LANE DETECTION IN PEDESTRIAN MOTION AND ENTROPY-BASED ORDER INDEX

Olivier Chabiron[1], Jérôme Fehrenbach[1], Pierre Degond[1], Mehdi Moussaïd[2,3], Julien Pettré[4] and Samuel Lemercier[4]

[1]*Toulouse Mathematics Institute, Paul Sabatier University*
118 route de Narbonne, Toulouse, France
[2]*Centre de Recherches sur la Cognition Animale, UMR-CNRS 5169, Université Paul Sabatier, Bât 4R3, 118 Route de Narbonne, 31062 Toulouse cedex 9, France.*
[3]*CNRS, Centre de Recherches sur la Cognition Animale, F-31062 Toulouse, France*
[4]*INRIA Rennes - Bretagne Atlantique, Campus de Beaulieu, F-35042 Rennes, France*

Keywords: Clustering, Classification, Motion and tracking, Detection, Graphs, Entropy, Complex systems.

Abstract: This paper proposes a distance measurement between pedestrian trajectories. This distance is used in a clustering method aiming to detect lanes of pedestrians in experimental data. The main ingredient is to take full advantage of the time sequence available. A study of the sensitivity of the clustering to the parameters shows it is possible to choose a stable set of parameters. We also define an order index based on the concept of entropy. The potential of this index is illustrated in the case of pedestrian lane detection.

1 INTRODUCTION

Human crowd is a complex system that exhibits the emergence of self-organized patterns. Several studies deal about the social and biological aspects of these phenomenons (Helbing and Molnar, 1995; Helbing et al., 2001; Helbing et al., 2005; Moussaïd et al., 2011; Sumpter, 2010; Couzin and Krause, 2003). A particular aspect of pedestrian traffic is the self-segregation into two or more lanes of opposite direction (Older, 1968; Yamori, 1998; Kretz et al., 2006). The formation of lanes can also be observed in other complex systems, like ants (Casellas et al., 2008) or charged particles (Rex and Lwen, 2007).

The present study addresses the quantitative aspects of lane formation, which were studied in army ants by (Couzin and Franks, 2003). This work proposes tools to process motion data in pedestrian crowds, these tools are general and the ideas presented here can also be applied to other data (e.g. ants, particles). The first objective of this work is to propose a robust method to detect lanes. The novelty of this method is the use of the temporal aspect of the data. The second objective is to define an order index that allows to quantify the orderliness of a pedestrian crowd. The index we propose is based on the notion of entropy. The lane detection algorithm and the order index are tested on experimental data of pedestrians walking in a ring-shaped arena.

2 PEDESTRIAN DATA ACQUISITION

Experimental Setup. Controlled experiments were conducted in May 2009 by INRIA in Rennes, France. A total of 119 participants took part in the study, which conformed to the Declaration of Helsinki. Pedestrians are walking in a ring-shaped arena, of inner radius 2 m and outer radius 4.5 m. Some are instructed to walk the arena clockwise, the others counter-clockwise. They are forbidden to change direction in the course of the session.

Participants wore 4 reflexive markers, one on the forehead, one on the left acromion, and two on the right acromion to easily distinguish the left shoulder from right one. Markers motion was reconstructed using Vicon IQ software. Participants motion was finally modeled as the one of the barycenter of the 4 markers projected onto the horizontal plane. The motion was recorded at the frequency of 10 pictures per second. Several configurations were experimented: 8 pedestrians, 10 pedestrians, 18 pedestrians, 30 pedestrians, 50 pedestrians and 60 pedestrians in the arena.

Chabiron O., Fehrenbach J., Degond P., Moussaïd M., Pettré J. and Lemercier S. (2012).
LANE DETECTION IN PEDESTRIAN MOTION AND ENTROPY-BASED ORDER INDEX.
In *Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods*, pages 192-197
DOI: 10.5220/0003762901920197

Figure 1: Experimental setup.

Post-processing. Pedestrians trajectories are extracted from the recordings. Trajectory data consists in every pedestrian's position and velocity at each capture time (i.e every 0.1 s). Positions are computed in a Cartesian coordinate system which origin is the center of the ring. In the end, for each session, we have access to:

- The number N of pedestrians taking part in the session.
- Their position in the arena plane. We denote by $(X_i(t), Y_i(t))$ the coordinates of pedestrian number i at time t.
- The direction each of them is walking in. The direction η_i takes a value of 1 for "clockwise" and -1 for "counter-clockwise".

3 LANE DETECTION

The technique that we propose here to detect pedestrian lanes is based on the definition of a distance between pedestrians and a graph-based clustering technique.

Qualitative Lane Definition

The first and most decisive step of our clustering method consists in defining a distance, which will be the one used in the computation of the distance matrix related to the pedestrians graph. This distance should give a measure of how far one pedestrian is from another, in terms of lane affiliation. Of course, taking only the physical (euclidean) distance on $\mathbb{R}^2$ into account is not sufficient to achieve lane detection. For this purpose, several things have been taken into account:

- A lane should contain only pedestrians following each other. Their positions should be close enough.
- Pedestrians walking in opposite directions should not be in the same lane, even if they are very close to one another.
- A pedestrian walking through a lane while going in the opposite direction should break that lane.
- A pedestrian's *trajectory over the next few moments* should be taken into account to determine its affiliation to a lane.

This last rule is the pillar and the main novelty upon which the method is built. By introducing a time dependency, we hope to achieve a better accuracy in the detection of the formation and breaking of lanes. The principle applied is as follows: if a pedestrian is following another, then he will walk right in the other's footsteps in the next second or so.

Non-alignment Penalization

In order to make sure that two pedestrians walking in opposite directions do not belong to the same lane, we define the following penalization term. This term depends on the sign of the product between the pedestrians direction.

Let i and j be two pedestrians. let us denote η_i the direction of i and η_j that of j. For every pedestrian, η is a constant and takes a value of 1 for "clockwise" and -1 for "counter-clockwise". The penalization term $P_{i,j}$ between pedestrians i and j at time t is:

$$P_{i,j} = \begin{cases} 0 & \text{if} \quad \eta_i\eta_j = 1 \\ \infty & \text{if} \quad \eta_i\eta_j = -1 \end{cases} \qquad (1)$$

Distance Definition

Since we are working with recorded data, the complete trajectory of every pedestrian is available. We take advantage of this information to detect if a pedestrian is following another. Let us denote by $d_{i,j}(t', t)$ the distance between i at time t' and j at time t. If we stop time at t for every pedestrian except one, let it be pedestrian number i, does i walk straight into the spot where pedestrian number j is standing (i.e was standing at time t)? This principle is illustrated in Figure 2.

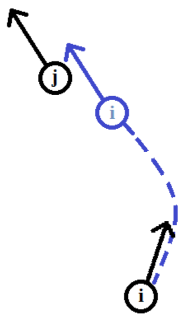

Figure 2: Trajectory alignment.

The procedure is as follows: We define τ a small time interval, that represent a pedestrian's "follow" time. At time t, we compute $d_{i,j}(t,t')$ for each time $t' \in [t\,,t+\tau]$.

$$d_{i,j}(t',t) = \sqrt{(X_i(t')-X_j(t))^2+(Y_i(t')-Y_j(t))^2} + \begin{cases} 0 & \text{if} \quad \eta_i\eta_j = 1 \\ \infty & \text{if} \quad \eta_i\eta_j = -1 \end{cases} \quad (2)$$

Then, we select the minimum distance over the interval as the distance between i and j :

$$d^{min}_{i,j}(t) = \min_{t\le t'\le t+\tau}(d_{i,j}(t',t)) \quad (3)$$

The "follow" time τ has been set to one second, which appeared to us as a good measure of the time needed for a "follower" pedestrian to literally *follow in the footsteps* of a "leader" pedestrian. Parameter exploration has shown this is an adequate value for τ (see section 4).

Distance Symmetrization

The definition in Equation (3) results in a non-symmetrical distance $d^{min}_{i,j}(t)$. In order to define lanes of pedestrians we define a symmetrical distance $D_{i,j}(t)$. We use a minimum-based symmetrization, and the $N\times N$ matrix $D(t)$ is defined by:

$$D_{ij}(t) = \min(d_{i,j}(t), d_{j,i}(t)) \quad (4)$$

Connected Components

At this step of the procedure we dispose of a full distance matrix. In order to define clusters we define an adjacency graph G in the network of pedestrians. Two pedestrians are connected in G if their distance is below a given threshold δ. The adjacency matrix $(G_{i,j})$ of the graph G is defined by:

$$G_{ij} = \begin{cases} 0 & \text{if} \quad D_{ij} \ge \delta \\ 1 & \text{if} \quad D_{ij} < \delta \end{cases} \quad (5)$$

The lanes of pedestrians are the connected components of the graph G.

To summarize this section, we present the algorithm that is used to compute the adjacency matrix $(G_{i,j})$ of the clustering graph.

Algorithm 1: Lane clustering algorithm.

```
Input:
N: number of pedestrians;
X(t),Y(t)_{0≤t≤T}: positions of the pedestrians;
η_i: walking directions of the pedestrians;
δ,τ: space and time thresholds;
Output:
G(t): clustering matrix at every time t.
begin
    for t ∈ [0..T − τ] do
        for i, j = 1..N do
            define d_{i,j}(t) by Equation (3) ;
        end
        for i, j = 1..N do
            define
              G_{i,j}(t) :=   0 if d_{i,j} ≥ δ and d_{j,i} ≥ δ
                              1 otherwise;
        end
    end
end
```

4 PARAMETER SENSITIVITY STUDY

Parameters for Pedestrians

Through parameter exploration, we set δ to 0.7m, because this value gives consistent results for every experimental configuration, i.e from eight to sixty pedestrians. There is also a biological justification for this value, which we can see if we note that the distance between two pedestrians walking in the same direction is actually euclidean. Then, 0.7 meters roughly correspond to the width of two bodies side to side. This value is relevant when a suitable time parameter τ is chosen. Indeed, both parameters together build this space-time corridor which selects only pedestrians who are "following footsteps". Figure 3 shows an example of lane detection accomplished with Algorithm 1 for a 60-pedestrians session.

The values chosen for our parameters, δ and τ, have been set mainly by biological considerations, data we possessed concerning the systems studied. However, one might not always dispose of such data, or might need improved stability. That is why a parametric sensitivity study is crucial to gather information about eventual ranges of parameter values of lesser sensitivity.

Parameter Range and Indicator

The study was performed for $\tau \in [0\,,2]$ (in seconds) and $\delta \in [0\,,1.25]$ (in meters, the data used being that of the pedestrian motion). The number of lanes was used as the main indicator. In order to be able to compare one value for each $\{\tau,\delta\}$ couple, we compute a mean of the indicator over the "meaningful" time of all experimental recordings with the same number of pedestrians. That is to say, from 5 seconds after the pedestrians start walking to 50 seconds later. We write $N^K_L(\tau,\delta)$ the mean number of lanes for the ensemble of all K-pedestrians experiments.

Graphical Study in the τ,δ Plane

Figure 4 shows the surfaces drawn by $N^{50}_L(\tau,\delta)$ and

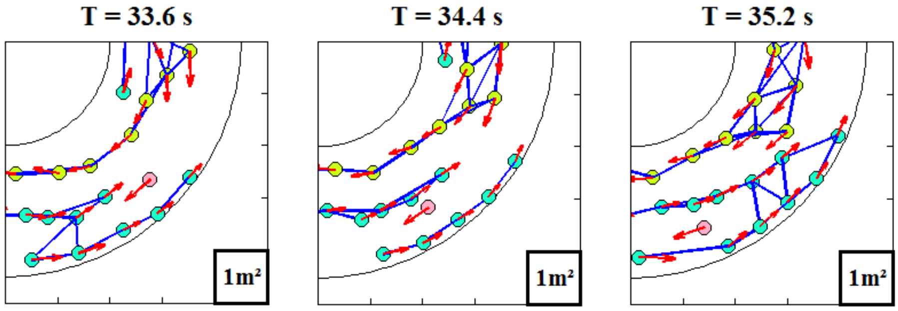

Figure 3: Example of lane detection with $N = 60$. We can see the pink pedestrian beginning to cross the cyan lane (left picture), forcing cyan pedestrians to evade him thus breaking the lane (center picture), before the lane reforms (right picture).

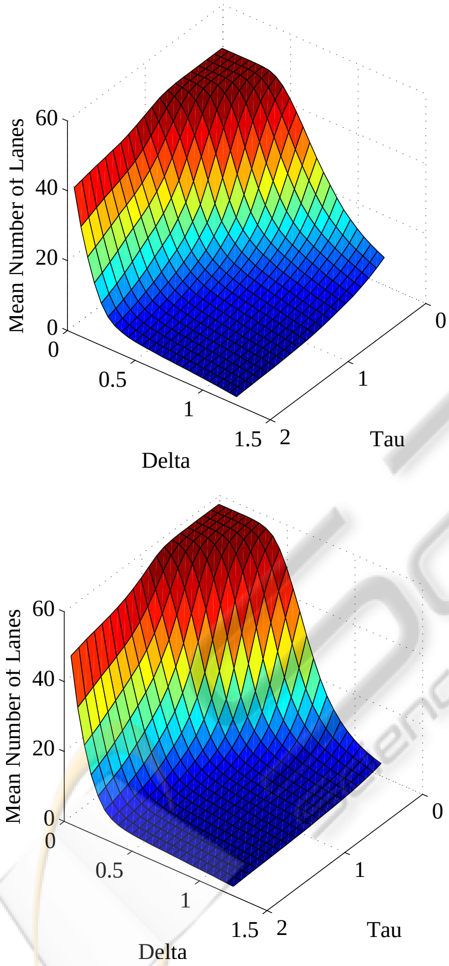

Figure 4: Surface plot of the mean number of lanes detected for 50 pedestrians experiments (top) and 60 pedestrians experiments (bottom).

$N_L^{60}(\tau,\delta)$. Since the study for small numbers of pedestrians shows abrupt variations, we focused our attention to the experiments with $N = 50$ and $N = 60$ pedestrians, those showing far smoother variations.

Lesser Sensitivity Window
Though this parametric study doesn't yield an ideal, universal set of values that would minimize parameter sensitivity for every kind of experiment, it does indicates that the values of both τ and δ must be high enough to ensure a low sensitivity. Very low values also show low sensitivity but the detection would be meaningless.

In the $N = 60$ pedestrians case, we observe a flat band in the τ,δ plane. In the $N = 50$ pedestrians case, this flat band is also relatively flat, though not as much. This result suggest that any set of parameters chosen in this area will yield a similar N_L function, which seems to be the case after testing.

However, more than minimizing parameter sensitivity, we need our detection of lanes to make sense. That is why we define a lesser sensitivity window that combines a low sensitivity to τ and δ and a somewhat realistic range of values for these same parameters. This is the following window:

$$1 \leq \quad \tau \quad \leq 1.6 \tag{6}$$

$$0.7 \leq \quad \delta \quad \leq 1.0 \tag{7}$$

Inside this window, the relative variation of N_L^{60} with both τ and δ does not exceed 7% and that of N_L^{50} do not exceed 10%. Though the choice of this window takes experiments with less pedestrians into account, these are more parameter-sensitive. Indeed, fewer agents naturally results in greater relative differences in clustering when parameters are changed.

5 ENTROPY-BASED ORDER INDEX

When studying complex systems, it is vital to be equipped with a trusted order index that should be both stable and efficient (relatively to some criteria).

Besides, an order index is rarely transferable to different systems. For these reasons, we worked on designing a new index, which should respect the following constraints:

- The index should not over-penalize a well-ordered system because of a few isolated agents.
- The index should present relatively smooth variations.
- The index should be transferable to a large panel of studies.

What is Entropy?
Entropy originates in thermodynamics and was introduced by Rudolph Clausius in the middle of the nineteenth century, but its significance was only highlighted around 1870 by Ludwig Boltzmann. It was first introduced in the complex systems field by (Wolfram, 1984). The entropy of a system is a function of the system's states which "measures its disorder". Let $\{\Omega, X, p\}$ be a finite state space, $|\Omega| = M$. Let $p(x)$ denote the probability that the system is in state x. Then the statistical entropy is given by:

$$S = -k \sum_{x \in X} \begin{cases} p(x) \ln p(x) & \text{if} \quad p(x) \neq 0 \\ 0 & \text{if} \quad p(x) = 0 \end{cases} \tag{8}$$

where k usually is a physical constant. An interpretation of this function consists in saying that entropy measures the logarithm of the number of states actually accessible for the system. To illustrate this, let us look at two opposite cases.

If a state x_0 is attained with certainty, i.e $p(x_0) = 1$, $p(x \neq x_0) = 0$, then:

$$S = S_{min} = 0 \tag{9}$$

If all states are equally likely, i.e $p(x) = 1/M \quad \forall x$, then:

$$S = S_{max} = -kM \frac{1}{M} \ln \frac{1}{M} = k \ln M. \tag{10}$$

Definition of the β Index. In order to adapt the concept of entropy to build an order index, we need to define a state space and write a normalized function based on Equation (8). Thanks to the clustering, we dispose of all the data concerning the lanes (their number, their sizes, and the identification number of the pedestrians who form them).

If we think of the arena as a box (an isolated system) containing a certain number of particles (pedestrians), we are close to describing the experiment as a thermodynamic system. The different states accessible to the system are the different kinds of particle aggregation, i.e each possible configuration of the lanes. The probability of being in a state-lane L, of size x_L, is replaced by the actual ratio of the number of pedestrians forming the lane over the total number of pedestrians, i.e x_L/N.

This allows us to define β_{tmp} a temporary index which measures disorder:

$$\beta_{tmp} = -\sum_{L=1}^{N_L} \begin{cases} \frac{x_L}{N} \ln \frac{x_L}{N} & \text{if} \quad x_L \neq 0 \\ 0 & \text{if} \quad x_L = 0 \end{cases} \tag{11}$$

In this formula, N_L denotes the total number of lanes. This number is obviously time-dependent, as is the size of each lane. Together these numbers rule the variations of the index.

Then, in the case of N lanes, each with a single pedestrian, the number of state-lanes accessible by the system is maximum and equal to N. In this configuration, all state-lanes of the system are of equal size. This corresponds to a maximum disorder situation.

In order to have an index that is normalized between 0 (no order at all) and 1 (perfectly ordered situation) an affine transformation is applied. This leads to define the β index by:

$$\beta = 1 - \frac{\beta_{tmp}}{\ln(N)} \tag{12}$$

Application to Pedestrian Traffic: Directional Order Index
In order to illustrate the capabilities of the β index, we show in Figure 5 an example where the order measurement is segregated. Indeed, in our experimental conditions, it is of interest to know if clockwise-walking pedestrians are as well ordered as counter-clockwise-walking pedestrians. If we compute a β index for each subsystem, we can transform visual insight into quantitative measurement. For this purpose, we define β_c the clockwise order index and β_{cc} its counter-clockwise counterpart.

Sensitivity of β to the Parameters
We conducted the same parametric study as in section 4, this time using a mean value of β. In the previously defined window (see (6) and (7)), the relative variation of the mean index does not exceed 2% (for 60 pedestrians) and 5% (for 50 pedestrians).

6 CONCLUSIONS

We have developed an effective clustering technique which realizes the detection of lanes in pedestrian crowds using only two parameters: a distance scale,

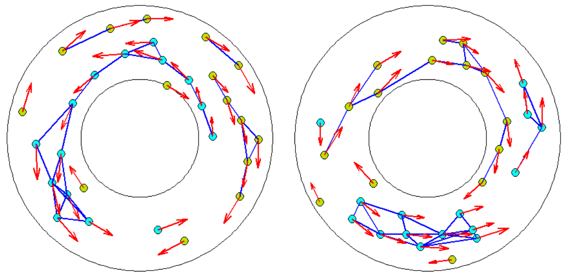

Figure 5: Comparison between two situations occurring within the same experimental session ($N = 30$ pedestrians). The first situation (left graph) is clearly asymmetric in its orderliness: counter-clockwise pedestrians are all in the same lane but one, whereas clockwise pedestrians form several small lanes. At this moment $\beta_c = 0.38$ and $\beta_{cc} = 0.91$. The second situation (right graph) is more balanced: $\beta_c = 0.73$ and $\beta_{cc} = 0.70$.

and a time scale. Indeed, its originality lies in taking time into account to detect the formation and the break-up of lanes. Moreover, sensitivity studies show that the method is robust to parameter variations as long as their values are high enough.

In addition, we designed a universal order index which can prove to be very useful in both experimental and numerical data of complex systems. Being based on the concept of statistical entropy, it ensures a measure of order in a very general sense, and is easily transferable to different contexts and studies in the complex systems field.

Future work will include the application of these tools to systems with a large number of agents, namely trail formation in simulated and experimental ant colonies.

ACKNOWLEDGEMENTS

This study was supported by a research grant from the PEDIGREE project funded by the French National Research Agency (Grant No ANR-08-SYSC-015).

REFERENCES

Casellas, E., Gautrais, J., Fournier, R., Blanco, S., Combe, M., Fourcassi, V., Theraulaz, G., and Jost, C. (2008). From individual to collective displacements in heterogenous environments. *Journal of Theoretical Biology, 250: 424-434*.

Couzin, I. and Franks, N. (2003). Self-organized lane formation and optimized traffic flow in army ants. *Proceedings of the Royal Society B: Biological Sciences 270(1511):139-146*.

Couzin, I. and Krause, J. (2003). Self-organization and collective behavior in vertebrates. *Advances in the Study of Behavior 32:1-75*.

Helbing, D., Buzna, L., Johansson, A., and Werner, T. (2005). Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation Science 39(1):1-24*.

Helbing, D. and Molnar, P. (1995). Social force model for pedestrian dynamics. *Physical Review E. 51(5):4282-4286*.

Helbing, D., Molnar, P., Farkas, I., and Bolay, K. (2001). Self-organizing pedestrian movement. *Environment and Planning B: Planning and Design 28:361-383*.

Kretz, T., Grünebohm, A., Kaufman, M., Mazur, F., and Schreckenberg, M. (2006). Experimental study of pedestrian counterflow in a corridor. *Journal of Statistical Mechanics: Theory and Experiments P10001*.

Moussaïd, M., Helbing, D., and Theraulaz, G. (2011). How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences 108(17):6884-6888*.

Older, S. (1968). Movement of pedestrians on footways in shopping streets. *Traffic Engineering and Control 10:160-163*.

Rex, M. and Lwen, H. (2007). Lane formation in oppositely charged colloids driven by an electric field: Chaining and two-dimensional crystallization. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) 75*.

Sumpter, D. (2010). Collective animal behavior. *Princeton University Press*.

Wolfram, S. (1984). Universality and complexity in cellular automata. *Physica 10D, 1-35*.

Yamori, K. (1998). Going with the flow: micro-macro dynamics in the macrobehavioral patterns of pedestrian crowds. *Psychological Review, Vol. 105, No.3, 530-557*.