The Korean Institute of Communications and Information Sciences

Available online at www.sciencedirect.com

ScienceDirect

ICT Express 7 (2021) 449–452

www.elsevier.com/locate/icte

Contextual analysis of pedestrian mobility in transport terminals

Joelma C.C. e Silva*, Flavio S.C. da Silva

Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, Brazil

Received 20 September 2020; accepted 21 February 2021
Available online 8 March 2021

Abstract

Data acquisition using security cameras for pedestrian tracking and counting is inherently inaccurate due to limited video quality and the dynamics of pedestrian mobility. In this article we explore how specific patterns of pedestrian behavior in transport terminals can improve accuracy in tracking and counting. We show the effectiveness of our proposed techniques using data from the New York Grand Central Station. The proposed tool improves the accuracy of pedestrian labeling based on information about the dynamics of pedestrian mobility in the specific context of transport terminals.

Keywords: Contextual analysis; Pedestrian behavior; Transport terminal monitoring

1. Introduction

In the present work, we improve quality of data about pedestrian mobility in transport terminals, grounded on heuristics based on the dynamics of movements of pedestrians in this specific context.[1]

Behavioral analysis of pedestrians is important for applications such as crowd management, design of public spaces, surveillance and the design of intelligent interactive environments [2]. The analysis of pedestrian behavior provides parameters for crowd management [3,4], disaster prevention [5] and the design of crossings [4,6], spaces for exhibits [7] and evacuation plans [8]. Information captured from security cameras has low cost and is largely available, but is also inaccurate. For this reason, a careful solution for processing, organizing and analyzing pedestrian mobility information is needed [9].

It can be challenging extracting individualized pedestrian information from videos when crowd density is big [10]. Some initiatives aim at understanding pedestrian behavior using simulations based on random variables. Realistic simulations are hard to build, however, because they rely on initialization parameters which can be hard to select [11–13]. A more promising alternative has been analysis using empirical data, usually based on aggregate crowd behavior rather than data at the individual level [9,14]. In the present article, we follow this alternative and introduce heuristics that explore contextual information to improve accuracy of data obtained from existing tracking algorithms. The proposed heuristics were tested using open data available about the New York Grand Central Station. Comparisons were made between the data obtained after application of the heuristics and the data obtained from one popular tracking algorithm. The results show that the proposed heuristics are effective in improvement of accuracy of available data.

2. Related work

Maheshwari and Heda [10] developed crowd analysis based on data from security cameras, organized as (1) people counting, (2) people tracking and (3) crowd behavior analysis. Information extraction followed three steps: (1) preprocessing to extract features from video; (2) object tracking to segment individuals from the crowd; and (3) behavior recognition to identify movement patterns based on attributes such as direction and speed.

Lam, Cheung and Lam [15] analyzed passenger clustering effects in Light Rail Transit (LRT) stations in Hong Kong. Surveys were used to collect opinions about comfort correlated

* Corresponding author.
E-mail addresses: joelmac@alumni.usp.br (J.C.C. e Silva), fcs@usp.br (F.S.C. da Silva).

[1] A detailed account of this work can be found in the PhD thesis of the first author [1].

Peer review under responsibility of The Korean Institute of Communications and Information Sciences (KICS).

https://doi.org/10.1016/j.icte.2021.02.007

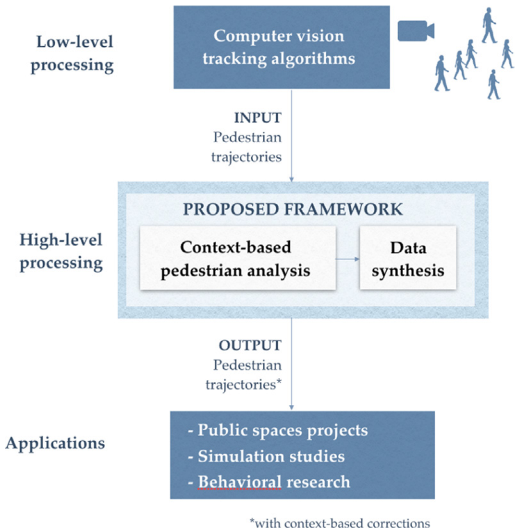

Fig. 1. Proposal to improve data accuracy of pedestrians.

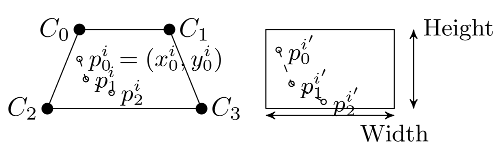

Fig. 2. Transformation to adjust perspective distortions.

to density of passengers. These results provided data to guide the design and planning of new LRT platforms.

Croft and Panchuk [16] studied pedestrian behavior and collision avoidance in correlation with eye contact and fluctuations in walking speed and direction, based on indoor lab experiments.

3. Methodology

In the present work we focus on improvements on data accuracy based on contextual information. We consider data from surveillance video cameras about behavior of pedestrians in transport terminals, as depicted in Fig. 1.

Video images are adjusted to correct geometric distortions due to perspective of video capture [17–19], as illustrated in Fig. 2. Frame sequences are transformed with respect to time using the frame-per-second ratio of the video being used, in order to have distances properly represented in meters and time intervals properly represented in seconds. Adjusted paths are then repaired in three steps, as detailed in Fig. 3. The output of this process is a collection of adjusted pedestrian paths, together with assessment reports to support design and planning of transport terminals.

The trajectory of a pedestrian i is denoted as t^i and consists of a list P^i of positions p^i. Each position is a triple (x_p^i, y_p^i, f_p^i) to denote coordinates (x_p^i, y_p^i) of the pedestrian in a video frame f_p^i. Trajectories are analyzed using a set of parameters as detailed in Table 1.

Component 1 corrects path distortions, improving pedestrian path labeling by correction of assignments of multiple paths to a single pedestrian. This problem occurs, for example,

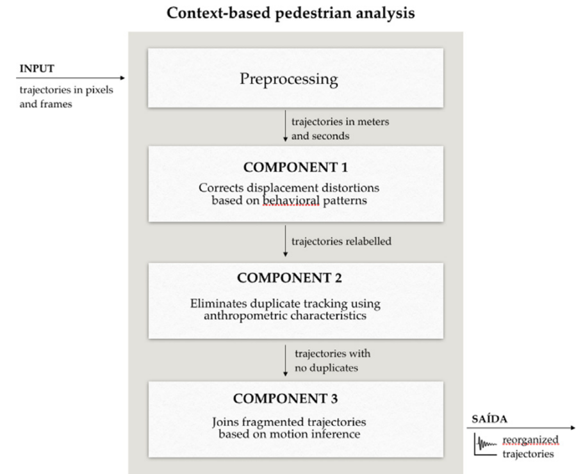

Fig. 3. Path repair in three steps.

Table 1
Contextual parameters by type and scope.

Parameter	Description Type - Scope - Reference
occlusionThreshold	Behavioral - 1 - θ
minDistThreshold	Anthropometric - 1, 2, 3 -
minTimeThreshold	Anthropometric - 1, 2, 3 - corporal space + comfort distance
widthThreshold	Anthropometric - 2, 3 -
heightThreshold	Anthropometric - 2, 3 -
speedThreshold	Behavioral - 2, 3 - Avg speed
dirChangeThreshold	Behavioral - 2, 3 - θ
similarityThreshold	Data capture - 3 -

when two people cross paths and the tracking algorithm fails to continue to track them independently. The proposed algorithm corrects the labels based on the identification of paths that have an angle that indicates an abrupt change of direction. The identification of the threshold angle to indicate abrupt changes of direction and, therefore, potential errors in path labeling, must take into consideration that tracking algorithms may not consistently identify an individual by a specific point in their body. For example, it can happen that an individual is identified by the head during part of a path and by the hands or feet during another part of the same path. To circumvent this problem, we adopt a minimum distance to calculate direction and speed.

Labels assigned to pedestrians by the tracking algorithm are updated based on the observation that a pedestrian tracking problem can occur due to occlusion during path crossing and that abrupt changes in paths violate the principle of least

Fig. 4. Correction based on abrupt change of direction.

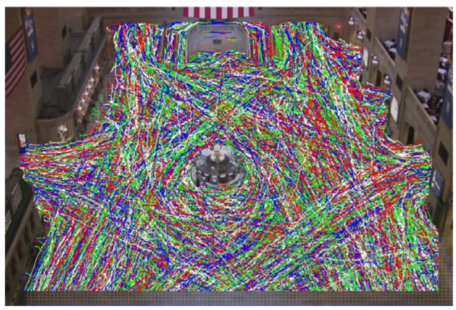

Fig. 5. Pedestrian trajectories — Grand Central Station NY.

effort that pedestrians choose to determine their paths. Fig. 4 illustrates the trajectory t^a with the abrupt point of change of direction p_3^a. As effect of Component 1, the original trajectory t^a is divided into two trajectories $t^{a'}$ and $t^{a''}$.

Component 2 eliminates duplicate paths using anthropometric attributes. Pedestrians can be counted twice if two different body parts – or a body part and e.g. a bag – are identified as two pedestrians. This error is reduced based on a minimum comfort space between pedestrians.

Component 3 joins fragmented paths based on motion analysis, based on two simple contextual assumptions: (1) all paths start at one of the terminal doors, and (2) the *Least Effort Hypothesis* [20], which assumes that people tend to move in rectilinear trajectories to minimize distance and effort. An iterative search is performed, to search for partial trajectories with starting points near the end of a potentially incomplete trajectory, therefore finding pairs of trajectories that violate assumption (1), and then to check whether pedestrian behavioral traits are preserved given assumption (2), in order to join these trajectories. Iterations are performed until no further trajectories can be joined.

4. Comparative tests

Many public data sets can be used to test the proposed tool [21]. In this section we illustrate an application of the proposed tool on data from the New York Grand Central Station [3]. The Grand Central Station has eight inbound/outbound regions, providing 56 different combinations of source–destination pairs, although there are some rare cases of pedestrians entering and leaving through the same door. All trajectories taken from the scene of the 33-minute video as featured in the literature [3] are presented in Fig. 5.

The labeling of trajectories in a data set is costly, time-consuming and not necessarily reliable. For these reasons, it is not viable to perform quantitative assessment of the effectiveness of our tool based on ground truth. In order to mitigate this problem, we introduce an indicator for approximate assessment of quality increase in trajectory labeling.

The indicator, named *Improvement Factor*, is influenced by *comprehensiveness*, i.e. the number of adjusted trajectories relative to the total trajectories in the original data set, and *correctness*, i.e. the extent to which adjustments improved labeling accuracy. It is defined for each Component c as $if_c = pc_c * pi_c$, where pc_c (comprehensiveness) is the percentage of adjusted trajectories that are identified and corrected by Component $c = 1, 2, 3$ and pi_c (correctness) is the percentage of adjustments that are estimated to be corrections, i.e. $pc_c = \frac{tc_c}{\tau}$, where tc_c is the total number of trajectories adjusted by component c, and τ is the total number of trajectories in the data set. Correctness pi_c is assessed based on a random sample of size n of adjusted trajectories, i.e. $pi_c = pr_c - pw_c$, where pr_c is the percentage of correct adjustments and pw_c the percentage of wrong adjustments, i.e. $pw_c = 1 - pr_c$, hence $pi_c = 2pr_c - 1$ and pr_c can be obtained by averaging the values calculated by $pr_c = \frac{1}{n}\sum_{k=1}^{n}\frac{rc_k}{tc_k}$.

Component 1 had accuracy of 93.3% in adjustments. The total amount of adjusted trajectories was 6.440(15.0%). These results suggest that the original data set featured high density of pedestrians and large number of trajectory crossings. The improvement factor $if_1 = 13.0\%$ indicates an improvement of approximately 5.566 trajectories. ($13.0\% * 42.821$). For Component 2, we had 6.950 (16, 2%) excluded trajectories, 92, 5% of which being correct indications, yielding $if_2 = 13.8\%$. These values indicate a significant reduction in pedestrian multiple counting (13.8%), calculated as the difference between the percentage of correctly deleted trajectories ($16.2\% * 92.5\% = 15.0\%$) and the mistakenly deleted trajectories ($16.2\% * 7.5\% = 1.2\%$). For Component 3, we had a relatively small number of trajectories that were adjoined (1.3%), however with good accuracy (90.0%), thus resulting in $if_3 = 1.0\%$. One possible reason for the low coverage of Component 3 is the density of pedestrians in the scenario under consideration, which hinders the capability of the proposed procedure to merge trajectories that are labeled as disconnected.

The *Total Improvement Factor - TIF*, obtained by summing the improvement factors of each component, was $TIF = if_1 + if_2 + if_3 = 13.0\% + 13.8\% + 1.0\% = 27.8\%$.

5. Conclusions

The benefits of using contextual information to increase the quality of pedestrian mobility data extracted from videos can only be achieved after proper calibration of parameters, which at the moment requires proper manual verification to be done properly. The results obtained using data from the Grand Central Station in NY indicate that the tool is feasible to use and can be effective to increase the accuracy of data. From the results presented here, several initiatives can be developed, such as the use of machine learning to optimize contextual parameters. It can be also interesting, in the future, to study patterns of variation of the contextual parameters with respect to different scenarios, to verify whether the optimized values of parameters themselves could unveil interesting aspects of specific contexts.

CRediT authorship contribution statement

Joelma C.C. e Silva: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft. **Flavio S.C. da Silva:** Conceptualization, Methodology, Supervision, Writing review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was developed as part of the PhD research work of the first author. We acknowledge support from the University of Sao Paulo and the research initiative INCT - Future Internet for Smart Cities.

References

[1] J.C.C. Silva, Análise Baseada em Contexto do Movimento de Pedestres em Terminais de Transport (Ph.D. thesis), Depto. de Ciencia da Computacao, USP, 2020.
[2] B. Zhan, D.N. Monekosso, P. Remagnino, S.A. Velastin, L.-Q. Xu, Crowd analysis: a survey, Mach. Vis. Appl. 19 (5) (2008) 345–357.
[3] J. Zhong, W. Cai, L. Luo, H. Yin, Learning behavior patterns from video : A data-driven framework for agent-based crowd modeling, AAMAS 2015 (ISSN: 15582914) (2015) 801–809.
[4] B. Zhou, X. Tang, X. Wang, Learning collective crowd behaviors with dynamic pedestrian-agents, Int. J. Comput. Vis. (ISSN: 15731405) 111 (1) (2014) 50–68.
[5] M. Moussaid, D. Helbing, G. Theraulaz, How simple rules determine pedestrian behavior and crowd disasters, Proc. Natl. Acad. Sci. 108 (17) (2011) 6884–6888.
[6] A. Gorrini, G. Vizzari, S. Bandini, Towards modelling pedestrian-vehicle interactions: Empirical study on urban unsignalized intersection, 2016, URL: http://arxiv.org/abs/1610.07892.
[7] L. Crociani, G. Lammel, G. Vizzari, Multi-scale simulation for crowd management: A case study in an urban scenario, in: Autonomous Agents and Multiagent Systems. AAMAS 2016, in: Lecture Notes in Computer Science, vol. 10002, Springer, Cham, 2016.
[8] M. Moussaïd, M. Kapadia, T. Thrash, R.W. Sumner, M. Gross, D. Helbing, C. Hölscher, Crowd behaviour during high-stress evacuations in an immersive virtual environment, J. R. Soc. Interface (ISSN: 1742-5689) 13 (122) (2016).
[9] M. Haghani, M. Sarvi, Crowd behaviour and motion: Empirical methods, Transp. Res. B 107 (2017).
[10] S. Maheshwari, S. Heda, A review on crowd behavior analysis methods for video surveillance, 2016, pp. 1–5.
[11] S. Kim, S.J. Guy, D. Manocha, M.C. Lin, Interactive simulation of dynamic crowd behaviors using general adaptation syndrome theory, in: Proc. of the ACM SIGGRAPH Symp. on Interactive 3D Graphics and Games, in: I3D '12, ACM, New York, NY, USA, ISBN: 978-1-4503-1194-6, 2012, pp. 55–62.
[12] S.J. Guy, S. Kim, M.C. Lin, D. Manocha, Simulating heterogeneous crowd behaviors using personality trait theory, in: Proc. of the 2011 ACM SIGGRAPH/Eurographics Symp. on Computer Animation, in: SCA '11, ACM, New York, NY, USA, ISBN: 978-1-4503-0923-3, 2011, pp. 43–52.
[13] F. Durupinar, U. Gudukbay, A. Aman, N.I. Badler, Psychological parameters for crowd simulation: From audiences to mobs, IEEE Trans. Vis. Comput. Graphics 22 (9) (2016) 2145–2159.
[14] K. Teknomo, Y. Takeyama, H. Inamura, Determination of pedestrian flow performance based on video tracking and microscopic simulations, CoRR (2016).
[15] W.H.K. Lam, C.-Y. Cheung, C.F. Lam, A study of crowding effects at the Hong Kong light rail transit stations, Transp. Res. A 33 (5) (1999) 401–415.
[16] J.L. Croft, D. Panchuk, Watch where you're going? Interferer velocity and visual behavior predicts avoidance strategy during pedestrian encounters, J. Motor Behav. 50 (4) (2017) 353–363.
[17] E.M. Mikhail, J.S. Bethel, J.C. McGlone, Introduction to Modern Photogrammetry, Vol. 31, 2001.
[18] L. Roberto, S.A. Lima, S.J.S. Sant'Anna, E.H. Shiguemori, Correção de distorção projetiva em imagens obtidas com câmera a bordo de VANT, in: XVIII Simpósio Brasileiro de Sensoriamento Remoto, 2017.
[19] V. OpenC, Geometric transformations of images, 2019.
[20] G. Zipf, Human Behaviour and the Principle of Least Effort, Addison-Wesley, 1949.
[21] M.S. Zitouni, H. Bhaskar, J. Dias, M. Al-Mualla, Advances and trends in visual crowd analysis: A systematic survey and evaluation of crowd modelling techniques, Neurocomputing (ISSN: 0925-2312) 186 (Supplement C) (2016) 139–159.