Do walking pedestrians stabily interact inside a large group? Analysis of group and sub-group spatial structure

Francesco Zanlungo (ZANLUNGO@Atr.Jp)

Takayuki Kanda (KANDA@Atr.Jp)

Intelligent Robotics and Communication Laboratories, ATR, Kyoto, Japan &
JSPS CREST, Tokyo, Japan

Abstract

We combine video recording and laser range tracking to analyse the geometrical structure of groups of walking pedestrians socially interacting. By recording their relative position and observing their social interaction for a large enough time span we can analyse the stability and universality of their spatial structure. We find that while 2-pedestrian and 3-pedestrian groups have a relatively "time stable" and "universal" geometrical structure (an abreast formation for pairs, and a "V" formation for triads, with the central pedestrian walking slightly behind), no such structure emerges for larger groups. Nevertheless, these larger groups result to be composed of time stable two or three people sub-groups with the same "universal" geometrical structure of isolated pairs and triads.

Keywords: Group dynamics; proxemics.

Introduction

The spatial relationship of socially interacting people, i.e proxemics, has been largely studied, starting from the seminal works of (Hall, 1969) and (Kendon, 1990) in which the distances between and spatial distribution of people participating in social activities have been investigated. At the same time other researchers have investigated the size of social groups (by size of a group we mean the number of its components) and the probability distribution of these sizes (James, 1953; Coleman & James, 1961). Many of the aforementioned studies are based on "ecological" observations, i.e. studies in which people are observed in their natural environment while reducing as much as possible the effect of observations on their behaviour. While these studies are obviously based on observations of people behaviour in public spaces, until recently they did not focus on one of the main components of public spaces population (at least in modern urban areas), i.e. walking pedestrians. Here by pedestrian we mean a person in a public space moving between two locations for practical or recreational purposes, or even "wandering around" an environment without any particular goal. Pedestrians are often part of social groups with a specific proxemics determined by their dynamical constraints (the fact that they are walking), but the study of these groups has been traditionally made difficult by the fact that they are moving and located inside a crowd, which makes the observation of their behaviour more troublesome. Nevertheless, lately a few works have focused on the behaviour of these groups (Moussaïd, Perozo, Garnier, Helbing, & Theraulaz, 2010; Costa, 2010), due also to the growing interest in crowd behaviour of which groups are a non negligible component (Aveni, 1977). This interest is due to the necessity of simulating crowd behaviour to design better pedestrian facilities (Helbing, Farkas, Molnar, & Vicsek, 2002), but also to reproduce faithfully the behaviour of virtual crowds for the entertainment industry (Karamouzas & Overmars, 2012).

While (Moussaïd et al., 2010) report that the spatial structure of a freely walking (i.e. not environmentally constrained) *n*-pedestrian group is a line of abreast walking pedestrians, that tends to be bent into "V" or "U" formations (i.e., the pedestrians on the sides walk ahead) when the crowd density grows, (Costa, 2010) reports different spatial structures, suggesting for example that the "V" structure is the most occurring one for three people groups (regardless of crowding), and that larger groups tend to split into smaller sub-groups. Nevertheless (Costa, 2010) does not analyse the possible effects of environmental constraints on observed behaviours (the width of the sidewalks pedestrians were observed in was comparable to the group spatial sizes), and does not provide a quantitative study of 2D space structures, nor follows groups for a time interval long enough to analyse their change in time.

The difference between these observations leads us to two related problems in walking group proxemics, to which we try to bring insight in this work:

- Do *n*-pedestrian groups (i.e. groups composed of *n* members) have a *prevalent* geometrical structure? Here by *prevalent* we mean *universal* (common to almost all groups, or at least present in a large majority of them) and *time stable* (i.e. the positions of pedestrians in an unconstrained group will be given by small oscillations around those of the *prevalent* structure).
- If such an overall structure does not exist for a *n*-pedestrian group, is it possible to find it at the sub-group level?

In order to analyse these issues, we have to observe pedestrians in a situation in which collision avoiding and environmental constraints are not very strong (otherwise it would be impossible to identify the "universal" structure). Furthermore, we have to combine the need to measure with good detail (and for long enough time) the position of pedestrians, with that of observing their social interactions, in order to analyse the group structure. If a large pedestrian group is divided into smaller sub-groups we may expect social interaction inside sub-groups to be more frequent than between different sub-groups, and for this reason in many cases the

belonging of sub-groups to a larger group structure may be determined only if the observation is long enough. To attain this goal we combine a laser range finder tracking technology, that allows us to determine with good precision the position and velocity of each pedestrian in a large public environment, with frontal view (face level) video recordings, that allow us to analyse their social interactions (Fig. 1). As a result, we can follow pedestrian groups for a relatively long time while examining both their social and spatial interactions from a (respectively) qualitative and quantitative point of view. We performed these observations in a large area completely dedicated to pedestrian motion, and in a location and time in which the pedestrian density was relatively low, in order to be able to observe the behaviour of "unconstrained" groups.

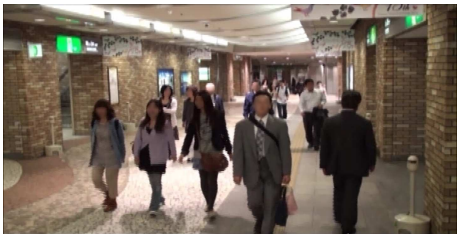

Figure 1: Video camera frame of the experimental area. A sensor pole is visible in the bottom-right corner.

Methodology and definitions

Data collection

We tracked pedestrian motion in two areas of a pedestrian underground facility in Umeda, Osaka (Japan), for a total time of 6 hours in each area. The pedestrian areas consist of a few corridors connecting a railway station to a shopping mall, each area being around 500 m^2. The environments are described in detail in (Zanlungo, Chigodo, Ikeda, & Kanda, 2012; Zanlungo, Ikeda, & Kanda, 2012), and the pedestrian tracking data are available at (Zanlungo, 2012). The average pedestrian density in the environment resulted to be ≈ 0.03 pedestrians per square meters while the width of the corridors varied between 4 and 7 meters, meaning that the average distance from a pedestrian to another pedestrian outside their group, or to a wall, is expected to be larger than the spatial size of the group (for example a group of 4 people walking in an abreast formation, assuming an interacting distance of 1 meter between first neighbours, should be 3 meters wide, compared to an expected distance between pedestrians > 5 meters at such a density). We can thus assume pedestrians to be fairly "unconstrained" by the environment and freely walk in their preferred spatial formation for most of the time.

We used 16 Hokuyo UTM sensors (situated on poles close to the environment walls not to hinder pedestrian motion, Fig. 1) and the tracking algorithm (Glas, Miyashita, Ishiguro, & Hagita, 2009) to determine pedestrian positions at times intervals $\delta t \approx 50$ ms with precision ≈ 50 mm. We smoothed the tracked positions on time windows $\Delta t = 500$ ms, to further improve the tracking precision. Pedestrian velocity is computed as the ratio of the displacement vector between two (smoothed) consecutive tracking positions (eq. 2), and has an expected precision $\approx 50 - 100$ mm/s. As we will see (see also the discussion in (Zanlungo, Chigodo, et al., 2012)) this tracking precision is negligible with respect to the typical interaction distances and velocities of pedestrians.

We also video recorded each experimental area with two different "frontal view" cameras (Fig. 1), located in such a way to allow observing the social interaction between the pedestrians for a sufficient long time (pedestrians are usually tracked and observed for a time of the order of tens of seconds). This camera based observation of social interaction was possible because the cameras are not needed for tracking and the density was relatively low. A "coder" (a non-technical staff member of our laboratory), was asked to identify all the pedestrian social groups in the environment and their members. In order to do that, she was asked to use all the information available from the videos, such as relative position, coherent motion, and social clues including conversation, gaze exchange and even age, sex and clothing (for example she identified a relatively large flock of coherent moving people as a single group because they were all dressed for and carrying similar hiking equipment). She was asked to identify only groups of which she could establish the nature without any reasonable doubt (i.e. she was asked to strongly avoid false positives, while false negatives were allowed). Furthermore, the coder was asked to annotate the groups, and the individuals in each group, for which she could without any doubt identify explicit social interaction clues (namely conversation, or explicit gaze exchange). Table 1 shows the number and size of labelled groups, distinguishing between "fully connected" groups (FCG) for which she could observe explicit social interaction between all the members, and "disconnected" groups (DG) that seemed to be related on the basis of other visual clues but for which explicit interaction could not be observed (or was observed only in smaller subgroups). To avoid false positives, only FCG are analysed in this paper. (The coder identified also six 5-pedestrian groups, six 6-pedestrian groups, one 7-pedestrian group and one 18-pedestrian group, not analysed in this work due to the small size of the samples).

Size	$n=2$	$n=3$	$n=4$
FCG	1126	114	17
DG	91	34	14

Table 1: Observed fully connected (FDG) and disconnected groups (DG) for each group size n.

Definitions

In order to identify the existence of an "universal" structure in a pedestrian group, it is necessary to study it in the correct reference frame. The most natural candidate is the group's

"centre of mass frame" (see Fig. 2). Let us call

$$\mathbf{x}^e(t_k) = (x^e(t_k), y^e(t_k)) \qquad t_k = k\Delta t \tag{1}$$

the position of a pedestrian in the "environment" reference frame, smoothed on regular $\Delta t = 500$ ms time windows, and

$$\mathbf{v}^e(t_k) \equiv \frac{\mathbf{x}^e(t_{k+1}) - \mathbf{x}^e(t_k)}{\Delta t} \tag{2}$$

the corresponding pedestrian velocity. Let us consider a *n*-pedestrian (from now on *n*-p) (sub-)group with position and velocities

$$\{\mathbf{x}_i^e(t_k), \mathbf{v}_i^e(t_k)\} \qquad i = 1, ..., n,$$

and define the group "centre of mass" position and velocity

$$\mathbf{X}(t_k) = \frac{\sum_i \mathbf{x}_i^e(t_k)}{n} \qquad \mathbf{V}(t_k) = \frac{\sum_i \mathbf{v}_i^e(t_k)}{n}. \tag{3}$$

Let us name GCMF the group centre of mass frame (at time t_k) with origin in $\mathbf{X}(t_k)$ and the y axis aligned to $\mathbf{V}(t_k)$, i.e. with axis versors

$$\hat{\mathbf{e}}_x(t_k) = \left(\frac{V_y(t_k)}{V(t_k)}, -\frac{V_x(t_k)}{V(t_k)}\right) \qquad \hat{\mathbf{e}}_y(t_k) = \frac{\mathbf{V}(t_k)}{V(t_k)} \tag{4}$$

(from now on we remove t from notation for simplicity's sake). The position of pedestrian i in the GCMF is then $\mathbf{x}_i = (x_i, y_i)$ with

$$x_i = (\mathbf{x}_i^e - \mathbf{X}) \cdot \hat{\mathbf{e}}_x \qquad y_i = (\mathbf{x}_i^e - \mathbf{X}) \cdot \hat{\mathbf{e}}_y. \tag{5}$$

We also define the polar coordinates (r_i, θ_i) through

$$x_i = r_i \sin\theta_i \qquad y_i = r_i \cos\theta_i, \tag{6}$$

where $r = \sqrt{x^2 + y^2}$ represents the distance of the pedestrian from the centre of mass (for 2-p groups $2r$ is the distance between pedestrians). Fig. 2 illustrates the previously defined quantities in the 2-p case, for which the following holds

$$r_2 = r_1, \quad x_2 = -x_1, \quad y_2 = -y_1, \quad \theta_2 = \pi + \theta_1. \tag{7}$$

It is important to quantitatively study the structure of the group in the GCMF, because the distinctive feature of walking groups is the presence of a centre of attention, namely the direction towards their current sub-goal, which is identified by the velocity of the group. The geometrical structure of the pedestrian groups is determined by the necessity to maintain the focus on the walking direction, and for this reason we will not consider non-moving pedestrian groups (i.e., data points in which at least a group member has $v_i(t_k) < 500$ ms, a threshold that corresponds to a velocity 3 standard deviations smaller than the typical pedestrian velocity (Daamen & Hoogendoorn, 2006); see also (Zanlungo, Chigodo, et al., 2012) for a discussion of this threshold). Since we collected data in a passing point between a station and a shopping centre without attraction points (Zanlungo, Chigodo, et al., 2012), only $\approx 5\%$ of data are not considered.

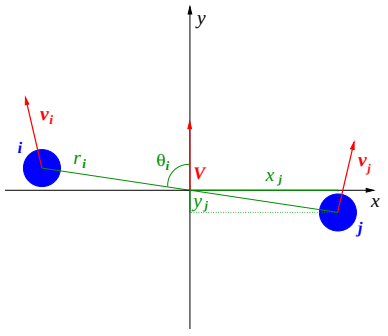

Figure 2: GCMF variables definition.

It is clear that if a *universal* and *time stable* *n*-p formation exists, then at (almost) all times and for (almost) all groups, the GCMF pedestrian positions should be close to those determined by such a structure. We will compute the empirical 2D probability distribution function (pdf) $\rho(x,y)$ for each n averaging on all groups, pedestrians i and and times t_k, and state that such a *universal* and *stable* formation exists only if $\rho(x,y)$ has n well defined maxima. The formation will be then empirically determined by the position of these maxima.

Results

Whole group GCMF structure for n-p groups

Fig. 3 shows the pedestrian pdf $\rho(x,y)$ for 2-p, 3-p and 4-p groups. The 2-p and 3-p groups have a well defined geometrical structure in the GCMF, i.e. their ρ has, respectively, 2 and 3 well defined maxima, one for each pedestrian. Such a structure is not present for 4-p groups, whose pdf has many local maxima. As we will see, a well defined structure emerges for 4-p only after the whole group is properly divided in two 2-p sub-groups.

2-p groups

Let us identify the leftmost ($x < 0$) pedestrian as P_1 and the rightmost one as P_2. Figs. 4 **a)** and 4 **b)** respectively show the $\rho(r_1)$ and $\rho(\theta_1)$ pdfs, while Table 2 shows the average values and standard deviations of all variables. While $\rho(y_1)$ and $\rho(\theta_1)$ are well described by a Normal distribution (i.e. a von Mises (Mardia & Jupp, 2009) one for the circular variable θ), $\rho(x_1)$ and $\rho(r_1)$ are not, and for this reason we report also the (approximate) value of their maxima. Our data show that 2-p groups walk abreast at a distance smaller than twice the average shoulder width (Pheasant, 1986). Such a

	$\rho(x_1)$	$\rho(y_1)$	$\rho(r_1)$	$\rho(\theta_1)$
$<>$	-387	-2	417	-89
σ	87	166	105	17
max	-360	0	365	-90

Table 2: Average values ($<>$), standard deviations (σ) and maxima for the $\rho(x_1)$, $\rho(y_1)$, $\rho(r_1)$ and $\rho(\theta_1)$ pdfs (linear variables in mm, circular in degrees).

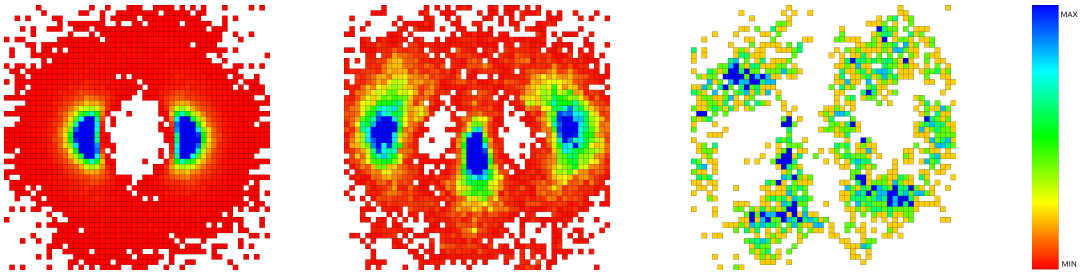

Figure 3: $\rho(x,y)$ for 2-p, 3-p and 4-p groups (respectively, from the left). Blue corresponds to maximum density, red to minimum density (colour bar on the right). Each figure covers a 2×2 meters area.

configuration is determined by the need of maintaining both partners in each other's field of view (or better at the border of it) while keeping the main attention focus on the walking direction. By walking abreast the partner is reachable for gaze exchange and conversation through a torsion of the neck (pedestrians can go on walking towards their goal with no gait modification) while the distance allows for conversation without collision or excessive proximity. This configuration is the most comfortable one for walking and interacting socially, but it cannot be extended to larger groups, because in a $n > 2$ abreast configuration the position of first neighbours would hinder gaze contact and conversation with second or larger neighbours. As clear from Fig. 3 and discussed below, this affects larger group configurations.

3-p groups

Let us name the pedestrians P_1, P_2 and P_3 starting from the leftmost to the rightmost ($x_1 < x_2 < x_3$). It is easier to understand the relation between 2-p and 3-p group structures if we analyse the 3-p variables in all possible 2-p sub-group GCMFs. A variable with subscript ij will denote the position of P_i in the (P_i, P_j) GCMF. In this way, for example, $2r_{12}$ is the relative distance between the leftmost pedestrian and the central one (first neighbours), and so on for each variable and pedestrian pair i, j with $i < j$ (the $i > j$ case can be obtained through eq. 7). Tables 3 and 4 report the values of all such variables, while Fig. 5 **a)** compares the $\rho(r_{12})$ and $\rho(r_{13})$ pdfs to the $\rho(r_1)$ distribution of the 2-p case. The same comparison is performed for θ in Fig. 5 **b)**.

The first neighbour distance distributions $\rho(r_{12})$, $\rho(r_{23})$ are very similar to the 2-p distance distribution, while the second neighbour distance distribution $\rho(r_{13})$ may be very well represented by a Normal one. Defining $\bar{r}$ as the value r for which $\rho(r)$ is maximum, we see that $\bar{r}_{13} \approx 2\bar{r}_{12} \approx 2\bar{r}_{23} \approx 2\bar{r}_1$, i.e. 3-p groups members try to maintain between them the same distances that occur between 2-p groups members, which we interpret as a strong sign of social interaction involving all

	$\rho(r_{12})$	$\rho(r_{13})$	$\rho(r_{23})$	$\rho(\theta_{12})$	$\rho(\theta_{13})$	$\rho(\theta_{23})$
$<>$	437	738	441	-74	-89	-105
σ	125	132	169	23	15	25
max	365	700	365	-80	-90	-110

Table 3: Average values ($<>$), standard deviations (σ) and maxima of the pdf ρ for the polar 3-p group variables (r in mm, θ in degrees).

	$\rho(x_{12})$	$\rho(x_{13})$	$\rho(x_{23})$	$\rho(y_{12})$	$\rho(y_{13})$	$\rho(y_{23})$
$<>$	-342	-686	-344	114	1	-112
σ	104	155	112	257	278	263
max	-350	-680	-350	70	0	-110

Table 4: Average values ($<>$), standard deviations (σ) and maxima of the pdf ρ for the Cartesian 3-p group variables (in mm).

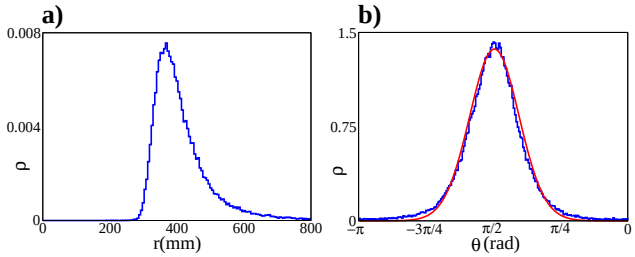

Figure 4: **a)**: 2-p pdf for r_1. **b)**: 2-p pdf for θ_1, compared to a best fit von Mises distribution.

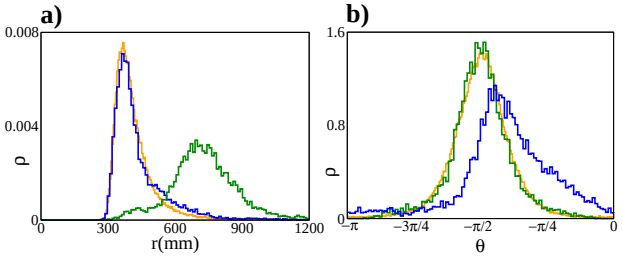

Figure 5: **a)**: $\rho(r_{12})$ distribution (blue) compared to the $\rho(r_{13})$ distribution in green and to the 2-p $\rho(r_1)$ distribution in orange. **b)**: $\rho(\theta_{12})$ distribution (blue) compared to the $\rho(\theta_{13})$ distribution in green and to the 2-p $\rho(\theta_1)$ distribution in orange.

three members. The similarity between the $\rho(\theta_{13})$ and the (2-p) $\rho(\theta_1)$ distributions is particularly striking, suggesting direct social interaction between P_1 and P_3. This interaction would hardly be possible if we had the same angle distribution for θ_{12} and θ_{23}, because the central pedestrian would hinder the communication. As a result, the central pedestrian steps slightly back (so that their partners remain in the vision field), and the angle between the three of them is, in average, ≈ 150 degrees. This "V" configuration had already been reported as the most often occurring one for walking triads (Costa, 2010). (Moussaïd et al., 2010) explain this configuration as the effect of a trade-off between easiness of communication and collision avoiding efficiency, assuming that the free-walking triads walk abreast. On the basis of our data, that as we already stated should not be strongly influenced by environmental constraints, we suggest that the "V" configuration is attained for maximum easiness of communication between the three partners, and occurs even for freely walking pedestrians.

4-p groups

(Moussaïd et al., 2010) report that freely walking 4-p groups assume an abreast configuration, that tends to bend in a "U" one with growing pedestrian density, in order to avoid collisions. According to this *abreast hypothesis*, we should see four clear maxima in a row in Fig. 3 on the right. Furthermore, if we name the pedestrians $P_1,...,P_4$ with $x_1 < x_2 < x_3 < x_4$ in the GCMF, the first neighbour variable distributions, such as

$$\rho_{fn}(\theta) = (\rho(\theta_{12}) + \rho(\theta_{23}) + \rho(\theta_{34}))/3 \qquad (8)$$

and the analogously defined $\rho_{fn}(r)$, $\rho_{fn}(x)$ and $\rho_{fn}(y)$, should resemble the 2-p group distributions. This hypothesis is clearly not supported by our data (Figs. 6, 7 **b)**). On the opposite, (Costa, 2010) reports different geometrical structures for 4-p groups, but none of these seems *prevalent* in our data (no clear maxima in Fig. 3 right).

We may then use a different *sub-group hypothesis*, assuming that the 4-p group may be divided in two sub-groups of 2 pedestrians, with "strong interaction" inside the sub-group and weaker interaction outside it. According to this hypothesis, we may find a *universal* and *time stable* structure only at the sub-group level. Let us rename the pedestrians in the following way. We name P_1 the pedestrian with the minimum x value in the 4-p GCMF, and compute the point

$$\mathbf{p}_2 = \mathbf{x}_1 + r_{int}\,\hat{\mathbf{e}}_x \qquad (9)$$

where $r_{int} = 730$ mm is the maximum for the pdf of distances for 2-p groups (i.e. twice the GCMF r value reported in Table 2). We then name P_2 the pedestrian whose euclidean distance to $\mathbf{p}_2$ is minimum, and P_3 and P_4 the remaining two. Let us finally name $\rho_{sg}(r)$, $\rho_{sg}(x)$, $\rho_{sg}(y)$ and $\rho_{sg}(\theta)$ the pdfs for the corresponding variables when averaged over all subgroups, as in

$$\rho_{sg}(r) = (\rho(r_{12}) + \rho(r_{34}))/2. \qquad (10)$$

The average values and standard deviations for these distributions are shown in Table 5, while Fig. 6 shows the $\rho_{sg}(x,y)$ 2D pdf, presenting two clear maxima. Fig. 7 **a)** shows the comparison between the pdf $\rho_{sg}(r)$ with the (2-p) r_1 distribution. Fig. 7 **b)** performs the same comparison for θ variables, including also the $\rho_{fn}(\theta)$ (*abreast hypothesis*) distribution. The presence of two clear maxima in $\rho_{sg}(x,y)$ suggests

	$\rho_{sg}(x)$	$\rho_{sg}(y)$	$\rho_{sg}(r)$	$\rho_{sg}(\theta)$
$<>$	-403	-54	530	-97
σ	180	347	195	33
max	-360	0	370	-90

Table 5: Average values ($<>$), standard deviations (σ) and maxima of the pdfs $\rho_{sg}(x)$, $\rho_{sg}(y)$, $\rho_{sg}(r)$ and $\rho_{sg}(\theta)$ (linear variables in mm, circular in degrees).

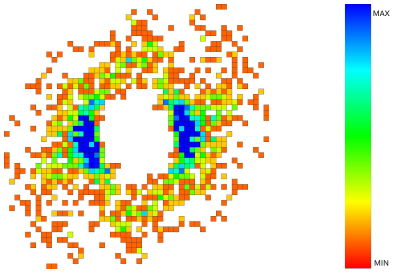

Figure 6: $\rho_{sg}(x,y)$ under the sub-groups hypothesis. The figure covers a 2×2 meters area.

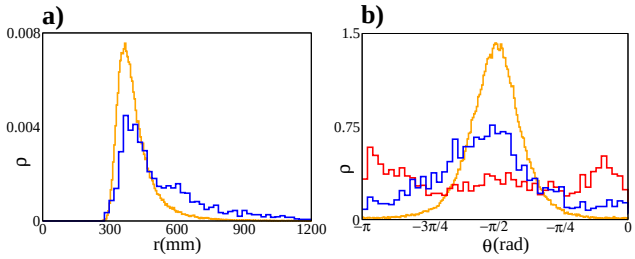

Figure 7: **a)**: Comparison between the 2-p pdf for $\rho(r_1)$ in orange and $\rho_{sg}(r_1)$ (*sub-group hypothesis*) in blue. **b)**: Comparison between the 2-p pdf for $\rho(\theta_1)$ in orange; $\rho_{fn}(\theta)$ (*abreast hypothesis*) in red; and $\rho_{sg}(\theta_1)$ (*sub-group hypothesis*) in blue.

that a *universal* and *stable* structure is indeed present at the sub-group level. The $\rho_{sg}(r)$, $\rho_{sg}(x)$, $\rho_{sg}(y)$ and $\rho_{sg}(\theta)$ distributions in a 4-p group result to be a "perturbed version" of the "proper" 2-p variable distributions, the perturbation being determined by the interaction with the members of the other sub-group. We can give different interpretations for the absence of a *universal* 4-p spatial configuration, that probably act as con-causes. Since the completely abreast configuration results to be uncomfortable even in the 3-p configuration, it

results to be even more problematic for 4 pedestrians. A solution is to, in a way similar to the "V" 3-p configuration, walk in a "U" or in a "staggered" configuration. Another solution would be to walk on two different abreast rows, for example roughly on the corners of a square. While this configuration has the shortcoming that the members on the back are not in the field of view of, nor can be comfortably watched by, the pedestrians on the front, it has the strong point that reduces the maximum distance between members of more than a factor 2, and it is more efficient in case of collision avoiding or other environmental influences. If we observe particular groups for a short time, as done by (Costa, 2010), we may observe all these occurrences, but, as shown by Fig. 3 (right) none of them is *universal* and *time stable*. What is stable (Fig. 6) is the association of pedestrians in pairs. This does not mean that this pairwise association is invariant (i.e. that the pair composition does not change in time) but the data suggest that this association is far more stable than the whole-group structures. Even though we do not report a quantitative analysis, we qualitatively observed a few 5-p and 6-p groups, and noticed also for these groups the tendency to part in stable 2-p or 3-p sub-groups.

Conclusions and future work

Our observations lead us to the conclusion that the dynamical constraints make social interaction between members of walking groups difficult to attain for subgroups larger than two units. For this reason, larger groups have a tendency to form relatively stable 2-p subgroups. Obviously, since odd-p groups cannot be divided in pairs without excluding a pedestrian, for 3-p (sub-)groups we find a configuration almost as stable as the 2-p one.

Regarding the difference between our observations and those of (Moussaïd et al., 2010), we could speculate on cultural and environmental features. Nevertheless we believe that the main difference may be operational, i.e. that our experimental setting using cameras and laser sensors allowed us to identify sub-groups of pedestrians as part of larger groups, even if their interaction was limited in time. On the contrary, observation methods based on shorter time windows may be biased towards large spatial configurations, because they would tend to consider sub-groups as separate entities. For this reason, despite these differences with previous works, and the limited amount of data for large groups, we may speculate that our work is universal in showing that 2-p and 3-p sub-groups are far more stable than larger configurations, with possible effects on the behaviour of crowds.

In our future work we plan to analyse the dynamical features that have been ignored in this work. In detail we want to study: 1) how pedestrians in (sub-)groups behave away from the equilibrium configuration, and 2) the time stability of, and interaction between, sub-groups inside a larger group. Such an analysis should bring further insight on the social dynamics of walking pedestrians and allow us to extend pedestrian group models in such a way to describe the findings of this work.

Acknowledgements

We thank Tetsushi Ikeda for the organisation of the data collection experiments and Kanako Tomita for labelling the social interactions of pedestrians from video recordings. We also thank the anonymous reviewers whose comments helped in improving the quality of our paper.

References

Aveni, A. (1977). The not-so-lonely crowd: Friendship groups in collective behavior. *Sociometry*, 96–99.

Coleman, J., & James, J. (1961). The equilibrium size distribution of freely-forming groups. *Sociometry*, *24*(1), 36–45.

Costa, M. (2010). Interpersonal distances in group walking. *Journal of Nonverbal Behavior*, *34*(1), 15–26.

Daamen, W., & Hoogendoorn, S. (2006). Free speed distributions for pedestrian traffic. In *TRB-annual meeting.*

Glas, D., Miyashita, T., Ishiguro, H., & Hagita, N. (2009). Laser-based tracking of human position and orientation using parametric shape modeling. *Advanced robotics*, *23*(4), 405–428.

Hall, E. (1969). *The hidden dimension*. Anchor Books.

Helbing, D., Farkas, I., Molnar, P., & Vicsek, T. (2002). Simulation of pedestrian crowds in normal and evacuation situations. *Pedestrian and evacuation dynamics*, *21*.

James, J. (1953). The distribution of free-forming small group size. *American Sociological Review;*.

Karamouzas, I., & Overmars, M. (2012). Simulating and evaluating the local behavior of small pedestrian groups. *Visualization and Computer Graphics, IEEE T.*, *18*(3), 394–406.

Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters* (Vol. 7). Cambridge.

Mardia, K., & Jupp, P. (2009). *Directional statistics* (Vol. 494). Wiley.

Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., & Theraulaz, G. (2010). The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PLoS One*, *5*(4), e10047.

Pheasant, S. (1986). *Bodyspace: anthropometry and design*. London: Taylor and Francis.

Zanlungo, F. (2012). Available from `https://sites.google.com/site/francescozanlungo/pedestriandata`

Zanlungo, F., Chigodo, Y., Ikeda, T., & Kanda, T. (2012). Experimental study and modelling of pedestrian space occupation and motion pattern in a real world environment. *Pedestrian and evacuation dynamics*.

Zanlungo, F., Ikeda, T., & Kanda, T. (2012). A microscopic social norm model to obtain realistic macroscopic velocity and density pedestrian distributions. *PloS one*, *7*(12), e50720.