

Experimental study and modelling of pedestrian space occupation and motion pattern in a real world environment

F. Zanlungo, Y. Chigodo, T. Ikeda and T. Kanda

June 13, 2012

Abstract

In this work we tackle the statistical study and modelling of the usage of space by pedestrians in a real world environment. A large amount of pedestrian trajectories is collected in a corridor used just as a transition place, and the density and velocity distributions are analysed as functions of the distance from the walls. The empirical data are fitted to a model assuming the density and velocity to be determined through a Boltzmann factor by a comfort function depending on the distance from the walls and assuming a maximum on the left side of the corridor (Japanese traffic convention). The empirical data are then compared to numerical simulations using pure collision avoidance models, to better analyse the influence of the environment on the pedestrian distribution and to investigate how to introduce in collision avoiding the bias that makes people walk preferentially on a given side of a corridor.

1 Introduction

There is still a gap in pedestrian studies between analyses based on controlled experiments [1] and real world data collection [2]. The former are performed in very simple environments, data are recorded with high precision and usually allow for an effective modelling of the results and comparison with simulations. Nevertheless these experiments may deviate strongly from real world behaviour, both for the artificial nature of the environments in which they are performed, and for the unnatural behaviour that subjects may exhibit. On the opposite real world data collection presents lower data quality, and real world environments are usually too complex to allow mathematical analysis and modelling, making a generalisation of the obtained information very difficult.

This work is an attempt to create a bridge between these two approaches, analysing pedestrian behaviour in a real world environment which is simple enough to be described with a mathematical model. We study the density and velocity patterns of pedestrians in corridors used only as transition places, checking that these quantities depend only on the distance from the walls. In Japan,

where data were collected, people walk on the left side of corridors [2], and our analysis, that divides them according to their walking direction, confirms this tendency. We found that the pedestrian density in each flow goes to zero close to the walls, and assumes a maximum in a point closer to the left wall than to the right one. The pedestrian velocity does not change more than 10% from the minimum value, assumed close to the walls, to the maximum, located on the right of the density maximum.

We model the data assuming that the pedestrian density distribution is given by a Boltzmann factor, whose “Hamiltonian” is a (dis)comfort function, assuming a minimum value where pedestrians can walk more easily. We also model the velocity distribution as a Gaussian centred around a preferred value that depends on the distance from the walls.

Pure collision avoiding behaviour leads to emergent self-organisation [1], and it has been proposed that a bias in the perceived position of other pedestrians may be a way to account for the (culture dependent) tendency to walk on a given side of a corridor [3]. We use computer simulations to understand to which extent the observed position and velocity distributions may be obtained using a pure collision avoiding model.

2 Environment

The purpose of this work is to study the behaviour of pedestrians in a real world environment with a simple geometry. Ideally, we would like to study a uniform corridor used just as a transition place, and since such an environment is expected to be invariant along the direction of the corridor, we want to study the density and velocity dependence on the only distance from the walls. We expect two flows to be present, i.e. identifying the x axis of the Cartesian system with the direction of the corridor, one flow will be given by pedestrians with a positive x component velocity, $v^x > 0$, the other by pedestrians with $v^x < 0$.

As an approximation to this ideal environment we studied an underground area in Umeda (Osaka) where some corridors connect a shopping area with a railway station. These corridors are quite uniform, without any shop, and used almost only by people transiting between the station and the shopping area, and thus their structure is quite similar to the ideal one. The pedestrian positions were recorded in two working day afternoons using 2D laser range finders, a technology that allows for tracking with an error of order 60 mm [4]. The tracked trajectories were smoothed in a time window $\delta t = 200$ ms, and the velocity obtained from the smoothed trajectories as $\mathbf{v}_i(t) = (\mathbf{x}_i(t + 2\delta t) - \mathbf{x}_i(t))/(2\delta t)$.

3 Observables

We divide the environment in 2D cells of size $\Delta = 0.25$ m and area $A = \Delta^2$, and record in each cell and for each time step the number of pedestrians detected and their velocity. Assuming the observation to be performed on a time span Δt ,

the number of time instants at which data are recorded is given by $T = \Delta t / \delta t$. Let us assume N_k pedestrian positions to be recorded on cell k , and $\{\mathbf{v}_i\}$ with $1 \leq i \leq N_k$ to be set of their velocities. Based on these microscopic observables, it is straightforward to define the macroscopic density ρ on k as

$$\rho_k = N_k / (TA) \quad (1)$$

The macroscopic vectorial velocity and speed are

$$\mathbf{V}_k = \sum_{i=1}^{N_k} \mathbf{v}_i / N_k \quad \mathcal{V}_k = \sum_{i=1}^{N_k} |\mathbf{v}_i| / N_k \quad (2)$$

In principle $\mathcal{V}_k \neq |\mathbf{V}_k|$. After fixing the orientation of our Cartesian system, we divide the pedestrian velocities in two sets,

$$P^+ = \{\mathbf{v}_i | v_i^x > 0\} \quad P^- = \{\mathbf{v}_i | v_i^x < 0\}$$

respectively with cardinalities N_k^+ and N_k^- , and define

$$\rho_k^+ = N_k^+ / (TA) \quad \rho_k^- = N_k^- / (TA) \quad (3)$$

$$\mathbf{V}_k^+ = \sum_{i=1}^{N_k^+} \mathbf{v}_i / N_k^+ \quad \mathbf{V}_k^- = \sum_{i=1}^{N_k^-} \mathbf{v}_i / N_k^- \quad (4)$$

$$\mathcal{V}_k^+ = \sum_{i=1}^{N_k^+} |\mathbf{v}_i| / N_k^+ \quad \mathcal{V}_k^- = \sum_{i=1}^{N_k^-} |\mathbf{v}_i| / N_k^- \quad (5)$$

3.0.1 Data filtering

We recorded data in two different corridors connecting the shopping area to the station (environments $E1$ and $E2$, fig. 1). Both recording campaigns lasted around 6 hours ($\approx 10^5$ time recordings). In $E1$ we tracked 17188 pedestrians ($\approx 1.8 \cdot 10^6$ recordings) while in $E2$ we tracked 11124 pedestrians ($\approx 1.1 \cdot 10^6$ events). Fig. 2a shows the $v = |\mathbf{v}|$ distribution (as total number of events in a 5 mm/s histogram) in $E1$. The empirical distribution can be described as the sum of a Gaussian distribution (pedestrians walking along the corridor) plus a Rayleigh distribution (standing pedestrians assuming white Normal noise in their x, y position recordings). We find, by best-fit, for the Gaussian velocity distribution $\sigma_G = 270$ mm/s and $\mu = 1260$ mm/s in $E1$, $\sigma_G = 260$ mm/s and $\mu = 1250$ mm/s in $E2$. For the Rayleigh distributions, we have $\sigma_R = 44$ mm/s in $E1$, $\sigma_R = 28$ mm/s in $E2$. Assuming also the Gaussian distribution of moving pedestrian velocities to be affected by noise with intensity $\approx \sigma_R$, we expect the estimation of σ_G not to be strongly affected by noise. Let us assume $\sigma_R / \sigma_G \approx 1/5$, and the observed velocity distribution to be given by the sum of two stochastic Gaussian processes, the actual velocity distribution plus the random noise. It follows for the ‘‘true’’ value $\sigma_T = (\sigma_G^2 - \sigma_R^2)^{\frac{1}{2}} \approx 49/50 \sigma_G$.

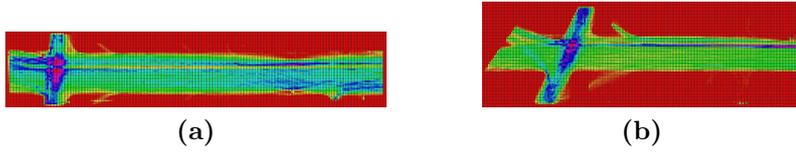


Figure 1: ρ_k for pedestrians with $|\mathbf{v}| > 500$ mm/s in $E1$ (a) and $E2$ (b). Red corresponds to $\rho = 0$, yellow to low ρ , violet to high ρ . $E1$ measures 60×12 m, $E2$ 52×17 m.

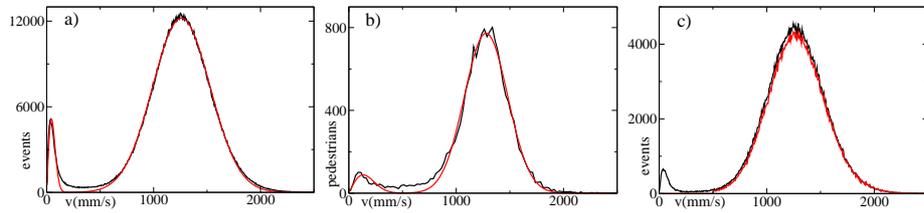


Figure 2: a) velocity distribution (black) in $E1$ compared to the best-fit to a model assuming the sum of a Gaussian plus a Rayleigh distribution (red). b) Average velocity distribution (black) in $E1$ compared to the best-fit to the same model as in a). c) Black line, velocity distribution in $E1a$ before filtering as in eq. (6). Red line, after filtering.

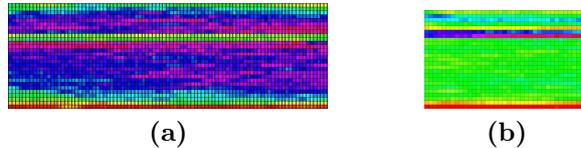


Figure 3: ρ_k in the “ideal corridors” $E1a$ (a, 22×7 m) and $E2a$ (b, 10×6 m).

Fig. 2b shows the empirical and best fitted distribution for average pedestrian velocities (i.e. the average value of v for each pedestrian over the tracked trajectory). We obtain $\sigma_G = 200$ mm/s, $\mu = 1280$ mm/s for both $E1$ and $E2$. Both environments present a larger corridor crossed by a smaller one on the left (fig. 1). On the right we observe areas in which the flow of pedestrians gets narrower, due to the presence of columns. In the central portion, the pedestrian density is almost invariant along the x axis, i.e. the environment corresponds to our definition of ideal corridor. We consider in our analysis only these portions (shown in fig. 3 and denoted as $E1a$ and $E2a$). By doing this we remove the portions in which the members of our staff, responsible for the largest part of the resting pedestrians peaks, were present. To further remove from our analysis standing pedestrians or pedestrians not behaving as commuters along the corridor, we filter our data keeping only velocities that satisfy

$$|v^x|/|v^y| > 3 \quad |\mathbf{v}| > 500 \text{ mm/s} \quad (6)$$

(from now on, the observables in eqs. (1-5) will be computed only for pedestrians satisfying eq. (6)). Our corridors can be considered close to be ideal only if the amount of data removed by filtering (6) is negligible. Fig. 2c shows the velocity distributions of pedestrians in $E1a$, before and after filtering. 8497 pedestrians are tracked in $E1a$, corresponding to $\approx 6 \cdot 10^5$ events ($\approx 5.5 \cdot 10^5$ after filtering) while 4586 pedestrians are tracked in $E1b$ ($\approx 1.5 \cdot 10^5$, $\approx 1.35 \cdot 10^5$ after filtering). We also clipped the areas corresponding to the small corridors on the left of fig. 1, and, after a proper rotation of the axes, filtered the data using eq. (6) (environments $E1b$ and $E2b$). These environments are far from being ideal corridors (in $E1b$ $1.0 \cdot 10^5$ events over $2.3 \cdot 10^5$ remain after filtering, 12452 pedestrians; $1.5 \cdot 10^5$ events over $2.5 \cdot 10^5$, 8751 pedestrians in $E2b$). Even if the amount of filtered data is large, since the density of pedestrians is low and inter pedestrians interactions are scarce, we assume the behaviour of filtered pedestrians to be close to the ideal one. $E1b$ measures 12×4 m and $E2b$ 17×4 m.

4 Empirical data

We integrate filtered data on the x direction to obtain 1D observables depending only on the distance from one of the walls, y . We notice the presence of an area,

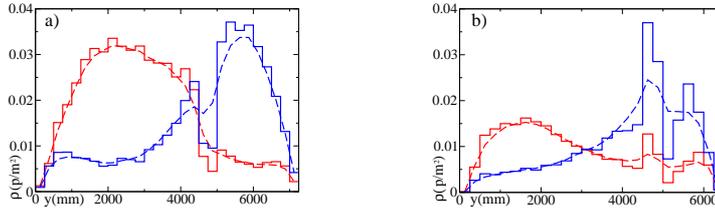


Figure 4: Comparison between the original data (continuous histogram) and their smoothed versions (dashed line) for $\rho^+(y)$ (blue) and $\rho^-(y)$ (red). a) $E1a$; b) $E2a$. The axes are chosen in such a way that the Japanese traffic condition corresponds to have the P^- flow on the left.

whose size is around 500 mm, in which ρ is much lower (see also fig. 3), corresponding to the location of tactile paving for the visually impaired, omnipresent in Japan. Since walking on these areas is uncomfortable, their presence causes a fluctuation in the density patterns. To simplify our analysis we will consider data smoothed with a filter of length higher than 500 mm, which removes almost completely their effect. Fig. 4 shows a comparison between the ρ^\pm distributions in $E1a$ and $E2a$ and their smoothed versions. The empirical data show clearly the presence of two distinct flows, since ρ^- presents a single peak located on the left, while ρ^+ presents a single peak located on the right. Density drops to zero close to the walls, and decreases quickly far from the maximum, until reaching an almost constant value. The \mathcal{V}^\pm and \mathbf{V}_x^\pm distributions (not shown) are convex functions assuming a maximum in the centre of the corridor and minima close to the walls, with differences between the maximum and the minima around 10%. The maximum is close to the centre of the corridor, i.e. on the right of the the ρ maximum, suggesting a tendency to overcome on the right. The analysed environments do not show enough variation in time to justify a time dependent analysis, and thus we will just perform a time-integrated analysis in order to use the maximum amount of data available. The stability in time of the observed patterns suggests that they represent some sort of equilibrium state that does not change significantly during the day and it is stable to the small density fluctuations observed in the environment.

5 Model

The empirical data present in every environment the separation between the two flows, together with a clear dependence of the density and velocity patterns on the distance from the walls. The observed environments do not present a wide variation nor in width (7.25 m in $E1a$, 6.5 m in $E2a$, 4 m in $E1b$ and $E2b$), average density (0.033 ped/m^2 in $E1a$, ≈ 0.02 elsewhere) or relative density between the flows (≈ 1 everywhere), nevertheless the similar shape and stability suggest some “universality” in the empirical distributions. We are going to

develop a model to describe these distributions, a model that is intrinsically empirical, since it is not derived by first principles about the pedestrian behaviour, but only as a function that tries to describe the observed distributions. Nevertheless, since these patterns arise from the behaviour of individual pedestrians, in describing the (macroscopic) empirical distributions, we will try to create a connection with the individual behaviour. We will assume that the pedestrian feels more or less (un)comfortable according to the position and velocity he has (this comfort depending also on the expectation of finding pedestrians walking in the opposite direction), and that this (dis)comfort can be expressed through an “Hamiltonian” function. The term Hamiltonian is used here because we make the hypothesis that, as in statistical mechanics, the microscopic behaviour and macroscopic observables are connected through a Boltzmann factor, i.e. the probability of observing pedestrians with given velocity and position is determined by a negative exponential of the Hamiltonian.

5.0.2 Density model

Denoting with $0 < y < L$ the distance of the pedestrian from one of the walls, we introduce an energy (or a (dis)comfort) function

$$U(y) = \frac{a}{y} + \frac{a}{L-y} + \left(\frac{\delta}{bL}\right)^2 \quad \delta = \begin{cases} y - cL & \text{if } |y - cL| < dL \\ dL & \text{if } |y - cL| \geq dL \end{cases} \quad (7)$$

$U(y)$ assumes a minimum (maximum comfort) in cL , while it increases (in a bounded way) with the distance from cL and diverges on the walls. Assuming that the probability distribution of pedestrians is given by a Boltzmann factor

$$p(y) \propto e^{-U(y)} \quad (8)$$

we expect results in good agreement with the ρ distribution in our data, since eqs. (7-8) describe a function that behaves as a Gaussian close to its maximum, drops to zero on the walls and reaches a constant value far enough from the walls and the maximum. Parameter c accounts for the maximum position, a regulates the distance to walls, b the width of the flow, while d is introduced to explain the finite probability of walking on the “wrong” side of the corridor.

Using environment and flow specific parameters, the best fits of eqs. (7-8) describe very well the empirical data (figure 5). To check the universality of the proposed law, it is more interesting to see to which extent the same parameters can describe all environments and ρ^\pm distributions (overall 8 density distributions). To avoid an over-fitting problem, we fixed $a = 300$ mm (which is very close to the value we found when we fitted the model to single curves, and is easily understandable as a tendency to maintain a distance from walls of the order of the human body size), and found the best fitted parameters to be $b=0.40$, $c = 0.28$, $d = 1.1$. The best fit curves are shown in fig. 6. Even if there is clearly a difference between the description of the larger ($E1a$ and $E2a$) and smaller ($E1b$ and $E2b$) environments, the model appears to be quite general.

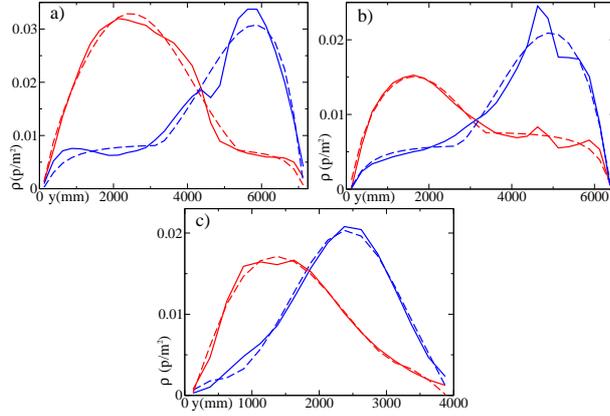


Figure 5: Best fit of ρ^\pm in: a) *E1a*; b) *E2a*; c) *E1b*.

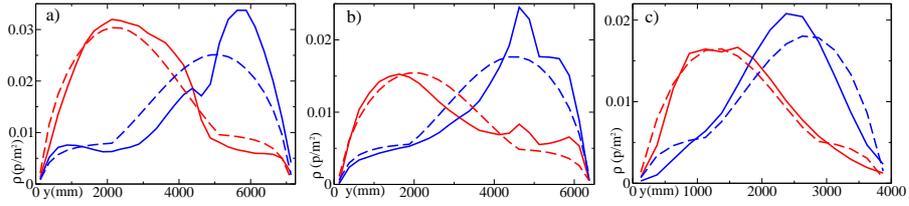


Figure 6: Comparison between the ρ^\pm distributions and the best fits of eq. (8) using $a = 300$ mm, $b=0.40$, $c = 0.28$, $d = 1.1$. a) *E1a*; b) *E2a*; c) *E1b*.

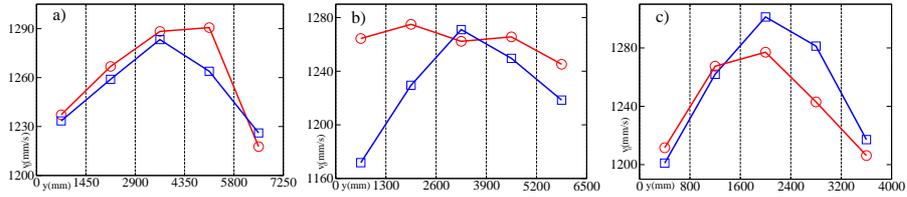


Figure 7: $v_0(y)$ for P^+ (blue) and P^- (red). a) $E1a$; b) $E2a$; c) $E1b$.

5.0.3 Velocity model

To generalise our (dis)comfort function to the velocity space, we must obtain velocity distributions as those shown in fig. 2 for different points of the corridor (i.e., values of y). The original discretisation using $\Delta = 250$ mm was too refined to have a significant statistics, so we divide each corridor in 5 lanes of width $L/5$ and obtain the velocity distribution for each lane (we use 5 lanes on the basis of the information obtained in the previous section, since the maximum of the density for each flow should be located in the centre of the second lane from left if $c \approx 0.3$). As reported in literature [2] and in fig. 2, the pedestrian velocity can be described with good approximation by a Gaussian function. According to this, we may extend the ‘‘Hamiltonian’’ for our pedestrian system introducing a quadratic term in the velocity, which takes in account the fact that our pedestrians have a preferred velocity directed along the corridor. We can assume the maximum of the Gaussian as given by a parameter v_0 depending on y , and extend eq. (7) as

$$H(y, \mathbf{v}) = U(y) + T(y, \mathbf{v}) \quad T(y, \mathbf{v}) = \frac{(v^x - v_0(y))^2}{2\sigma_x^2} + \frac{(v^y)^2}{2\sigma_y^2} \quad (9)$$

v_0 is the pedestrians’ preferred velocity, directed along the corridor (the macroscopic velocity distribution in y will be given by the negative exponential of T , as for eq. (8)). In eq. (9) we have considered the possibility that the spread in velocity along and orthogonal to the corridor may be different. We calibrate Gaussian distributions to each v^x and v^y distribution in all lanes to obtain the v_0 , σ_x and σ_y parameters of eq. (9), whose values are shown in table 1 and figure 7. With the exception of P^- in $E2a$, all data show a tendency to have maximum velocity on the right of the ρ maximum cL . We notice that $\sigma_x > \sigma_y$, due to the fact that while σ_y includes only deviations from the preferred velocity (since for all pedestrians $v_0^y = 0$) σ_x includes also individual variation in v_0^x , which can be estimated to have a standard deviation around 200 mm/s (fig. 2b). Assuming the individual variation to be ≈ 150 mm/s as for v_0^y , the overall deviation for v_0^x should be ≈ 250 mm/s, in agreement with the data).

Table 1: Parameters calibrated using a best fit method. v_0 and σ in mm/s.

<i>E1a</i>	v_0^+	σ_x^+	σ_y^+	v_0^-	σ_x^-	σ_y^-
1	1233	275	171	1237	255	156
2	1259	255	186	1267	252	158
3	1283	265	171	1288	250	160
4	1264	259	159	1291	263	167
5	1226	259	158	1218	280	163

<i>E2a</i>	v_0^+	σ_x^+	σ_y^+	v_0^-	σ_x^-	σ_y^-
1	1172	252	160	1264	250	157
2	1229	257	177	1275	254	164
3	1271	254	174	1262	250	174
4	1250	249	164	1266	225	170
5	1218	238	163	1245	274	155

<i>E1b</i>	v_0^+	σ_x^+	σ_y^+	v_0^-	σ_x^-	σ_y^-
1	1201	350	223	1212	313	170
2	1262	314	224	1267	299	194
3	1301	286	181	1277	283	182
4	1281	280	183	1243	308	204
5	1217	286	202	1206	318	226

<i>E2b</i>	v_0^+	σ_x^+	σ_y^+	v_0^-	σ_x^-	σ_y^-
1	1203	313	207	1228	269	197
2	1253	272	229	1265	255	211
3	1260	260	215	1264	265	208
4	1260	256	208	1267	263	223
5	1239	266	207	1202	290	223

6 Simulations

Microscopic collision avoiding behaviour is known to give rise to pedestrian organisation in lanes [1, 3]. Even if a realistic collision avoiding behaviour could hardly generate an organised pattern in the small environments and at the low densities observed in this work, these environments are part of a pedestrian area that may be large enough to generate such patterns even at low densities. In this section we use numerical simulations to investigate this possibility.

6.0.4 Collision avoiding models

In simulations we use a corridor with periodic boundary conditions and study the equilibrium distributions as density and velocity patterns that appear to be stable on the time scale relevant for pedestrian studies (10^3 - 10^4 s). In our paper [5] we investigated the extension and calibration of the Social Force Model (SFM) to describe low pedestrian densities, and verified that the Elliptical specifications using relative velocities [6] are those that better describe low densities, provided that a large enough parameter τ is used. In the original work this parameter was presented as the time length of a stride (≈ 0.5 seconds) but our calibration suggested a value close to 2 seconds to describe low densities. We hypothesised that τ actually represents the average time to the next collision, and proposed a new specification, called Collision Prediction, that explicitly computes such time (more precisely, the time to the maximum approach between pedestrians). This specification outperformed the Elliptical specifications at low density, and does not need tuning τ to be used at different densities. In this section we use the Elliptical Specification 2 (ES) and Collision Prediction (CP) models as described in [5]. We introduced a modification to CP as a maximum and minimum value for the time to the next collision (t_i in [5]), i.e. $\Delta t \leq t_i \leq t_{max}$, where Δt

is the integration time of the method (0.2 seconds in this paper), while t_{max} is a new parameter. The introduction of these constraints is not relevant to the current work, but it is useful to extend CP to higher densities and to more realistic simulations. Since in CP the force of the pedestrian is modulated by the term v_i/t_i , i.e. it is strong enough to stop in t_i , $t_i \geq \Delta t$ is an obvious request for numerical stability (it disappears in the continuous limit). The t_{max} constraint is useful to describe pedestrians walking in the same direction at similar speeds and it is important to describe, for example, the fundamental diagram using our method.

The interaction with the walls is realised in ES through a velocity independent repulsion force (as in the original Circular Specification of the SFM, [6]), while in CP the collision time to the walls is explicitly computed in the same way as it is for pedestrians (in both methods the values of the parameters for the interaction with walls are possibly different from the pedestrian interaction ones). Furthermore we introduce in both methods a maximum distance of interaction, r_I , to reduce possible interaction effects between pedestrians in different lanes, which seem unrealistic and could represent a difficulty in obtaining the empirical distributions. We notice that in the CP specification this limit can be introduced in an elegant way as acting on the predicted distance at the time of collision, i.e. it is realised as a "cognitive process" more than a limit in perception.

6.0.5 Cultural bias

Being symmetrical, the models in [5] cannot describe the cultural bias that leads Japanese people to walk on the left side of the corridor. By simplifying the model proposed in [3], we can introduce such a bias tilting the perceived relative position of other pedestrians. While this bias could describe well the observed lane formation, we don't expect it to describe the observed velocity distribution, since, as shown in fig. 8, it leads not only to avoiding on the left, but also to overcome on the left. We thus propose also a different mechanism, based on rotating not the relative position but the other pedestrian's velocity, which should lead to avoiding on the left while overcoming on the right, see fig. 8. This rotation corresponds to the expectation that the "opponent" applies the same cultural bias avoiding on the left or moving to the left if walking slower. We refer to the position tilt bias as TP, while to the velocity one as TV, and notice that the left-side bias of Japanese walkers is obtained by rotating position of a clockwise angle θ_t in TP, while it is obtained using a counter clockwise angle, $-\theta_t$, in TV. For overcoming to work properly in TV as used in this paper the pedestrian on the front should interact weakly ($\lambda \approx 1$, [6]), while a possible improvement in TV could be to multiply the tilt perceived by pedestrian i by $(\mathbf{d}_{ji} \cdot \mathbf{v}_i)/(|\mathbf{d}_{ji}||\mathbf{v}_i|)$ to obtain a more realistic behaviour in any setting.

6.0.6 Noise

We calibrated models in [5] assuming them to be deterministic, i.e. virtual pedestrians always behave in the same way given the same stimulus. Obviously

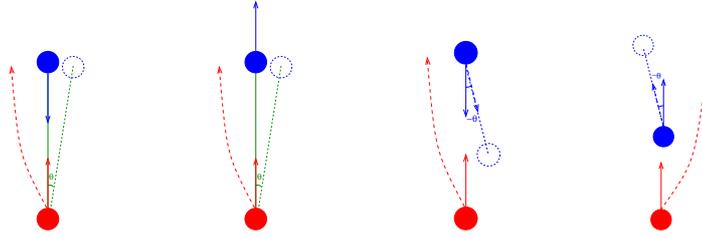


Figure 8: Tilting the perceived position with respect to the real one leads to a bias towards left in both avoiding and overcoming, while tilting the perceived velocity leads to avoiding on the left while overcoming on the right.

this is not true for real pedestrians, who have limited precision perception capabilities and arbitrary response. A good deterministic pedestrian model should provide the average behaviour, while the complex mechanism that leads to the richness of human response can be (statistically) approximated adding the right amount of noise to the model output. To estimate the amount of noise necessary to reproduce the variety of human behaviour, we first ran simulations using the parameter values of [5], assuming the preferred velocity distribution to be centred in 1.28 m/s with deviation 0.2 m/s (see fig. 2b). The resulting standard deviations in velocity distributions were $\sigma_x \approx 0.2$ and $\sigma_y \approx 0.03$, much lower than the observed ones (table 1). To obtain values similar to the empirical ones, we added in all simulations to both components of the models' velocity output a Gaussian white noise with standard deviation 0.12 m/s.

6.0.7 Calibration

Our calibration process consists in simulating environments $E1a$, $E2a$ and $E1b$ at the observed densities and to find the parameters that better describe the observed distributions (we omit $E2b$ due to its similarity with $E1b$). In calibration we allow the collision avoiding parameters to assume stronger values than in [5], since in that paper we didn't use noise, and freely calibrate the wall interaction parameters and other parameters not calibrated in the previous work. For simulations, we use 3 different binary conditions, resulting in 8 possible settings: the model condition (ES or CP), the bias condition (TP or TV) and the interaction with wall condition (W0, no interaction with walls, and W1). We introduce this latter condition to differentiate between the distribution given just by collision avoiding between pedestrians and the effect of walls which is to some extent "environmental". To calibrate parameters we use a genetic algorithm, with 30 generations and 30 genomes. Each genome (possible solution) is tested on all 3 environments 100 times. Each repetition lasts 5000 seconds, and the "equilibrium distribution" is defined as the average density distribution in the last 2500 seconds, averaged over all repetitions. All repetitions use the same number of pedestrians, chosen in such a way that the overall density (sum of the two direc-

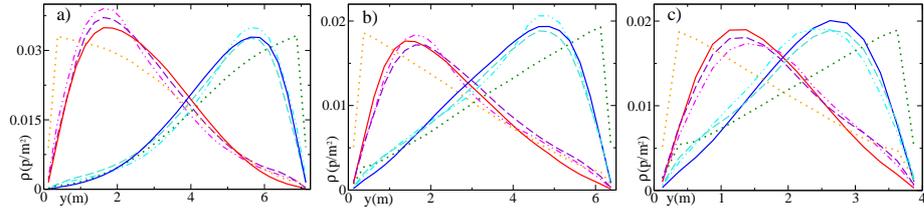


Figure 9: Comparison between the simulated distributions using different methods and conditions. Continuous: CP, TV, W1. Dotted: CP, TV, W0. Dashed: CP, TP, W1. Dash-dotted: ES, TV, W1. a) $E1a$; b) $E2a$; c) $E1b$.

tions) in each environment is always equal to the experimental average density in each environment. Nevertheless, the number of pedestrians in each direction is chosen in a probabilistic way, such that the average density over repetitions in each direction approaches the observed one, but the number of pedestrian for each direction present at any given time may be different (which is a more realistic setting that using always a balanced distribution). We run the GA 8 times for each condition, and for each best solution (i.e. the best solution for each condition over the 8 runs) we perform a test on 1000 repetitions to obtain better statistical significance. This test is performed also using repetitions of length 10000 seconds, to check the stability of the simulated distribution. The fitness function was defined as follows. Let $\rho_i^\pm(j\Delta)^s$ be the simulated density distribution in environment i (1 stands for $E1a$, 2 for $E2a$ and 3 for $E1b$) and $\rho_i^\pm(j\Delta)^e$ the experimental one, $\langle \rho_i^\pm \rangle$ the average density and L_i the width of the 3 environments; the fitness, a mean square distance weighted in such a way that all environments, regardless of size and density, contribute the same, is

$$\mathcal{F} = \sqrt{\left(\sum_{i=1,3;k=\pm} \left(\sum_j \left(\frac{\rho_i^k(j\Delta)^s - \rho_i^k(j\Delta)^e}{\langle \rho_i^k \rangle} \right)^2 \right) / \frac{L_i}{\Delta} \right) / 6} \quad (10)$$

This fitness function, taking in account only the density distributions, was used also to calibrate the model (8), for which we obtained as a best solution 0.248. Once the parameters that best fit the density distribution are found, the velocity distribution generated by these parameters is tested.

6.0.8 Results

Table 2 shows the values of \mathcal{F} and θ_t under different conditions, while figs. 9 and 10 compare some of the best solutions between themselves and with observed and theoretical curves. Solutions resulted to be very stable under the increase of simulation time from 5000 to 10000 seconds, i.e. they describe an “equilibrium distribution” at the time scales of interest. Table 2 shows that under condition W1 the simulations achieve a quantitative performance

Table 2: \mathcal{F} and θ_t for different methods and conditions after calibration (average and standard deviation are over different GA repetitions).

ES	TV W0	TV W1	TP W0	TP W1
Average \mathcal{F}	0.4562 ± 0.0008	0.2305 ± 0.006	0.4654 ± 0.005	0.2495 ± 0.007
Best \mathcal{F}	0.4546	0.2210	0.4572	0.2377
Average θ_t	-0.290 ± 0.016	-0.281 ± 0.022	0.139 ± 0.030	0.112 ± 0.047

CP	TV W0	TV W1	TP W0	TP W1
Average \mathcal{F}	0.4251 ± 0.002	0.2304 ± 0.021	0.4428 ± 0.011	0.2630 ± 0.020
Best \mathcal{F}	0.4218	0.2073	0.4278	0.2335
Average θ_t	-0.221 ± 0.099	-0.137 ± 0.106	0.038 ± 0.022	0.042 ± 0.030

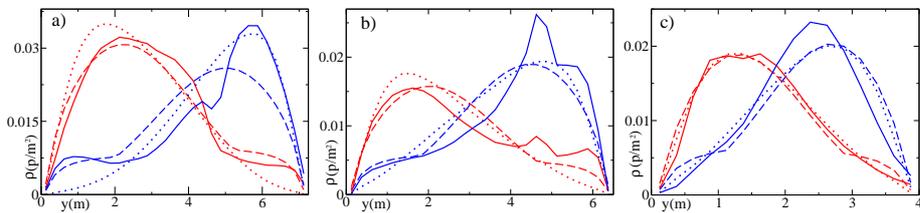


Figure 10: Comparison between the empirical distributions (continuous), the simulated ones (dotted) and the theoretical ones (dashed). For simulated densities, we show the overall best simulation (CP,TV,W1). a) $E1a$; b) $E2a$; c) $E1b$. P^+ in blue, P^- in red.

similar to that of the theoretical model, even outperforming it under the TV condition. From a qualitative point of view, the simulated distributions have problems in describing both large spreads around peaks and relatively high density corresponding to the location of the other flow's peak. The difference between W0 and W1 (table 2 and fig. 10) conditions shows the importance of the interaction with walls in obtaining a distribution with a significant spread around a peak at a significant distance from the wall. This means that this pattern can be obtained only by introducing some kind of environmental effect, since only collision avoiding leads, at these low densities, to the formation of lanes very close to the walls. There is also a weaker but clear tendency of TV to outperform TP, which is once again (figure 10) given by a better ability of describing the position and spread of the peak. This effect is probably due to the tendency of faster pedestrians to walk closer to the middle of the corridor in TV.

Both in ES and CP the best solutions had low values of r_I , probably to reduce the interaction between the two lanes. While in ES r_I represents a limit for the pedestrians' perception, and thus it cannot be set to values smaller than a few meters (we used 4 m as a limit in our calibrations), in CP it can be

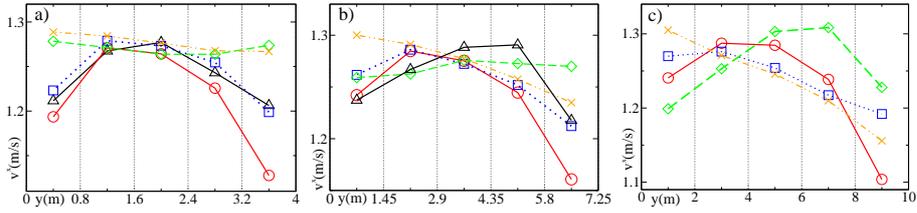


Figure 11: Average values of v^x for P^- under W1, compared to the results of fig. 7 (black, triangles). Red, circles and continuous: CP, TV. Blue, squares and dotted: CP, TP. Green, diamonds and dashed: ES, TV. Orange, crosses and dot-dashed: ES, TP. a) $E1a$; b) $E2a$; c) $E3$ (virtual environment, no experimental data).

set to values close to the size of the human body (or of their social space) to discriminate between behaviours leading to actual collisions or just approaching behaviours that cannot cause collisions [7]. The other parameters did not seem to be relevant for the problem under study.

6.0.9 Velocity

The velocity distributions for best solutions under all conditions present values of σ_x between 0.25 and 0.3 m/s and of σ_y between 0.15 and 0.23 m/s, in good agreement with the values in table 1. Fig. 11 shows the average value of v^x in different “lanes” (defined as and compared to the results of fig. 7) in P^- under W1. In $E1b$ there is almost no qualitative difference between the TV and TP conditions, while there is a clear difference between CP and ES, since the latter describes a flat distribution over the corridor. Our hypothesis is that at this low density overcoming occurs very seldom and the mechanism described in fig. 8 does not produce macroscopic effects, while the difference between CP and ES is due to the fact that the former uses a velocity dependent interaction with the walls, and thus the quickest pedestrians have a tendency to stay farther from the walls. In $E1a$, due to the higher density ($\langle \rho \rangle \approx 0.021$ in $E1b$, $\langle \rho \rangle \approx 0.033$ in $E1a$) and larger size (i.e. more overcoming and less influence from walls) the difference between TV and TP is more clear. To further test this hypothesis we simulate also virtual environment $E3$ ($\langle \rho \rangle \approx 0.09$, $L = 10$) in which the macroscopic effect of the different bias conditions is evident. We expect the differences between the experimental data and the TV condition to be solved with the proposed enhancement $\theta'_t = \theta_t(\mathbf{d}_{ji} \cdot \mathbf{v}_i) / (|\mathbf{d}_{ji}| |\mathbf{v}_i|)$.

7 Discussion and conclusions

Simulations based on collision avoiding reproduce quite well the observed density distributions, even if the “environmental” role played by interaction with walls is important in the attainment of such a task. Also the tendency of faster

pedestrians to walk in the middle of the corridor can be obtained using a proper velocity dependent interaction with the walls. The values of the bias θ_t was much higher than the one reported in [3]. This could be due to the fact that they observed an artificial environment in which such a bias is not necessary, and also could be related to the fact that this bias is particularly strong in Japanese culture. Nevertheless such a strong value could be unrealistic, suggesting that the observed patterns cannot be obtained by using pure collision avoiding. We leave for future work the comparison between the microscopic behaviour of real and simulated pedestrians to clarify this point.

If pedestrians have, as we suggested, a tendency to overcome on the right, this can be simulated by using TV, while TP could lead to unrealistic behaviours, in particular at high densities where overcoming happens more often.

Based on these results, we believe that a pure collision avoiding behaviour using the TV bias (with values of θ_t lower than those reported in this paper) could describe properly the behaviour of pedestrians in (relatively) high density corridors, while at low density their distribution is given by a non emergent behaviour that can be expressed using eq. (7). The fact that a strong bias leads to distributions similar to those described by eq. (8) suggests according to us that the pedestrians use at low density a behaviour similar to the one that they “have learnt” at higher ones (see also figs. 7 and 11, walking with high velocity on the right is probably necessary only at high densities, but pedestrians have this tendency also at low ones). To clarify these points we want to collect data at different densities and compare microscopic and macroscopic behaviours to better understand the consistency of the θ_t values. We expect the microscopic bias values to be compatible with the observed macroscopic patterns only at densities considerably higher than those observed in this work.

The model (8) may be used as a boundary condition for the simulation of more complex environments, assuming for example that if pedestrians are entering the environment from some kind of corridor, the given distribution can be considered as the spatial distribution of entering pedestrians, supposing no further knowledge is available. From a microscopic point of view, this function could be seen as “the external potential” acting on the pedestrian (influence of the environment) and thus the force perceived by the pedestrians from the environment could be modelled as the negative gradient of U .

We also plan to observe the effect of group behaviour on simulations.

Acknowledgements This work was supported by JST, CREST.

References

- [1] T. Kretz, A. Grünebohm, M. Kaufman, F. Mazur, and M. Schreckenberg, *Experimental study of pedestrian counterflow in a corridor*, Jour. Stat. Mech. Theory and Experiment P10001 (2006).
- [2] D. Helbing, I. J. Farks, P. Molnr, and T. Vicsek (2002) *Simulation of pedestrian crowds in normal and evacuation situations*, pg. 21-58 in Pedestrian

and Evacuation Dynamics (2002).

- [3] M. Moussaïd, D. Helbing, S. Garnier, A. Johansson, M. Combe and G. Theraulaz, *Experimental study of the behavioural mechanisms underlying self-organization in human crowds*, Proc. Roy. Soc. B: Biological Sciences, **276**, 1668, 2755-2762, (2009)
- [4] D.F. Glas, T. Miyashita, H. Ishiguro, N. Hagita, *Laser-Based Tracking of Human Position and Orientation Using Parametric Shape Modeling*, Advanced Robotics, **23** (4), 405-428 (2009)
- [5] F. Zanlungo, T. Ikeda and T. Kanda, *Social Force Model with explicit collision prediction*, Europhysics Letters, **93**, 68005 (2011).
- [6] D. Helbing, A. Johansson (2010) *Pedestrian, Crowd and Evacuation Dynamics*. Encyclopedia of Complexity and Systems Science 16, 6476-6495.
- [7] F. Zanlungo, *Microscopic dynamics of artificial life systems*, Ph.D Thesis in Physics, University of Bologna, 2007