians and the stability of a crowd, a novel motion structure analysis model is established based on purposiveness, and is used to describe the continuity of pedestrians' pursuing their own goals. We represent the crowd with self-driven particles using a destination-driven analysis method. These self-driven particles are trackable feature points detected from human bodies. Then we use trajectories to calculate these self-driven particles' purposiveness and select trajectories with high purposiveness to estimate the common destinations and the inherent structure of the crowd. Finally, we use these common destinations and the crowd structure to evaluate the behavior of newly arriving pedestrians and crowd stability. Our studies show that the purposiveness parameter is a suitable descriptor for middle-density human crowds, and that the proposed destination-driven analysis method is capable of representing complex crowd motion behaviors. Experiments using synthetic and real data and videos of both human and animal crowds have been conducted to validate the proposed method.

Key words: Crowd modeling; Intelligent video surveillance; Crowd stability
https://doi.org/10.1631/FITEE.2000312 **CLC number:** TP391.9

1 Introduction

To realize surveillance and forewarning of crowd congestion, stampede accidents, and other dangerous accidents in abnormal crowded scenarios, it is imperative that we understand crowd motion (Ihaddadene and Djeraba, 2008). Crowd modeling is an important and active research topic, and accuracy and suitability of the model directly influence crowd motion understanding. Many essential factors have been discovered and analyzed, such as collectiveness (Zhou BL et al., 2013) and coherent neighbor invariance (Zhou BL et al., 2012a). However, until now, special cases like pedestrian cross cases and cases with high noise levels have not been well understood. To gain insight into crowd motion from middle density to high density and in various types of stressful situations, we propose to adapt purposiveness parameters, including asymmetry (Asym) (Huet et al., 2006; Helmuth et al., 2007; Nurgaliev et al., 2013), cosine similarity (CS) (Karnik et al., 2007; Ye, 2011; Naseer et al., 2018), mobility (Mob), and variance-to-mean ratio (VMR) parameters (Zhou MY and Carin, 2015). These are invariant for rigid-body

‡ Corresponding author

* Project supported by the Shenzhen Science and Technology Innovation Council (No. JCYJ20170410171923840), the National Key R&D Program of China (Nos. 2019YFB1310403 and 2019YFB1310402), the National Natural Science Foundation of China (Nos. U1613226 and U1813216), the Chinese University of Hong Kong, Shenzhen (No. PF.01.000143), and Shenzhen Institute of Artificial Intelligence and Robotics for Society

ORCID: Ning DING, https://orcid.org/0000-0001-5618-6359; Huihuan QIAN, https://orcid.org/0000-0001-8269-0882

rotations and translations of an entire trajectory. For every step, parameters are calculated to describe the purposive state of a pedestrian at both global and local levels. After preliminary experiments are conducted to determine the impact of these parameters, we combine three of them to calculate the new purposiveness parameter for individual pedestrians' trajectories at the local level. We define the mean value of all the pedestrians' global purposiveness as the purposiveness of a crowd system, which is a crucial feature for determining the stability of the crowd.

Once the purposiveness is computed, we apply the proposed destination-driven crowd analysis method to study the trajectories with high purposiveness and estimate the destinations of pedestrians. We model pedestrians as solid self-driven particles, which are obvious feature points extracted from the image frames of videos. Crowd motion is like multiphase flow (Brennen, 2005), and for some aspects, it is the demonstration of the MATLAB-based finite element method solver for large problems (Dabrowski et al., 2008). The critical difference between crowd flow and fluid flow is that every pedestrian in the crowd is an autonomous and self-driven particle that knows its own destination at the very beginning. Many pedestrians share a common destination, which generates a force field that attracts corresponding pedestrians from many directions. Multiple destinations generate an overlapping multilayer force field. To represent this, we establish a multilayer force field (MFF) to model the crowd. A trajectory is the path of a particle that is driven by force from its destination. If the density is not high enough, pedestrians are crossing each other. When pedestrians move toward a common destination, this generates a collective phenomenon with a robust structural characteristic. When a crowd is in panic, the purposiveness value decreases and people tend to follow their neighbors. The crowd structure is deformed, and possibly even disintegrates because the initially common destinations have changed drastically. Therefore, the crowd structure model can be used to describe the crowd. We establish a structure model to comprehensively describe the stability of the crowd and predict the destinations of new arrivals.

2 Related work

Crowd motion analysis (Zhang et al., 2018; Sreenu and Durai, 2019) is of great significance for risk monitoring and early warning in crowded cases. Such analysis has experienced much development in recent years. The research community has made valuable contributions to the study of collective behaviors, such as crowd modeling (Helbing et al., 2002; Zhou BL et al., 2012a, 2012b, 2013), convolutional neural networks based crowd behavior studies (Rodriguez et al., 2011; Tripathi et al., 2018), crowd simulation (Zhao et al., 2018), and crowd behavior under emergency conditions for non-human biological entities (Shiwakoti et al., 2014). In this study, we focus on the crowd modeling study.

Crowd modeling approaches (Ijaz et al., 2015) are mainly zone-based models (Anh et al., 2011), layer-based models (Xiong et al., 2013), and sequential models (Patil et al., 2010). High-density crowd applications that take individual behaviors into account become possible when the various techniques are combined across the macroscopic and microscopic models. However, zone-based models rely heavily on the definition of the physical region boundaries, and layer-based models and sequential models meet the demand on computational efficiency when executing the microscopic model for the whole crowd. Multi-resolution approach (Xiong et al., 2009) has been applied to improve simulation efficiency by not applying models for the whole crowd. However, the present methods still have limitations such as incomplete factor consideration and poor adaptability (Ijaz et al., 2015). To solve the problem of poor algorithm adaptability caused by increasing population density, our method is based on MMF and purposiveness to conduct crowd modeling. At the same time, with the development of modern hardware technology, the computing power has been significantly improved.

In crowd modeling, parameter design is essential because it influences the pedestrian behavior estimation directly. Collectiveness (Zhou BL et al., 2013) is an instantaneous behavioral similarity descriptor for pedestrians that compares neighbors in a short period of time. It is suitable for measuring a high-density crowd when neighboring particles have similar motion patterns and short distances. However, this heavily relies on short distances among pedestrians and neglects their destinations. We find that there are common motion phenomena for which measuring collectiveness does not work well. Examples of these phenomena include the situation in which a

few scattered individuals move toward one destination from arbitrary directions or two groups of pedestrians move in different directions that cross each other. The coherence neighbor invariance based approaches, such as the coherence filter (CF) (Zhou BL et al., 2012a) method, also have limitations because they cannot extract sparse trajectories from data with high noise levels. Moreover, the CF method considers only adjacent neighbors with similar motion patterns as a group. However, when two groups encounter each other, the main group is divided into several smaller groups, although they have the same destination. Therefore, the CF method does not apply in the case where two groups encounter each other. In this work, purposiveness is proposed to estimate the trajectories and destinations of pedestrians with a range of densities and noise. The impact of stressful situations (i.e., an emergency) on the stability of the crowd is also explored.

3 Crowd modeling

To model crowd collective behavior with a certain pattern, first, we use particles to represent pedestrians and estimate their trajectories. Then, the purposiveness of the trajectories is computed; the trajectories with a high purposiveness are chosen to estimate the common destinations and the inherent structure of the crowd. Finally, according to these common destinations and the crowd structure, we evaluate the behavior of new arrivals and their impact on crowd stability.

3.1 Pedestrian trajectory

Ideally, a trajectory is the path of a pedestrian moving from a starting point to a destination in a straight line. However, in real cases, social forces (Yoon and Ayalew, 2018) from a neighboring environment cause random interferences among pedestrians and make them deviate from their original orientations. This increases the difficulty in modeling this complicated, yet common crowd phenomenon. So, it is necessary to analyze the effect of social forces on a trajectory before calculating the purposiveness. A trajectory $T = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{t-1}, \boldsymbol{x}_t, \boldsymbol{x}_{t+1}, \ldots, \boldsymbol{x}_n\}$ is a sequence of a pedestrian's positions through time t. Every step from x_t to x_{t+1} has a velocity vector with $t \in [1, n-1]$. So, the velocity during a time interval Δt is

$$\boldsymbol{v}_t = \frac{\Delta \boldsymbol{x}_t}{\Delta t} = \frac{\boldsymbol{x}_{t+1} - \boldsymbol{x}_t}{\Delta t}, \tag{1}$$

where $\Delta \boldsymbol{x}$ is the displacement during the time interval Δt. In Fig. 1b, there are 42 821 trajectories; the majority is fragmented due to object occlusion during the Kanade–Lucas–Tomasi (KLT) (Shi and Tomasi, 1994) feature extraction process, in which template image alignment techniques are used and trajectories are bent randomly by social forces during a half-hour video capture at the grand central station in New York City. To remove the projective distortion effect, every dot of a trajectory needs to be normalized according to the spatial relationship between the camera and the pedestrian. The trajectory filter function Φ (Helmuth et al., 2007) is used to reduce the zigzag phenomenon of the trajectory caused by the KLT extraction process. The operator is

$$\Phi\colon T \to \tilde{T}, \quad \tilde{\boldsymbol{x}}_t = \frac{1}{w} \sum_{j=(i-1)w+1}^{iw} \boldsymbol{x}_j, \tag{2}$$

where w is the window size, and i is the order of the window size. There is a window size sequence $1w, 2w, \ldots, iw, (i+1)w, \ldots, nw$ for every trajectory.

In Fig. 2a, there are displacements of the actual trajectory relative to the desired trajectory, which

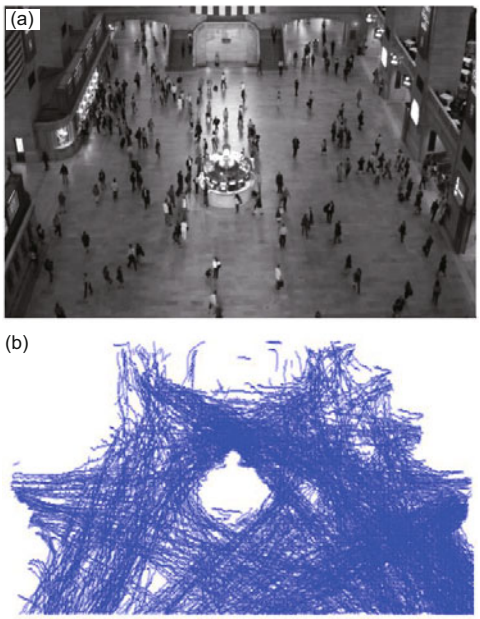

Fig. 1 Trajectories at the grand central station in New York City: (a) sample image in the ground station video; (b) computed trajectories of the pedestrians at the grand central station during the half-hour video

can be considered to be caused by social forces. x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively. Every actual velocity is the composition of social forces and the pedestrian's original desired force, which is to arrive at the destination. The displacements can be estimated as proportional to the composition of social forces from the neighboring environment. In Fig. 2b, there is an obvious turning point on the trajectory. We need to detect this significant behavior which may hide vital information, especially if this phenomenon is a collective response.

A typical trajectory consists of the starting region, the waypoint (WP), and the final destination. In normal situations, the positions of entrances and exits will not change. Hence, the purpose of the pedestrians is fixed. On the way to the final destination, many turning points on a trajectory can be detected by computing a purposiveness value. Turning points show that a pedestrian might change his/her mind and head to a new local destination. From the study of distribution for turning points, we find that there are regions with a higher density of turning points than other areas, and we call them WP. These are the local destinations of pedestrians on the way to their final destinations. In most cases, pedestrians avoid some fixed obstacles and go to the final destinations through paths with certain shapes. If there is a region with a high probability of turning points, the implication is that this region might be an interesting region, or WP, for many pedestrians. If the turning points are just randomly distributed, it means that these turning points are more likely caused by random social forces and there is no need for concern.

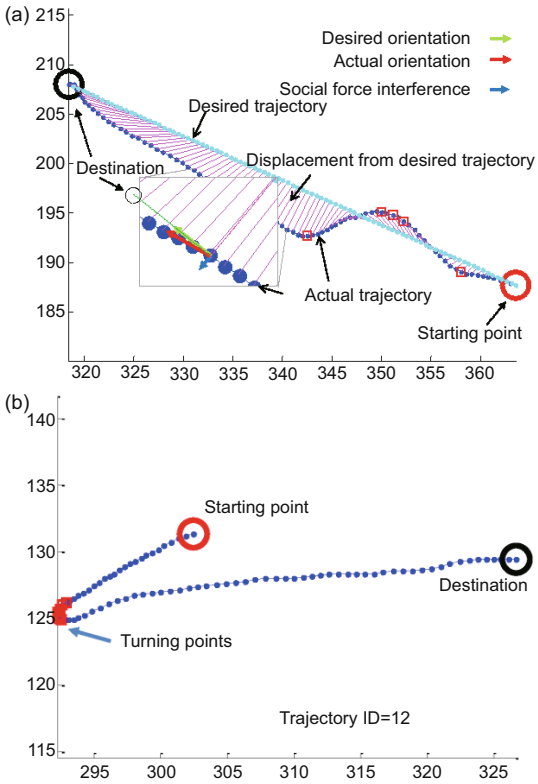

Fig. 2 Effect of social forces on trajectories: (a) the social forces change the trajectory of a pedestrian; (b) the pedestrian changes his/her mind and moves in the opposite direction due to social forces

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively

3.2 Purposiveness calculation

We calculate the purposiveness parameters of individual pedestrians and the crowd at two levels: global and local. The global level parameters are calculated from the entire trajectory. The local level describes the state of the recent period, and represents the smooth level of a trajectory. Here, the window size $w = 10$.

Asym was previously proposed as a feature for the detection of the directed motion. In this study, we use the two-dimensional (2D) equivalent, accounting for non-cylindrically symmetric point distributions.

$$p_{\text{Asym},i} = -\ln\left(1 - \frac{(\lambda_{1,i} - \lambda_{2,i})^2}{2(\lambda_{1,i} + \lambda_{2,i})^2}\right), \qquad (3)$$

where λ_1 and λ_2 are the eigenvalues of $\boldsymbol{R}$, and $\boldsymbol{R}$ is the 2D radius of the gyration tensor of the set of all points of a trajectory. High values of Asym indicate a greater tendency for a preferred direction. For a straight line, $p_{\text{Asym}} = -\ln(1/2) = 0.6931$, whereas for a Brownian random walk, the Asym value is close to 0.242. Obviously, the tendency of the straight line is larger. $p_{\text{Asym},i}$ is normalized in the range of [0, 1].

CS represents the similarity between the velocity of the current step and all the historical vectors within the window w. It is an inner product space that measures the cosine of the angle between them,

which provides a judgment of orientation.

$$p_{\mathrm{CS},i}=\frac{\sum_{j=1}^{w}\boldsymbol{v}_{i+j}\cdot\boldsymbol{v}_i}{\sqrt{\sum_{j=1}^{w}\|\boldsymbol{v}_{i+j}\|^2}\sqrt{\sum_{j=1}^{w}\|\boldsymbol{v}_i\|^2}}. \tag{4}$$

If there are two vectors with the same orientation, then $p_{\mathrm{CS},i}=1$; if the vectors are at 90°, then $p_{\mathrm{CS},i}=0$; if the vectors are diametrically opposed, then $p_{\mathrm{CS},i}=-1$. $p_{\mathrm{CS},i}$ is normalized in the range of [0, 1].

Mob is the ratio of the net displacement to the actual length of the trajectory. It is a description of the efficiency of particle movements.

$$p_{\mathrm{Mob},i}=\frac{\|\tilde{\boldsymbol{x}}_{(i-1)w}-\tilde{\boldsymbol{x}}_{iw}\|}{\sum_{j=1}^{w}\|\tilde{\boldsymbol{x}}_{(i-1)w+j}-\tilde{\boldsymbol{x}}_{(i-1)w+j-1}\|}. \tag{5}$$

For a trajectory with a straight line, $p_{\mathrm{Mob},i}=1$ (high value); for a circular trajectory, when turning back to the starting point, $p_{\mathrm{Mob},i}=0$ (low value); a random walk trajectory maintains $p_{\mathrm{Mob},i}=0$. $p_{\mathrm{Mob},i}$ is normalized in the range of [0, 1].

VMR is used to measure the dispersion of speed $\boldsymbol{v}$. It is defined as the ratio of the variance σ^2 to the mean μ.

$$p_{\mathrm{VMR},i}=\frac{\sigma_i^2}{\mu_i}. \tag{6}$$

When the trajectory orientation is fixed, the stability of the trajectory magnitude $\|\boldsymbol{v}\|$ should be of concern. When $\|\boldsymbol{v}\|$ is a constant variable, then $p_{\mathrm{VMR},i}=0$. When $p_{\mathrm{VMR},i}$ is in the range $[0,1]$, $\|\boldsymbol{v}\|$ follows a binomial distribution. When $p_{\mathrm{VMR},i}=1$, $\|\boldsymbol{v}\|$ follows a Poisson distribution. When $p_{\mathrm{VMR},i}>1$, it means that the variance is larger than the mean value and that $\|\boldsymbol{v}\|$ is over-dispersed. Thus, we use the formula to confine $p_{\mathrm{VMR},i}$ in the range of [0, 1].

If the trajectory is a straight line with steady speed, its purposiveness value is almost 1. If the trajectory is a smooth curve with a low curvature, it means that the pedestrian may be avoiding an obstacle and that the purposiveness value should be very close to 1. If there are many bends on the trajectory, but the orientation is still stable and the pedestrian keeps moving forward, it means that the pedestrian is determined and has a strong will to follow the goal despite any interference. In this case, the global purposiveness value is high, but the local purposiveness may be low. There are five experiments for the virtual dataset which contains five kinds of synthetic trajectories to analyze the relationship of parameters and trajectory behaviors.

In Fig. 3a, when the speed is stable, all parameters are close to 1, but when the speed is unstable, we find that the global VMR is much lower than other parameters, because some of the velocity angles are negative. The pedestrian moves by several steps back toward the starting point. In Fig. 3b, sinusoidal curves have stable local and global parameters, but when adding some noise, the local parameters are greatly affected. After a period, the global Mob value slowly increases, which means that the pedestrian is moving forward globally. The global VMR is almost stable. In Fig. 3c, the pedestrian on a circular trajectory turns back to the starting point no matter how smooth or irregular the trajectory is. The global VMR is almost stable. Globally, the three other parameters decrease to zero. Therefore, the pedestrian gives up their original destination. In Fig. 3d, the trajectories are random walks, which means that the pedestrian does not have a fixed goal. We suppose that the pedestrians are trying to follow their neighbors but always change their route quickly. Asym is close to 0 and VMR is almost stable. In Fig. 3e, the global VMR does not change, but other parameters have decreased due to the change of the trajectory.

From the above experiments on virtual datasets, we find that the local parameters describe the smoothness of the trajectory, and that more importantly, they can be used to detect turning points. We also observe that the global parameters are more stable than the local parameters. The value of VMR with the range [0.6, 0.9] is stable. In the following experiments using real cases, we explore the purposive computations at global and local levels to determine their stability for use in real situations.

There are two experiments using real datasets. In Fig. 4a, the 19^{th} trajectory exhibits a typical high purposiveness behavior. There is a group of social forces that causes a significant bending. The local and global parameters are high and stable. The 12^{th} trajectory has an obvious turning point detected by local parameter values. It also affects the global parameters due to a drastic change in its orientation. After a while, the global value starts increasing, but it takes a long time to achieve the initial altitude due to this immense change. In Fig. 4b, the 5261^{st} trajectory has two turning points in the later part,

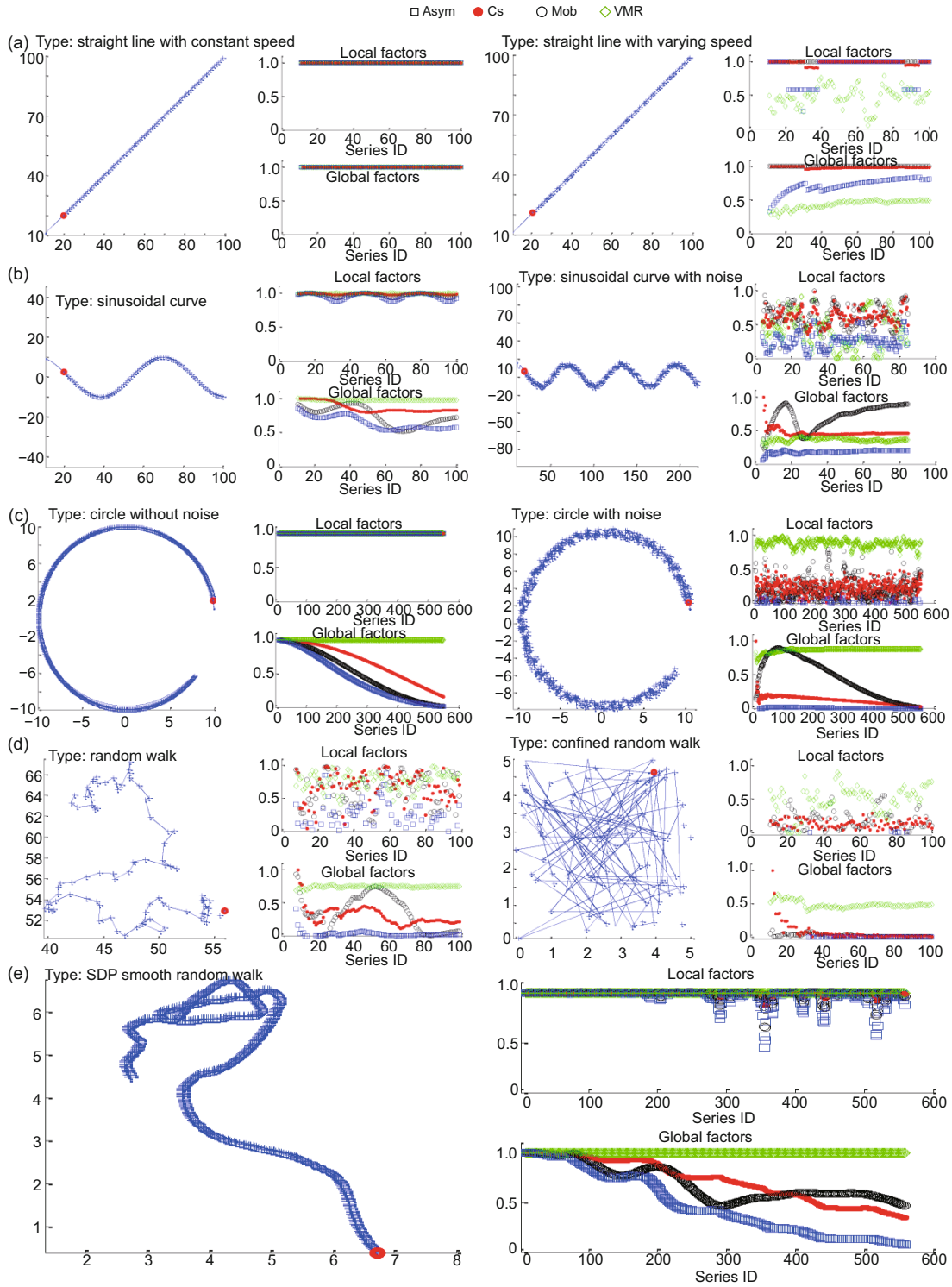

Fig. 3 Purposiveness evaluation of the synthetic dataset: (a) a line trajectory with a stable speed (left) and a line trajectory with an unstable speed (right); (b) a sinusoidal curve (left) and a sinusoidal curve trajectory with heavy noise (right); (c) a smooth circle trajectory (left) and a smooth circle trajectory with noise (right); (d) a Brownian random walk trajectory (left) and a confined random walk trajectory (right); (e) a self-driven particle's trajectory

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively

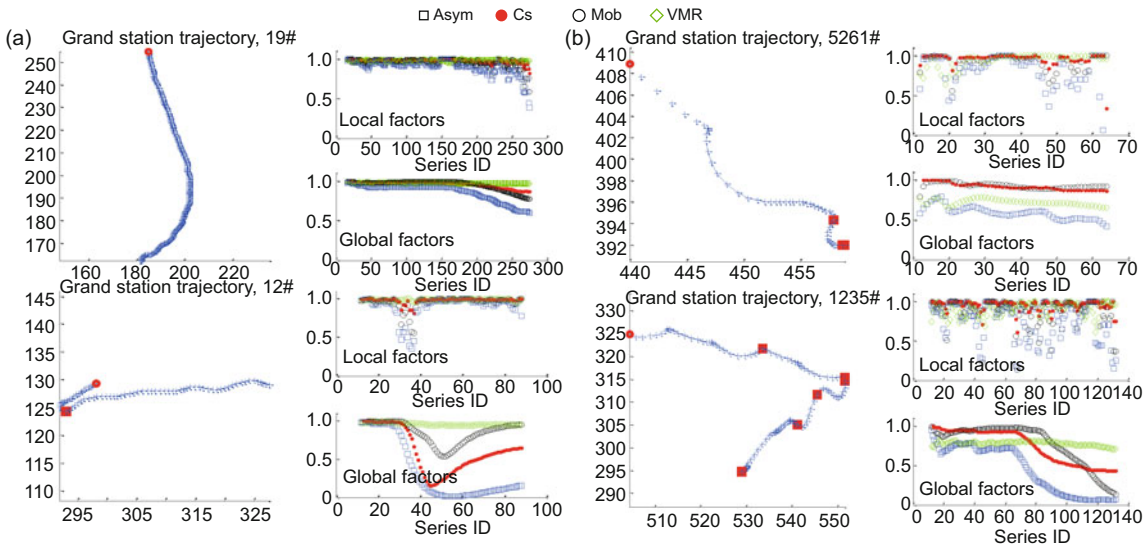

Fig. 4 Purposiveness evaluation of the real dataset: (a) the 19$^{\text{th}}$ trajectory (upper) and the 12$^{\text{th}}$ trajectory (lower); (b) the 5261$^{\text{st}}$ trajectory (upper) and the 1235$^{\text{th}}$ trajectory (lower)

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively. In the 19$^{\text{th}}$ trajectory, both local and global parameters are high and stable. The 12$^{\text{th}}$ trajectory has a distinct turning point, and VMR is stable, but there is a corresponding turning trend for other local and global parameters. The 5261$^{\text{st}}$ trajectory has two turning points in the later part, and the global VMR is stable. The 1235$^{\text{th}}$ trajectory has several turning points, and VMR is still stable, but the three other parameters decrease slowly

but it has a high purposiveness value. The 1235$^{\text{th}}$ trajectory has several turning points. The global parameters decrease significantly when the orientation of a trajectory changes several times.

According to the experiments using synthetic and real data described above, we can conclude that local purposiveness can be used for turning point detection, and that the global parameters are more suitable for determining the real purposiveness, especially Asym, CS, and Mob. In the experiments of real trajectory, it is obvious that the value of VMR within range [0.6, 0.9] is stable; its contribution is less than the three other parameters. Thus, we combine three other parameters with the most impact to obtain purposiveness p_i as

$$p_i = p_{\text{Asym},i} \cdot p_{\text{CS},i} \cdot p_{\text{Mob},i}. \tag{7}$$

The mean of all windows can be used to compute the purposiveness of a trajectory:

$$p = \frac{1}{|n|}\sum_{i=1}^{n} p_i. \tag{8}$$

Then the purposiveness of a crowd system C can be defined as the mean value of all the pedestrians' purposiveness values, which can be explicitly written in a closed form as

$$P = \frac{1}{|N|}\sum_{i=1}^{|N|} p_i, \tag{9}$$

where N is the number of all pedestrians.

3.3 Destination-driven crowd analysis

3.3.1 Multilayer force field

Many pedestrians share a common destination, which can be viewed as a source that generates a force field that attracts pedestrians from many directions. Each pedestrian has a trajectory—the path of a particle driven by force from its destination. Multiple destinations generate an overlapping MFF. Every destination can be considered as a phase of the human crowd flow. Consequently, estimating destinations is a critical step for crowd modeling. In this study, we use the trajectories with high purposiveness values to establish an MFF and then derive the destination from MFF.

1. Force field of a single trajectory

Crowd motion is a large-scale spatial and long-term temporal transition process of pedestrians. Many researchers study crowds with a motion field derived from an optical flow method (Cao et al.,

2009; Mehran et al., 2009), which uses partial deviation with respect to the spatial and temporal coordinates. Optical flow represents only the current motion information. However, in a middle-density crowd, it is difficult to derive a stable historical motion pattern (temporal) or the structure of the crowd flow (spatial) from the optical motion field. We employ trajectories within a period to obtain a dense force field. For a single trajectory with high purposiveness, assuming that the destination produces a dense force field, the pedestrian is attracted and moves toward it. Assume that the distributed force field is surrounded by the particles normally. Thus, we use a 2D Gaussian function to obtain a single force field $\boldsymbol{f}_{t,i}$ in the current position:

$$\boldsymbol{f}_{t,i}=\sum_{t=1}^{n}\boldsymbol{v}_{t,i}\cdot\exp\left(-\left(\frac{(x_{t,i}-x_0)^2}{2\sigma_x^2}+\frac{(y_{t,i}-y_0)^2}{2\sigma_y^2}\right)\right), \tag{10}$$

where $\boldsymbol{v}_{t,i}$ is the velocity vector from time $t-1$ to t. The current position (x_0, y_0) is used as the center, and a fixed σ^2 value is set as the standard deviation, $\sigma_x^2=\sigma_y^2$. The force field of a trajectory is the accumulation of vectors of all paces.

2. Force field of multilayer trajectories

We use a force field combination method to combine force fields. Adjacent trajectories with similar motion patterns are combined into one main field. We add the unit velocity vector to the overlapping region of two force fields and then check the status of this region. The maximum magnitude of the combined velocity vectors is 2, and the minimum is 0. When the average magnitude of the combined velocity vectors increases, it means that most of the vectors for two single force fields are similar. Otherwise, a decrease of the average magnitude of the combined velocity vectors means that the two force fields belong to different layers. In Fig. 5, the force field is derived from two trajectories. We can see that there is an overlapping region. The force fields of the 19$^{\text{th}}$ and 50$^{\text{th}}$ trajectories need to be separated because they have different orientations in the overlapping region.

The size of the overlapping region is also a key parameter. The larger size area contains more information about the original force fields and will become the key trajectory. We suppose that the trajectories always try to find similar longer trajectories to identify themselves. So, the combination condition is defined as the product of two factors, $c_i=c_{1,i}\cdot c_{2,i}$.

(1) The area ratio of the overlapping region with respect to the entire region of the smaller one is

$$c_{1,i}=\frac{\text{area}(\boldsymbol{f}_{t,i}\cap\boldsymbol{f}_{t,j})}{\text{area}(\boldsymbol{f}_{t,j})}. \tag{11}$$

(2) The average magnitude of overlapping regions is

$$c_{2,i}=\|\text{norm}(\boldsymbol{f}_{t,i})+\text{norm}(\boldsymbol{f}_{t,j})\|/2. \tag{12}$$

At time t, the MFF $\boldsymbol{F}_t$ is defined as

$$\boldsymbol{F}_t=\sum_{k=1}^{K}\boldsymbol{f}_{t,i}, \tag{13}$$

where $\boldsymbol{f}_{t,i}$ are the single force fields that satisfy the combination conditions, assuming that the number of single force fields that satisfy the combination conditions is K, counting from $k=1$.

We use an iterative process that starts from the longest trajectory and matches other shorter ones in turn. If the purposiveness is high enough, then a shorter trajectory matches the longest one selected; otherwise, the approach moves on to the next trajectory. As the process to combine trajectories continues, some trajectories will combine numerous single fields and become the main field. If a single force field is not combined with all other force fields, then it will be on standby and wait to match other force fields in the next turn. After a while, several main

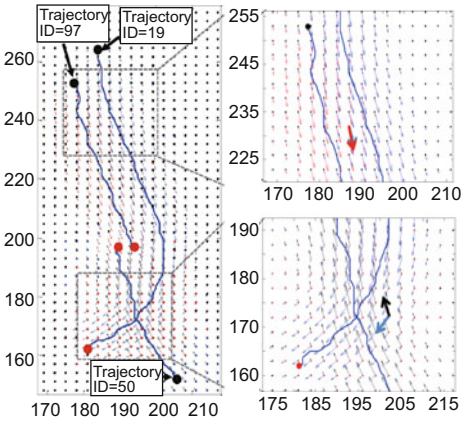

Fig. 5 Force field of three trajectories

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively. The 19$^{\text{th}}$ and 97$^{\text{th}}$ trajectories have the same orientations. The 19$^{\text{th}}$ and 50$^{\text{th}}$ trajectories have different orientations in the overlapping region

layers form and stabilize. Those layers that can contain only one or a few single force fields are discarded. From the main force field, we can observe the motion pattern of the crowd.

If there is only one main layer, most of the adjacent trajectories have similar motion patterns. Fig. 6a is the motion field produced by optical flow for a well-known phenomenon: wildebeest migration. There is no obvious multiphase phenomenon and the flow fields are fragmented. In Fig. 6b, the upper red part indicates that most of the particles are moving from the left side to the right side. In addition, there is an apparent blue region near the center, which means that there is an intensive flow moving from the bottom to the upper space. From MFF, we can see clear movement in the blue region. This clearly indicates that wildebeests are trying to join the mainstream in the upper space and are struggling to squeeze through a narrow passageway. The enlarged fields show the smoothness and evenness of the proposed force field. However, in Fig. 7, there is an obvious crossing phenomenon of two motion fields. Two groups of people are walking through each other toward different destinations. It is possible for MFF to describe such phenomena. 3. MFF adaptive updating

The multilayer dense motion field contains large spatial and long-term temporal crow motion information. MFF is updated with newly arriving trajectories in every frame. In a stable crowd, there will be more and more trajectories moving along similar paths toward the same destination, which reinforces the main layers and destinations. If there is a path that seldom has trajectories passing through, it will be forgotten gradually. We use the Gaussian adaptive approach (Muchoney and Williamson, 2001) and a factor α to describe this phenomenon:

$$\mathbb{F}_{t+1} = \alpha \boldsymbol{F}_{t+1} + (1-\alpha)\mathbb{F}_t, \tag{14}$$

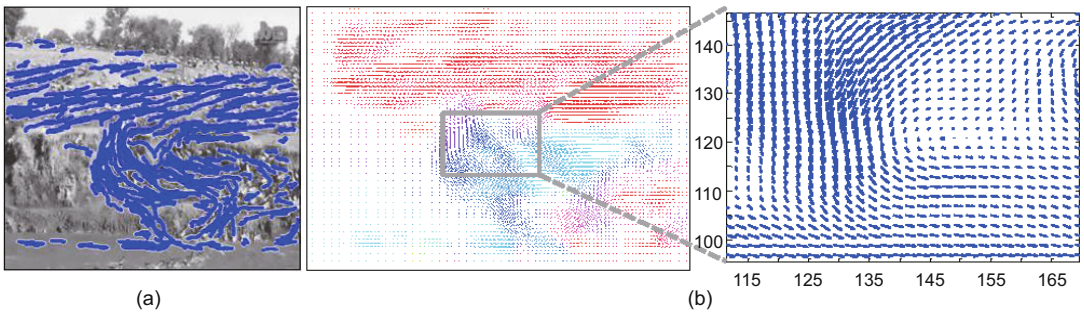

Fig. 6 MFF of the wildebeest migration video: (a) a motion field made by optical flow; (b) MFF

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively. In (b), red means the direction of the velocity vector is ranged in $0^\circ - 60^\circ$, and blue is $120^\circ - 180^\circ$. References to color refer to the online version of this figure

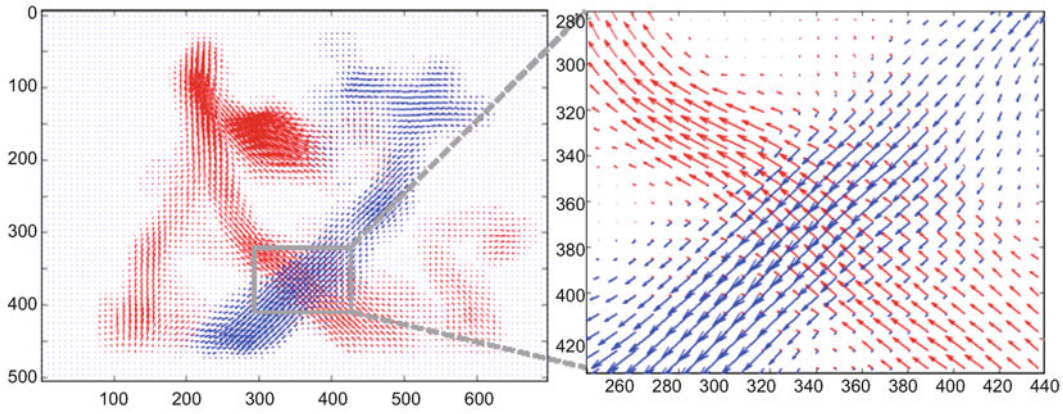

Fig. 7 Two separated overlapping force fields

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively

where α is a time factor and $\alpha = 1/(\text{time interval})$. If we assume a video frame rate of 20 frames/s, this results in $\alpha = 0.05$; here, frames/s multiplied by α means a one-second video. With time t, the force field of every layer is dynamically computed. A larger α makes the force field be more sensitive to the crowd motion. A smaller α value provides a more stable and robust representation. Using this approach, we obtain a series of dense and smooth fields $\boldsymbol{F}_t$ of crowd motions. We will use them to study the stability of the crowd.

3.3.2 Destination estimation

An MFF is a dense field. We use divergence to locate the starting region of the forces. The divergence measures the magnitude of a vector field, i.e., the extent to which the vector field flow behaves like a source or a sink at a given point. Formally, the divergence of a vector field $\mathbb{F}$ at a point p of a region V is defined as

$$\mathrm{div}\mathbb{F}_t(p) = \lim_{V\to\{p\}} \iint_{S(V)} \frac{\mathbb{F}_t \cdot \boldsymbol{n}}{|V|} \mathrm{d}S, \qquad (15)$$

where $|V|$ is the volume of V, $S(V)$ is the boundary of V, the integral is a surface integral, and $\boldsymbol{n}$ is the outward unit normal to the surface. div$\mathbb{F}$ is a function of p. From this definition, div$\mathbb{F}$ can be seen as the source density of the flux of $\mathbb{F}$.

Numerous positive regions (divergence) and negative regions (convergence) are evident in the divergence map in Fig. 8a. A threshold is set to remove the small regions, which results in Fig. 8b. We put particles at the peak of all the starting regions and allow the force field to drive them to move forward. We consider these trajectories as ideal trajectories without noise or interference. We can predict the future trajectory as follows. If a pedestrian appears in the starting region, the matching layer that is most aligned with their current orientation is determined, and their trajectories will be combined with the layer. As more force fields with a single trajectory are combined, every layer of the force field becomes more stable. Therefore, when free particles are dispensed on the field, and force is applied to drive particles forward, we can obtain many ideal trajectories, as shown in Fig. 8b. WP and the destination regions are a convergence region. We can easily classify the WP region and the final destination. The destination is a kind of WP without an outlet. In Fig. 8b, we can see three positive regions and they are all correct starting regions. The wildebeest has to crush through two narrow passageways on the abrupt slope to join the main herd. The wider passageway is the main port for most wildebeests

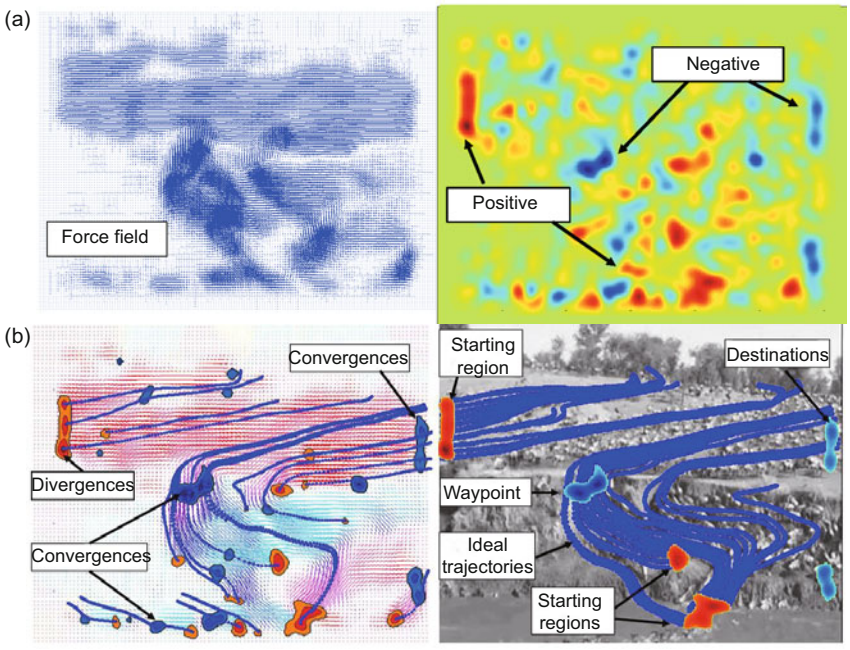

Fig. 8 The divergence map and the ideal trajectories: (a) divergence map of an MFF; (b) ideal trajectories from the divergence map

moving up and it becomes WP. There are another two negative regions located on the right side of the image. The correct destination region is the upper one, to which the main herd is heading. We can use the interesting regions as nodes and use the ideal trajectories as links to represent the structure of the crowd.

Helmholtz decomposition of MFF: The crowd flow can be regarded as an incompressible and irrotational fluid. The force field can be decomposed into the incompressible part and the irrotational part with Fourier transforms. Using the Helmholtz decomposition theorem (Bladel, 1959), we can obtain the potential functions:

$$\phi(x,y)=\phi_0+\frac{1}{2}\int_0^x(u_{\mathrm{r}}(s,y)+u_{\mathrm{r}}(s,0))\mathrm{d}s+\frac{1}{2}\int_0^y(v_{\mathrm{r}}(x,s)+v_{\mathrm{r}}(0,s))\mathrm{d}s, \tag{16}$$

$$\psi(x,y)=\psi_0+\frac{1}{2}\int_0^y(u_{\mathrm{c}}(x,s)+u_{\mathrm{c}}(0,s))\mathrm{d}s-\frac{1}{2}\int_0^x(v_{\mathrm{c}}(s,y)+v_{\mathrm{c}}(s,0))\mathrm{d}s, \tag{17}$$

where the force field $\boldsymbol{\Omega}=(u,v)^{\mathrm{T}}$ denotes a planar vector field. We can use the Helmholtz decomposition theorem $\boldsymbol{\Omega}=\boldsymbol{\Omega}_{\mathrm{c}}+\boldsymbol{\Omega}_{\mathrm{r}}$, where $\boldsymbol{\Omega}_{\mathrm{c}}$ and $\boldsymbol{\Omega}_{\mathrm{r}}$ denote the incompressible and irrotational parts of the vector field, respectively. The incompressible part is divergence free, $\boldsymbol{\nabla}\cdot\boldsymbol{\Omega}=0$, and the irrotational part is curl free, $\boldsymbol{\nabla}\times\boldsymbol{\Omega}=\mathbf{0}$. Thus, there are functions ϕ and φ known as the stream function and the velocity potential, respectively, satisfying $\boldsymbol{\Omega}_{\mathrm{c}}=\boldsymbol{\nabla}\phi$ and $\boldsymbol{\Omega}_{\mathrm{r}}=\boldsymbol{\nabla}\varphi$, where s is the boundary of x and y. In Fig. 9, the force field has been decomposed into two maps, the velocity potential map and the divergence map. We use the velocity potential map to derive the interesting regions. It is smoother than the divergence map, and there are fewer peaks in this map. In the divergence map, we place free particles at the peaks of the starting regions to obtain the ideal trajectories, which illustrate the crowd flow in a simplified manner.

3.3.3 Destination-driven crowd structure establishment and parameterization

The structure of crowd motion is the dynamic motion pattern of the entire crowd flow. Ali and Shah (2007) used Lagrangian coherent structures (LCS) to describe high-density crowd motion patterns and segmentation. Zhou BL et al. (2013) used collectiveness to detect motion patterns from highly dense and noisy data. Our crowd structure is based on the ideal trajectories of MFF. We unify the release position of the free particles and the step number of the ideal trajectories to obtain a generalized representation of the crowd structure. First, a certain number of free particles are scattered on the force field evenly with a certain pixel interval, e.g., 30 pixels. Second, at every time t of every frame, those free particles are driven by the force according to its current position in the field $\boldsymbol{F}_t$ and stopped within a certain number of steps, e.g., 50. Finally, we compute the time series of the ideal trajectory. Every trajectory in the time series has the same starting point and number of steps, but they have different shapes. The ideal trajectory looks like an elastic curve on the field. A stronger force makes it move faster, and vice versa. Its orientation is the

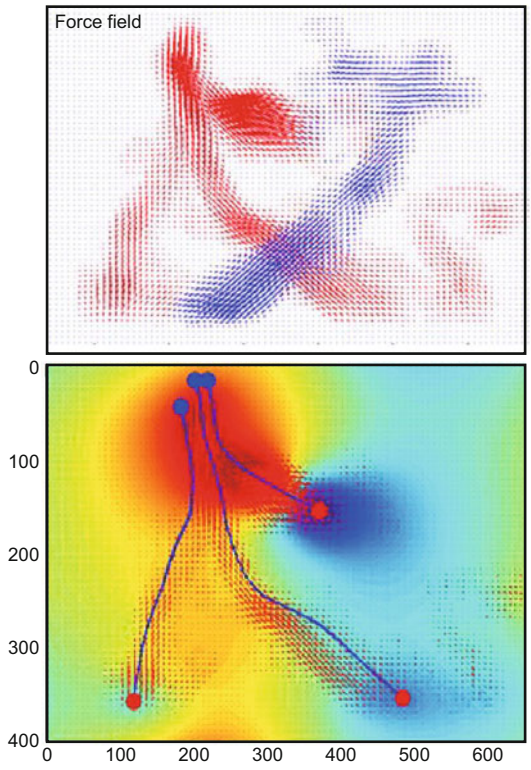

Fig. 9 Ideal trajectories from the decomposed divergence map

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively. The red point represents the starting point, and the blue point represents the destination. References to color refer to the online version of this figure

composition of neighboring force vectors.

Fig. 10a shows the unified ideal trajectories of the two-layer force field of a grand central station video sample. Every ideal trajectory has a fixed starting point. The blue trajectories are stretched and deformed by the force field; the red points are the destinations of the trajectories. We can parameterize the unified ideal trajectories of the crowd structure with 2D polynomial curve fitting, in which the algorithm derives the coefficient of the fitting result with a certain polynomial order according to the complexity of the curve, e.g., 20 orders. The fitting process attempts to minimize the orthogonal distance, including both axes of displacement to two trajectories with total least squares. Then, we obtain a polynomial parameter matrix of every frame, which can be used to reproduce the crowd scenario in the crowd simulation platform. Fig. 10b shows the result of parameterization. With the unified ideal trajectories, it is easy to evaluate the stability of the crowd. There are many ways to measure the extent of trajectories varying over time. A drastic variation from the ideal trajectories means that the structure is unstable. Displacements of the corresponding dots on the trajectories from the same starting point are the direct representation of trajectory deformation. We use the mean value and standard deviation of displacements to describe this crowd structure variation.

4 Experiments

4.1 Datasets

4.1.1 Synthetic datasets

Set 1 in Fig. 11a contains four motion patterns generated by the parabola, circle, line, and disk with high levels of noise. Set 2 contains three typical collective motion patterns in a heavy noisy crowd: (1) In Fig. 11b, there are two groups of particles; each

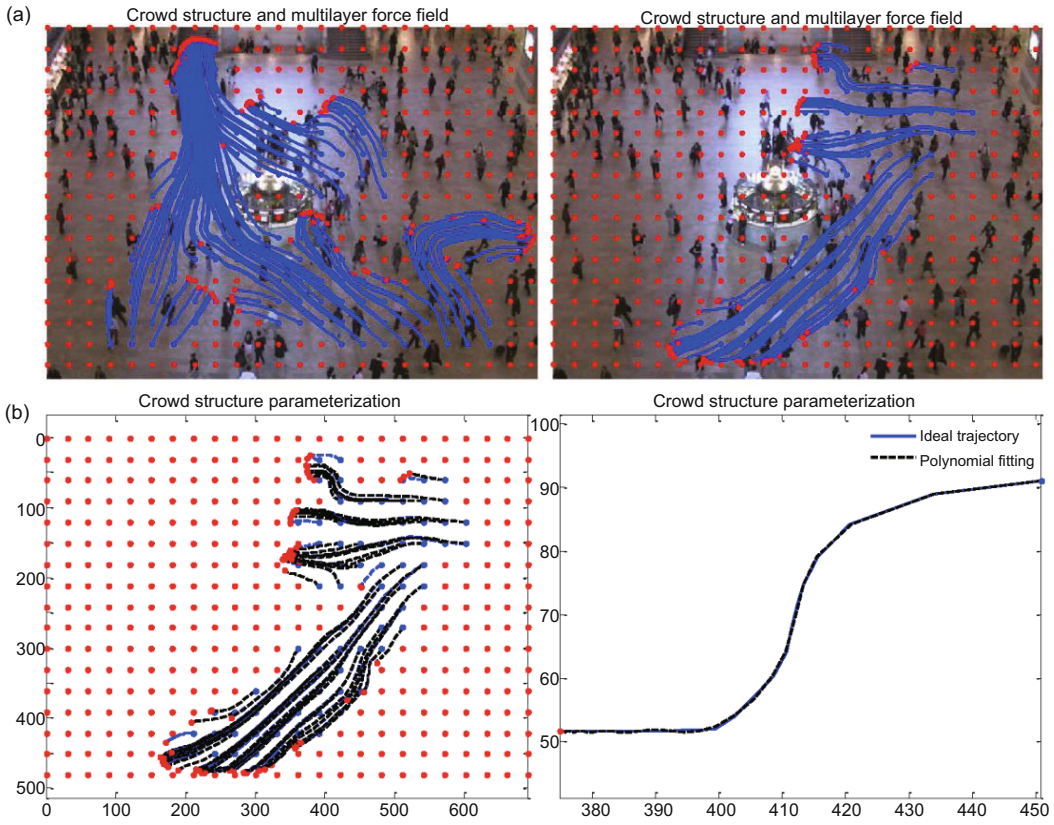

Fig. 10 Ideal trajectory parameterization of the grand central station: (a) unified ideal trajectories for the representation of crowd structure; (b) ideal trajectory parameterization

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively. References to color refer to the online version of this figure

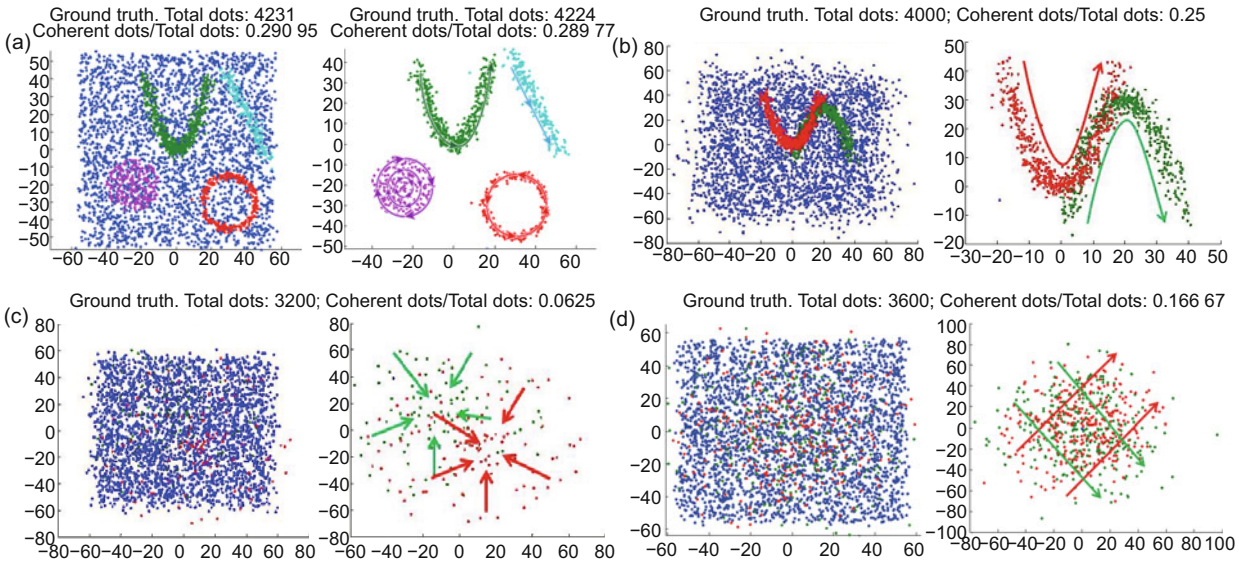

Fig. 11 Synthetic datasets: (a) there are four motion patterns generated by the parabola, circle, line, and disk with heavy noise (left) and without noise (right); (b) there are two groups of particles with double-V patterns with high levels of noise (left) and without noise (right); (c) there are two groups of sparse particles with high levels of noise (left) and without noise (right); (d) two hundred sparse particles gather toward two destination regions with high levels of noise (left) and without noise (right)

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively

group has 300 particles and the groups cross through one another. (2) In Fig. 11c, there are two groups of sparse particles; every group with 300 particles moves toward two orientations. (3) Two hundred sparse particles gather toward two destination regions in Fig. 11d. The initial positions of the coherent dots are randomly sampled from the Gaussian distribution along the traces of the coherent motion patterns.

4.1.2 Real case datasets

To quantitatively evaluate the proposed destination-driven approach, we collect additional data from two kinds of real crowd video footage: (1) raw videos from surveillance systems, e.g., Olympic game stadium entrance, Beijing's Tiananmen Square, and a bus station in Shenzhen; (2) stampede and riot events video downloaded from YouTube, e.g., the stampede at Jagannath Rath Yatra Festival in India in 2009, the Folsom Prison Riot in 2002, and the shooting event and panic in a Mexican Soccer Stadium in 2011. There are 22 clips in total, four of which are shown in Fig. 12.

The ground truth is the average rate of the manually labeled results acquired from three independent subjects. Each subject is asked to rate the stability of a crowd in a video using two states: stable and unstable. We use the correlation between the human-labeled ground truth and the result of the proposed approach as the performance evaluation criterion.

4.2 Crowd motion detection and analysis

4.2.1 For synthetic data

Using set 1, this task consists of two steps: (1) detecting trajectories with four typical collective motion patterns from neighboring trajectories with

Fig. 12 Real case datasets

Raw videos from a surveillance system (upper two figures) and stampede and riot event videos downloaded from YouTube (lower two figures)

high levels of noise; (2) estimating the number of motion patterns. We use the proposed purposiveness parameter (pur-value) to fulfill this task. To evaluate the performance, four representative clustering methods, i.e., normalized cuts (Ncuts), K-means, Mean-shift, and CF, are selected for comparison. Particles are clustered based on pur-values or collectiveness values (col-values). For Ncuts and K-means methods, the motion pattern amount is chosen as 5 and 3; Mean-shift automatically determines the cluster size. The time interval d is set to 4, which means that the trajectory of each step has six samples as the inputs to all the algorithms. The quantitative result is measured by the normalized mutual information (NMI), where a larger NMI indicates better clustering performance. As shown in Table 1, our approach obtains the best NMI score. As shown in the detection results in Fig. 13, particles moving with four motion patterns have higher global purposiveness values than neighboring random walk trajectories for the first 10 frames. We use the depth of color to illustrate the pur-value. The pur-value of the solid circular motion decreases over time because trajectories of this pattern always turn back to the starting region; its mobility efficiency is low. The pur-value of straight lines is maintained at the highest level. The result clearly distinguishes between coherent dots and incoherent noisy dots.

Essentially, the first step of this task is a binary classification, which is used to determine whether the trajectory is a random walk. As long as the noisy particles are removed, the pur-value and CF can all achieve a good result at the classification step. However, it is challenging for CF and other approaches when the target particles are sparsely scattered in the region because it is difficult to find neighbors with a similar motion pattern surrounding the target trajectories. Under these conditions, our approach can detect target trajectories in the first step with a higher pur-value. The next step is establishing the MFF and estimating the destinations for motion pattern recognition. Fig. 14 shows the results of our approach; all motion patterns have been detected successfully.

The most important thing for classification is that the motion pattern should be recovered from the ideal trajectories of the different layers. We set a free particle on the peak of the divergence map and allow the force to drive it to the destination; it stops at the valley bottom automatically. The polynomial equation of the ideal trajectory is the parameterized motion pattern. We use the col-value to cluster the overlapping and double-V crossing motion, and

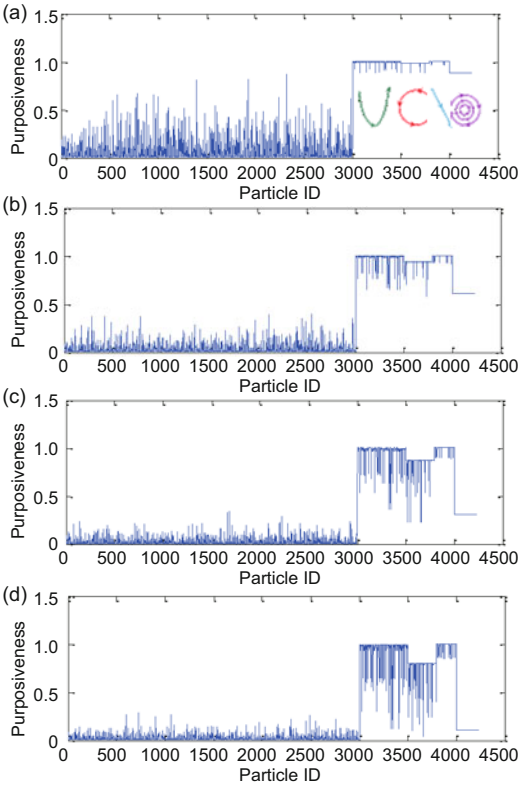

Fig. 13 Purposiveness evaluation of the synthetic dataset: (a) 10#; (b) 20#; (c) 30#; (d) 40#

Table 1 Crowd pattern detection results (synthetic cases dataset)

Approach	Normalized mutual information			
	Set 1, 4 shapes, 4 patterns	Set 2, double V, 2 patterns	Set 2, crossing, 2 patterns	Set 2, gathering, 4 patterns
Ncuts	0.28	0.622	0.385	0.226
K-means	0.21	0.766	0.113	0.261
Mean-shift	0.20	0.692	0.255	0.155
Coherence filter	0.890 (4 patterns)	0.919 (3 patterns)	0.324 (31 patterns)	0.189 (6 patterns)
Ours	0.971 (1 layer, 4 patterns)	0.991 (2 layers, 2 patterns)	0.933 (2 layers, 2 patterns)	0.8201 (1 layer, 2 patterns)

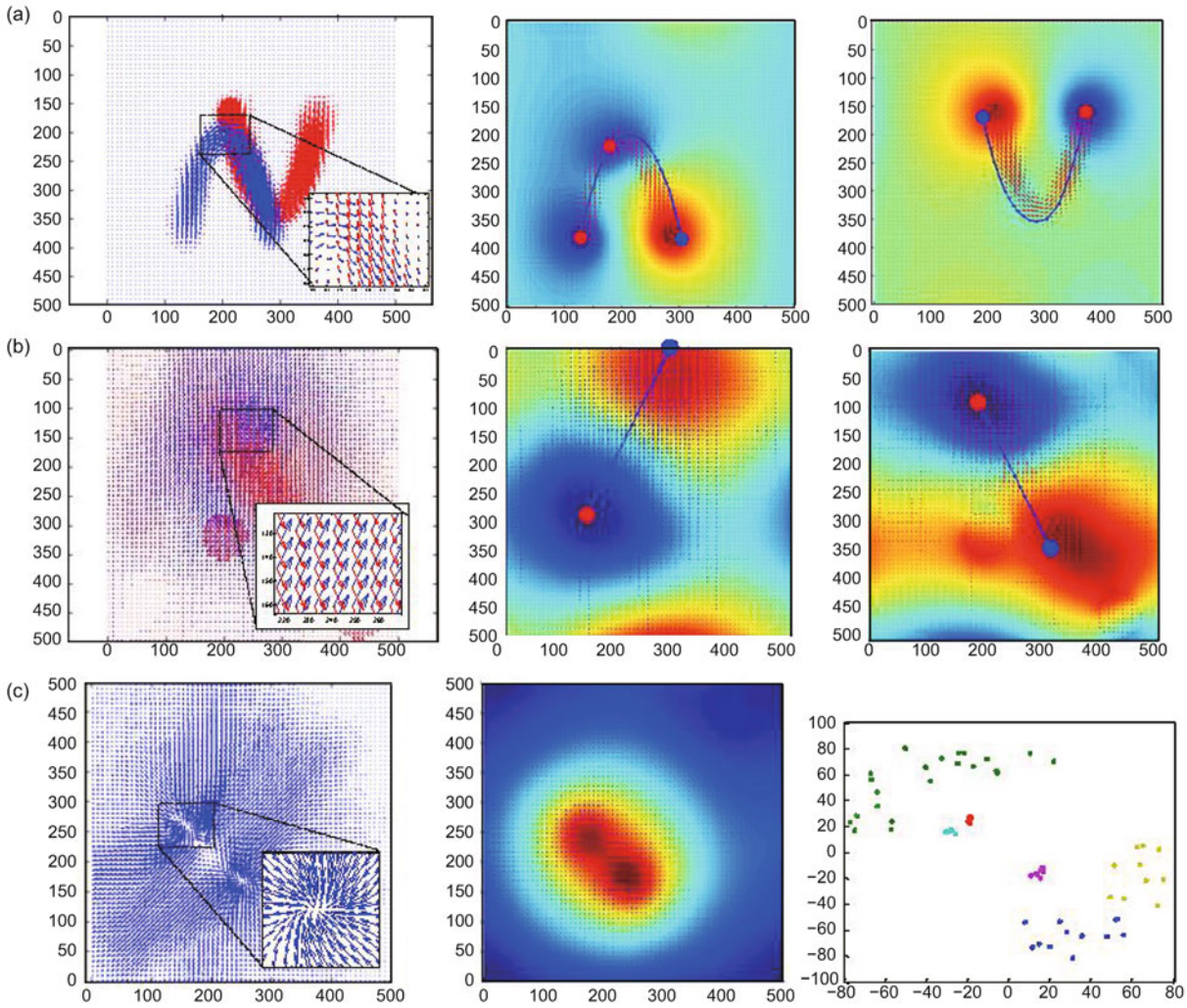

Fig. 14 Results of the synthetic motion pattern dataset: (a) MFF and the two trajectories of the double-V crossing motion pattern; (b) MFF and the two trajectories of the crossing motion pattern; (c) MFF and the trajectories of the gathering motion pattern

x_{img} and y_{img} denote the horizontal and vertical axes of the image plane in the unit of pixel, respectively

obtain three clusters (the ground truth is two clusters) and an NMI score of 0.919. From Fig. 14a, we can see that the branches of the two V shapes overlap and that the two groups of particles cross each other with very small distances. With our approach, a two-layer force field has been formed, two destinations have been found, and the NMI score of 0.991 is very close to a perfect score.

Behaviors in which particles gather and cross sparsely are too difficult for the CF method. The proposed destination-driven model is a generalized approach for solving these problems. We obtain a two-layer force field from the crossing tasks, and it is obvious that there is a destination on every layer. The NMI score surpasses the alternative methods.

In the experiments using set 1, the NMI values rely mainly on the correction rate of particles for the trajectories with noise. The CF method performs well. However, in the experiments using set 2, the CF method performance drops drastically. The reason is that this approach cannot extract sparse trajectories from data with high noise levels. It has a good NMI score in the double-V data because the two groups of data are in close proximity. However, it performs poorly in the pattern estimation step. The proposed purposiveness based destination-driven crowd analysis approach is suitable for handling this situation and it achieves the best performance.

4.2.2 For real cases

To quantitatively evaluate the collective motion detection performance in real cases, we employ the people-counting-based evaluation method of Zhou BL et al. (2012b). The experiment has three steps: counting the number of pedestrians at each key frame, calculating the average count error, and calculating the detection rate (DR) and false alarm rate (FAR) as the evaluation criteria. Table 2 shows the performances of the ALDENTE method (Rabaud and Belongie, 2006), the Bayesian detection method (Brostow and Cipolla, 2006), the CF method, and our method in the pedestrian detection experiments. Fig. 15 shows the corresponding results for the four methods, where one color represents a kind of motion pattern. The CF method and our approach achieve better performances. The performances of the ALDENTE method and the Bayesian detection method are worse. However, the CF approach does not perform well if an MFF exists; e.g., two groups of pedestrians are crossing the road. There might be some interference when the two groups encounter and cross each other; our approach has successfully demonstrated this phenomenon.

Table 2 Crowd pattern detection results using a real case dataset

Method	DR	FAR	Count error
Our method	0.77	0.12	1.62
Coherence filter	0.62	0.22	1.95
Bayesian detection	0.57	0.20	2.88
ALDENTE	0.41	0.34	3.15

DR: detection rate; FAR: false alarm rate

There are 400 frames and 2973 unfiltered trajectories in our crossing street videos. There is heavy fragmentation in the trajectories especially when two groups of pedestrians encounter and occlude each other. It is hard to derive the ground truth to evaluate the performance qualitatively, so we roughly establish the ground truth that there are two collective motion groups moving in opposite directions. First, 1973 trajectories are selected randomly to establish an MFF. Then, for the remaining 1000 test trajectories, we evaluate their similarity with every layer. Finally, every layer has a group of most similar trajectories, i.e., a classification process. Here, we obtain 63 force field layers. There are two main layers that represent the mainstream crowd motion. The rest is produced by fragmented trajectories, which will be neglected if there are too few trajectories in this layer. Fig. 16 shows the observation results, where the CF method considers only adjacent neighbors with a similar motion pattern as a collective team. Thus, when two groups encounter one another, the main group is separated into several small groups, despite the fact that they have the same destination. In our method, there are two main force fields and pedestrians have been classified into two groups that are moving toward two destinations.

4.3 Crowd stability evaluation

We extract 10 stable clips from the real case videos and manually label them with three subjects separately. There are two typical crowd motions observed in Figs. 17a and 17b. The four snapshots in Fig. 17a are from the wildebeest migration video. We can see that there is an obvious deformation

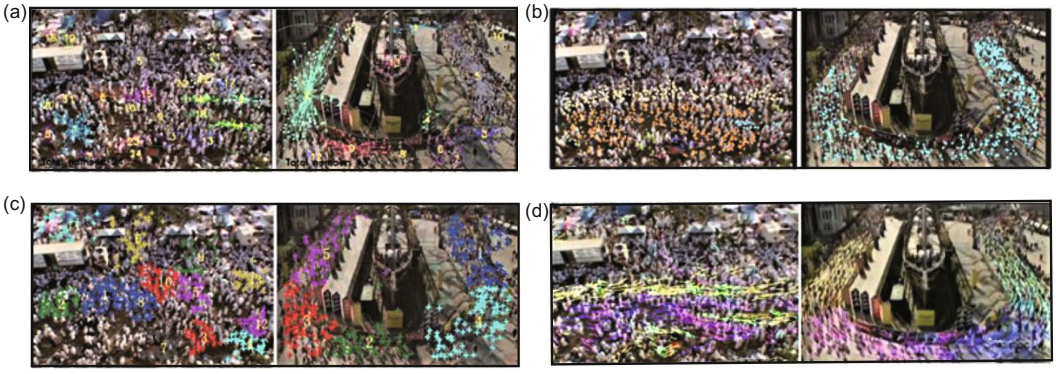

Fig. 15 Results of motion pattern detection on real cases (high-density crowd): (a) Bayesian detection; (b) coherence filter; (c) ALDENTE; (d) our method

References to color refer to the online version of this figure

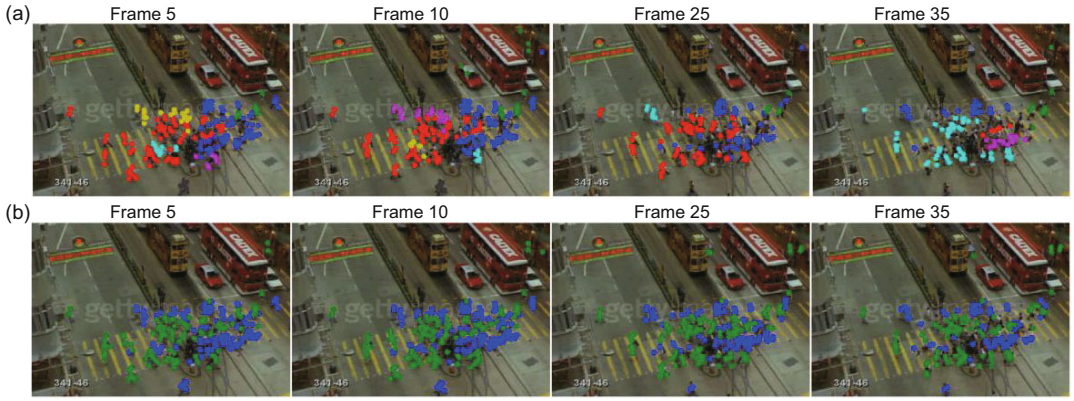

Fig. 16 Results of motion pattern detection in real cases (crossing a street): (a) results with the coherence filter method; (b) results with our method

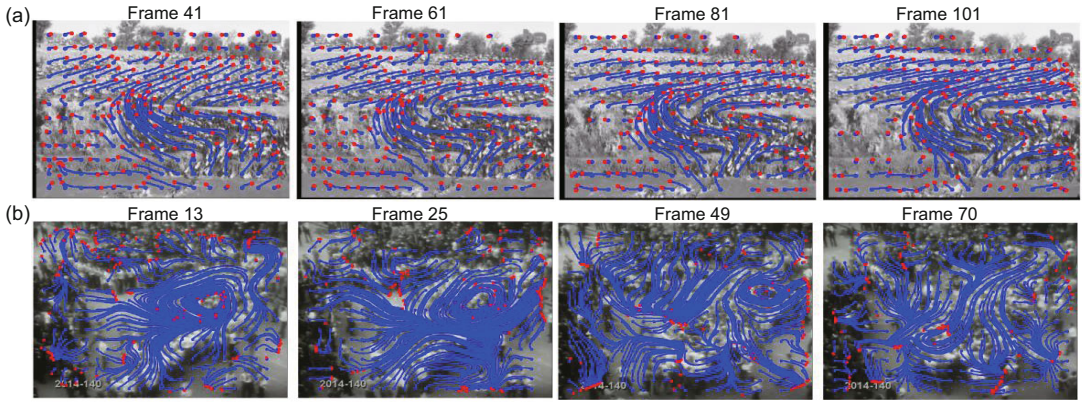

Fig. 17 Parameterized crowd structures of wildebeest migration (a) and a riot event (b)

pattern of the trajectories, and that there are only a few changes to this pattern from the 41st to the 101st frame. This means the crowd flow is stable. The four snapshots in Fig. 17b are from a video capturing a violent riot. We can observe that there is a vortex pattern in the first two figures, which become completely different in the latter two figures. This crowd flow is not stable. Fig. 18 shows the mean and standard deviation for all the displacements of the ideal trajectories. The curves show the results of wildebeest migration and a violent riot. From the smaller mean and standard deviation of displacement, we can observe that there are fewer trajectory deformations and that the crowd structure is stable for wildebeest motion. The final correlation between the result and the labeled ground truth is 0.71.

5 Conclusions

Collective and fluid properties are two characteristics of a high-density crowd, but they do not represent the stability characteristics of crowd motion. The proposed purposiveness is a suitable description of a middle-density human crowd. Combining the collectiveness and other motion representation methods, our novel approach to crowd modeling and stability estimation is demonstrated to be both feasible and accurate. The experiments

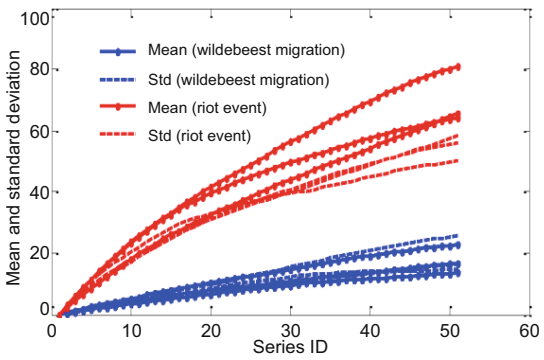

Fig. 18 Comparison of two typical crowd motions

References to color refer to the online version of this figure

have been separated into two parts: (1) Use synthetic data to detect and analyze the motion of the crowd. In the experiments using set 1, with the purposiveness parameter, we extract all motion patterns from trajectories with high noise levels and successfully estimate the number of patterns. We find there are higher global purposiveness values for particles in the four motion patterns than for random trajectories. Both the pur-value and CF methods achieve good results in these experiments. In the experiments using set 2, although particles gathering and crossing sparely are two difficult tasks for CF, our destination-driven model exhibits good performance in detecting target trajectories, and it achieves the best NMI value. (2) Evaluate the collective motion detection performance quantitatively in real cases using four methods. We detect the collective motion in situations with a high-density crowd and groups crossing the street, and where the entities in the crowd gather and cross within close proximity. Our proposed approach achieves the best performance (the largest DR, the smallest FAR, and the smallest count error) among four methods. For the stability of a crowd, the wildebeest migration crowd exhibits the deformation pattern of trajectories and few changes of the mainstream pattern, demonstrating the reality that the wildebeest migration crowd is stable with smaller mean and standard deviation. The riot event crowd initially presents a clear vortex pattern, but it then changes drastically over time, demonstrating that the riot event crowd is unstable with larger mean and standard deviation.

Crowd modeling is an important field worthy of continuous study. We believe it is important to note that multiple field layers exist at the same time. From the fluid field aspect, destinations are local minima that attract particles, or it is a global energy minimum process that occurs when pedestrians are getting close to their destination. In many cases, the destinations are fixed and determined by the geometric layout of the scene; locally, pedestrians turn to other destinations as they process to their destination, thereby increasing the global energy. It is necessary to learn more about the destination and global energy.

Contributors

Ning DING designed the algorithm, conducted experiments, and drafted the manuscript. Weimin QI helped organize the manuscript. Huihuan QIAN supervised the research. Ning DING, Weimin QI, and Huihuan QIAN revised and finalized the paper.

Compliance with ethics guidelines

Ning DING, Weimin QI, and Huihuan QIAN declare that they have no conflict of interest.

References

Ali S, Shah M, 2007. A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. IEEE Conf on Computer Vision and Pattern Recognition, p.1-6. https://doi.org/10.1109/CVPR.2007.382977

Anh NTN, Daniel ZJ, Du NH, et al., 2011. A hybrid macro-micro pedestrians evacuation model to speed up simulation in road networks. Int Conf on Autonomous Agents and Multiagent Systems, p.371-383. https://doi.org/10.1007/978-3-642-27216-5_28

Bladel J, 1959. On Helmholtz's theorem in finite regions. *IRE Trans Antenn Propag*, 7(5):119. https://doi.org/10.1109/TAP.1959.1144767

Brennen CE, 2005. Fundamentals of Multiphase Flow. Cambridge University Press, Cambridge, UK, p.386.

Brostow GJ, Cipolla R, 2006. Unsupervised Bayesian detection of independent motion in crowds. IEEE Conf on Computer Vision and Pattern Recognition, p.594-601. https://doi.org/10.1109/CVPR.2006.320

Cao T, Wu XY, Guo JN, et al., 2009. Abnormal crowd motion analysis. IEEE Int Conf on Robotics and Biomimetics, p.1709-1714. https://doi.org/10.1109/ROBIO.2009.5420408

Dabrowski M, Krotkiewski M, Schmid DW, 2008. Milamin: Matlab-based finite element method solver for large problems. *Geochem Geophys Geosyst*, 9(4):Q04030. https://doi.org/10.1029/2007GC001719

Helbing D, Farkas IJ, Molnar P, et al., 2002. Simulation of pedestrian crowds in normal and evacuation situations. In: Schreckenberg M, Sharma SD (Eds.), Pedestrian and Evacuation Dynamics. Springer, Berlin, p.21-58.

Helmuth JA, Burckhardt CJ, Koumoutsakos P, et al., 2007. A novel supervised trajectory segmentation algorithm identifies distinct types of human adenovirus motion in host cells. *J Struct Biol*, 159(3):347-358. https://doi.org/10.1016/j.jsb.2007.04.003

Huet S, Karatekin E, Tran VS, et al., 2006. Analysis of transient behavior in complex trajectories: application to secretory vesicle dynamics. *Biophys J*, 91(9):3542-3559. https://doi.org/10.1529/biophysj.105.080622

Ihaddadene N, Djeraba C, 2008. Real-time crowd motion analysis. 19th Int Conf on Pattern Recognition, p.1-4. https://doi.org/10.1109/ICPR.2008.4761041

Ijaz K, Sohail S, Hashish S, 2015. A survey of latest approaches for crowd simulation and modeling using hybrid techniques. 17th UKSIM-AMSS Int Conf on Modelling and Simulation, p.111-116. https://doi.org/10.1109/UKSim.2015.46

Karnik A, Goswami S, Guha R, 2007. Detecting obfuscated viruses using cosine similarity analysis. 1st Asia Int Conf on Modelling and Simulation, p.165-170. https://doi.org/10.1109/AMS.2007.31

Mehran R, Oyama A, Shah M, 2009. Abnormal crowd behavior detection using social force model. IEEE Conf on Computer Vision and Pattern Recognition, p.935-942. https://doi.org/10.1109/CVPR.2009.5206641

Muchoney D, Williamson J, 2001. A Gaussian adaptive resonance theory neural network classification algorithm applied to supervised land cover mapping using multitemporal vegetation index data. *IEEE Trans Geosci Remote Sens*, 39(9):1969-1977. https://doi.org/10.1109/36.951087

Naseer T, Burgard W, Stachniss C, 2018. Robust visual localization across seasons. *IEEE Trans Rob*, 34(2):289-302. https://doi.org/10.1109/TRO.2017.2788045

Nurgaliev D, McDonald M, Benson B, et al., 2013. A robust quantification of galaxy cluster morphology using asymmetry and central concentration. *Astrophys J*, 779(2): 112. https://doi.org/10.1088/0004-637X/779/2/112

Patil S, van den Berg J, Curtis S, et al., 2010. Directing crowd simulations using navigation fields. *IEEE Trans Vis Comput Graph*, 17(2):244-254. https://doi.org/10.1109/TVCG.2010.33

Rabaud V, Belongie S, 2006. Counting crowded moving objects. IEEE Conf on Computer Vision and Pattern Recognition, p.705-711. https://doi.org/10.1109/CVPR.2006.92

Rodriguez M, Sivic J, Laptev I, et al., 2011. Data-driven crowd analysis in videos. Int Conf on Computer Vision, p.1235-1242. https://doi.org/10.1109/ICCV.2011.6126374

Shi JB, Tomasi C, 1994. Good features to track. IEEE Conf on Computer Vision and Pattern Recognition, p.593-600. https://doi.org/10.1109/CVPR.1994.323794

Shiwakoti N, Sarvi M, Burd M, 2014. Using non-human biological entities to understand pedestrian crowd behaviour under emergency conditions. *Safety Sci*, 66:1-8. https://doi.org/10.1016/j.ssci.2014.01.010

Sreenu G, Durai MAS, 2019. Intelligent video surveillance: a review through deep learning techniques for crowd analysis. *J Big Data*, 6(1):48. https://doi.org/10.1186/s40537-019-0212-5

Tripathi G, Singh K, Vishwakarma DK, 2018. Convolutional neural networks for crowd behaviour analysis: a survey. *Vis Comput*, 35(5):753-776. https://doi.org/10.1007/s00371-018-1499-5

Xiong MZ, Cai WT, Zhou SP, et al., 2009. A case study of multi-resolution modeling for crowd simulation. Proc Spring Simulation Multiconf, p.1-8. https://doi.org/10.5555/1639809.1639827

Xiong MZ, Tang SY, Zhao D, 2013. A hybrid model for simulating crowd evacuation. *New Gener Comput*, 31(3):211-235. https://doi.org/10.1007/s00354-013-0304-2

Ye J, 2011. Cosine similarity measures for intuitionistic fuzzy sets and their applications. *Math Comput Modell*, 53(1-2):91-97. https://doi.org/10.1016/j.mcm.2010.07.022

Yoon DD, Ayalew B, 2018. Social force control for human-like autonomous driving. ASME Int Design Engineering Technical Conf and Computers and Information in Engineering Conf, Article V003T01A003. https://doi.org/10.1115/DETC2018-86288

Zhang XG, Yu QN, Yu H, 2018. Physics inspired methods for crowd video surveillance and analysis: a survey. *IEEE Access*, 6:66816-66830. https://doi.org/10.1109/ACCESS.2018.2878733

Zhao M, Cai W, Turner SJ, 2018. CLUST: simulating realistic crowd behaviour by mining pattern from crowd videos. *Comput Graph Forum*, 37(1):184-201. https://doi.org/10.1111/cgf.13259

Zhou BL, Tang XO, Wang XG, 2012a. Coherent filtering: detecting coherent motions from crowd clutters. European Conf on Computer Vision, p.857-871. https://doi.org/10.1007/978-3-642-33709-3_61

Zhou BL, Wang XG, Tang XO, 2012b. Understanding collective crowd behaviors: learning a mixture model of dynamic pedestrian-agents. IEEE Conf on Computer Vision and Pattern Recognition, p.2871-2878. https://doi.org/10.1109/CVPR.2012.6248013

Zhou BL, Tang XO, Wang XG, 2013. Measuring crowd collectiveness. IEEE Conf on Computer Vision and Pattern Recognition, p.3049-3056. https://doi.org/10.1109/CVPR.2013.392

Zhou MY, Carin L, 2015. Negative binomial process count and mixture modeling. *IEEE Trans Patt Anal Mach Intell*, 37(2):307-320. https://doi.org/10.1109/TPAMI.2013.211