Modeling the impact of interaction on pedestrian group motion

Z. Yücel, F. Zanlungo & M. Shiomi

View supplementary material

Published online: 15 Jan 2018.

Submit your article to this journal

View related articles

View Crossmark data

RSJ

Taylor & Francis
Taylor & Francis Group

FULL PAPER

Check for updates

Modeling the impact of interaction on pedestrian group motion

Z. Yücel[a], F. Zanlungo[b] and M. Shiomi[b]

[a]Department of Computer Science, Division of Industrial Innovation Sciences, Graduate School of Natural Science and Technology, Okayama University, Okayama, Japan; [b]Intelligent Robotics and Communication Laboratories, ATR International, Kyoto, Japan

ABSTRACT

Mobile social robots aimed at interacting with and assisting humans in pedestrian areas need to understand the dynamics of pedestrian social interaction. In this work, we investigate the effect of interaction on pedestrian group motion by defining three motion models to represent (1) interpersonal-distance, (2) relative orientation and (3) absolute difference of velocities; and model them using a dataset of 12000+ pedestrian trajectories recorded in uncontrolled settings. Our contributions include: (i) Demonstrating that interaction has a prominent effect on the empirical distributions of the proposed joint motion attributes, where increasing levels of interaction lead to more regular behavior (ii) Developing analytic motion models of such distributions and reflect the effect of interaction on model parameters (iii) Detecting the social groups in a crowd with almost perfect accuracy utilizing the proposed models, despite the constant flow direction in the environment which causes unrelated pedestrians to move in a correlated way, and thus makes group recognition more difficult (iv) Estimating the level of intensity with considerable rates utilizing the proposed models

1. Introduction and motivation

Social mobile robots, i.e. mobile robots designed to assist and accompany humans in pedestrian facilities such as shopping malls and stations [1–6], need to understand and replicate human (pedestrian) behavior in order to move safely in the crowd and properly assist people [7]. Since a very large portion of the crowd is composed of social groups [8], it is particularly important for robots to be able (i) to recognize groups to avoid hindering or properly approach/assist them, and (ii) to replicate their behavior to be able to move along them [9].

Although pedestrian dynamics has been studied by physicists, computer scientists, and psychologists in the last decades, most works focus on basic pedestrian movement, simulation, and evacuation [10]. The crowd is considered as a combination of individuals and crowd behavior is represented in terms of the collective locomotion of individuals rather than groups [11,12]. Therefore, dynamics of social groups is not well-explored yet. Even though some recent works [8,13] focus on group dynamics, pointing also to its dependence on density [14] and group composition [15], the role of *intensity of interaction* is not investigated yet.

According to McPhail and Wohlstein, *social groups* are groups of people engaged in a social relation to one or more pedestrians moving together toward a common goal [16]. Although these groups are engaged in a social relation (e.g. family, friends), they do not always actively interact.[1] Interaction in small groups is regulated by several non-verbal factors such as visual, auditory, temporal, and spatial elements in addition to verbal communication [18]. In particular, eye contact is an essential non-verbal factor affecting equilibrium of physical proximity, since it helps estimating the reactions of peers and anticipating their actions [19,20]. In order to regulate their interaction, pedestrians in social groups tend to keep a matching pace and a comfortable view of their peers, which is often done in an unconscious manner.

This study explores the possibility of pointing out the (unconscious) changes on locomotion behavior in a quantitative way in the absence and presence of (various levels of) interaction, in order to provide robots with the ability to recognize interacting groups, and properly assist them, and with the ability to change their dynamical pattern corresponding to the level of interaction with their human companions.

To that end, we consider three joint motion attributes as essential indicators, i.e. the interpersonal distance, relative orientation and absolute difference of velocities. We pose the following hypotheses:

(H-i) Pedestrians belonging to the same group have a bonding that keeps them together making the proposed joint motion attributes stay around certain values.

(H-ii) Pedestrians who are not in the same group move under the influence of the environment (architecture, pedestrian density, flow direction etc) and have a disperse structure in terms of the attributes.

(H-iii) In pedestrian groups, interaction helps improving the regularity of the proposed attributes. In other words, the distribution of the attributes becomes more concentrated as the intensity of interaction increases.

Several models are employed to quantify the regularity of the proposed attributes and the relating statistics derived from the empirical observations are demonstrated to verify our hypotheses. Moreover, by modeling the behavior of a large set of pedestrians, it is shown that we can determine the interaction status of first neighbors with a good accuracy by looking only at their trajectories.

2. Related work

Proxemics, which is introduced by Hall in 1963, has long been studied to investigate social relations. However, most works on proxemics are carried out basically for standing people or individual pedestrians [21]. For instance, Cristani et al., propose adopting a social signal standpoint [22] using the non-verbal cues, whereas [23–25] consider social groups in poster presentation or meeting room scenarios, where there is no locomotion.

Regarding actively moving pedestrians, floor fields are particularly interesting. Mehran et al. track people in densely populated environments [26] with a large scale interaction between a person and the crowd rather than between two individuals in a social group. Similarly, [27] forecasts trajectories, taking in account collision avoidance oriented 'interaction'. Shao et al. employ deep learning and carry out crowd scene understanding for video classification, or action recognition [28]. Habe et al. consider a similar problem formulated at a lower density investigating the influence which unrelated pedestrians exert on each other [29]. Costa studies the statistics of interpersonal space with respect to personal features such as height, age, gender [30], whereas [31–33] propose methods for detecting the social groups without accounting for their interaction.

The interaction of pedestrians in a social group with other individuals, groups or, in general, the crowd, is crucial for various application domains such as crowd simulation, robotics, public space design, or event organization [34,35]. A common approach in these domains is treating the crowd as a collection of individual agents and using the microscopic motion models as a building block of crowd level activity. An important study in modeling of individuals' behavior is [36], which has long series of follow-up works, mainly aiming at adapting the model to various scenarios such as escape or panic situations [37], detecting abnormal activities in a crowd [26], or collective animal behavior [38] among others.

But the crowd is constituted of heterogeneous elements, i.e. individuals and groups, whose composition needs to be treated carefully. Profiling of groups is important for understanding the crowd level activities, for instance, for detecting stability, collectiveness or conflict [39,40]. In addition, collective behaviors in crowds are treated as flow of pedestrians, who are not necessarily socially related but participate in the same gathering or assembly actions by [41]. The term crowd may refer to people participating in large scale events like pilgrimage or music festivals [42] as well as in moderate densities such as a station, street or mall. Particularly, in the latter case, the effect of social relation between members of a social group emerge as an important factor of pedestrian motion. Since numerous factors, such as gender composition, purpose, social relation/hierarchy play a determining role in group motion [15,43,44], models based only on individual level activity cannot be directly applied. Fang et al. treat hierarchical relation of group members (in particular leader-follower relation) [45] for building a realistic crowd simulator investigating social collective behavior during egress.

Social interaction within the group makes group level activity particularly complex. However, often social interaction between pedestrians is confined to the social norms or socially acceptable behavior [46] and very few works address interaction of pedestrians inside a group. For instance, Li et al. study dynamic spatial-temporal trajectory features without clearly pointing out to the interaction state [47], whereas Jaklin et al. study interaction within groups but in a simulation environment, where the agents present different behaviors according to state of interaction. Similarly, Fu et al. develop pedestrian simulators addressing group interaction but they evaluate their model considering collision avoidance performance rather than estimation of interaction [48]. Velocity and angle consistency are used to determine pedestrian groups in [49] but the reasons for variable levels of motion consistency, which probably are related to the group-level interaction among other phenomena, are not discussed. Therefore, to the best of our knowledge, this is the first study investigating effects of interaction on locomotion of mobile groups on real world data.

Table 1. (a) The number of 2p, 3p groups and pedestrians who are not in any group (non-groups). (b) The number of pedestrian pairs interacting with varying intensities.

(a)

	N	N_{pair}
2p	634	634
3p	75	225
Non-groups	10892	34243

(b)

	Intensity			
	0	1	2	3
2p	66	97	386	85
3p	57	65	90	13

3. Data-set

We use a data-set [13,50] recorded in an underground pedestrian area in Umeda, the main commercial and business district of Osaka, Japan. The experimentation is reviewed and approved for studies involving human participants by ATR ethics board.[2] As the pedestrians move naturally (i.e. in an uninstructed way) in the environment, we recorded range data (for deriving the trajectories) and video data (for establishing the ground truth for group relation and interaction status). Human coders were asked to mark, based on their subjective evaluation, the intensity of interaction $\mathcal{I}$ between the pedestrians in the groups, on a scale going from $\mathcal{I} = 0$ to $\mathcal{I} = 3$.

In total, the data-set includes 10892 pedestrians not moving in groups, 642 2 people (2p)[3] and 75 3p groups, and 28 4p or larger groups (not analyzed in this work). The number of pedestrian *pairs*, i.e. dyads that were simultaneously present in the environment at a distance of 3 meters or less, and their interaction levels are reported in Table 1.

Further details about the nature of the data-set and labeling of group belonging and $\mathcal{I}$ are described in [51].

4. Motion models

In this section, we develop from an empirical and a mathematical standpoint, the dynamical models of group locomotion. These models lie in the basis of our recognition algorithms, and may be used to replicate group behavior in social mobile robots.

Figure 1 illustrates an example of a 3p group composed of pedestrians p_i, p_j and p_k.

We term a set of two pedestrians, possibly in the same group, as a pair. In this case, there are three pairs: $\{p_i, p_j\}$, $\{p_j, p_k\}$ and $\{p_i, p_k\}$.

To have a better insight into the structure of 3p groups, we define a *degree of neighborhood* for each pair. Namely, the group structure is expressed as a minimum spanning tree [52] and the degree of neighborhood concerning a
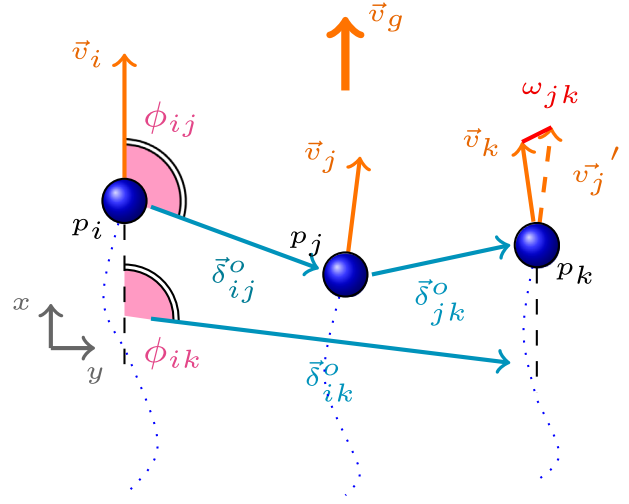

Figure 1. Parameters used in modeling joint motion illustrated on a 3p group.

pair is defined by the number of edges along the shortest path. According to this definition, $\{p_i, p_j\}$ and $\{p_j, p_k\}$ have a degree of neighborhood that equals 1 (*first neighbors*), whereas $\{p_i, p_k\}$ are *second neighbors* since they are connected through p_j.

In examining the joint behavior, we focus particularly on three joint motion attributes:

(1) *Interpersonal distance*, δ, is defined as the distance between pedestrians in a pair.
(2) *Relative orientation*, ϕ_{ij}, of p_j with respect to the motion direction of p_i, is defined as the angle between the velocity vector of p_i and the displacement vector from p_i to p_j, i.e. $\angle \vec{v}_{ij}, \delta_{ij}$. To represent relative orientation of a pair $\{p_i, p_j\}$, we use $\{\phi_{ij}, \phi_{ji}\}$.
(3) *Absolute difference of velocity vectors*, ω_{ij}, is defined as the magnitude of the difference vector,

$$\omega_{ij} = \left| \vec{v}_i - \vec{v}_j \right|.$$

Figure 1 demonstrates the absolute difference of velocity for the pair $\{p_j, p_k\}$, ω_{jk}.

4.1. Empirical distributions and analysis of variance

This section demonstrates empirical distributions of δ, ϕ, and ω for various group relations (2p, 3p, and non-groups).

First we contrast the empirical distributions of 2p and 3p groups with non-groups for all three variables. Groups are shown to illustrate the effect of bonding in terms of the distributions presented in Figure 2-(a-b), which proves the first hypothesis H-i presented in Section 1. Besides, the dynamics of non-groups is mostly determined just by the environment and pedestrian flow as presented in

Figure 2-(c) supporting H-ii of Section 1. As a result, the group and non-group distributions differ significantly in all three cases, which indicates that the proposed joint motion attributes can serve useful in identification of group relation.

Next, we consider 2p group motion for varying $\mathcal{I}$ (See Figure 3). δ and ω assume lower values, and all three attributes are less spread with increasing $\mathcal{I}$, suggesting more regular behavior for higher intensity, supporting H-iii.

Similar phenomena are sought in the motion patterns of 3p groups and it is observed that the distinction is present and similar to those of the 2p groups, yet not as prominent. Although the effect is not as obvious as in the case of 2p groups, it can be seen easily when the two extremes, namely $\mathcal{I}_0$ and $\mathcal{I}_3$, are compared.

In [51], we confirm these qualitative observations in a quantitative way using an analysis of variance. As expected, concerning group relation all three attributes have quite small p-values. Furthermore, also concerning $\mathcal{I}$ in 2p groups, the p-values suggest a statistically significant difference for all of δ, ϕ and ω, whereas in 3p groups, ω appears to be the best indicator in distinction of $\mathcal{I}$.

4.2. Modeling interpersonal distance

Since interpersonal distance depends to a great extent on the degree of neighborhood, we treat 2p groups and the first neighbors in 3p groups in the same manner, whereas second neighbors in 3p groups and non-groups are handled in different ways.

4.2.1. Interpersonal distance for groups
The work of [13] introduces a potential that describes the position dynamics of two pedestrians in a socially interacting group. According to their analysis, based on statistical physics methods, the distribution of δ is given by,

$$p(\delta|\beta, r_0) = \delta \exp\left(\frac{-\beta(\delta - r_0)^2}{\delta r_0}\right). \tag{1}$$

Here, r_0 is the preferred distance for social interaction, while β is, in the statistical physics terminology, the *inverse temperature* (namely, the larger β is the more the pedestrian system will be ordered and the distribution narrowly centered on r_0).

According to [13], the interaction of first neighbors in a 3p group is given by the same potential. For second neighbors, by assuming the group to walk mostly abreast [8], the distance between second neighbors can be modeled as the convolution of the unimodal model with itself,

$$p(\delta|\beta, r_0) = (p * p)\left(\delta \exp\left(\frac{-\beta(\delta - r_0)^2}{\delta r_0}\right)\right). \tag{2}$$

4.2.2. Interpersonal distance for non-groups
Positions of non-interacting pedestrians are assumed to be randomly distributed. Nevertheless, it is known [21] that there is a lower probability to find standing pedestrians at very short distances. In the case of pedestrians moving inside a corridor characterized by a strong directional flow, we found (see [51]) that there is an area, on the front and back of the pedestrian, and with a width comparable to body size, with extremely low probability of finding another pedestrian. We thus assume that the variables δ_x and δ_y have uniform distributions subject to the conditions $\delta = \sqrt{\delta_x^2 + \delta_y^2} \leq R$ and $\delta_x \geq r$, where $R = 3$ meters is the maximum distance we consider and $r = 0.5$ is the 'human body size'. The resulting pdf is,

$$p(\delta|R, r)$$
$$= \frac{4}{\pi R^2 - 2R^2 \arcsin\left(\frac{r}{R}\right) - 2r\sqrt{R^2 - r^2}} \delta \arccos\left(\frac{r}{\delta}\right). \tag{3}$$

4.3. Modeling relative orientation

For ϕ, generalizing the results of [13], we assume, without paying regard to the group size or the degree of neighborhood, a *von Mises* distribution [53], i.e. the generalization of a Gaussian to a circular random variable (angle),

$$p(\phi|\nu, \kappa) = \frac{\exp\left(\kappa \cos\left(\phi - \nu\right)\right)}{2\pi I_0(\kappa)}, \tag{4}$$

where ν denotes the mean value, κ is analogous of $1/\sigma^2$ in the normal distribution, and I_0 is the modified Bessel function of the first kind with order zero [54].

Since we do not pay regard to be positioned on the right or left side with respect to motion direction, both pedestrians in a pair are treated in the same manner, which leads to a distribution with two peaks. Due to the left-right symmetry of the problem, the final model can be expressed as

$$\Phi(\phi|\nu, \kappa) = 0.5p(\phi|\nu, \kappa) + 0.5p(\phi|\nu + \pi, \kappa). \tag{5}$$

4.4. Modeling absolute difference of velocities

Pedestrians in a group move toward the same target independent of group size and neighborhood. Even though there may be slight deviations from the target direction due to avoidance or interaction, the expected value is always around 0. Therefore, we approximate the x and y components of ω with a Gaussian with 0 mean, which makes ω come from a Rayleigh distribution,
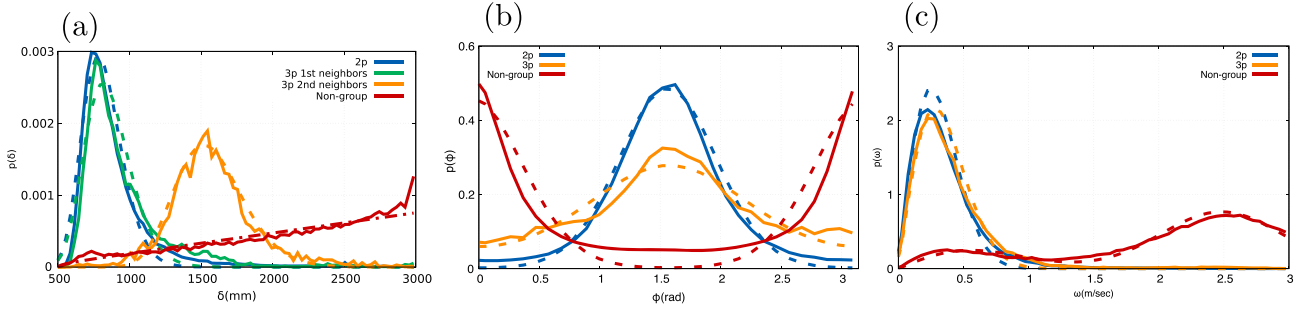
Figure 2. Empirical distributions and models of 2p groups, 3p groups and non-groups for (a) δ, (b) ϕ and (c) ω. Solid lines represent the observations and dashed lines demonstrate the models.
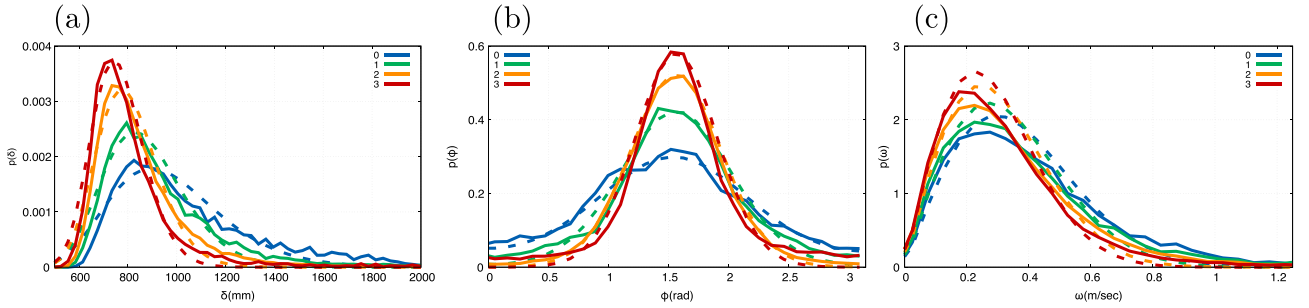

Figure 3. Empirical distributions and models of 2p groups with various intensity of interaction for (a) δ, (a) ϕ, (a) ω. Solid lines represent the observations and dashed lines demonstrate the models.

$$p(\omega|\sigma) = \frac{\omega}{\sigma^2} \exp\left(\frac{-\omega^2}{2\sigma^2}\right). \tag{6}$$

For non-groups, we expect the distributions of the x and y components of the velocity difference to have non-zero mean value ν, and this leads to a Rice distribution,

$$p(\omega|\nu,\sigma) = \frac{\omega}{\sigma^2} \exp\left(\frac{-\omega^2 - \nu^2}{2\sigma^2}\right) I_0\left(\frac{\omega \nu}{\sigma^2}\right). \tag{7}$$

Furthermore, we need to distinguish the case of pedestrians moving in the same direction (for which we expect $\nu \ll \nu$, ν being the pedestrian average velocity) and pedestrians moving in opposite directions (for which we expect $\nu \approx 2\nu$),

$$p(\omega|C_{1,2}, \nu_{1,2}, \sigma_{1,2})$$

$$= \sum_{i=1,2} C_i \frac{\omega}{\sigma_i^2} \exp\left(\frac{-\omega^2 - \nu_i^2}{2\sigma_i^2}\right) I_0\left(\frac{\omega \nu_i}{\sigma_i^2}\right). \tag{8}$$

5. Models for identification of groups and intensity of interaction

All the models proposed in Sections 4.2, 4.3 and 4.4 are calibrated by selecting 20% of the set and optimizing the parameters that fit the empirical distributions best, a process that is repeated 50 times to evaluate the variation

in the values of such parameters. More details may be found in [51].

5.1. Models for identification of groups

The resulting parameters for the models of Sections 4.2–4.4 concerning non-groups and 2p and 3p groups are shown to differ in a significant way verifying H-i and H-ii in Section 1 and proving that the proposed models can be used in estimation of group relation. Moreover, the parameters are shown not to be sensitive to the changes in the set of observations which is used to estimate them. This indicates that the models can stably be constructed from any subset of observations (of a reasonable size). The details of these statements are presented in [51].

5.2. Models for varying levels of intensity of interaction

As explained in Section 1, according to H-iii, for increasing intensities of interaction, we expect to see a shift toward more regular behavior in terms of all joint motion attributes. Namely, as $\mathcal{I}$ grows,

(1) Interpersonal distance δ decreases (decreasing r_0) and its variation drops pointing to more steady group structure (increasing β).
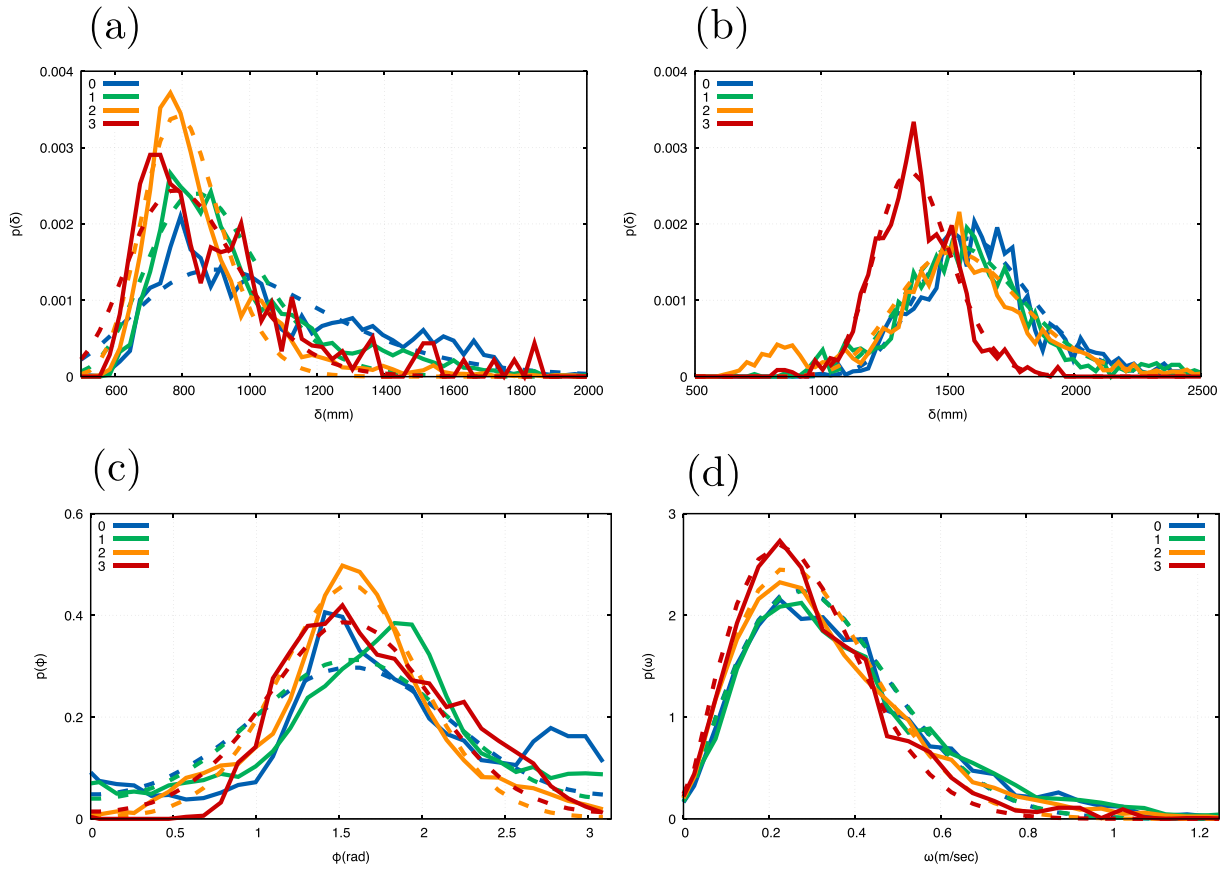
Figure 4. Empirical distributions and models of 3p groups with various intensity of interaction for (a) δ regarding first neighbors, (b) δ regarding second neighbors, (c) ϕ, (d) ω. Solid lines represent the observations and dashed lines demonstrate the models.

(2) The dispersion of ϕ around its average value $\pm\pi/2$ decreases (increasing κ), suggesting more abreast formation.

(3) Absolute difference of velocities ω decreases and adopts a more compact distribution (decreasing σ).

As shown empirically on Figures 3 and 4 and demonstrated in terms of model parameters in [51], these statements are fundamentally true for all motion attributes, in particular in the case of 2p people groups for which the statistics is performed on a larger sample and thus is more stable. These findings support H-iii of Section 1.

An Anova on the model parameters [51], reveals that neighboring interaction intensities, namely $\mathcal{I} = i$ and $\mathcal{I} = i + 1$, present a quite similar behavior, which is not completely surprising, given the arbitrary and subjective nature of these labels. Nevertheless, the difference between the two extremes, $\mathcal{I} = 0$ and $\mathcal{I} = 3$, is much larger than the parameter variation. For this reason, we will use such extremes as benchmarks for interaction recognition in Section 7.

6. Identification of groups

For identifying the group relation of a particular pair of pedestrians, we use a similar approach to [55], which aims at integrating two indicators of group motion. Although each of the two indicators can be used on its own for identification, their integration through an evaluation of individual uncertainties is shown to enhance their respective performances.

In explicit terms, first the likelihood that a particular observation comes from a certain distribution is computed. Consider that $\Delta = \left\{ |\vec{\delta}_{ij}| \right\}$, $\Phi = \left\{ \phi_{ij}, \phi_{ji} \right\}$, $\Omega = \left\{ \omega_{ij} \right\}$ are three sets of observations concerning a particular pair $\{p_i, p_j\}$. Let us also call the δ attribute models of group and non-group relation given in Section 4.2 as Γ_G^Δ and Γ_N^Δ.[4] We define the likelihood ratio for δ as in Equation (9), where similar approaches are adopted for ϕ and ω with corresponding models,

$$L^\delta = \log \left(\frac{\prod_{\forall \delta \in \Delta} \Gamma_G^\Delta \left(\delta | \nu, \sigma \right)}{\prod_{\forall \delta \in \Delta} \Gamma_N^\Delta \left(\delta \right)} \right). \tag{9}$$

Provided that $L^\delta > 1$, the pair is determined to be a group and otherwise a non-group based on δ. If both of L^δ and L^ϕ are >1 or <1, the two indicators are said to indicate to the same decision. However, if they contradict, the uncertainty of the individual decisions are compared and the one which has a smaller uncertainty is chosen. The uncertainty that a particular set of observations Δ comes from the distribution Γ_G^Δ is given by the divergence term,

$$D_G^\delta(\Delta||\Gamma_G^\Delta) = \max_\delta \left\{ \Delta(\delta) \log\left(\frac{\Delta(\delta)}{\Gamma_G^\Delta(\delta|\kappa)} \right) \right\}, \quad (10)$$

which implies that the divergence value that indicates the maximum dissimilarity is accounted for.[5]

Using the likelihood ratios given by Equation (9), we obtain the performance rates for identification of groups as in Table 2. We integrate the indicators of δ and ϕ mimicking [55] and contrast it with our alternative combination of δ and ω, in Table 2. It is clear that the proposed models and replacement of ϕ with ω improve the performance of group detection, where all 2p, 3p groups and non-groups are detected with almost perfect accuracy.

We compare these to the performance rates of two alternative methods of [56] and [55] in the last two columns of Table 2. The proposed motion models outperform the alternatives both at the individual level (i.e. using a single indicator of Δ, Φ or Ω) and at the integration level (i.e. using combinations of $\Delta + \Omega$ or $\Delta + \Theta$).

In addition, it is shown that, [56] achieves particularly low detection rates in case of 3p groups (0.70), since it is based on hard thresholds and cannot account for 2p and 3p groups at the same time. Moreover, although the overall performance rate of [56] appears to be higher than [55] and comparable to our method, it is principally due to the fact that non-groups dominate the samples by far and a high detection rate of non-groups is reflected in the overall detection rate quite clearly, whereas 2p and 3p groups do not have a big impact, even though they lie in the essence of the proposed problem. Therefore, our method gives more valuable estimates about group relation, treating all 2p and 3p groups and non-groups in more balanced manner.

7. Estimation of intensity of interaction

Intensity of interaction is in a sense very different from group belonging, since it is inherently time dependent.[6] Although such time dependence of $\mathcal{I}$ was not explicitly annotated, it is better to develop a detection (or estimation) system that may cope with it. For this reason, we use a Bayesian approach inspired by [57].

For this Bayesian approach, we decided to use the indicators of δ and ω omitting the observations of ϕ. The reason for this choice is two-fold:

(1) Employing ω instead of ϕ in detection of group relation is shown to improve the performance rates. Therefore, also in estimation of $\mathcal{I}$, we adopt a similar strategy.
(2) The δ and ω distributions have the peak locations at different positions as well as quite different variations around the peaks, as seen in Figure 3 and quantitatively discussed in [51]. On the other hand (Figure 4 and [51]), the peak of the ϕ distributions are always around $\pi/2$ and $-\pi/2$. Since all these distributions have the bulk of the observations around the same values, once a likelihood based approach is used, most observations would appear to be coming from the distribution with the highest peak.

Moreover, in order to obtain a sharper contrast between the distributions, we omit the distributions for medium intensities $\mathcal{I} = 1, 2$ and consider only the two extremes of 'lack of interaction' and 'intense interaction'.[7] In other words, we use these two models as boundaries and evaluate the probability of each pair to be in interaction relation with respect to them. Here, we expect to have high probabilities for pairs with $\mathcal{I} = 3$, a gradual decrease of probabilities for pairs with $\mathcal{I} = 2$ and $\mathcal{I} = 1$ and the lowest values for pairs with $\mathcal{I} = 0$.

Given an observation set $\{\delta, \omega\}$, the motion models and prior probabilities; we have the post probability of having interaction as,

$$P(i|\delta, \omega) = \frac{P(\delta, \omega|i)P(i)}{P(\delta, \omega)}, \quad (11)$$

where i stands for intensity value.

For assessing the likelihood, we use, the aforementioned models of $\mathcal{I} = 0$ and $\mathcal{I} = 3$. In addition, we assume that they are independent,

$$P(\delta, \omega|i) = P(\delta|i)P(\omega, i).$$

To compute the denominator of Equation (11), we make use of the fact that the sum of the post probabilities for all intensities add up to 1 and scale the numerator. Additionally, as for an initial value for our prior belief of having interaction, we adopt an equal probability to avoid a bias,

$$P_0(i) = \begin{pmatrix} 0.5 & 0.5 \end{pmatrix}.$$

Then we contrast three schemes where (i) we do not update the priors, or (ii) update them using a linear combination of the initial value and last computed probability

Table 2. Group identification performance using individual motion attributes and various combinations of them. The highest performance rates for 2p, 3p groups, non-groups, and highest overall performance value are presented in bold.

	N_p	Δ	Φ	Ω	$\Delta + \Phi$	$\Delta + \Omega$	[56]	[55]
2p	642	**0.98 ± 0.003**	0.93 ± 0.01	0.94 ± 0.007	0.97 ± 0.008	0.97 ± 0.004	0.90	0.89 ± 0.04
3p	225	0.95 ± 0.04	0.78 ± 0.03	0.91 ± 0.01	0.96 ± 0.01	**0.97 ± 0.009**	0.70	0.83 ± 0.07
Groups (2p+3p)	867	0.97 ± 0.01	0.89 ± 0.01	0.93 ± 0.007	0.97 ± 0.008	**0.97 ± 0.005**	0.84	0.87 ± 0.04
Non-group	10892	0.82 ± 0.06	0.85 ± 0.02	0.83 ± 0.01	0.93 ± 0.02	**0.96 ± 0.01**	0.96	0.95 ± 0.07
Total		0.83 ± 0.05	0.85 ± 0.02	0.84 ± 0.01	0.93 ± 0.02	**0.96 ± 0.009**	0.95	0.94

Table 3. The probability of being in interaction for 2p groups with respect to varying intensities of interaction and various values of α.

	Intensity			
	0	1	2	3
$\alpha = 0$	0.13 ± 0.09	0.30 ± 0.14	0.48 ± 0.12	0.64 ± 0.08
$\alpha = 0.5$	0.18 ± 0.12	0.35 ± 0.15	0.59 ± 0.14	0.74 ± 0.07
$\alpha = 1$	0.17 ± 0.11	0.33 ± 0.14	0.56 ± 0.15	0.72 ± 0.07
[56]	0.19	0.29	0.44	0.02
[55]	0.39	0.27	0.40	0.07

Table 4. The probability of being in interaction with respect to varying intensities of interaction for various values of α in 3p groups.

		Intensity			
		0	1	2	3
First neighbors	$\alpha = 0$	0.460 ± 0.501	0.457 ± 0.499	0.636 ± 0.482	0.400 ± 0.498
	$\alpha = 0.5$	0.500 ± 0.503	0.496 ± 0.501	0.642 ± 0.480	0.667 ± 0.479
	$\alpha = 1$	0.500 ± 0.503	0.435 ± 0.497	0.619 ± 0.486	0.667 ± 0.479
	[56]	0.00	0.00	0.23	0.86
	[55]	0.11	0.25	0.14	0.50
Second neighbors	$\alpha = 0$	0.23 ± 0.43	0.30 ± 0.46	0.39 ± 0.49	1.00 ± 0.00
	$\alpha = 0.5$	0.29 ± 0.46	0.30 ± 0.46	0.37 ± 0.48	1.00 ± 0.00
	$\alpha = 1$	0.29 ± 0.46	0.30 ± 0.41	0.37 ± 0.48	1.00 ± 0.00
	[56]	0.00	0.00	0.21	0.5
	[55]	0.00	0.83	0.00	0.33

value or (iii) adopt the last computed value as the prior for the next step. This can be expressed with

$$P_{t+1}(i) = \alpha P_0(i) + (1 - \alpha)P_t(i).$$

and the 3 cases described above realized using $\alpha = \{0, 0.5, 1\}$, respectively.

As an evaluation strategy, we make a binary decision of having interaction or not at every time instant of the observation duration of a particular pair. By this means, we find the ratio of interaction time over the entire observation, for the pairs labeled with $\mathcal{I} = 0, 1, 2, 3$ as in Table 3.[8]

As seen in Table 3, for all values of α we have an increasing probability of being in interaction relation as $\mathcal{I}$ increases. For instance in the $\alpha = 0$ case, pairs labeled as non-interacting are detected as interacting only for 13% of the observation duration, and thus as non-interacting in the remaining 87%. Moreover, 64–74% of

time, the intensely interacting pairs are detected to be in interaction relation, which we regard as satisfactory.

Note that, for instance, in the case of $\mathcal{I} = 3$, a rate of 74% does not mean that the method correctly identifies interaction 74% of the time and fails in the remaining 26%. This is due to the fact that even though a pair may be labeled as having intense interaction, they do not necessarily interact at every instant of their trajectory. It rather means that they are strongly involved in their interaction as long as they interact [58]. Therefore, we expect higher rates for larger intensities, but a 100% does not necessarily define perfect performance. Nevertheless, we expect the interaction to last longer as it gets more intense. Table 4 reports the performance for 3p groups, which yields similar observations.

Since, to our knowledge, this is the first work in estimating intensity of interaction, we cannot provide a comparison to an exact alternative method. However, we adapted the methods of [56] and [55] to account for interaction and computed the detection rates as in

Tables 3 and 4. Due to their formulations, the modified methods estimate $\mathcal{I}$ and not the possibility of being in interaction. Therefore, in contrast to our performance rates, in their formulation, higher values indicate better accuracy. Tables 3 and 4 clearly indicate a failure in estimating $\mathcal{I}$.

8. Conclusion and future work

This paper investigates the effect of varying levels of interaction on group dynamics, with the purpose of developing both a group recognition algorithm and a behavioral model to be deployed in social mobile robots.

To that end, we propose to analyze interpersonal distance, relative orientation and absolute difference of velocities. The empirical observations show that proposed attributes present differentiating characteristics in the cumulative sense with respect to group relation as well as intensity of interaction. Subsequently, we propose three models for each attribute, all of which are demonstrated to characterize the impact of group and interaction relation with significant accuracy.

We also treat each pair of pedestrians in the 2p and 3p groups separately and evaluate their group relation. These pairs are further treated to investigate the level of interaction intensity for assessing $\mathcal{I}$ with the proposed models at every time step. A course of probability of being in interaction relation is built for the entire trajectory through a Bayesian framework. The results indicate that estimation of $\mathcal{I}$ is very challenging but we achieve satisfactory performance rates particularly for 2p groups and first neighbors in 3p groups.

Thereby, we not only improve the group detection algorithm by proposing new indicators and their corresponding models, but also treat groups in more detail addressing varying levels of interaction, which stayed as an unexplored field to this day.

Although in this work we do not explicitly implement interaction behavior in robots, the proposed models could be used to develop a choice mechanism in robot aimed to maximize interaction, either adopting the models as cost functions on a grid of possible choices, or through gradient methods.

Notes

1. The term interaction refers to oral communication possibly accompanied with non-verbal elements such as gestures or gaze exchange, as defined by [17].
2. With document number (10-502-1).
3. For 8 of these groups interaction labels are missing
4. The corresponding models for ϕ and ω will be denoted similarly with Γ_G^Φ, Γ_N^Φ, Γ_G^Ω and Γ_N^Ω.
5. D_G^ϕ and D_G^ω are computed similarly with corresponding observations and models.
6. Obviously, groups can form and separate, but they usually exists as groups during the whole tracking process, while group interaction intensity naturally changes with time.
7. As noted in Section 5.2, the models for $\mathcal{I} = 0$ and $\mathcal{I} = 3$ are more clearly separated, their parameter variances being much smaller than the difference in the parameter values, than models for neighboring categories, such as $\mathcal{I} = 1$ and $\mathcal{I} = 2$.
8. Note that the pairs are labeled according to their intensity of interaction only during their course of interaction.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Z. Yücel got her BS degree from Bogazici University, Istanbul, Turkey in 2003 in Electrical and Electronics Engineering, and her MS and PhD degrees from Bilkent University, Ankara, Turkey in 2005 and 2010. She worked as a postdoctoral researcher at ATR labs in Kyoto, Japan for 5 years. In 2016 she was awarded a JSPS fellowship. Currently she serves as an assistant professor at Okayama University. Her research interests are investigation of human behavior and social interaction robots through signal processing, computer vision, pattern recognition techniques.

F. Zanlungo got his master in Physics from Milan University, Italy, in 2002, and his PhD in Physics from Bologna University, Italy, in 2007. From 2007 and 2009 he was a postdoctoral researcher at Bologna University, with a visiting fellowship at the Theoretical Physics Center of Marseilles, France. Since 2009 he is a researcher at ATR labs in Kyoto, Japan, and during the 2015–2016 academic year he was a Lecturer at Kingston University London, UK. His research interests concern the mathematical and numerical modelisation of complex system, with a particular interest in pedestrian crowd phenomena.

M. Shiomi received MEng and PhD degrees in engineering from Osaka University in 2004 and 2007. From 2004 to 2007, he was an intern researcher at the Intelligent Robotics and Communication Laboratories (IRC), Advanced Telecommunications Research Institute International (ATR). He is currently a group leader in the Agent Interaction Design department at ATR-IRC. His research interests include human-robot interaction, social touch, robotics for childcare, networked robots, and field trials.

References

[1] Burgard W, Cremers AB, Fox D, et al. The interactive museum tour-guide robot. In: Aaai/iaai. Madison, WI; 1998. p. 11–18.

[2] Thrun S, Bennewitz M, Burgard W, et al. Minerva: a second-generation museum tour-guide robot. In: ICRA. Vol. 3; Detroit (MI): IEEE; 1999.

[3] Gross HM, Boehme H, Schroeter C, et al. Toomas: interactive shopping guide robots in everyday use-final implementation and experiences from long-term field trials. In: IROS. St. Louis (MO): IEEE; 2009. p. 2005–2012.

[4] Kanda T, Shiomi M, Miyashita Z, et al. A communication robot in a shopping mall. IEEE Trans Rob. 2010;26:897–913.

[5] Huttenrauch H, Eklundh KS. Fetch-and-carry with cero: observations from a long-term user study with a service robot. In: Workshop on RHIC. Berlin, Germany: IEEE; 2002. p. 158–163.

[6] Garrell A, Sanfeliu A. Cooperative social robots to accompany groups of people. Int J Rob Res. 2012;31:1675–1701.

[7] Shiomi M, Zanlungo F, Hayashi K, et al. Towards a socially acceptable collision avoidance for a mobile robot navigating among pedestrians using a pedestrian model. Int J Soc Rob. 2014;6:443–455.

[8] Moussaïd M, Perozo N, Garnier S, et al. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. PLoS ONE. 2010;5:e10047.

[9] Murakami R, Morales Saiki LY, Satake S, et al. Destination unknown: walking side-by-side without knowing the goal. In: HRI. Bielefeld, Germany: ACM; 2014. p. 471–478.

[10] Templeton A, Drury J, Philippides A. From mindless masses to small groups: conceptualizing collective behavior in crowd modeling. Rev General Psychol. 2015;19:215–229.

[11] Voloshin D, Puzyreva K, Derevitsky I, et al. Groups and collectivities in crowd modeling: critical evaluation of the state-of-the-art and suggestions for further studies. In: Chugunov AV, Bulgov R, Kabanov Y, et al., editors. Digital transformation and global society. St. Petersburg, Russia: Springer; 2016. p. 492–497.

[12] Liu W, Lau RW, Wang X, et al. Exemplar-amms: recognizing crowd movements from pedestrian trajectories. IEEE Trans Multimedia. 2016;18:2398–2406.

[13] Zanlungo F, Ikeda T, Kanda T. Potential for the dynamics of pedestrians in a socially interacting group. Phys Rev E. 2014;89:012811.

[14] Zanlungo F, Brščić D, Kanda T. Spatial-size scaling of pedestrian groups under growing density conditions. Phys Rev E. 2015;91:062810.

[15] Zanlungo F, Yucel Z, Brscic D, et al. Intrinsic group behaviour: dependence of pedestrian dyad dynamics on principal social and personal features, 2017, preprint arXiv:170302672.

[16] McPhail C, Wohlstein R. Using film to analyze pedestrian behavior. Sociol Methods Res. 1982;10:347–375.

[17] Knapp M, Hall J. Nonverbal communication in human interaction. Belmont (CA): Wadsworth Publishing Company; 2009.

[18] Ketrow S. Nonverbal aspects of group communication. In: Frey LR, Gouran DS, Poole MS, editors. The handbook of group communication theory and research. California: SAGE; 1999. p. 251–287.

[19] Kleinke C. Gaze and eye contact: a research review. Psychol Bull. 1986;100:78–100.

[20] Argyle M, Dean J. Eye-contact, distance and affiliation. Sociometry. 1965;61(3):289–304.

[21] Hall ET. The hidden dimension. Vol. 6. Garden City (NY): Doubleday; 1963.

[22] Cristani M, Raghavendra R, Del Bue A, et al. Human behavior analysis in video surveillance: a social signal processing perspective. Neurocomputing. 2012;100:86–97.

[23] Hung H, Kröse B. Detecting f-formations as dominant sets. In: ICMI. Alicante, Spain: ACM; 2011. p. 231–238.

[24] Brdiczka O, Maisonnasse J, Reignier P. Automatic detection of interaction groups. In: ICMI; ACM; 2005. p. 32–36.

[25] Yi S, Li H, Wang X. Understanding pedestrian behaviors from stationary crowd groups. In: CVPR. Boston, MA; 2015. p. 3488–3496.

[26] Mehran R, Oyama A, Shah M. Abnormal crowd behavior detection using social force model. In: CVPR. Miami, FL; 2009. p. 935–942.

[27] Robicquet A, Alahi A, Sadeghian A, et al. Forecasting social navigation in crowded complex scenes; 2016, preprint arXiv:160100998.

[28] Shao J, Kang K, Loy CC, et al. Deeply learned attributes for crowded scene understanding. In: CVPR. Boston (MA): IEEE; 2015. p. 4657–4666.

[29] Habe H, Honda K, Kidode M. Human interaction analysis based on walking pattern transitions. In: ICDSC. Como, Italy: IEEE; 2009. p. 1–8.

[30] Costa M. Interpersonal distances in group walking. J Nonverbal Behav. 2010;34:15–26.

[31] Ge W, Collins RT, Ruback B. Vision-based analysis of small groups in pedestrian crowds. IEEE TPAMI. 2012;34:1003–1016.

[32] Sandikci S, Zinger S, de With PHN. Detection of human groups in videos. In: ACIVS. Ghent, Belgium; 2011. p. 507–518.

[33] Chandran AK, Poh LA, Vadakkepat P. Identifying social groups in pedestrian crowd videos. In: ICAPR. Kolkata, India: IEEE; 2015. p. 1–6.

[34] Yogameena B, Nagananthini C. Computer vision based crowd disaster avoidance system: a survey. Int J Disaster Risk Reduction. 2017;22:95–129.

[35] Kumagai K, Ehiro I, Ono K. Numerical simulation model of group walking for tsunami evacuees. In: Song W, Ma J, Fu L, editors. Pedestrian and evacuation dynamics 2016. Hefei, China: University of Science and Technology of China Press; 2016. p. 348–358.

[36] Helbing D. Models for pedestrian behavior, 1998; arXiv preprint cond-mat/9805089.

[37] Helbing D, Farkas I, Vicsek T. Simulating dynamical features of escape panic. Nature. 2000;407:487–490.

[38] Sumpter DJ. The principles of collective animal behaviour. Philos Trans Roy Soc London B: Biol Sci. 2006;361:5–22.

[39] Shao J, Change Loy C, Wang X. Scene-independent group profiling in crowd. In: CVPR. Columbus, OH; 2014. p. 2219–2226.

[40] von Krüchten C, Schadschneider A. Empirical study on social groups in pedestrian evacuation dynamics. Phys A: Stat Mech Appl. 2017;475:129–141.

[41] Sieben A, Schumann J, Seyfried A. Collective phenomena in crowds-where pedestrian dynamics need social psychology; 2017, preprint arXiv:170207012.

[42] Kok VJ, Lim MK, Chan CS. Crowd behavior analysis: a review where physics meets biology. Neurocomputing. 2016;177:342–362.

[43] Zanlungo F, Yucel Z, Kanda T. The effect of social roles on group behaviour; 2017, preprint arXiv:170203081.

[44] Gorrini A, Vizzari G, Bandini S. Age and group-driven pedestrian behaviour: from observations to simulations. Collective Dyn. 2016;1:1–16.

[45] Fang J, El-Tawil S, Aguirre B. Leader-follower model for agent based simulation of social collective behavior during egress. Safety Sci. 2016;83:40–47.

[46] Kawamoto K, Tomura Y, Okamoto K. Learning pedestrian dynamics with kriging. In: ICIS. Kochi, India: IEEE; 2016. p. 1–4.

[47] Li X, Xiong S, Duan P, et al. A study on the dynamic spatial-temporal trajectory features of pedestrian small group. In: DCIT; IEEE; 2015. p. 112–116.

[48] Fu YW, Li M, Liang JH, et al. Modeling and simulating the walking behavior of small pedestrian groups. In: Asiasim. Kitakyushu, Japan: Springer; 2014. p. 1–14.

[49] Li N, Zhang Y, Luo W, et al. Instant coherent group motion filtering by group motion representations. Neurocomputing. 2017;266:304–314.

[50] Zanlungo F, Bršćić TDI, Kanda T. Dataset: pedestrian tracking with group annotations; 2017. Available from: http://www.irc.atr.jp/sets/groups/

[51] Yücel Z, Zanlungo F, Shiomi M. Supplemental material; 2017. Available from: http://www.irc.atr.jp/zeynep/pub/supp.pdf (to be published online after acceptance).

[52] Diestel R. Graph theory. Vol. 173, Graduate texts in mathematics. New York (NY): Springer-Verlag Berlin and Heidelberg GmbH; 2000.

[53] Mardia KV, Jupp PE. Directional statistics. Hoboken (NY): Wiley; 2000.

[54] Abramowitz M, Stegun I. Handbook of mathematical functions with formulas, graphs, and mathematical tables. Vol. 55. New York (NY): Dover publications; 1964.

[55] Yücel Z, Zanlungo F, Ikeda T, et al. Deciphering the crowd: modeling and identification of pedestrian group motion. Sensors. 2013;13(1):875–897.

[56] Yücel Z, Ikeda T, Miyashita T, et al. Identification of mobile entities based on trajectory and shape information. In: IROS. San Francisco, CA; 2011. p. 3589–3594.

[57] Dražen Bršćić ZF. Kanda T. Modelling of pedestrian groups and application to group recognitio. In: MIPRO. Opatija, Croatia; 2017.

[58] Yücel Z. Detection examples; 2017. Available from: https://youtu.be/qkuQC1OJU3Q