Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 207 (2022) 2364–2373

Procedia Computer Science

www.elsevier.com/locate/procedia

26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

A Novel Spatial–Temporal Analysis Approach to Pedestrian Groups Detection

Claudia Cavallaro[a,*], Giuseppe Vizzari[b]

[a] *Department of Mathematics and Computer Science, University of Catania, Viale Andrea Doria 5, Catania 95125,Italy*
[b] *Department of Informatics, Systems and Communication, University of Milano-Bicocca, Viale Sarca 336/14, Milano 20126, Italy*

Abstract

The growing availability of geo–referred data describing human behaviour, at different scales and levels of granularity, represents an opportunity for the development and application of data analysis algorithms, whose usage can range from security, to traffic, to architectural design and planning, and even marketing. Focusing on pedestrian generated trajectories, the presence of groups within an analyzed population can influence overall dynamics, from microscopic perspective, and it can provide significant indications. Several approaches for video footage analyses are available, but they generally focus on microscopic features of videos and trajectories and they are generally not suited to scale to the analysis of relatively large datasets of trajectories. The present work proposes a novel approach to spatial–temporal analysis of pedestrian trajectories aimed at detecting groups of pedestrians within large datasets and having minimal assumptions on the nature of these groups.

Peer-review under responsibility of the scientific committee of the 26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Keywords:
Groups; Analysis of pedestrian trajectories; Clustering

1. Introduction

The ubiquitous presence of digital technological devices, both personal/mobile and infrastructural/fixed, created an unprecedented opportunity for acquiring data describing human behaviour in social contexts. This new technological setting, coupled with urbanization (by 2050 66% of the world's population will reside in cities [17]), pushes innovative perspectives like Urban Informatics [14] and Smart Cities [12], and more generally highlights possibilities for both innovations, growth of diffusion, and even completely novel applications of data mining techniques. In particular, the growing availability of geo–referred data describing human behaviour, at different scales and levels of granularity, represents an opportunity for the development and application of data analysis algorithms, whose usage can range

* Corresponding author.
E-mail address: claudia.cavallaro@unict.it

1877-0509
Peer-review under responsibility of the scientific committee of the 26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)
10.1016/j.procs.2022.09.295

from security (leveraging traditional video surveillance, but also scaling up to district level zone), to traffic and vehicle oriented applications, to architectural design and planning, but also recommendation systems and even marketing. Focusing on pedestrian generated trajectories (acquired directly via GPS or through infrastructural monitoring, like a network of video cameras), the presence of groups within an analyzed population can influence overall dynamics (and this represents a useful information for generating or calibrating realistic simulations of pedestrians in the premises [1, 7]), and it can also represent useful information *per se* (e.g. to better characterize the structure of the environment and its internal areas).

Several approaches for video footage analyses, also due to the body of research and results available from the surveillance community, were proposed in the last few years, but they generally focus on microscopic features of videos and trajectories and they are generally not suited to scale to the analysis of relatively large datasets of trajectories. The present work introduces a novel approach to spatial–temporal analysis of pedestrian trajectories aimed at detecting groups of pedestrians within large datasets and having minimal assumptions on the nature of these groups: as such, we can expect to have a potentially lower accuracy, but a wider applicability. The proposed approach constructs a spatial–temporal analysis around results of QuickBundles (QB) [9], a cluster analysis algorithm designed to analyze tractographies, essentially large sets of *streamlines*, i.e. sequences of points in 3D spaces generated by Magnetic Resonance scans of the brain. QB is used to identify *bundles*, i.e., particular sets of streamlines that are in close proximity (according to some distance metric) and that could have some relevance for the analysis of dynamics taking place in the brain. Besides the domain specific details, meaning, and significance of the approach, what is relevant from our perspective is the possibility to apply the algorithm to the analysis of trajectories (streamlines have a quite similar semantics to trajectories) in an extremely effective way: QB, in fact, has an average linear cost in the number of trajectories. Due to the fact that trajectories can be generated in distant time frames, though, we need to perform additional spatial–temporal analyses to produce results to our problem, but the overall results are promising both from the perspective of effectiveness (i.e. plausibly discriminate group vs non–group trajectories) and efficiency.

After introducing and discussing a selection of relevant related works, Section 3 will present a dataset we employed for supporting both the development and the experimental evaluation of the proposed method. This method will be presented in Section 4. Achieved results and their quality will be described and discussed in Section 5, while conclusions and future developments will end the paper.

2. Related works

Data about the behaviour of people in urban settings, and especially their position in time, is ever easier to obtain thanks to personal devices with GPS sensors on board. Of course there are privacy issues, but in principle (and in practice - when users opt-in systems gathering location data), data mining techniques can be used, for instance to identify and characterize Points of Interest [4], and more specific techniques can be successfully adopted to recommend locations even considering congestion issues [13] (a topic that at times of COVID seems even more significant). From the perspective of applications more aimed at societal rather than individual benefits, trajectories gathered in urban settings might be analyzed through data mining techniques to identify patterns of traffic useful to restructure public transport [5].

Clustering techniques can also be employed to analyze data gathered at a more microscopic scale, also in indoor settings, in which even video footage can be processed to achieve pedestrian trajectories. Ad-hoc clustering techniques were defined to identify and characterize lanes in pedestrian flows [6], a phenomenon characterizing human crowds' dynamics and having significant impact (it substantially improves the capacity of corridors in public transport stations, see, e.g., [8]). Analyses of real-world human behaviour in the built environment can be therefore highly beneficial for supporting simulation studies and therefore supporting crowd dynamics aware architectural design.

This also applies to the presence of groups in a population of pedestrians. Groups have peculiar effects on collective dynamics at the microscopic level (see, e.g., the already cited [1] or also [10]), due to the reduced walking speed and tendency to form relatively stable patterns. While initial works identified groups through the work of human annotators (provided with guidelines and checklists to support the identification of groups, whose intuitive semantics is easily grasped, but rather fuzzy and informal), later on different automated approaches were defined for analyzing video footage and related trajectories. From initial works (such as [11]) that provided relatively simple approaches (i.e. members of groups have similar initial and final locations in most time windows an overall video footage is subdivided

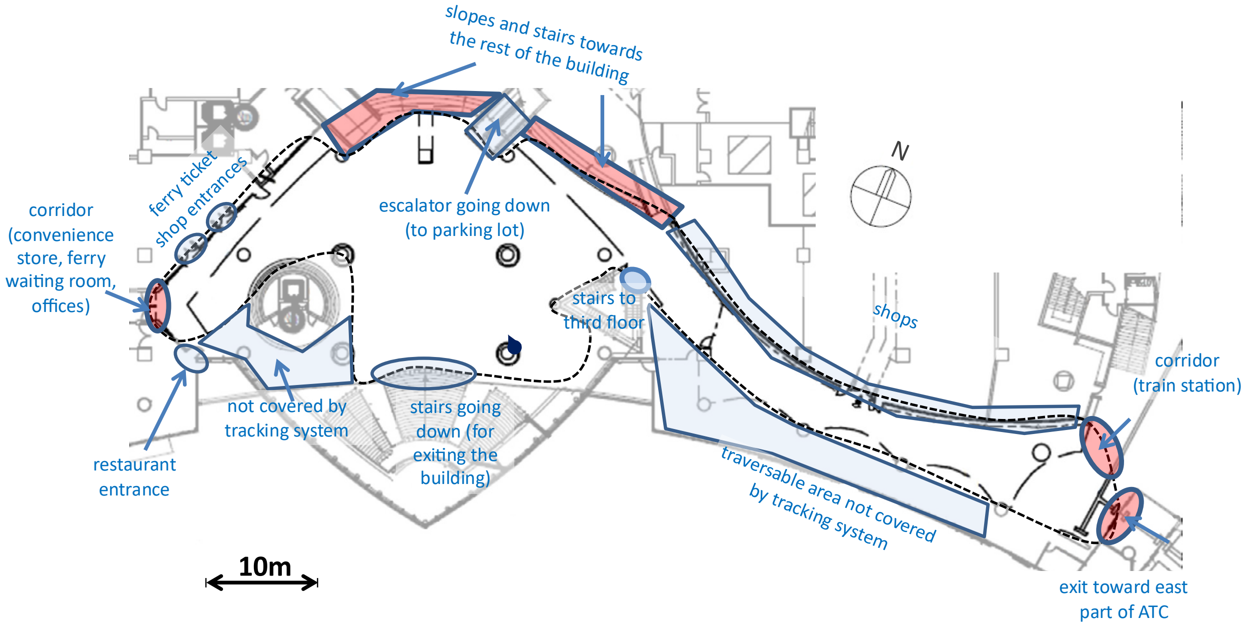


Fig. 1: Map of the shopping center area monitored by sensors and details on the exits. Figure appears courtesy of authors of [2].

into), more complex approaches considering a combination of factors (not just spatial patterns but also proxemics and sometimes even an account of social structure of the groups) to cluster trajectories and identify groups [15, 18]. Within this work we move back to relatively simple models, with minimal assumptions on group structure and behaviour, trying to balance effectiveness in the detection of groups and efficiency of the proposed technique, so as to allow the analysis of relatively large datasets of trajectories and to generate more robust statistics about group frequency in time, dimension, and other characteristics within a specific premise or classes of premises.

3. The dataset

The dataset we analyze consists of pedestrian tracking inside the Asia and Pacific Trade Center (ATC), a shopping mall in Osaka, Japan[1]. A long-term continuous monitoring system composed of a network of 49 sensors was installed in ATC, covering an area of about 900 m^2, as shown in Figure 1. The $3D$ data collection (see [2] for technical details on both the sensor network and tracking algorithms) was carried out in the period between October 2012 and November 2013, generally on Wednesdays and Sundays during the opening hours of the building to the public.

A description of the sensor network and overall pedestrian tracking system and algorithms is out of the scope of the paper, we just report here some information that is relevant for sake of understanding our approach and the achieved results. In particular, the area is mostly a corridor connecting different parts of a mall, with exits on the Eastern side and a (mostly tracked) relatively large square on the Western side, with additional exits. The corridor has several shops on its Northern side and some areas not covered by the tracking system on the Southern side. While some of the trajectories are therefore necessarily incomplete or segmented, the overall flow of people in the area assures that many pedestrians (and groups) are tracked for a relatively long time, and throughout almost all of the covered area. Authors of the observation also carried out interesting analyses [3], from the point of view of someone trying to understand the way the area is really used by the public (e.g. regularities in changes of velocity between hours of the day and types of days – workdays and weekends). Authors of the observation also carried out investigations on selected days to identify and characterize groups [19]: *social groups* (i.e. groups in which at least one of the members was coded as

[1] Available here: https://irc.atr.jp/crest2010_HRI/ATC_dataset/

having social interaction with all the other members, so not necessarily identifiable by just looking at trajectories) were manually identified by human coders, and later on additional analyses were carried out to characterize the observable features of groups in different contextual conditions (e.g. density)[2].

The tracking information is technically available as .CSV files providing the daily monitoring, in particular the sensor records belong to the following time slots: $[10:00, 11:00]$, $[12:00, 13:00]$, $[15:00, 16:00]$ and $[19:00, 20:00]$. Each line of the file corresponds to the position of a person in a given instant, and it contains the following information: timestamp (in seconds), ID, position x (measured in millimeters), position y (mm), height estimation (mm), velocity (mm/s), angle of motion (rad), facing angle (rad). The average error in the estimated position is generally lower than 125 mm, and the frequency of detection is no lower than 10 Hz (again reported in [2]), so the data is more than sufficient for a good characterization of the dynamics of walking (even very quickly) pedestrians.

4. The proposed method

Thanks to the rich and high-quality information provided by the dataset describing pedestrians' movements in the ATC shopping center, we were able to reconstruct a high number of trajectories in several days of the observation. The amount of data associated to trajectories is quite significant, so we considered the application of unsupervised machine learning algorithms to positioning data with fine granularity, in order to extract in a very efficient way correlations among trajectories, for supporting the understanding the behaviour of pedestrians in that area. After some initial tests, we identified and adopted QuickBundles (QB) [9], an approach conceived and experimented in the area of so-called "tractography algorithms", for the automatic extraction of fiber bundles in the brain. Tractographies are essentially large sets of *streamlines*, sequences of points in a 3D space, generated by Magnetic Resonance scans of the brain. Besides the domain specific significance of these data, streamlines are very much related to our trajectories, that are also sequences of points (albeit in a 2D space). QB essentially proposes an efficient way to cluster streamlines based on their similarity in shape and proximity. Efficiency is key for QB since a tractography can contain a very large number of flow lines (up to 10^6), and this also applied to our situation (i.e. a single day of tracking can generate about 10 millions of points, as we will discuss in Section 5). Undoubtedly, the main advantage of QB is its speed, but besides the general similarity in the problem we are facing (i.e. grouping similar sequences of points) we need to consider that the semantics of the points is different. QB, in fact, essentially does not consider the temporal difference among streamlines: sequences (associated to streamlines) that are similar in shape and whose points are spatially close pairwise will be grouped together, irrespectively of the time they were generated. In our case, this would imply the risk of considering a group pedestrians that follow very similar trajectories but maybe in different moments in time (and this happens surprisingly frequently). To this end, we needed to post-process results of QB to add a phase of temporal analysis, to further filter provided results.

Before describing the overall method, however, we want to discuss some features of QB that are relevant to the decision of adopting it for our problem. In preliminary phases of analysis of the problem, we have considered more basic point-by-point clustering techniques, and particularly those not requiring a preliminary knowledge of the number of clusters to be identified in the analyzed dataset, such as density based approaches, but we excluded them due to computational costs, as well as to the difficulty to scale them up to deal with trajectories. Most unsupervised learning algorithms have $O(N^2)$ complexity, at best, where N is the total number of flow lines. They also require the computation of the distances between all the pairs of flow lines to create a distance matrix.

Among the many distance metrics used to determine the similarity between curves, we mention for example the famous Hausdorff distance. The two main disadvantages are that (i) it ignores the sequential nature of flow lines and it treats each trajectory as a cloud of points (ii) its computation requires that every point on the first flow line is compared to every point on the second discrete curve and vice versa. Therefore, the Hausdorff distance has a complexity $O(K * L)$ when comparing the lines with a number of points K and L. Quickbundles uses a particular distance measure called MDF (Minimum average Direct Flip) and requires as input a distance threshold chosen by us. The MDF measure provides that each streamline is resampled to have K points, in any case given the flexibility

[2] Data available here: https://irc.atr.jp/sets/groups/

Table 1: Hyperparameters of the method for the described experiments.

	STEP 1	STEP 2		STEP 3			STEP 4
Hyperparameter	MaxDist	DistThr	TimeThr	DistTol	TimeTol	PctClose	InOutTol
Value	0.60 *m*	1 *m*	1 *s*	0.60 *m*	2 *s*	80%	5 *m*

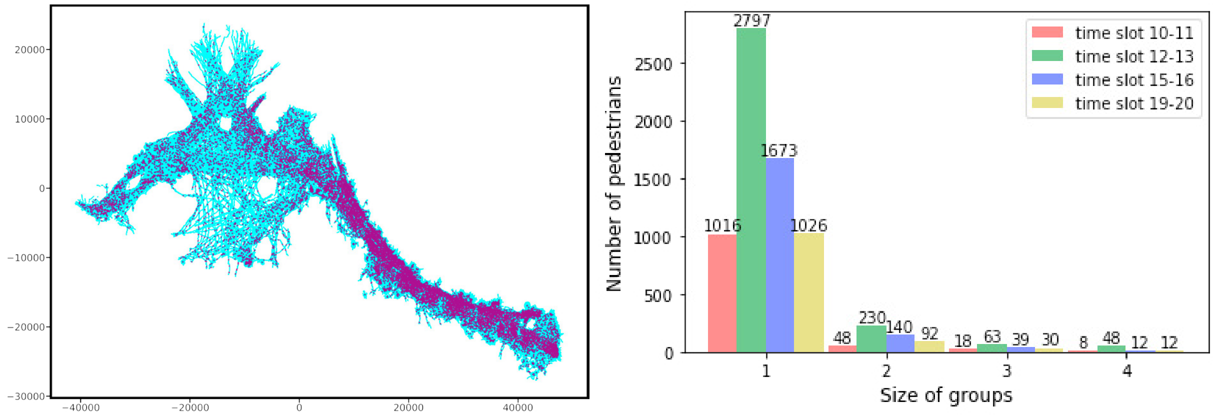
of the algorithm it adapts to any distance measure: for example in the case of GPS points (represented by latitude and longitude) it will be possible to choose the Haversine formula which takes into account the Earth's curvature. The clustering algorithm internally uses a representation of flow lines that are all automatically downsampled / resampled with the same number of points, maintaining the order of the points.

In this work we apply the algorithm using as input the discrete trajectories of pedestrians, so it extracts the length of the arc (i.e. the sum of the length of each segment for a given flow line) and computes the distance between the paths. The MDF function used also takes into account the distance in "length" between the trajectories and not only in "width". So, another important strength for our analysis is that it separates the short flow lines from the long flow lines, grouping them together. The execution time of QB is influenced by the spatial threshold chosen in input, which also controls the heterogeneity of the clusters and by the number of trajectories on which the clustering will be performed. When this threshold is high, fewer clusters are assembled (including heterogeneous paths), and vice versa when it is low, a larger number of clusters (including more homogeneous paths) are created. The complexity of QB is, in the best case, O(N) with N number of trajectories or, in the worst case, O(N^2) when each cluster contains only one flow line. The average case is $O(M * N)$ where M is the number of clusters produced: since M is usually much smaller than N, we can neglect M and denote the complexity as $O(N)$. We can therefore say that QB is particularly efficient for processing / analyzing large datasets.

QB therefore can generate reasonable cluster of *candidate* trajectories whose associated pedestrians could be part of a group, based on spatial similarity criteria. We need to discriminate these candidates introducing relevant temporal considerations to properly identify plausible groups of pedestrians walking together. To this end, we propose an analysis of the instantaneous walking speed of each individual, obtaining the *Hesitation Points* (*HPs*) of each path. An *HP* is defined as the centroid of the area in which a pedestrian is situated (within a fixed spatial threshold) for more than a chosen threshold time. To be able to synthesize the trajectories as *HP* sequences, so that they were significant, we chose 1 *m* and 1 *s* as parameters. The concept of *HP* is essentially a generalization of a *StayPoint (SP)* [20]: *HPs* represent a way to capture a cadence in the movement, without the need of having a complete stop in a *SP* area, which occurs for example when a tourist is attracted to a building, shop, or a point of interest and slows down to the point of standing still for some time. Generally speaking, identifying *HPs* associated to a trajectory represents a way to simplify it, providing a more compact and discrete representation: if the spatial extension is sufficiently large and the time window is relatively small, almost all pedestrian trajectories (unless they are actually related to jogging or running) will be associated to *HPs*. The smaller the area and larger the time window, the less *HPs* will be identified, and they will be really related to points in which pedestrians really slow down. The general intuition, however, is that when two or more pedestrians have a certain number of *HPs* that are close in both space and time they are essentially moving together and they might be member of a group. However, this intuition needs additional support, especially to find candidates similar trajectories to be compared by means of this concept in an efficient way.

The original trajectories, for members of candidate groups have been transformed into *HP* sequences, which are associated with each individual's x and y position, arrival time and leaving time for areas associated to those *HPs*. For each cluster identified by QB, we perform an additional filter by extracting only the paths for which at least a certain share of *HPs* are close to each other in space (within a margin of 60 *cm* from the nearest pedestrian on the right or left). About the temporal dimension, we also impose that the *HPs* of the pedestrians moving together must be offset by a maximum of distance of 2 seconds.

The last operation of the proposed method consists, for the trajectories selected and considered in their entirety not only as *HP* sequences, in controlling the origin and destination points of the trajectories of candidate members of groups. Considering the extremes of the trajectories, we consider that they must be compatibly close in space: when a group enters the hallway entrance and when heading for an exit. This is, of course, an arbitrary decision, as several others in the defined procedure: its rationale is to select *stable* groups in the analysed scenario. If the goal is to



(a) Recorded trajectories in the time slot 10-11 in cyan and the detected *HPs* colored in purple.

(b) Frequency of identified groups according to their types (groups of size 1 means individuals).

Fig. 2: Overall analyzed trajectories and identified groups distribution.

identify instead potential groups *forming* or *separating* in the scenario, this choice would likely be wrong. All steps of the proposed method are shown as pseudocode in Figure 4 in the Appendix.

5. Results and discussion

The results we are going to describe are related to analyses carried out with the proposed method using specific values for the hyperparameters summarized in Table 1: the specific values were defined based on expertise in the analysis of pedestrian and crowd dynamics, and considering preliminary tests to evaluate the achieved results. We did not perform a sensitivity analysis yet, but we can give qualitative indications about the effect of changing these values.

- STEP 1 (QB): lowering MaxDist pushes QB to be more selective when trying to aggregate trajectories; this means that lowering this value decreases the number of detected groups (particularly groups of more than 2 members, which are inevitably less compact), whereas increasing it could lead to false positives.
- STEP 2 (*HP* identification): as suggested in Section 4, DistThr represents the side of a square centered in the *HP* and TimeThr represents the time a pedestrian needs to spend within this square; therefore, increasing DistThr and descreasing TimeThr makes the procedure less strict, and leads to the detection of more *HPs* (that would be however less and less related to an actual reduction of walking speed). These values for the parameters lead to the detection of *HPs* for over 80% of pedestrians, with on average of 8-10 *HPs* per trajectory.
- STEP 3 (candidate groups filtering based on *HP*): DistTol and TimeTol represent the distance, respectively in space and time, for pairs of *HPs* in the trajectories of group members, whereas PctClose represents the minimum share of *HPs* that need to be within DistTol (in space) and TimeTol (in time) for the pedestrians associated to the corresponding trajectories to be considered parts of a group. The rationale is that members of a group should have similar patterns in walking speed, slowing down (and maybe changing direction) in points that are close in space and time. The semantics of these two hyperparameters is similar to DistThr and TimeThr: increasing DistTol and decreasing TimeTol, and correspondingly decreasing PctClose leads to the detection of more groups, but increases the possibility of false positives; nonetheless, having a moderate tolerance grants robustness and the ability to accommodate for temporary divergences among the trajectories that might be due to contextual conditions (i.e. obstacles, other pedestrians).
- STEP 4 (entrance and exit points close in space): InOutTol represents the maximum distance between extreme points of the trajectories (entrance points and exit points in the analyzed area).
  We do not consider a temporal distance since in some of the exits from the scenario there might be queues, leading members of a group to actually exit a scenario in moments in time that might even be quite different although being relatively close in space (but some members might just be out of the tracking area).

Table 2: Test results details

Time slot	Total points	Total trajectories	QB clusters	*HPs* of candidate groups	Detected groups
$10:00-11:00$	$1,611,268$	$1,090$	837	850	32
$12:00-13:00$	$2,759,663$	$3,138$	$2,337$	$3,886$	148
$15:00-16:00$	$2,369,129$	$1,864$	$1,349$	$2,090$	86
$19:00-20:00$	$1,877,048$	$1,160$	750	$1,361$	59

Adopting these values for the hyperparameters, and taking as reference the .CSV files associated to January 9, 2013 (a winter working day), Table 2 shows some aggregated results obtained with the proposed analysis method. The execution time of QB respectively took 2.16 s, 2.05 s, 1.07 s, and 701 ms in the four time slots. Figure 2a shows all trajectories in a time slot of morning, together with their *HPs*; each trajectory is associated to a sequence of about 8-10 *HPs* on average. The execution time of the *HP* detection procedure is respectively 7 minutes and 46 seconds, 24 *min* 26 s, 8 *min* 49 s, and 9 *min* 21 s. QB is therefore significantly more efficient than the *HP* detection, that moreover was only performed for trajectories associated to members of group candidates. All the tests were performed in Python 3 language, on an Intel Core i5 at 1GHz having 16 GB of RAM.

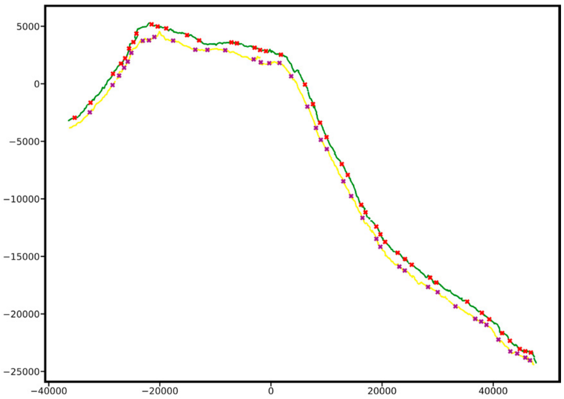
To discuss the effectiveness of the proposed method, we show some details about detected groups. Figure 2b shows the distribution of pedestrians by group size and by time slots. Pedestrians belong to a group between about 15% of the cases (time slot $10:00-11:00$) and about 24% of the cases (time slot $19:00-20:00$); smaller groups are more frequent. The distribution of groups by size is relatively in line with observations from the literature (e.g. [10]), but the current calibration of the method is probably a bit too selective considering results of a manual analysis of the dataset described in [19].

The *context* of the observed dynamics plays a fundamental role, and here we take a multi–faceted acceptance of the term context: cultural aspects, the type of environment and time of day of the observation (since they can imply that pedestrians have a very different motivation to be there, and motivation can lead to very different walking speed and style), even the weather conditions (if pedestrians are outdoor). The ATC dataset seems associated to a situation in which groups are less frequent than in several other datasets available in the literature (starting from the above cited [10]), in which the presence of groups seems to be preponderant over individuals.

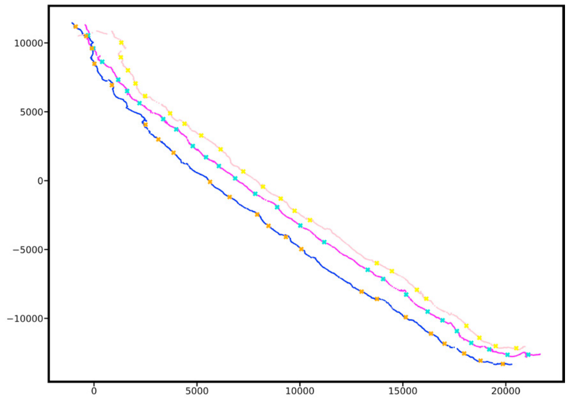
It should be clear that we are considering almost inevitable a margin of error, whose estimation will require additional work to take into consideration also additional datasets, to be able to evaluate the method also in different contexts and to have, therefore, a more well–rounded and robust estimation of the quality of group detection. Also, once again, a more thorough analysis of the sensitivity of the method to the variation of the hyperparameters is still ongoing and it could lead to improved quality of the overall group detection.

Let us now analyze four examples of trajectories that were considered to be associated to pedestrians belonging to a group, in good accordance with manual analyses described in [19]: Figure 3 shows two pairs, a group of 3, and a group of 4 pedestrians. Figure 3a shows very limited distance among the paths taken by the couple of pedestrians, that do not only have points of entrance end exit from the scenario that are close in space, but they also have almost the same way to bend and change direction throughout their tracked movement. By considering the trajectories it seems totally reasonable to consider them a group, although they might show no other sign of social interaction, and that would certainly put a human annotator in some difficulty when deciding if the pedestrians were to be considered a pair or simply two individuals walking consistently side by side.

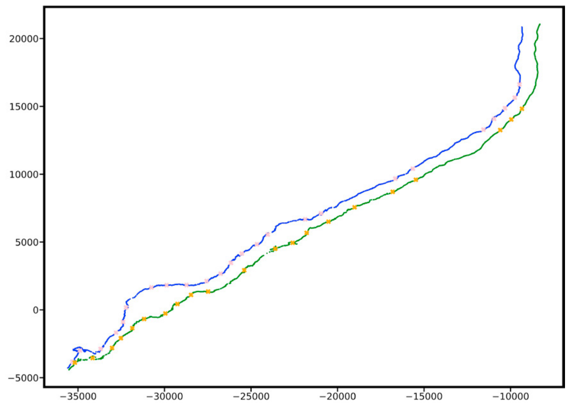
Figure 3b shows a case in which the group of 3 members almost consistently moves in a very compact formation, but we can also see a particular situation in the upper left corner: depending on the fact that it is an entrance or exit from the scenario, we could interpret it as a dynamics of *formation* or *dissolution* of the group pattern. This also represents an interesting possibility offered by the structure of the method, in which there is a first phase of purely spatial analysis, followed by a spatial and temporal analysis. This example shows that QB can be configured to be sufficiently inclusive to group together trajectories having a relatively large portion somewhat close to each other, despite having extremes even departing in a visible way. This could open up the possibility to focus more on these



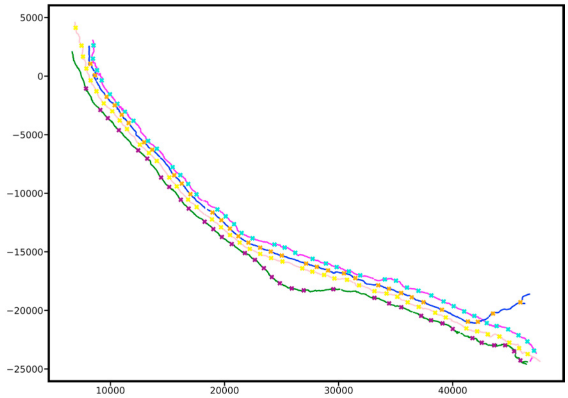
(a) A group of 2 pedestrians detected in time slot 10-11.



(b) Paths of 3 individuals walking together in time slot 19-20.



(c) A couple of pedestrians and their *HPs*, detected in time slot 12-13.



(d) Trajectories of a group of 4 people with their *HPs*, in time slot 15-16.

Fig. 3: Four examples of detected groups.

dynamical phases of groups, which are some of the aspects of groups that are still object of research and attention also from the perspective of social sciences and psychology [16].

Figure 3c shows two trajectories in which the two pedestrians do take some distance (for reasons that are not apparent by considering just the trajectories, which might be related to the presence of other pedestrians moving in a conflicting way being avoided by one of the members of the group) but they eventually stay close to each other most of the time. This shows that the defined procedure can accommodate for some differences, thanks to the hyperparameters associated to the STEP 3 (DistTol and TimeTol), but this also shows the risk of increasing these values too much, i.e. increasing the number of false positives.

Figure 3d finally shows how pedestrians take a "walk abreast" walking pattern (i.e. the try to walk side by side) but due to some local contextual condition two of them change one member takes some distance from the rest of the group, but comes again closer to the others. In the lower right corner, however, another pattern of group formation / dissolution can be identified.

6. Conclusions and future developments

The paper has introduced a novel spatial temporal analysis approach to support the detection of groups of pedestrians within datasets of trajectories. The defined method takes minimal assumptions on group dynamics, and does not require human intervention or supervision, being based on an unsupervised machine learning approach. In particular, it employs QB to perform a first spatial analysis phase producing clusters of trajectories that could be considered as being generated by pedestrians that are members of a group, that are object of subsequent spatial temporal analyses, employing the notion of Hesitation Point. We described the method and its internal working, as well as results on the analysis of an available dataset of trajectories: results are promising, on the side of effectiveness, and they are compatible with the required efficiency in managing large quantities of trajectories in a tractable way. Open points for future works are, for instance, a sensitivity analysis for the hyperparameters of the defined method, that could also lead to an improved quality of group detection. To improve the quality of group detection, we could also consider additional information available in the ATC dataset describing pedestrian facing direction: looking at each other can imply some form of social interaction, but this kind of information is not necessarily available in other datasets of trajectories considering groups, so this extension might be of little general use. The overall approach could also be employed to study phases of group dynamics that are so far not well characterized qualitatively and quantitatively.

References

[1] Bode, N.W.F., Holl, S., Mehner, W., Seyfried, A., 2015. Disentangling the impact of social groups on response times and movement dynamics in evacuations. PLOS ONE 10, 1–14.

[2] Brscic, D., Kanda, T., Ikeda, T., Miyashita, T., 2013. Person tracking in large public spaces using 3-d range sensors. IEEE Transactions on Human-Machine Systems 43, 522–534. doi:10.1109/thms.2013.2283945.

[3] Bršćić, D., Kanda, T., 2015. Changes in usage of an indoor public space: Analysis of one year of person tracking. IEEE Transactions on Human-Machine Systems 45, 228–237. doi:10.1109/THMS.2014.2374172.

[4] Cavallaro, C., Verga, G., Tramontana, E., Muscato, O., 2020. Eliciting cities points of interest from people movements and suggesting effective itineraries. Intelligenza Artificiale 14, 75–87. URL: https://doi.org/10.3233/IA-190040, doi:10.3233/IA-190040.

[5] Cavallaro, C., Vitrià, J., 2020. Corridor detection from large gps trajectories datasets. Applied Sciences 10. URL: https://www.mdpi.com/2076-3417/10/14/5003, doi:10.3390/app10145003.

[6] Crociani, L., Vizzari, G., Gorrini, A., Bandini, S., 2018a. Identification and characterization of lanes in pedestrian flows through a clustering approach, in: Ghidini, C., Magnini, B., Passerini, A., Traverso, P. (Eds.), AI*IA 2018 – Advances in Artificial Intelligence, Springer International Publishing, Cham. pp. 71–82.

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

[8] Feliciani, C., Nishinari, K., 2016. An improved cellular automata model to simulate the behavior of high density crowd and validation by experimental data. Physica A: Statistical Mechanics and its Applications 451, 135–148. URL: https://www.sciencedirect.com/science/article/pii/S0378437116001047, doi:https://doi.org/10.1016/j.physa.2016.01.057.

[9] Garyfallidis, E., Brett, M., Correia, M.M., Williams, G.B., Nimmo-Smith, I., 2012. QuickBundles, a method for tractography simplification. Frontiers in Neuroscience 6, 175. doi:10.3389/fnins.2012.00175.

[10] Gorrini, A., Vizzari, G., Bandini, S., 2016. Age and group-driven pedestrian behaviour: from observations to simulations. Collective Dynamics 1, 1–16. URL: https://collective-dynamics.eu/index.php/cod/article/view/A3, doi:10.17815/CD.2016.3.

[11] Khan, S.D., Vizzari, G., Bandini, S., Basalamah, S.M., 2015. Detection of social groups in pedestrian crowds using computer vision, in: Battiato, S., Blanc-Talon, J., Gallo, G., Philips, W., Popescu, D.C., Scheunders, P. (Eds.), Advanced Concepts for Intelligent Vision Systems - 16th International Conference, ACIVS 2015, Catania, Italy, October 26-29, 2015, Proceedings, Springer. pp. 249–260. URL: https://doi.org/10.1007/978-3-319-25903-1_22, doi:10.1007/978-3-319-25903-1_22.

[12] Khatoun, R., Zeadally, S., 2016. Smart cities: Concepts, architectures, research opportunities. Commun. ACM 59, 46–57. URL: https://doi.org/10.1145/2858789, doi:10.1145/2858789.

[13] de Nijs, F., Theocharous, G., Vlassis, N., de Weerdt, M.M., Spaan, M.T.J., 2018. Capacity-aware sequential recommendations, in: André, E., Koenig, S., Dastani, M., Sukthankar, G. (Eds.), Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2018, Stockholm, Sweden, July 10-15, 2018, International Foundation for Autonomous Agents and Multiagent Systems Richland, SC, USA / ACM. pp. 416–424. URL: http://dl.acm.org/citation.cfm?id=3237448.

[14] Pelechrinis, K., Quercia, D., 2015. Urban informatics and the web, in: Gangemi, A., Leonardi, S., Panconesi, A. (Eds.), Proceedings of the 24th International Conference on World Wide Web Companion, WWW 2015, Florence, Italy, May 18-22, 2015 - Companion Volume, ACM. p. 1547. URL: https://doi.org/10.1145/2740908.2741983, doi:10.1145/2740908.2741983.

[15] Solera, F., Calderara, S., Cucchiara, R., 2016. Socially constrained structural learning for groups detection in crowd. IEEE Trans. Pattern Anal. Mach. Intell. 38, 995–1008. URL: https://doi.org/10.1109/TPAMI.2015.2470658, doi:10.1109/TPAMI.2015.2470658.

[16] Templeton, A., Drury, J., Philippides, A., 2015. From mindless masses to small groups: Conceptualizing collective behavior in crowd modeling. Review of General Psychology 19, 215–229. URL: https://doi.org/10.1037/gpr0000032, doi:10.1037/gpr0000032, arXiv:https://doi.org/10.1037/gpr0000032. pMID: 26388685.

[17] United Nations, 2014. World Urbanization Prospects: The 2014 Revision. United Nations, Department of Economic and Social Affairs, Population Division.

[18] Yücel, Z., Zanlungo, F., Ikeda, T., Miyashita, T., Hagita, N., 2013. Deciphering the crowd: Modeling and identification of pedestrian group motion. Sensors 13, 875–897. URL: https://www.mdpi.com/1424-8220/13/1/875, doi:10.3390/s130100875.

[19] Zanlungo, F., Brščić, D.c.v., Kanda, T., 2015. Spatial-size scaling of pedestrian groups under growing density conditions. Phys. Rev. E 91, 062810. URL: https://link.aps.org/doi/10.1103/PhysRevE.91.062810, doi:10.1103/PhysRevE.91.062810.

[20] Zheng, Y., Zhang, L., Xie, X., Ma, W.Y., 2009. Mining correlation between locations using human location history, in: Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Association for Computing Machinery, New York, NY, USA. p. 472–475. URL: https://doi.org/10.1145/1653771.1653847, doi:10.1145/1653771.1653847.

Appendix

```
% QuickBundles is applied to all trajectories of the time slot.
STEP 1:
Input: {A set of trajectories D, which are sequences of 2D points, and a maximum distance threshold.}
Output: {A partition of the subset of trajectories given in input, formed by clusters.}

        QuickBundles(D, MaxDist)
        return {C_i} #clusters

% Hesitation Points for trajectories, which are contained in clusters obtained in
% STEP 1 and with cardinality greater than 1, are extracted.
STEP 2
Input: {A set D' of N trajectory D'={T_k}_{k=1,...N}, where a trajectory T_k is a sequence of timestamped points
T_k={(x_0, y_0, t_0) ,..., (x_n, y_n, t_n)}, a distance threshold (DistThr) and time threshold (TimeThr).}
Output: {A set S of sequences of Hesitation Points HPs, where for each pedestrian
T'_k ={HP_1, HP_2,..., HP_s}, HP_m=(x_mean_m, y_mean_m, time_arrive_m, time_left_m).}

for each trajectory T in D':
    S={}
    i=0, cardinality_traj=|T|
    while i<cardinality_traj
            j=i+1;
            while j<cardinality_traj:
                dist=distance((x_i,y_i), (x_j, y_j))
                if dist>DistThr
                    deltatime=t_j-t_i
                    if deltatime>TimeThr
                            S.x_mean=Mean({x_k | i<k<j})
                            S.y_mean=Mean({y_k | i<k<j})
                            S.time_arrive=t_i
                            S.time_left=t_j
                            HPs.append(S)
                            break
            j=j+1
        i=j
    return HPs for each trajectory ID.

% For every achieved cluster with cardinality higher than 1 we examine trajectories, considering their HPs obtained
% in STEP 2, saving only pedestrian IDs whose HPs are close in time and space, that could be members of a group.
% Trajectories which have not any HPs are not analyzed: they are considered associated to pedestrians
% moving too quickly to be members of a group
STEP 3:
Input: {Clusters C_i returned by STEP 1.}
Output: {Groups of pedestrians' trajectories for HPs near in space and time.}

groups={} # empty set
for each C_i, with cardinality(C_i)>1:
    group={} #empty set
    for each pair of trajectories Tr=(T_j,T_k) in C_i:
        if (number of HPs of T_j)>(number of HPs of T_k):
            Tr=(T_k,T_j)
        NearHP={} # empty set
        for each HP of Tr_0:
            for each HP' of Tr_1:
                if dist(HP, HP')<DistTol &&
                |timestamp(HP)-timestamp(HP')|<TimeTol:
                    NearHP.append((HP,HP'))
                    break
        if cardinality(NearHP)>=PctClose*(number of HPs of Tr_0):
            group.append(T_j,T_k)
    if cardinality(group)>0:
        groups.append(group)

% For candidate groups, we look at origins and destinations; if those points are close, we confirm group membership
STEP 4:
Input: {Groups saved in the previous step.}
Output: {Confirmed groups of pedestrians walking together, with common origin and
destination.}

groups_confirmed={} #empty set
for each group:
    for each pair of IDs of a same group:
        if dist(origin_1, origin_2)<=InOutTol &&
        dist(destination_1, destination_2)<=InOutTol:
            groups_confirmed.append({IDs})
```

Fig. 4: Pseudocode of the proposed approach.