Transportation Research Part C 152 (2023) 104154

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

A force-based model for adaptively controlling the spatial configuration of pedestrian subgroups at non-extreme densities

Wenhan Wu, Wenfeng Yi, Xiaolu Wang, Xiaoping Zheng *

Department of Automation, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Keywords:
Pedestrian subgroup
Crowd dynamics
Spatial configuration
Force-based model
Adaptive control
Walking behavior

ABSTRACT

The study of crowd dynamics has provided new insights into the understanding of human collective motion. However, most previous studies treated large-scale crowds as consisting of isolated individuals, but ignored the fact that pedestrian subgroups are prevalent in reality. With the increasing advancement of computer simulations, the critical role of subgroups in crowd modeling has been gradually recognized. Here, we develop a force-based model to reproduce the walking behaviors of pedestrian subgroups, in which the quantitative laws extracted from public datasets are incorporated for controlling their spatial configurations at different non-extreme densities. Numerical simulations indicate that our model achieves the simulation of pedestrian subgroups that coincides with empirical observations, and presents better simulation performance than existing subgroup models. The presence of pedestrian subgroups has also been confirmed to have certain effects on both fundamental diagrams and lane formation. Overall, this paper contributes a valuable framework to the modeling of pedestrian subgroups for guiding relevant implementations in potential application areas.

1. Introduction

Collective motion is an important topic emerging in complex social systems, from small-scale organisms such as cells and bacteria to large-scale ones such as schooling fish and sheep herds (Vicsek and Zafeiris, 2012). With the increasing frequency of mass activities, recent trends in collective motion have led to a proliferation of studies involving human crowds (Bain and Bartolo, 2019; Xu et al., 2021). The modeling of human collective motion has been attracting a lot of interest, since the development of computational social science enables researchers to use computer simulations to reproduce various empirically observed crowd behaviors (Lazer et al., 2020; Zhou et al., 2019). Despite a lot of crowd models merely considering the interaction between isolated individuals, the realistic evidence denotes that most pedestrians do not walk alone (Nicolas and Hassan, 2023). As an intermediate layer from isolated individuals to human crowds, pedestrian subgroups are prevalent and represent those who move together in a congregated form based on social relationships. Therefore, establishing human motion models involving pedestrian subgroups is a central concern, which plays a key role in understanding crowd dynamics (Mukherjee et al., 2015) and managing large-scale events (Helbing et al., 2001).

A considerable amount of literature has been published on the modeling of pedestrian subgroups, these studies mainly focus on the improvement of several classical human motion models: social force models (SFMs) (Helbing and Molnár, 1995; Helbing et al., 2000), cellular automata models (CAMs) (Burstedde et al., 2001), and agent-based models (ABMs) (Bonabeau, 2002). The SFMs are continuous in spatio-temporal dimensions, and the group forces originating from socio-psychological definitions are generally

* Corresponding author.
E-mail address: asean@mail.tsinghua.edu.cn (X. Zheng).

https://doi.org/10.1016/j.trc.2023.104154
Received 5 December 2022; Received in revised form 3 April 2023; Accepted 26 April 2023
Available online 11 May 2023

added to the equations of motion (Moussaïd et al., 2010; Turgut and Bozdag, 2021), for the purpose of reflecting the characteristics of pedestrian subgroups. The CAMs are discrete in spatio-temporal dimensions, which formulate a series of local state update rules by introducing aggregation dynamics (Crociani et al., 2018) and behavioral features (Lu et al., 2017) into the ground fields to model the movement of pedestrian subgroups. The ABMs can be continuous or discrete in spatio-temporal dimensions, and often adopt the structure matrix (Qiu and Hu, 2010), behavioral rules (Hussein and Sayed, 2016), attribute sets (Crociani et al., 2013), and avoidance mechanisms (Karamouzas and Overmars, 2012) to simulate subgroup behaviors. Despite that CAMs and ABMs respectively have the advantages of fast computing efficiency (Li et al., 2019) and high flexible scalability (Bonabeau, 2002), we still prefer the SFM inspired by Newtonian mechanics as a basic model in this paper. The reasons are mainly reflected in two aspects of reality and fineness: One is that SFMs describe pedestrian motion by a set of forces reflecting external influences and internal motivations, rather than directly defining the rules of behaviors, which can more realistically characterize the decision process of individuals among different movement choices. The other is that SFMs are continuous and have infinitely high spatio-temporal resolution compared with CAMs and most ABMs, which can more finely reproduce human continuous behaviors in real physical space.

Most studies of subgroup modeling based on the SFM have been developed following two aspects. One aspect suggests that the leader–follower principle can be introduced to model the motion of pedestrian subgroups. Li et al. (2017) incorporated the cohesive force generated by social relations and the attractive force from group leaders into the SFM, and simulated a more realistic evacuation process. Zhang et al. (2018) proposed a two-layer SFM considering the interaction between leaders and members to reproduce the group behavior in earthquake evacuation. Xie et al. (2021) introduced the social identity theory of leadership into the group force derived from Lennard-Jones potential, and related simulations denoted that the group effect promotes the evacuation performance. Although the leader–follower pattern appears to be common in emergency evacuations, it is obviously different from the walking behaviors under normal conditions. The other aspect argues that pedestrian subgroups maintain coordinated movement through cohesive interactions. Xu and Duh (2010) improved the SFM by including the bonding effects of group cohesion to explain the observed phenomena of walking delay and overtaking. Moussaïd et al. (2010) combined a vision field, an attraction of the mass center, and a repulsion of group members into the SFM, whose simulations are consistent with empirical observations. Huang et al. (2018) designed a social group force model (SGFM) based on the behaviors of inter-group coordination and intra-group avoidance, and verified it by collected videos and virtual reality (VR) experiments. However, the behavioral patterns of pedestrian subgroups presented by these models are relatively fixed (i.e., strongly dependent on parameter settings before simulation), and cannot be adjusted adaptively with the surrounding environment as in realistic situations.

To address the above critical problems of these models, we develop a force-based model to simulate the walking behaviors of pedestrian subgroups. Based on data extracted from six public datasets, the quantitative laws of spatial configurations of pedestrian subgroups are discovered. The mathematical functions of these laws are incorporated into the expression of the group control force to adjust the relative distance and relative angle between subgroup members. Numerical simulations indicate that our model presents the organized patterns of pedestrian subgroups at different non-extreme densities, and also exhibits stronger reproducibility than existing subgroup models. In addition, we respectively conduct numerical simulations on the motion of pedestrian subgroups in unidirectional and bidirectional flows. Simulation snapshots and quantitative analysis demonstrate that the presence of pedestrian subgroups has non-negligible effects on both fundamental diagrams and lane formation. These findings provide valuable guidance for some potential application areas such as crowd management, facility design, and animation production.

The rest of this paper is organized as follows. Section 2 describes the extraction of quantitative laws of spatial configurations. The mathematical model of pedestrian subgroups is proposed in Section 3. Section 4 provides numerical simulations and corresponding analysis details. In Section 5, the primary conclusions and future prospects are summarized.

2. Extracting quantitative laws

2.1. Pedestrian subgroups identified from public datasets

To reveal the quantitative laws implicit in spatial configurations of pedestrian subgroups, six public datasets from various outdoor environments are selected for subsequent analysis. The former two datasets labeled Seq_eth and Seq_hotel were recorded from a bird-eye view with a frame rate of 2.5FPS, which include walking pedestrians in busy scenarios outside a college building and a bus station (Pellegrini et al., 2009). The following three datasets named Crowds_zara01, Crowds_zara02, and Students003 were taken in public spaces from the top-view at 25FPS, and captured the movement of pedestrians in a shopping street and a college campus (Lerner et al., 2007). The last dataset called VEG_gall is a five-minute video sequence collected with a frame rate of 8FPS, covering pedestrian crowds with density changes in Vittorio Emanuele II Gallery (Bandini et al., 2014). It is worth noting that the crowd densities in Students003 and VEG_gall are relatively high, whereas those in the other four datasets are relatively low. Table 1 presents further details of the above datasets, involving the year, location, number of pedestrians, flow type, and density range (i.e., the minimum and maximum numbers of pedestrians divided by the estimated area of the collection region).

In these datasets, pedestrian subgroups are discovered based on an automatic algorithm developed by our team (Wu et al., 2023), and a few misidentified cases have also been manually corrected to improve the identification accuracy. Fig. 1(a) displays the identification examples from different public datasets, in which pedestrians belonging to the same subgroup are marked with red circles. Those pedestrians standing together (e.g., chatting) are not taken into account, because we mainly focus on the walking behaviors of pedestrian subgroups. Totally, we extract the pedestrian labels for 1824 isolated individuals, 586 subgroups of 2 members, 113 subgroups of 3 members, 19 subgroups of 4 members, and 12 subgroups of more than 5 members. As shown in

Table 1
Further details of public datasets.

Dataset	Year	Location	No. Ped	Flow type	Density range
Seq_eth	2009	Zurich, Switzerland	360	Bidirectional	0.007 ~ 0.181 m^{-2}
Seq_hotel	2009	Zurich, Switzerland	389	Bidirectional	0.011 ~ 0.192 m^{-2}
Crowds_zara01	2007	Nicosia, Cyprus	148	Bidirectional	0.012 ~ 0.239 m^{-2}
Crowds_zara02	2007	Nicosia, Cyprus	204	Bidirectional	0.012 ~ 0.227 m^{-2}
Students003	2007	Tel Aviv, Israel	434	Multidirectional	0.105 ~ 0.410 m^{-2}
VEG_gall	2014	Milan, Italy	630	Multidirectional	0.057 ~ 0.403 m^{-2}

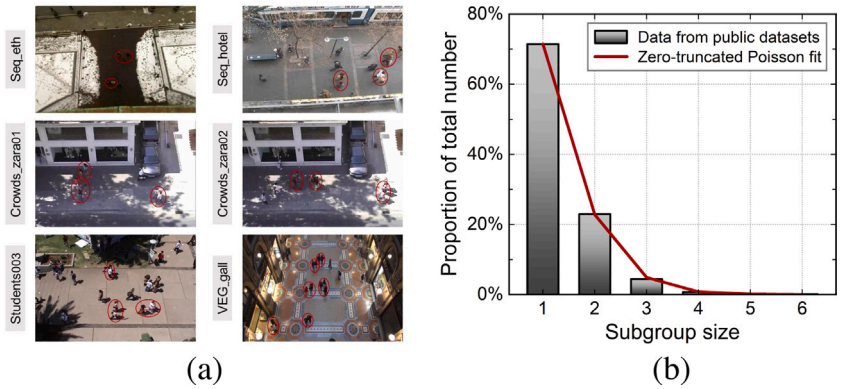

Fig. 1. Pedestrian subgroups identified from public datasets. (a) Identification examples of pedestrian subgroups, those belonging to the same subgroup are marked with red circles. (b) The proportion distribution of subgroup size, and the red line corresponds to a zero-truncated Poisson fit.

Fig. 1(b), the proportion decreases rapidly as the subgroup size increases, which can be approximated with a zero-truncated Poisson distribution, as given by the following equation:

$$P(s) = \frac{\exp(-\lambda) \cdot \lambda^s}{s!\left[1 - \exp(-\lambda)\right]} \tag{1}$$

where s indicates the subgroup size, and the parameter is estimated as $\lambda = 0.64$. It tallies with the distribution characteristics of subgroup size (Coleman and James, 1961), which demonstrates the reliability of subgroup identification in these datasets. Due to the fact that the subgroup size exceeding 4 is rare, those subgroups of 2 to 4 members are regarded as the main research objects.

2.2. Definition of state variables for pedestrian subgroups

From the above identification results, we are able to extract the trajectories of subgroup members according to their IDs. These trajectories imply the relative positions of subgroup members at different times, which contain 144804, 20928, and 1943 samples for subgroups of 2, 3, and 4 members, respectively. This number of samples is sufficient to effectively observe the organized patterns of pedestrian subgroups when walking. Fig. 2(a) performs the observed spatial configurations for pedestrian subgroups with different sizes, and a darker area reflects a higher frequency of members appearing at this position. The red filled circles represent the average positions of subgroup members, which are connected to reveal the average patterns of pedestrian subgroups. For subgroups of 2 members, they prefer to form linear formations orthogonal to the walking direction. In subgroups of sizes 3 and 4, the pedestrians in middle positions stand slightly behind to those on both sides to form bending formations. These results suggest that the walking behaviors of pedestrian subgroups can give rise to special patterns of spatial organization.

The previous empirical research has demonstrated that at very low density, subgroup members tend to walk side by side due to the compatibility of walking faster and facilitating social interactions, but as the density reaches a moderate level, the growing importance of social interactions would lead to “V”-like or “U”-like spatial configurations (Moussaïd et al., 2010). However, little is known about how the increasing density quantitatively affects the transition of spatial configurations from horizontal to “V”-like or “U”-like formations, we therefore define several state variables to describe pedestrian subgroups in a mathematical form. Fig. 2(b) illustrates the relative distances and relative angles between subgroup members, defined by the nearest neighbor on the right-hand side of the individual. These state variables allow us to create more practical parameters, one is the average relative distance within subgroup G, which is expressed as follows:

$$\langle d_G \rangle = \frac{1}{n_G - 1} \sum_{i=1}^{n_G - 1} d_{i,i+1} \tag{2}$$

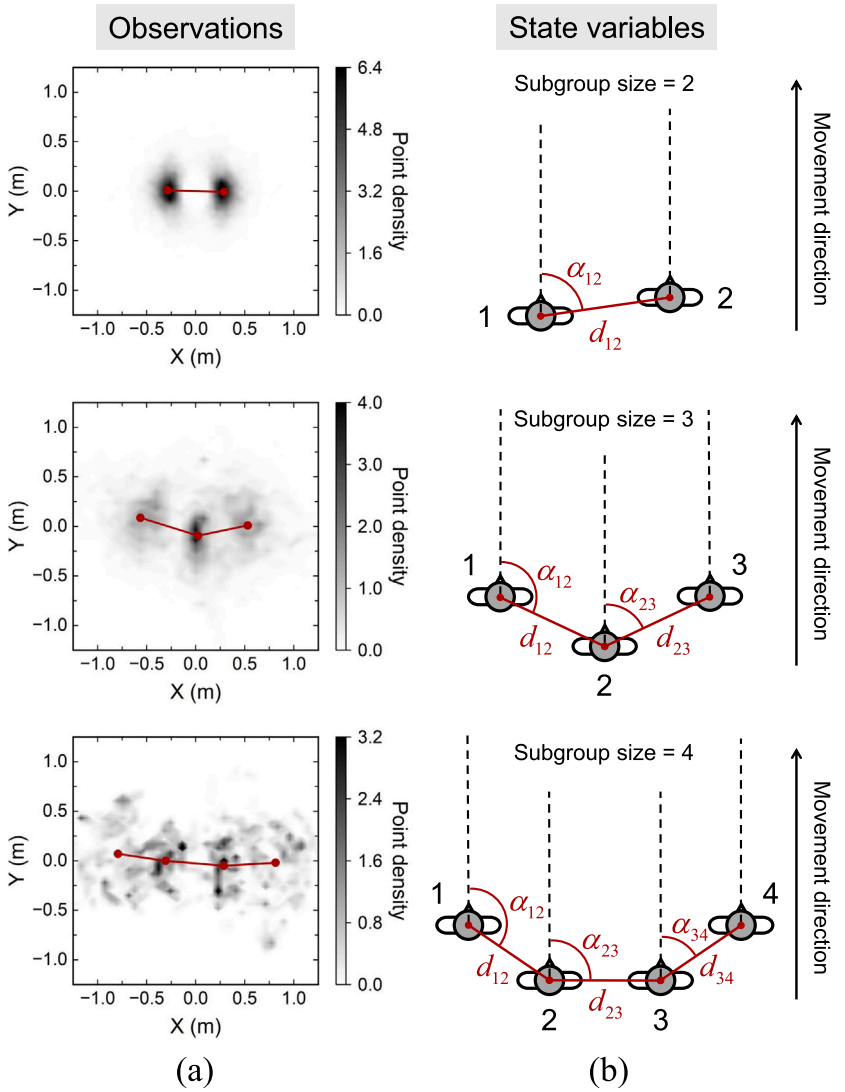

Fig. 2. Spatial configurations for pedestrian subgroups with different sizes. (a) Organized patterns of pedestrian subgroups from the observed data, where the red filled circles represent the average positions of subgroup members. (b) The definition of state variables, including the relative distances and relative angles between subgroup members.

where n_G is the number of subgroup members. The other is the average relative angle that characterizes the concavity degree of subgroup G, which is given by:

$$\langle \alpha_G \rangle = \begin{cases} \alpha_{12}, n_G = 2 \\ \left[\alpha_{12} + \left(180° - \alpha_{n_G-1,n_G}\right)\right]/2, n_G = 3,4 \end{cases} \tag{3}$$

For subgroups of 2 members, α_{12} is sufficient to describe their organized patterns. In subgroups of sizes 3 and 4, α_{12} and $180°-\alpha_{n_G-1,n_G}$ correspond to the angles of the two outer members with respect to the middle pedestrians, the average of which can be adopted to measure the concavity of "V"-like and "U"-like formations.

2.3. Quantitative laws of spatial configurations

Our goal here is to explore the quantitative laws of spatial configurations at different non-extreme densities. The classical density is the number of pedestrians divided by the area of the region, but this calculation comes at the cost of resolution since large jumps of exact placements are reduced by taking averages (Steffen and Seyfried, 2010). From this, we consider that the density defined

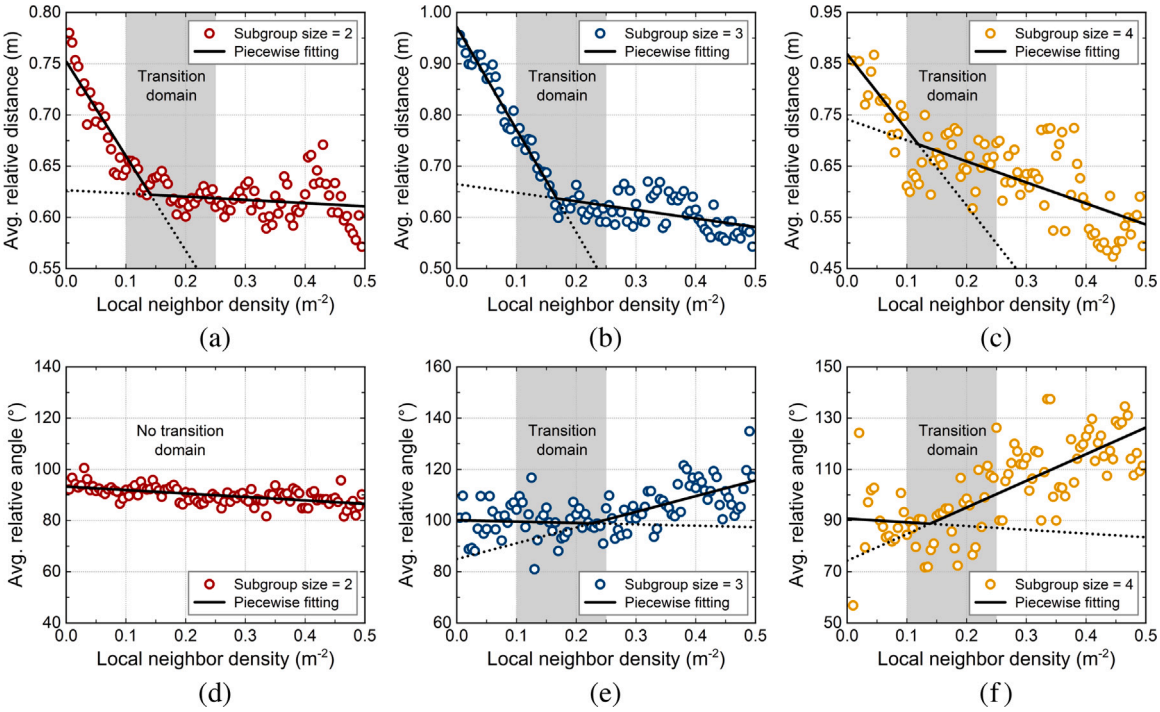

Fig. 3. Quantitative laws of spatial configurations under different local neighbor densities. (a)–(c) The average relative distance as a function of the local neighbor density for subgroup size = 2, 3, and 4. (d)–(f) The average relative angle as a function of the local neighbor density for subgroup size = 2, 3, and 4. These circles represent the observed data from datasets, the solid and dotted lines indicate the fitting curves and extension lines, respectively. The gray shaded areas mark the transition domains, ranging from 0.1 m^{-2} to 0.25 m^{-2}.

within a certain range around the subgroup is more accurate, which reflects the influence of local environment. The local neighbor density of subgroup G is therefore defined as follows:

$$\rho_G\left(\mathbf{x}_G\right) = \sum_{j \notin G} f\left(\mathbf{x}_j - \mathbf{x}_G\right) \tag{4}$$

Here, $\mathbf{x}_G$ stands for the position of the mass center of subgroup G, $\mathbf{x}_j$ are the positions of other pedestrians $j \notin G$ in the surrounding of $\mathbf{x}_G$. Note that subgroup members $j \in G$ are not included as they account for a large proportion in the calculation and weaken the impact of surrounding pedestrians. The weight function $f_w(d_{ij})$ decreases as distance d_{ij} increases, which is represented in the form of a Gaussian kernel function:

$$f\left(\mathbf{x}_j - \mathbf{x}_G\right) = \frac{1}{\pi R^2} \exp\left(-\left\|\mathbf{x}_j - \mathbf{x}_G\right\|^2 / R^2\right) \tag{5}$$

where R is a parameter reflecting the size of the influence area, and we take $R = 3.66$ m as a reasonable evaluation, in accordance with the maximum radius of the social zone (Bethel and Murphy, 2008).

Based on the extracted samples, we calculate the local neighbor density, average relative distance, and average relative angle for each pedestrian subgroup. To reveal the organized patterns from a statistical perspective, $\langle d_G \rangle$ and $\langle \alpha_G \rangle$ are averaged within certain intervals, which are obtained by equally dividing $\rho_G(\mathbf{x}_G)$. Fig. 3 presents the quantitative laws of spatial configurations under different local neighbor densities, these circles with red, blue, and yellow colors represent the data corresponding to subgroups of 2, 3, and 4 members. As $\rho_G(\mathbf{x}_G)$ increases in Fig. 3(a)–(c), $\langle d_G \rangle$ first declines with a larger slope and changes slowly after a transition point. The explanation might be that even if the surrounding is crowded, subgroup members who walk almost close to each other can no longer reduce their distances due to the constraints of physical space. Besides, Fig. 3(d)–(f) indicate that $\langle \alpha_G \rangle$ almost remains around 90° for subgroups of 2 members, while it first tends to be stable and then increases at a faster rate for subgroups of 3 and 4 members. This is attributed to the fact that subgroup members have to coordinate social interactions by enhancing the concavity degree of their structures, which is noticeable after $\rho_G(\mathbf{x}_G)$ exceeding a transition point.

Apart from Fig. 3(d), we notice that obvious transition points exist for both $\langle d_G \rangle$ and $\langle \alpha_G \rangle$, which are distributed in the transition domain from 0.1 m^{-2} to 0.25 m^{-2}. The linear slopes before and after the transition point are apparently different in these figures, whereby the piecewise fitting is employed to extract the mathematical functions. Here, these fitting curves are treated as the explicit relationships between preferred relative distance $d_G^*(\rho_G)$ or preferred relative angle $\alpha_G^*(\rho_G)$ with local neighbor density $\rho_G(\mathbf{x}_G)$, which

Table 2
Parameter values in piecewise linear functions.

Linear function	Subgroup size	Piecewise 1		Piecewise 2		Transition point
		k_1	b_1	k_2	b_2	ρ_G^{tp}
$d_G^*(\rho_G)$	2	−0.922	0.752	−0.032	0.627	0.141 m^{-2}
$d_G^*(\rho_G)$	3	−2.001	0.971	−0.167	0.665	0.167 m^{-2}
$d_G^*(\rho_G)$	4	−1.478	0.869	−0.411	0.742	0.119 m^{-2}
$\alpha_G^*(\rho_G)$	2	−13.478	93.271	−13.478	93.271	–
$\alpha_G^*(\rho_G)$	3	−5.448	100.143	61.408	84.987	0.227 m^{-2}
$\alpha_G^*(\rho_G)$	4	−14.561	90.766	103.942	74.338	0.139 m^{-2}

are expressed by the following piecewise linear functions:

$$\boldsymbol{\Phi}_G^*\left(\rho_G\right) = \begin{cases} k_1 \cdot \rho_G\left(\mathbf{x}_G\right) + b_1, 0 \le \rho_G\left(\mathbf{x}_G\right) \le \rho_G^{tp} \\ k_2 \cdot \rho_G\left(\mathbf{x}_G\right) + b_2, \rho_G^{tp} < \rho_G\left(\mathbf{x}_G\right) \le \rho_G^{\max} \end{cases} \tag{6}$$

where $\boldsymbol{\Phi}$ is a symbol denoting distance ($\boldsymbol{\Phi} = d$) or angle ($\boldsymbol{\Phi} = \alpha$), k_1 and k_2 are the slopes of piecewise linear functions, b_1 and b_2 are corresponding intercepts, ρ_G^{tp} and ρ_G^{max} are the transition point and maximum value of the local neighbor density, respectively. The above equation reveals the quantitative laws of spatial configurations changing with the density, which is beneficial for establishing a model to reproduce more realistic subgroup behaviors.

The values of k_1, k_2, b_1, b_2, and ρ_G^{tp} in piecewise linear functions are listed in Table 2, which can be easily calculated from the extracted data. However, due to the limited range of the local neighbor density in these datasets, the values of ρ_G^{max} need to be taken by extrapolating the explicit relationships in Fig. 3. Here, two extrapolation conditions should be satisfied simultaneously: First, the preferred relative distance cannot be lower than the pedestrian radius (see Table 3) under the tolerance of partial body overlap (e.g., side-shoulder behavior) of neighboring members (i.e., $d_G^*\left(\rho_G\right) \ge 0.25$ m). Second, the preferred relative angle cannot destroy the order of subgroup members (see Fig. 2(b)) defined in orthogonal to the walking direction (i.e., $0° \le \alpha_G^*\left(\rho_G\right) \le 180°$). For subgroups with a specific size, the values of ρ_G^{max} for the preferred relative distance and preferred relative angle should be consistent, which is determined as the minimum derived from the two above conditions. From this, we have $\rho_G^{max} = 6.920$ m^{-2}, 1.547 m^{-2}, and 1.017 m^{-2} for subgroup size = 2, 3, and 4, respectively.

3. Mathematical model

3.1. Modeling the motion of pedestrian subgroups

The SFM adopts the nonlinear coupled Langevin equation to describe the change of pedestrian acceleration with the interaction of multiple forces (Helbing et al., 2000). On this basis, a group control force is introduced to simulate the movement process of pedestrian subgroups, and pedestrian i belonging to subgroup G satisfies the following differential equation:

$$m_i \frac{d\mathbf{v}_i(t)}{dt} = \mathbf{f}_{id} + \sum_{j \notin G} \mathbf{f}_{ij} + \sum_{W} \mathbf{f}_{iW} + \sum_{q \in G} \mathbf{f}_{iq}^G \tag{7}$$

in which the motion of pedestrian i is driven by the four forces on the right side of the above equation: (1) $\mathbf{f}_{id}$ is a force driving pedestrians to move towards the target at a certain velocity; (2) $\mathbf{f}_{ij}$ is an interaction force with pedestrians outside subgroup G; (3) $\mathbf{f}_{iW}$ is an interaction force with obstacles or walls; (4) $\mathbf{f}_{iq}^G$ is a group control force with neighboring members within subgroup G.

The mathematical form corresponding to the self-driven force $\mathbf{f}_{id}$ of pedestrian i is given by:

$$\mathbf{f}_{id} = m_i \frac{v_i^0 \mathbf{e}_i^0 - \mathbf{v}_i(t)}{\tau_i} \tag{8}$$

where m_i is the mass of pedestrian i, v_i^0 and $\mathbf{e}_i^0$ respectively represent the desired speed and desired direction, τ_i denotes a time constant, which is related to the relaxation time for adapting to his or her actual velocity $\mathbf{v}_i(t)$.

The function of the interaction force $\mathbf{f}_{ij}$ with pedestrian j outside subgroup G is defined as:

$$\mathbf{f}_{ij} = A_i \exp\left[\left(r_{ij} - d_{ij}\right)/B_i\right]\mathbf{n}_{ij} + kg\left(r_{ij} - d_{ij}\right)\mathbf{n}_{ij} + \kappa g\left(r_{ij} - d_{ij}\right)\Delta v_{ji}^t \mathbf{t}_{ij} \tag{9}$$

Here, the psychological tendency of pedestrians i and j to move away from each other is reflected by the repulsive interaction $A_i \exp\left[\left(r_{ij} - d_{ij}\right)/B_i\right]\mathbf{n}_{ij}$, where A_i and B_i are constants, d_{ij} is the distance between the centroids of two pedestrians, r_{ij} is the sum of their radii r_i and r_j, $\mathbf{n}_{ij}$ is the normalized vector pointing from pedestrian j to i. Besides, two additional contact forces are "body force" $kg\left(r_{ij} - d_{ij}\right)\mathbf{n}_{ij}$ and "sliding friction force" $\kappa g\left(r_{ij} - d_{ij}\right)\Delta v_{ji}^t \mathbf{t}_{ij}$, in which $g(x)$ is zero if two pedestrians do not contact each other ($r_{ij} < d_{ij}$), otherwise it is equal to x. k and κ are body elasticity and sliding friction coefficients standing for the obstruction effects, $\mathbf{t}_{ij} = \left(-\mathbf{n}_{ij}^2, \mathbf{n}_{ij}^1\right)$ holds the tangential direction and $\Delta v_{ji}^t = \left(\mathbf{v}_j - \mathbf{v}_i\right) \cdot \mathbf{t}_{ij}$ is the tangential velocity difference.

The expression of the interaction force $\mathbf{f}_{iW}$ with obstacles or walls is similar to the above equation, which reads:

$$\mathbf{f}_{iW} = A_i \exp\left[\left(r_i - d_{iW}\right)/B_i\right]\mathbf{n}_{iW} + kg\left(r_i - d_{iW}\right)\mathbf{n}_{iW} - \kappa g\left(r_i - d_{iW}\right)\left(\mathbf{v}_i \cdot \mathbf{t}_{iW}\right)\mathbf{t}_{iW} \tag{10}$$

where d_{iW} is the distance from the centroid of pedestrian i to the obstacle or wall, $\mathbf{n}_{iW}$ denotes the normalized vector perpendicular to it, and $\mathbf{t}_{iW}$ is the direction tangential to it.

The group control force $\mathbf{f}_{iq}^{G}$ between member i and its neighboring member j on the right-hand side within subgroup G is composed of $\mathbf{f}_{iq}^{G\_d}$ and $\mathbf{f}_{iq}^{G\_\alpha}$, which control the relative distance and relative angle in the connecting and orthogonal directions, respectively. To simplify the modeling process, the spatial configurations for subgroups of 3 and 4 members are assumed to be symmetrical with respect to the centerline. In this case, the relative distances between neighboring members are equal, and the outer angles satisfy $\alpha_{12} = 180° - \alpha_{n_G-1,n_G}$.

On the one hand, the interaction between neighboring members within subgroup G should be long-range attractive and short-range repulsive. The potential describing realistic intermolecular interactions is considered here, whose general form is written as:

$$V\left(d_{iq}\right) = 4\varepsilon\left[\left(\frac{\sigma}{d_{iq}}\right)^{2n} - \left(\frac{\sigma}{d_{iq}}\right)^{n}\right] \tag{11}$$

Here, as one of the most widely used intermolecular potentials, the Lennard-Jones potential ($n = 6$) has been applied to control the relative distance between subgroup members (Xie et al., 2021). The first half term is the repulsive interaction (Pauli repulsion), and the second half term is the attractive interaction (London dispersion), d_{iq} represents the distance between the centroids of members i and q, σ is a critical distance at which the potential is zero, ε is the well depth. If $V'(d_{iq}) = 0$, $d_{iq}^{e} = 2^{1/6}\sigma$ corresponds to the equilibrium distance between attractive and repulsive forces (i.e., the potential reaches a minimum). However, there are two issues with the subgroup force based on Lennard-Jones potential: First, the growth rate of repulsion is extremely fast when $d_{iq} < d_{iq}^{e}$, it probably pushes the two members away from each other instantaneously. Second, the attraction first increases to the maximum with d_{iq} and then converges to zero at a faster rate when $d_{iq} > d_{iq}^{e}$, making it hard for farther members to approach. For this, the exponential factor is reduced to $n = 1$ for alleviating these issues, and the distance control force $\mathbf{f}_{iq}^{G\_d}$ is therefore given by:

$$\mathbf{f}_{iq}^{G\_d} = -\nabla_{\mathbf{d}_{iq}} V\left(d_{iq}\right) = 4\varepsilon\left(2\frac{\sigma^2}{d_{iq}^3} - \frac{\sigma}{d_{iq}^2}\right)\mathbf{n}_{iq} \tag{12}$$

where $\mathbf{n}_{iq}$ is the normalized vector pointing from member q to i, and the equilibrium distance satisfying $\|\mathbf{f}_{iq}^{G\_d}\| = 0$ is equal to 2σ. For this, we regard $d_{iq}^{e}(\rho_G)$ as consistent with $d_G^{*}(\rho_G)$, which also depends on local neighbor density ρ_G. This equation is then rewritten as below:

$$\mathbf{f}_{iq}^{G\_d} = \lambda_d\left[\frac{d_{iq}^{e}\left(\rho_G\right)^2}{d_{iq}^3} - \frac{d_{iq}^{e}\left(\rho_G\right)}{d_{iq}^2}\right]\mathbf{n}_{iq} \tag{13}$$

where λ_d represents the regulating parameter for distance control force.

On the other hand, subgroup members tend to maintain specific spatial configurations to balance their social interactions (Moussaïd et al., 2010), whereby $\alpha_{iq}^{e}(\rho_G)$ is assumed to be a density-dependent equilibrium angle between members i and q. We therefore argue that, $\alpha_{iq}^{e}(\rho_G)$ is the same as $\alpha_G^{*}(\rho_G)$ if the two members are biased to the left side of the centerline, otherwise it equals to $180° - \alpha_G^{*}(\rho_G)$. The magnitude of relative angle deviation is defined as $\Delta\alpha_{iq} = |\alpha_{iq} - \alpha_{iq}^{e}(\rho_G)|$ if their actual relative angle is α_{iq}. In general, a larger deviation $\Delta\alpha_{iq}$ corresponds to a stronger tendency $T_{iq}(\Delta\alpha_{iq})$ for members adjusting to their preferred relative angles. However, the incremental relationship between $T_{iq}(\Delta\alpha_{iq})$ and $\Delta\alpha_{iq}$ can take many functional forms (e.g., linear, S-shaped, and inverse-S-shaped). Here, it is assumed to follow a classical Sigmoid function (S-shaped) for the following reasons: One is that when $\Delta\alpha_{iq}$ is small, the adjustment tendency is similar to the mode of "fine-tuning". The other is that when $\Delta\alpha_{iq}$ is large, this tendency tends to be saturated, since it is restricted by the physical acceleration of pedestrians. As a result, $T_{iq}(\Delta\alpha_{iq})$ as a function of $\Delta\alpha_{iq}$ is obtained as follows:

$$T_{iq}\left(\Delta\alpha_{iq}\right) = \frac{1}{1 + \exp\left[-k_\alpha\left(\Delta\alpha_{iq} - \Delta\alpha_f\right)\right]} \tag{14}$$

where k_α denotes the ascending gradient that controls the slope change, and $\Delta\alpha_f$ indicates a critical angle deviation at which the slope increases the fastest. Owing to the fact that the adjustment tendency has directionality, a sign function $\mathrm{sgn}(\cdot)$ is integrated into $T_{iq}(\Delta\alpha_{iq})$ to make it symmetrical with respect to the origin. In this case, the angle control force can be formally defined as $\mathbf{f}_{iq}^{G\_\alpha} = \lambda_\alpha \mathrm{sgn}\left(\alpha_{iq} - \alpha_{iq}^{e}\left(\rho_G\right)\right)T_{iq}\left(\Delta\alpha_{iq}\right)\mathbf{t}_{iq}$, whose explicit expression is given by:

$$\mathbf{f}_{iq}^{G\_\alpha} = \lambda_\alpha \frac{\mathrm{sgn}\left(\alpha_{iq} - \alpha_{iq}^{e}\left(\rho_G\right)\right)}{1 + \exp\left[-k_\alpha\left(\Delta\alpha_{iq} - \Delta\alpha_f\right)\right]}\mathbf{t}_{iq} \tag{15}$$

where λ_α corresponds to the regulating parameter for angle control force, and $\mathbf{t}_{iq} = (-n_{iq}^2, n_{iq}^1)$ stands for the tangential direction perpendicular to $\mathbf{n}_{iq}$.

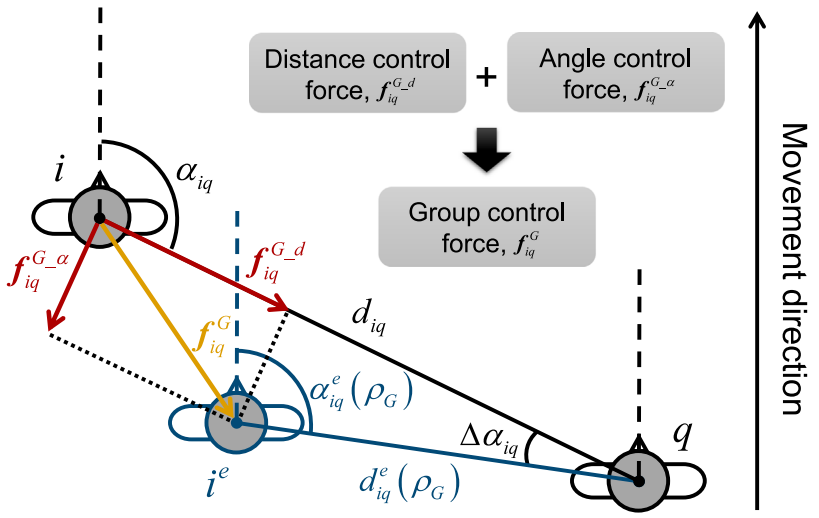

Fig. 4. Modeling schematic of the group control force. The group control force is composed of the distance control and angle control forces, which act along the connecting and orthogonal directions, respectively.

Based on these discussions and derivations, the final expression of group control force $\mathbf{f}_{iq}^{G}$ is organized as follows:

$$\mathbf{f}_{iq}^{G}=\lambda_d\left[\frac{d_{iq}^{e}(\rho_G)^2}{d_{iq}^{3}}-\frac{d_{iq}^{e}(\rho_G)}{d_{iq}^{2}}\right]\mathbf{n}_{iq}+\lambda_\alpha\frac{\operatorname{sgn}\left(\alpha_{iq}-\alpha_{iq}^{e}(\rho_G)\right)}{1+\exp\left[-k_\alpha\left(\Delta\alpha_{iq}-\Delta\alpha_f\right)\right]}\mathbf{t}_{iq} \tag{16}$$

For further aspects regarding the above definitions see the modeling schematic in Fig. 4. The group control force consists of the distance control and angle control forces, which act along the connecting and orthogonal directions, respectively. This force drives member i to move towards the expected position of invisible individual i^e, at which both relative distance d_{iq} and relative angle α_{iq} reach the equilibrium values.

3.2. Optimal selection of regulating parameters

The next section is concerned with the optimal selection of regulating parameters in this model. This aims to minimize the artifactual dynamics when subgroup members adapt to their preferred positions, which includes ensuring faster transition periods and reducing unrealistic back-and-forth oscillations. The process of adjusting the relative distance and relative angle is similar to that in typical control systems (Ang et al., 2005). Hence, two crucial dynamic performance indicators are selected as the evaluation criteria: One is the peak time required for the response to reach the first peak value:

$$t_p=\arg\max c(t) \tag{17}$$

where $c(t)$ represents the time response function, and its maximum value generally corresponds to the first peak value. The other is the overshoot denoting the deviation of the response at peak time from the final response divided by the final response value, which is expressed in the following form:

$$\sigma_p(\%)=\frac{c(t_p)-c(\infty)}{c(\infty)}\times 100\% \tag{18}$$

where $c(t_p)$ is the response value at peak time, and $c(\infty)$ is the final response value. To sum up, peak time t_p and overshoot σ_p evaluate the response speed and damping degree of the system, respectively.

These two indicators are then utilized to ascertain the optimal domains of regulating parameters λ_d and λ_α. Under the conditions that the relative angle (or distance) remains unchanged, we calculate the corresponding t_p and σ_p for a certain value of λ_d (or λ_α) at different relative distances (or angles), which are taken averaged to reflect statistical characteristics. Fig. 5(a) and 5(b) both illustrate that peak time t_p is a decreasing function of regulating parameters λ_d and λ_α, while overshoot σ_p is an increasing function of them. This implies that a larger regulating parameter makes the response speed faster, but increases the oscillation between individuals; however, a smaller regulating parameter would slow down the response speed, even though it makes the adjustment process more stable. For this, we attempt to find a trade-off range where the response speed and damping degree are moderate, $\lambda_d\in[500,650]$ and $\lambda_\alpha\in[350,425]$ are finally estimated as the optimal domains for these two regulating parameters.

For the sake of unification, $\lambda_d=600$ and $\lambda_\alpha=400$ are selected as the optimal values in subsequent simulations, this also allows us to present the features of distance control force $\mathbf{f}_{iq}^{G\_d}$ and angle control force $\mathbf{f}_{iq}^{G\_\alpha}$ in Fig. 6. It is supposed that the equilibrium

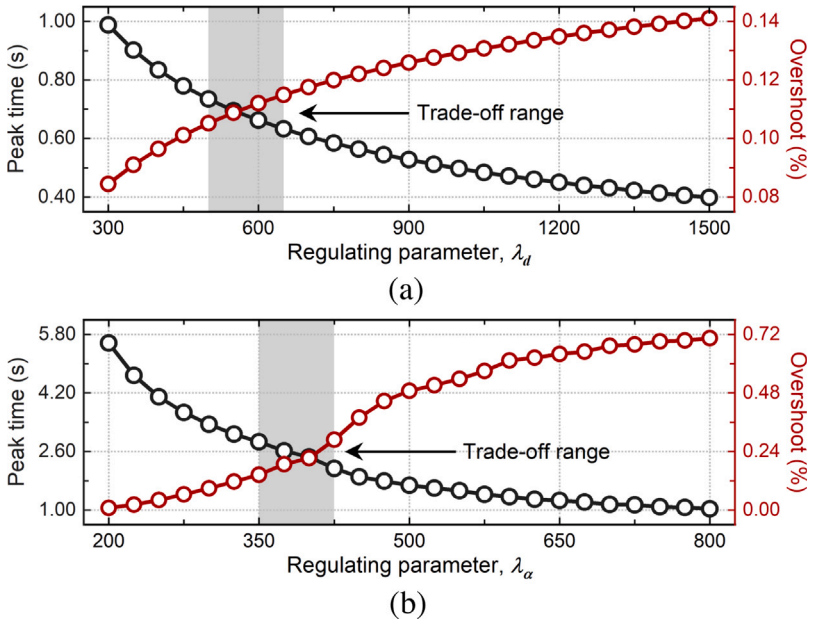

Fig. 5. Optimal selection of regulating parameters. (a) Peak time t_p and overshoot σ_p as functions of regulating parameter λ_d. (b) Peak time t_p and overshoot σ_p as functions of regulating parameter λ_α. The optimal domains are marked by gray shaded areas.

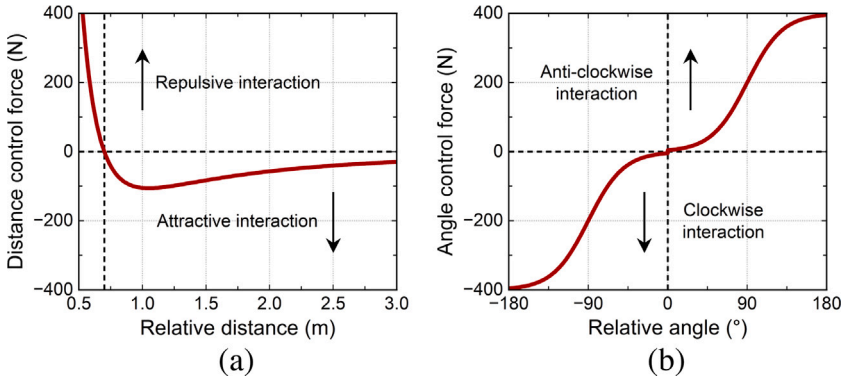

Fig. 6. Illustration of the distance control and angle control forces in our model. (a) The scalar of distance control force as a function of the relative distance. (b) The scalar of angle control force as a function of the relative angle. The dashed lines in the vertical direction represent the equilibrium distance and equilibrium angle, respectively.

distance is $d_{iq}^e = 0.7$ m, as shown in Fig. 6(a), we depict the scalar of distance control force as a function of the relative distance traversing from 0.5 m to 3.0 m. The repulsive interaction appears if $d_{iq} < d_{iq}^e$, while the attractive interaction dominates if $d_{iq} > d_{iq}^e$, whose degree of change is significantly less than that of the repulsive interaction. Similarly, we assume that the equilibrium angle is $\alpha_{iq}^e = 0°$, Fig. 6(b) draws the scalar of angle control force as a function of the relative angle traversing from $-180°$ to $180°$. The force exhibits a clockwise interaction if $\alpha_{iq} < \alpha_{iq}^e$, while it performs an anti-clockwise interaction if $\alpha_{iq} > \alpha_{iq}^e$, which grows in an S-shaped form as the relative angle deviation increases. Taken together, Fig. 6 provides an intuitive insight into the adjustment process of the distance control and angle control forces.

4. Numerical simulations

4.1. Simulation setup

As one of the most classical crowd movement scenes in real situations, walkways have been widely used to analyze pedestrian flow characteristics and complex behavioral patterns (Shi et al., 2018). Therefore, as shown in Fig. 7, numerical simulations are expected to be performed in a walkway scene of 50 m length and 10 m width, from which the simulation scene is proportionally

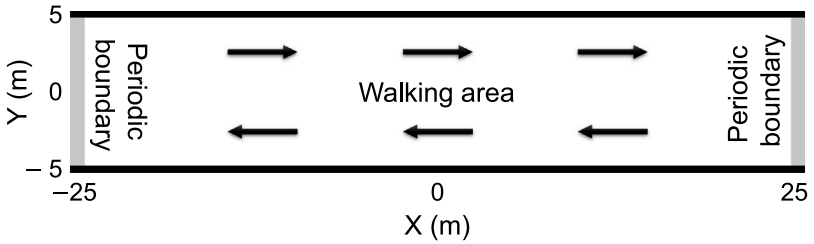

Fig. 7. Schematic illustration of the walkway scene for numerical simulations. The real size is 50 m length and 10 m width, with each pixel corresponding to a 0.1 m-side square.

Table 3
Parameter settings in our model.

Symbol	Description	Value
v_i^0	Desired speed	$\mathcal{N}(1.3, 0.04)$ ms^{-1}
m_i	Pedestrian mass	80 kg
r_i	Pedestrian radius	0.25 m
τ_i	Relaxation time	0.5 s
A_i	Constant 1	$2 \cdot 10^3$ N
B_i	Constant 2	0.08 m
k	Body elasticity coefficient	$1.2 \cdot 10^5$ kgs^{-2}
κ	Sliding friction coefficient	$2.4 \cdot 10^5$ kgm^{-1}s^{-1}
λ_d	Regulating parameter 1	600
λ_α	Regulating parameter 2	400
k_α	Ascending gradient	0.05
$\Delta\alpha_f$	Critical angle deviation	90°

scaled, with each pixel corresponding to a 0.1 m-side square. The pedestrians are randomly distributed in the walking area at the initial time, after which they will walk in a particular direction (right or left), depending on the settings of relevant simulations. The gray strips on both sides of the walkway are periodic boundary areas, this means an identical pedestrian will enter from the other side when a pedestrian leaves either side. The reason for this is to make pedestrian flows more stable without being disrupted at the ends of the walkway by randomly entering pedestrians (Helbing, 1991)

Turning now to the parameter settings in numerical simulations, whose symbols, descriptions, and values are summarized in Table 3. The desired speed v_i^0 is assumed to follow a Gaussian distribution with mean value 1.3 ms^{-1} and standard deviation 0.2 ms^{-1}, tallying with the heterogeneity of walking speed from controlled experiments (Moussaïd et al., 2009). These values of pedestrian mass m_i, pedestrian radius r_i, relaxation time τ_i, constants A_i and B_i, body elasticity coefficient k, and sliding friction coefficient κ are given in accordance with the classical SFM (Helbing et al., 2000). Besides, several parameters involved in the group control force are also illustrated as follows: regulating parameters λ_d and λ_α are taken from the previous section, with ascending gradient $k_\alpha = 0.05$ and critical angle deviation $\Delta\alpha_f = 90°$ at the maximum slope we can generate a reasonable dynamic tendency of the related function. It is notable that those parameters not mentioned in Table 3 are indicated in subsequent simulations.

4.2. Simulation performance of the proposed model

In the first part of our simulations, it is worth exploring the evolution process of spatial configurations at different density levels in the surrounding environment. For this, we divide the local neighbor density into five levels, at which the walking behaviors of pedestrian subgroups are simulated in the walkway scene according to the previous simulation setup. Fig. 8 illustrates the typical spatial configurations of pedestrian subgroups at five levels of the local neighbor density, in which subgroup size = 2, 3, and 4 respectively correspond to red, blue, and yellow hollow circles, and the circle radius is consistent with the given pedestrian radius. According to the simulation results, it can be seen that the spatial configurations of pedestrian subgroups are exactly emerged from the quantitative laws found in Fig. 3, that is, subgroup members spontaneously adjust their positions to accommodate the spatial impact of changes in local neighbor density. This indicates that our model successfully reproduces the spatial preferences of walking patterns for pedestrian subgroups, which is also in good agreement with empirical observations.

Let us investigate the simulation performance of our model compared with existing subgroup models, proposed by Xu and Duh (2010), Xie et al. (2021), and Moussaïd et al. (2010), respectively. These models all incorporate the group force on the basis of the SFM, whose modeling processes are similar to that in this paper. With the same initial conditions of all pedestrians guaranteed, we program these models to simulate subgroup movements according to their given formulas and parameters, and the corresponding snapshots of unidirectional pedestrian flows are exhibited in Fig. 9(a). The model of Xu and Duh (2010) can only ensure the relative distances between any pair of subgroup members are equal, resulting in the shape of equilateral polygons. Due to the leader–follower effect, most pedestrian subgroups simulated by Xie et al. (2021) present "queue-like" structures, in which members follow each other along the movement direction. However, the walking patterns from these two models obviously differ from those in empirical

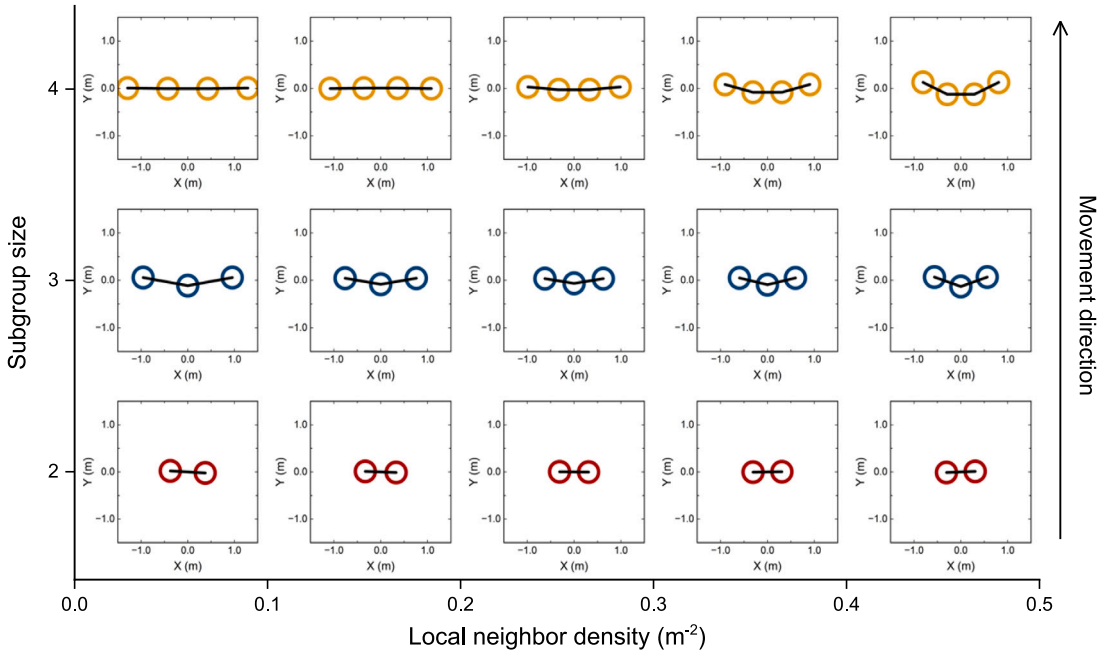

Fig. 8. Typical spatial configurations of pedestrian subgroups at five levels of the local neighbor density from simulation results. The vertical direction denotes different subgroup sizes, and the horizontal direction indicates various local neighbor densities.

observations, even if they succeed in making subgroup members gather together. Turning to the model of Moussaïd et al. (2010), although it emerges the organized patterns very close to those in reality, it only determines the specific shape of pedestrian subgroups by setting parameters before simulation, rather than adaptively adjusting it with density changes in the surrounding environment as our model does. As a result, our model is more realistic than existing subgroup models, and also reflects the heterogeneity features of spatial configurations.

To compare the reproducibility of these subgroup models, the Jensen–Shannon (JS) divergence is selected to assess the similarity between the distributions of simulation data and real data. Before its concrete form is given, it necessitates declaring the Kullback–Leibler (KL) divergence $KL(P\|Q)$, which is a measurement of the difference between probability distributions $P_X(x)$ and $Q_X(x)$:

$$KL(P\|Q) = \sum_{x\in\Theta} P_X(x) \log \frac{P_X(x)}{Q_X(x)} \tag{19}$$

where X is a random variable, and Θ means the probability space. However, a major problem with the KL divergence is asymmetry, that is, $KL(P\|Q) \neq KL(Q\|P)$. As a variant of the KL divergence, the JS divergence solves this problem, which is given as below:

$$JS(P\|Q) = \frac{1}{2}KL\left(P\|\frac{P+Q}{2}\right) + \frac{1}{2}KL\left(Q\|\frac{P+Q}{2}\right) \tag{20}$$

Note that $JS(P\|Q) \in [0,1]$ is symmetrical, with its value closer to 0 corresponding to the greater similarity between $P_X(x)$ and $Q_X(x)$. For relative observables, we calculate the JS divergence of their distributions simulated by various subgroup models with respect to the distribution of real data extracted from public datasets. These results of our model are closest to those obtained from empirical observations in Fig. 9(b), which strongly confirms that our model has better reproducibility than existing subgroup models.

4.3. Effect of pedestrian subgroups on fundamental diagrams

The fundamental diagrams, as one of the most common quantitative indicators, have received considerable attention in crowd dynamics. This dependency of density and speed (or flow) can effectively measure the characteristics of moving crowds, and allows the evaluation of human motion models and facility capacity (Seyfried et al., 2005). Extensive studies on fundamental diagrams show that the speed decreases monotonically with the increasing density (Dong et al., 2020), the explanation might be that pedestrians would slow down to avoid potential collisions in crowded spaces (Parisi et al., 2021). Nonetheless, subgroup factors are often ignored in pedestrian flows, and little is known about how they affect fundamental diagrams. To investigate this issue, we design a series of unidirectional flow simulations covering three possible cases: (1) No subgroups (pedestrians are all isolated individuals); (2) Normal proportion of subgroups (the proportion is consistent with that in public datasets); (3) High proportion of subgroups (the proportion is twice as much as that in public datasets). Based on this, different numbers of total pedestrians are further given to create the environmental conditions at various density levels.

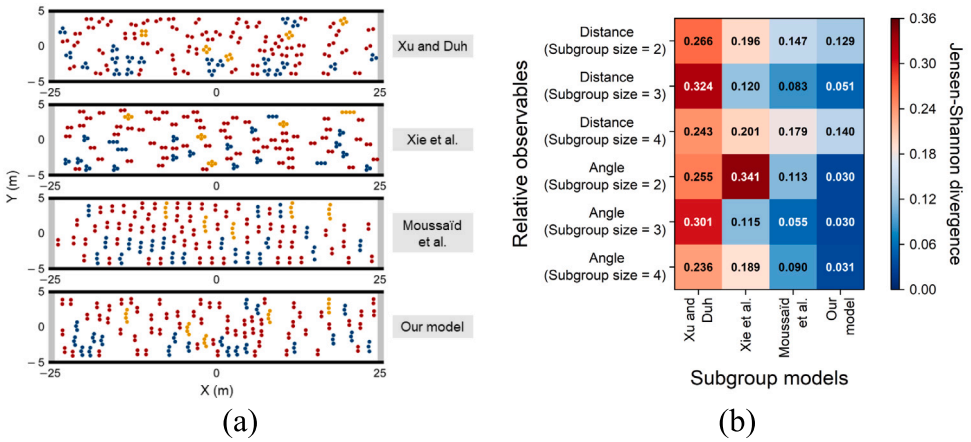

Fig. 9. Comparison of simulation performance with existing subgroup models. (a) Snapshots of unidirectional pedestrian flows simulated by various subgroup models. (b) Jensen–Shannon divergence of the distributions simulated by various subgroup models with respect to the distribution of real data extracted from public datasets.

Under the three cases of subgroup proportions, Fig. 10 displays the snapshots of unidirectional pedestrian flows at various density levels. The direction of pedestrian flow is from left to right, with black, red, blue, and yellow solid circles representing isolated individuals and pedestrian subgroups of 2, 3, and 4 members, respectively. At low densities ($N = 80$ and $N = 160$), most pedestrian subgroups maintain walking side by side, with the exception of a few constrained by the surrounding space adjust their spatial configurations. No matter for individuals or subgroups, the local space is generally sufficient for them to move at a near-free speed. As the density raises ($N = 240$ and $N = 320$), pedestrian subgroups adapt to more stringent physical constraints by adjusting their spatial configurations with a stronger degree of concavity. The crowds are found to be non-uniform distributed in the scene (i.e., partial areas are crowded, while others are empty), and the phenomenon becomes more apparent as the proportion of subgroups increases. These findings intuitively indicate that pedestrian subgroups have a certain influence on unidirectional flows, and it is also expected to further explore the underlying features reflected by this phenomenon.

To quantitatively analyze the speed–density and flow–density relations for the three cases, there are multiple methods to measure the fundamental quantities. The local density, local speed, and local flow proposed by Helbing et al. (2007) are chosen here, because the results derived by this method show more continuity and less fluctuation than those of classical methods. By combining the local density at specific positions with corresponding local speed and local flow, the fundamental diagrams of unidirectional pedestrian flows are illustrated in Fig. 11. As the local density increases, Fig. 11(a) reveals that the local speed first remains constant and then steadily decreases, accompanied by the local flow raising first and then reaching saturation and falling in Fig. 11(b). In particular, the local speed–density relation can be well fitted with the Weidmann function (Weidmann, 1993), as given by the following equation:

$$v(\rho) = v_{free}\left\{1 - \exp\left[-\gamma\left(\frac{1}{\rho} - \frac{1}{\rho_{\max}}\right)\right]\right\} \tag{21}$$

where v_{free} is the free speed at near-zero density, $\rho_{\max}$ is the maximum density at which the speed drops to zero, and γ corresponds to an adaptation parameter. Likewise, the local flow–density relation can also be derived by multiplying the above equation with the local density:

$$f(\rho) = \rho \cdot v_{free}\left\{1 - \exp\left[-\gamma\left(\frac{1}{\rho} - \frac{1}{\rho_{\max}}\right)\right]\right\} \tag{22}$$

These fitting functions for the three cases are shown by solid lines with different colors. Interestingly, the presence of pedestrian subgroups hardly affects fundamental diagrams at relatively low density, but a high proportion of them enhances the speed and flow of crowds as the density increases, in good agreement with a recent empirical study (Hu et al., 2021). The explanation might be that pedestrian subgroups are more likely to form local aggregation (i.e., tolerate closer social distances) in crowded spaces due to the mutual familiarity of members, which gives rise to more walkable areas and promotes more unblocked pedestrian flows.

4.4. Effect of pedestrian subgroups on lane formation

Let us move to the next topic regarding lane formation in crowds, which is an important self-organization phenomenon that reveals the macroscopic collective motion emerging from individual anticipatory interactions (Murakami et al., 2021). The oppositely moving pedestrians randomly distributed at the initial time will form spontaneously separated flows after a transition period. The underlying principle is that oncoming pedestrians evade left or right with equal probability and enter a queue in the same direction as oneself, this improves the efficiency of pedestrian flow by reducing the frequency of avoidance behaviors (Helbing et al., 2001).

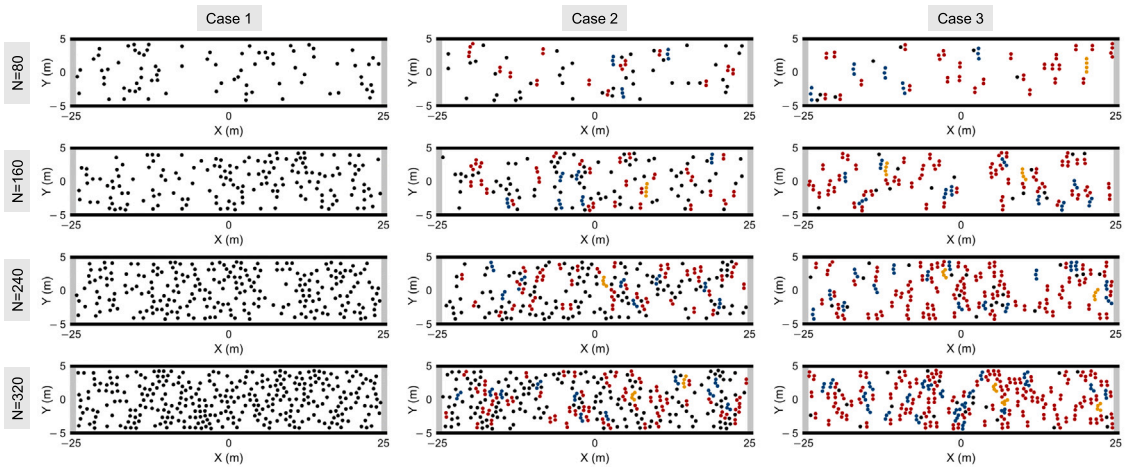

Fig. 10. Snapshots of unidirectional pedestrian flows at various density levels under the three cases of subgroup proportions. The vertical direction represents different numbers of total pedestrians, and the horizontal direction indicates the three cases of subgroup proportions.

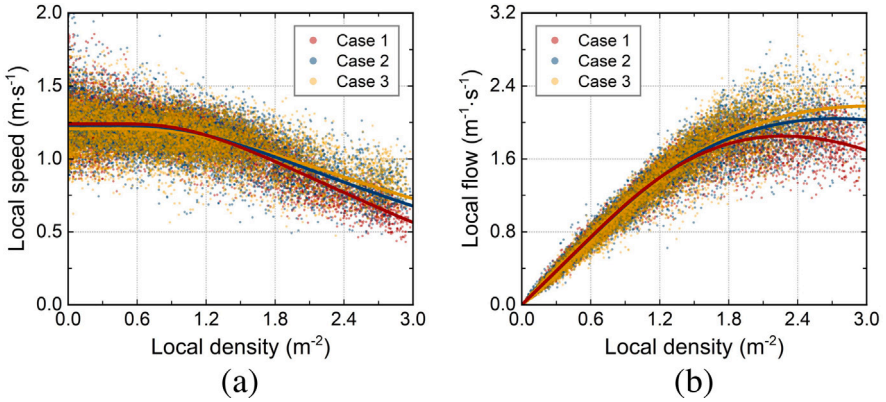

Fig. 11. Fundamental diagrams in unidirectional pedestrian flows under the three cases of subgroup proportions. (a) Local speed–density relation. (b) Local flow–density relation. These circle points and solid lines represent the simulation data and fitting functions, respectively.

However, although many experiments and models with regard to this topic have been performed, the effect of pedestrian subgroups on lane formation still remains poorly understood. As a consequence, the simulations of bidirectional flows are carried out to explore this issue, and the three cases of subgroup proportions are in line with those in the previous section. Incidentally, the number of total pedestrians is fixed as $N = 240$ to ensure a clear lane formation phenomenon from the visual perception.

Under the three cases of subgroup proportions, Fig. 12 illustrates the snapshots of bidirectional pedestrian flows over time. Isolated individuals and pedestrian subgroups of 2, 3, and 4 members are also represented by black, red, blue, and yellow circles, in which solid circles are those walking from left to right, and hollow ones are those walking in the opposite direction. For Case 1 composed entirely of isolated pedestrians, four distinct lanes can be found in the walkway scene at just 50 s, and they remain high stability in the subsequent stage. If we now turn to Case 2, the situation seems a little chaotic at 50 s, and four lanes with non-uniform continuity are gradually formed after 100 s. Compared with the former two cases, Case 3 presents more chaotic bidirectional flows, and three lanes interspersed with partial oppositely moving pedestrians can be reluctantly recognized until 150 s. These results preliminarily indicate that the presence of pedestrian subgroups interferes with the process of lane formation, and such disruption degree becomes more significant as the subgroup proportion increases.

Furthermore, the quantitative analysis of the aforementioned phenomenon is beneficial for better understanding the effect of pedestrian subgroups on lane formation. The order of crowds in lane formation can be characterized by Yamori's band index $Y(t)$, which measures the segregation of opposite pedestrian flows (Yamori, 1998), as given by this equation:

$$Y(t) = \frac{1}{B}\sum_{b=1}^{B}\frac{\left|n_1^b(t) - n_2^b(t)\right|}{n_1^b(t) + n_2^b(t)} \tag{23}$$

Here, the walkway is divided into B bands at equal intervals of width d_y in the vertical direction, $n_2^b(t)$ is the number of pedestrians walking in one direction within band b, and $n_2^b(t)$ is the number of pedestrians walking reversely. That is, $Y(t) \to 1$ implies that the

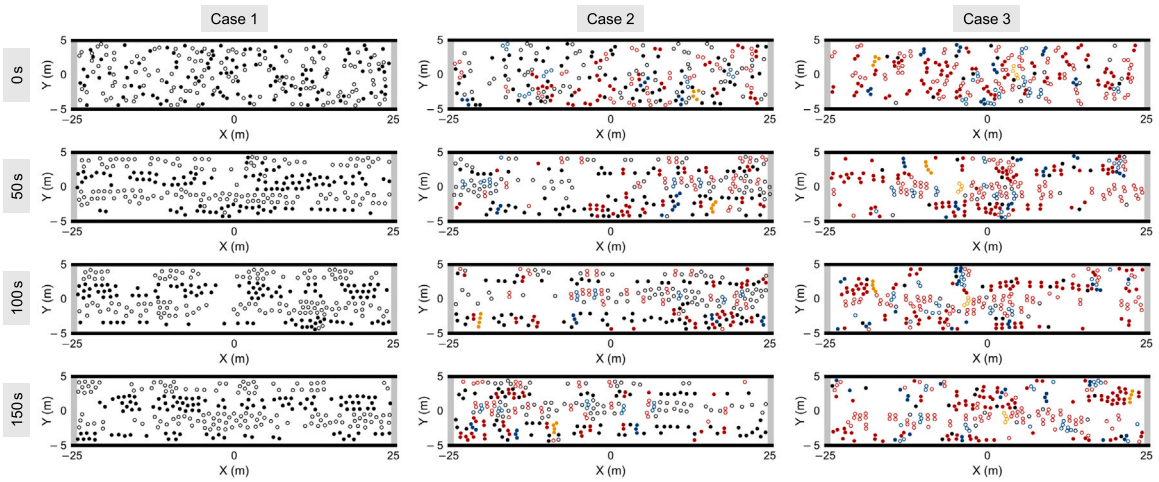

Fig. 12. Snapshots of bidirectional pedestrian flows over time under the three cases of subgroup proportions. The vertical direction represents different moments of the simulations, and the horizontal direction indicates the three cases of subgroup proportions.

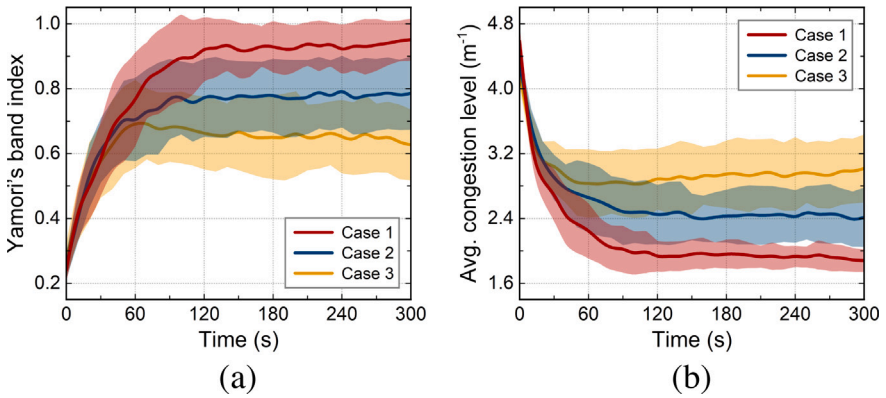

Fig. 13. Quantitative analysis of lane formation in bidirectional pedestrian flows under the three cases of subgroup proportions. (a) Yamori's band index as a function of time. (b) Average of the congestion level as a function of time. These solid lines and error bands denote the mean values and standard deviations in 50 trials, respectively.

segregation is proximately perfect, while the opposite flows are extremely chaotic if $Y(t) \to 0$. We set $d_y = 0.5$ m for the measurement, as shown in Fig. 13(a), the separation time to reach saturation is shortened sequentially from Cases 1 to 3, but the order degree of the final state becomes worse in turn. Besides, we are interested in the congestion level during lane formation, the average of which in space is defined as below:

$$Cl\,(t) = \langle Cl\,(\mathbf{x}, t)\rangle_{\mathbf{x}\in S} \tag{24}$$

where $Cl(\mathbf{x}, t)$ is the congestion level at position $\mathbf{x}$ belonging to the whole walking area S, regarding further details of the calculation process see Feliciani and Nishinari (2018). Fig. 13(b) demonstrates that the congestion is notably exacerbated from Cases 1 to 3, which reveals the negative impact caused by unstable lanes. In summary, a higher proportion of pedestrian subgroups makes the lane formation more disordered and congested, due to the fact that subgroups with longer boundary space orthogonal to the walking direction require more time to accomplish larger position offsets during avoidance interactions, which triggers a chain reaction for the crowd behind (i.e., longer waiting times and more potential collisions) to further deepen the chaos at the macroscopic level.

5. Conclusion

In this paper, we discover the quantitative laws of spatial configurations of pedestrian subgroups at non-extreme densities, whose mathematical forms are incorporated into a force-based model to simulate corresponding behaviors and phenomena. By achieving numerical simulations in a walkway scene, the main conclusions are summarized below: (1) The preferred relative distance and preferred relative angle between subgroup members both exhibit piecewise linear relationships with the local neighbor density, which can also be extrapolated to relatively higher density levels. (2) This model reproduces the organized patterns of pedestrian

subgroups more in accordance with empirical observations, and presents better simulation performance than existing subgroup models. (3) In unidirectional flows, pedestrian subgroups barely affect fundamental diagrams at low density but promote the speed and flow as the density increases. (4) In bidirectional flows, lane formation can be significantly interfered by the presence of pedestrian subgroups, which further exacerbates the degree of disorder and congestion.

These findings from this study may facilitate the understanding of pedestrian subgroups and make contributions to several practical application areas. For the design of pedestrian facilities, it has been pointed out that subgroup factors need to be considered when addressing space requirements (Willis et al., 2004), thereby adjusting the design scheme by simulating pedestrian subgroups based on our model is conducive to reaching a higher level of efficiency and safety. Regarding the analysis of crowd behaviors, the quantitative laws of spatial configurations can help computer scientists analyze the interaction between subgroup members from a large number of surveillance videos, which can serve important tasks such as the detection of abnormal crowds by security departments (Li et al., 2014; Gao and Liu, 2017). With respect to animation production, typical organized patterns of pedestrian subgroups are well presented by our model, which can be embedded in the animation production for more realistic visual rendering and motion control (Peters and Ennis, 2009). As a consequence, relevant studies on pedestrian subgroups are expected to bring encouraging inspirations into more potential fields.

It should be indicated that our model might have limitations in two aspects. First, the quantitative laws extracted from public datasets just involve normal density levels and can also only be extrapolated to relatively higher (non-extreme) density levels. From this, our model is not suitable for reproducing subgroup behaviors at extremely high densities such as bottlenecks (Seyfried et al., 2009), thereby future work may focus on exploring the quantitative laws of subgroup configurations (e.g., "river-like" formations (Helbing et al., 2005)) at such densities. Nevertheless, it has been reported that subgroups are not salient in very high-density crowds (Yang et al., 2018), whether these potential studies are necessary remains debatable. Second, although incorporating the extracted quantitative laws into the model can more precisely describe the spatial configurations of pedestrian subgroups at different densities, it inevitably leaves the model with the issue of forcing individuals to adapt to preferred positions. From this, future work can attempt to improve this model from other perspectives (e.g., visual information (Moussaïd et al., 2011)) to help enhance its interpretability (e.g., explain how these preferred positions arise) and advance the understanding of subgroup behaviors to the next stage. In spite of its limitations, this work certainly achieves a breakthrough in describing the laws of subgroup configurations from qualitative to quantitative aspects, and also provides new insights for developing a more realistic and elaborate model to reproduce subgroup behaviors in crowd dynamics.

CRediT authorship contribution statement

Wenhan Wu: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Wenfeng Yi:** Formal analysis, Methodology, Software, Writing – review & editing. **Xiaolu Wang:** Funding acquisition, Project administration, Resources, Writing – review & editing. **Xiaoping Zheng:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

This work was supported in part by the National Major Scientific Research Instrument Development Project (Grant No. 61927804); and in part by the National Key Research and Development Program of China (Grant No. 2021YFB3301100).

References

Ang, K.H., Chong, G., Li, Y., 2005. PID control system analysis, design, and technology. IEEE Trans. Control Syst. Technol. 13 (4), 559–576. http://dx.doi.org/10.1109/tcst.2005.847331.

Bain, N., Bartolo, D., 2019. Dynamic response and hydrodynamics of polarized crowds. Science 363 (6422), 46–49. http://dx.doi.org/10.1126/science.aat9891.

Bandini, S., Gorrini, A., Vizzari, G., 2014. Towards an integrated approach to crowd analysis and crowd synthesis: A case study and first results. Pattern Recognit. Lett. 44, 16–29. http://dx.doi.org/10.1016/j.patrec.2013.10.003.

Bethel, C., Murphy, R., 2008. Survey of non-facial/non-verbal affective expressions for appearance-constrained robots. IEEE Trans. Syst. Man Cybern. C (Appl. Rev.) 38 (1), 83–92. http://dx.doi.org/10.1109/tsmcc.2007.905845.

Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. 99 (suppl_3), 7280–7287. http://dx.doi.org/10.1073/pnas.082080899.

Burstedde, C., Klauck, K., Schadschneider, A., Zittartz, J., 2001. Simulation of pedestrian dynamics using a two-dimensional cellular automaton. Physica A 295 (3–4), 507–525. http://dx.doi.org/10.1016/s0378-4371(01)00141-8.

Coleman, J.S., James, J., 1961. The equilibrium size distribution of freely-forming groups. Sociometry 24 (1), 36. http://dx.doi.org/10.2307/2785927.

Crociani, L., Gorrini, A., Vizzari, G., 2013. Pedestrian dynamics in presence of groups: An agent-based model applied to a real world case study. Complex Adapt. Syst. Model 1.

Crociani, L., Zeng, Y., Vizzari, G., Bandini, S., 2018. Shape matters: Modelling, calibrating and validating pedestrian movement considering groups. Simul. Model. Pract. Theory 87, 73–91. http://dx.doi.org/10.1016/j.simpat.2018.06.001.

Dong, H., Zhou, M., Wang, Q., Yang, X., Wang, F.-Y., 2020. State-of-the-art pedestrian and evacuation dynamics. IEEE Trans. Intell. Transp. Syst. 21 (5), 1849–1866. http://dx.doi.org/10.1109/tits.2019.2915014.

Feliciani, C., Nishinari, K., 2018. Measurement of congestion and intrinsic risk in pedestrian crowds. Transp. Res. C 91, 124–155. http://dx.doi.org/10.1016/j.trc.2018.03.027.
Gao, C., Liu, J., 2017. Network-based modeling for characterizing human collective behaviors during extreme events. IEEE Trans. Syst. Man Cybern.: Syst. 47 (1), 171–183. http://dx.doi.org/10.1109/tsmc.2016.2608658.
Helbing, D., 1991. A mathematical model for the behavior of pedestrians. Behav. Sci. 36 (4), 298–310. http://dx.doi.org/10.1002/bs.3830360405.
Helbing, D., Buzna, L., Johansson, A., Werner, T., 2005. Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. Transp. Sci. 39 (1), 1–24. http://dx.doi.org/10.1287/trsc.1040.0108.
Helbing, D., Farkas, I., Vicsek, T., 2000. Simulating dynamical features of escape panic. Nature 407 (6803), 487–490. http://dx.doi.org/10.1038/35035023.
Helbing, D., Johansson, A., Al-Abideen, H.Z., 2007. Dynamics of crowd disasters: An empirical study. Phys. Rev. E 75 (4), 046109. http://dx.doi.org/10.1103/physreve.75.046109.
Helbing, D., Molnár, P., 1995. Social force model for pedestrian dynamics. Phys. Rev. E 51 (5), 4282–4286. http://dx.doi.org/10.1103/physreve.51.4282.
Helbing, D., Molnár, P., Farkas, I.J., Bolay, K., 2001. Self-organizing pedestrian movement. Environ. Plan. B: Plann. Des. 28 (3), 361–383. http://dx.doi.org/10.1068/b2697.
Hu, Y., Zhang, J., Song, W., Bode, N.W., 2021. Social groups barely change the speed-density relationship in unidirectional pedestrian flow, but affect operational behaviours. Saf. Sci. 139, 105259. http://dx.doi.org/10.1016/j.ssci.2021.105259.
Huang, L., Gong, J., Li, W., Xu, T., Shen, S., Liang, J., Feng, Q., Zhang, D., Sun, J., 2018. Social force model-based group behavior simulation in virtual geographic environments. ISPRS Int. J. Geo-Inf. 7 (2), 79. http://dx.doi.org/10.3390/ijgi7020079.
Hussein, M., Sayed, T., 2016. A bi-directional agent-based pedestrian microscopic model. Transportmetrica A: Transp. Sci. 13 (4), 326–355. http://dx.doi.org/10.1080/23249935.2016.1266531.
Karamouzas, I., Overmars, M., 2012. Simulating and evaluating the local behavior of small pedestrian groups. IEEE Trans. Visual. Comput. Graph. 18 (3), 394–406. http://dx.doi.org/10.1109/tvcg.2011.133.
Lazer, D.M.J., Pentland, A., Watts, D.J., Aral, S., Athey, S., Contractor, N., Freelon, D., Gonzalez-Bailon, S., King, G., Margetts, H., Nelson, A., Salganik, M.J., Strohmaier, M., Vespignani, A., Wagner, C., 2020. Computational social science: Obstacles and opportunities. Science 369 (6507), 1060–1062. http://dx.doi.org/10.1126/science.aaz8170.
Lerner, A., Chrysanthou, Y., Lischinski, D., 2007. Crowds by example. Comput. Graph. Forum 26 (3), 655–664. http://dx.doi.org/10.1111/j.1467-8659.2007.01089.x.
Li, Y., Chen, M., Dou, Z., Zheng, X., Cheng, Y., Mebarki, A., 2019. A review of cellular automata models for crowd evacuation. Physica A 526, 120752. http://dx.doi.org/10.1016/j.physa.2019.03.117.
Li, Y., Liu, H., peng Liu, G., Li, L., Moore, P., Hu, B., 2017. A grouping method based on grid density and relationship for crowd evacuation simulation. Physica A 473, 319–336. http://dx.doi.org/10.1016/j.physa.2017.01.008.
Li, W., Mahadevan, V., Vasconcelos, N., 2014. Anomaly detection and localization in crowded scenes. IEEE Trans. Pattern Anal. Mach. Intell. 36 (1), 18–32. http://dx.doi.org/10.1109/tpami.2013.111.
Lu, L., Chan, C.-Y., Wang, J., Wang, W., 2017. A study of pedestrian group behaviors in crowd evacuation based on an extended floor field cellular automaton model. Transp. Res. C 81, 317–329. http://dx.doi.org/10.1016/j.trc.2016.08.018.
Moussaïd, M., Helbing, D., Garnier, S., Johansson, A., Combe, M., Theraulaz, G., 2009. Experimental study of the behavioural mechanisms underlying self-organization in human crowds. Proc. R. Soc. B: Biol. Sci. 276 (1668), 2755–2762. http://dx.doi.org/10.1098/rspb.2009.0405.
Moussaïd, M., Helbing, D., Theraulaz, G., 2011. How simple rules determine pedestrian behavior and crowd disasters. Proc. Natl. Acad. Sci. 108 (17), 6884–6888. http://dx.doi.org/10.1073/pnas.1016507108.
Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., Theraulaz, G., 2010. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. In: Chirico, G. (Ed.), PLoS One 5 (4), e10047. http://dx.doi.org/10.1371/journal.pone.0010047.
Mukherjee, S., Goswami, D., Chatterjee, S., 2015. A Lagrangian approach to modeling and analysis of a crowd dynamics. IEEE Trans. Syst. Man Cybern.: Syst. 45 (6), 865–876. http://dx.doi.org/10.1109/tsmc.2015.2389763.
Murakami, H., Feliciani, C., Nishiyama, Y., Nishinari, K., 2021. Mutual anticipation can contribute to self-organization in human crowds. Sci. Adv. 7 (12), http://dx.doi.org/10.1126/sciadv.abe7758.
Nicolas, A., Hassan, F.H., 2023. Social groups in pedestrian crowds: Review of their influence on the dynamics and their modelling. Transportmetrica A: Transp. Sci. 19 (1), 1970651. http://dx.doi.org/10.1080/23249935.2021.1970651.
Parisi, D.R., Sartorio, A.G., Colonnello, J.R., Garcimartín, A., Pugnaloni, L.A., Zuriguel, I., 2021. Pedestrian dynamics at the running of the bulls evidence an inaccessible region in the fundamental diagram. Proc. Natl. Acad. Sci. 118 (50), http://dx.doi.org/10.1073/pnas.2107827118.
Pellegrini, S., Ess, A., Schindler, K., van Gool, L., 2009. You'll never walk alone: Modeling social behavior for multi-target tracking. In: 2009 IEEE 12th International Conference on Computer Vision. IEEE, http://dx.doi.org/10.1109/iccv.2009.5459260.
Peters, C., Ennis, C., 2009. Modeling groups of plausible virtual pedestrians. IEEE Comput. Graph. Appl. 29 (4), 54–63. http://dx.doi.org/10.1109/mcg.2009.69.
Qiu, F., Hu, X., 2010. Modeling group structures in pedestrian crowd simulation. Simul. Model. Pract. Theory 18 (2), 190–205. http://dx.doi.org/10.1016/j.simpat.2009.10.005.
Seyfried, A., Passon, O., Steffen, B., Boltes, M., Rupprecht, T., Klingsch, W., 2009. New insights into pedestrian flow through bottlenecks. Transp. Sci. 43 (3), 395–406. http://dx.doi.org/10.1287/trsc.1090.0263.
Seyfried, A., Steffen, B., Klingsch, W., Boltes, M., 2005. The fundamental diagram of pedestrian movement revisited. J. Stat. Mech. Theory Exp. 2005 (10), P10002. http://dx.doi.org/10.1088/1742-5468/2005/10/p10002.
Shi, X., Ye, Z., Shiwakoti, N., Grembek, O., 2018. A state-of-the-art review on empirical data collection for external governed pedestrians complex movement. J. Adv. Transp. 2018, 1–42. http://dx.doi.org/10.1155/2018/1063043.
Steffen, B., Seyfried, A., 2010. Methods for measuring pedestrian density, flow, speed and direction with minimal scatter. Physica A 389 (9), 1902–1910. http://dx.doi.org/10.1016/j.physa.2009.12.015.
Turgut, Y., Bozdag, C.E., 2021. Modeling pedestrian group behavior in crowd evacuations. Fire Mater. 46 (2), 420–442. http://dx.doi.org/10.1002/fam.2978.
Vicsek, T., Zafeiris, A., 2012. Collective motion. Phys. Rep. 517 (3–4), 71–140. http://dx.doi.org/10.1016/j.physrep.2012.03.004.
Weidmann, U., 1993. Transporttechnik der Fussgänger: Transporttechnische Eigenschaften des Fussgängerverkehrs, Literaturauswertung. Tech. Rep., ETH Zurich, http://dx.doi.org/10.3929/ETHZ-B-000242008.
Willis, A., Gjersoe, N., Havard, C., Kerridge, J., Kukla, R., 2004. Human movement behaviour in urban spaces: Implications for the design and modelling of effective pedestrian environments. Environ. Plan. B: Plann. Des. 31 (6), 805–828. http://dx.doi.org/10.1068/b3060.
Wu, W., Yi, W., Li, J., Chen, M., Zheng, X., 2023. Automatic identification of human subgroups in time-dependent pedestrian flow networks. IEEE Trans. Multimedia 1–12. http://dx.doi.org/10.1109/TMM.2023.3262975.
Xie, W., Lee, E.W.M., Li, T., Shi, M., Cao, R., Zhang, Y., 2021. A study of group effects in pedestrian crowd evacuation: Experiments, modelling and simulation. Saf. Sci. 133, 105029. http://dx.doi.org/10.1016/j.ssci.2020.105029.
Xu, S., Duh, H.-L., 2010. A simulation of bonding effects and their impacts on pedestrian dynamics. IEEE Trans. Intell. Transp. Syst. 11 (1), 153–161. http://dx.doi.org/10.1109/tits.2009.2036152.
Xu, M., Xie, X., Lv, P., Niu, J., Wang, H., Li, C., Zhu, R., Deng, Z., Zhou, B., 2021. Crowd behavior simulation with emotional contagion in unexpected multihazard situations. IEEE Trans. Syst. Man Cybern.: Syst. 51 (3), 1567–1581. http://dx.doi.org/10.1109/tsmc.2019.2899047.

Yamori, K., 1998. Going with the flow: Micro–macro dynamics in the macrobehavioral patterns of pedestrian crowds. Psychol. Rev. 105 (3), 530–557. http://dx.doi.org/10.1037/0033-295x.105.3.530.

Yang, F., Shabo, J., Qureshi, A., Peters, C., 2018. Do you see groups? The impact of crowd density and viewpoint on the perception of groups. In: Proceedings of the 18th International Conference on Intelligent Virtual Agents. ACM, http://dx.doi.org/10.1145/3267851.3267877.

Zhang, H., Liu, H., Qin, X., Liu, B., 2018. Modified two-layer social force model for emergency earthquake evacuation. Physica A 492, 1107–1119. http://dx.doi.org/10.1016/j.physa.2017.11.041.

Zhou, C., Han, M., Liang, Q., Hu, Y.-F., Kuai, S.-G., 2019. A social interaction field model accurately identifies static and dynamic social groupings. Nat. Hum. Behav. 3 (8), 847–855. http://dx.doi.org/10.1038/s41562-019-0618-2.