Collective Dynamics

Are depth field cameras preserving anonymity?

Cécile Appert-Rolland[1] · Sami Habet[1]

[1] Université Paris-Saclay, CNRS/IN2P3, IJCLab Orsay, France
E-mail: Cecile.Appert-Rolland@ijclab.in2p3.fr, sami.habet@synchrotron-soleil.fr

Received: 26 October 2023 / Last revision received: 30 April 2024 / Accepted: 05 May 2024
DOI: 10.17815/CD.2024.176

Abstract This paper presents a preliminary study to assess the degree of anonymization provided by the use of depth field camera, for various degrees of pixelization.

First the passage of 24 participants under a depth field camera was recorded. Each of the corresponding video was degraded with various levels of pixelization. Then the videos were shown to a subset of 6 participants, using a dedicated software which presents the videos in random order, starting with the lowest resolution. Each participant had to recognize themself, and in order to achieve this goal, could progressively improve the resolution.

Our results question the fact that pixelization is the proper way to improve anonymity. Actually recognition seems to a large extend to be based on dynamic features rather than on the resolution of the picture. Besides we identify mostly 2 groups of responses: either the person can identify him/herself whatever the pixelization, or the recognition task is out of reach. Thus, the ability to use dynamic features could be person dependent. Further exploration would be useful to confirm this observation.

Keywords Pedestrians · tracking · depth field cameras · anonymity preserving

1 Introduction

While fluid behavior is strongly constrained by mass and momentum conservation, pedestrians dynamics only obey number conservation, leaving much freedom for the choice of their walking behavior. Compared to cars, pedestrians have much less inertia, and can change their speed and direction almost instantaneously. Besides, individuals can have different goals in the 2D world within which they evolve. As a result, pedestrian behavior can hardly be derived from general principles, and one needs data in order to gain insight into their dynamics [1–3].

The development of crowd science owes a lot to the progress in individual tracking. At the same time, tracking may raise privacy concerns. Among the various devices allowing to track pedestrians' dynamics, cameras are usually a good choice to collect data in real environments, as they don't assume any equipment to be carried along. In particular, video cameras allow to obtain individual trajectories - among other features, at least for not too dense crowds. However, they raise an issue about privacy protection. Indeed, regular videos provide in general images that allow to recognize people.

This is in favor of rather using some depth field cameras, as it was done in several experiments [4–6]. Depth field cameras, as for example stereo cameras or Kinect ones, measure the distance from the first obstacle to the camera plane. If placed vertically, a pedestrian head appears as a point with maximum height (or minimum distance from the camera) that can be tracked to provide the pedestrian trajectory. Though this technique was first usable only at rather low density, machine learning based algorithms allow now to obtain trajectories for densities up to 3 ped/m^2 [7].

It seems that depth field measurements protect more the anonymity of people than regular RGB images. We wanted nevertheless to test whether pixelization was needed to ensure sufficiently the anonymity of tracked people. Our underlying a priori assumption was that more pixelization would make it more difficult to recognize a particular person, and we wanted to establish which level of pixelization would guarantee anonymity, in the spirit of [8].

2 Experiment description

2.1 Data collection

An experiment was performed, in which various members of our laboratory were invited to walk under a Kinect[1] camera hanging on the ceiling. Part of the region seen by the camera was a stair, part was a flat floor (see Fig. 1). The distance between the camera and the floor was 282 cm. Volunteers were invited to go once upstairs and once downstairs. Only depth-field images were recorded [2].

We collected in total passages from 25 volunteers under the Kinect camera. One of the videos, in which the volunteer was hopping on one foot throughout his passage, was discarded.

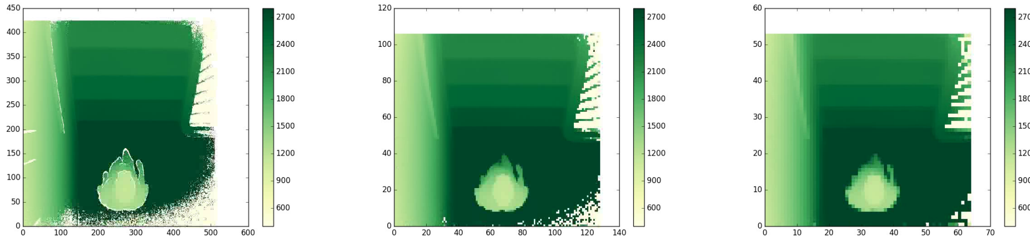
Half of the volunteers were permanent staff (so people that know each other quite well), one quarter were regular visitor or PhD students, and one quarter were trainees (most of them not known from other teams).

Volunteers arrived at random times, and were mostly ignoring who else took part in the experiment, except in some cases for the 2 or 3 others arriving at the same time.

[1] We used a Xbox One Kinect camera, model number 1520. The depth pictures provided by the camera are 512×424 pixels.

[2] Kinect cameras allow to record not only depth-field images, but also regular RGB images.

Figure 1 Experimental place. The camera was hanging from the upper floor.



(a) (b) (c)

Figure 2 Image with 3 levels of pixelization: (a) $q = 1$ (original picture), (b) $q = 4$, (c) $q = 8$. Color bar gives the depth in mm. Undefined pixels are white.

2.2 Pixelization

Some post-treatment was then applied to the images in order to group the pixels. Various degrees of aggregation were used, characterized by a number q. For each value of q, the pixels in a square of size qxq were grouped together into a single pixel with averaged depth value. Fig. 2 shows an example of an image with various degrees of pixelization. Only 3 levels of pixelization are shown but actually q could take any integer value between 1 and 8.

2.3 Recognition test

The aim of the experiment was to test whether volunteers could identify themselves[3] among the set of collected videos. A graphical interface was developed so that volunteers could explore freely the set of videos, presented by the software in a random order (see Fig. 3).

As our prior hypothesis was that more pixelization would make it more difficult to

[3]As participants did not necessarily know each other, we asked them to recognize themselves, as it would put everybody on an equal footing.

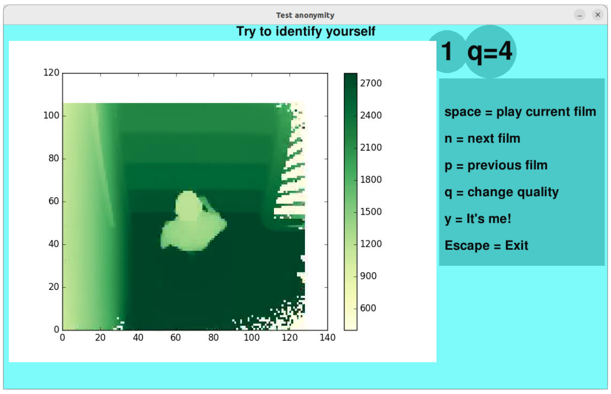


Figure 3 Graphics interface allowing to test the impact of various degrees of pixelization onto the ability to identify oneself. The software is available at [9].

recognize oneself, videos were first presented with the higher degree of pixelization ($q = 8$).

If the person could not recognize themself, or made a wrong choice, they could decrease the pixelization, and explore again the videos, until identification succeeds or until the lowest pixelization was reached.

We also took note of the remarks done by the volunteers during this identification task, as we wanted also to identify the criteria that people were using, and the cognitive process at work.

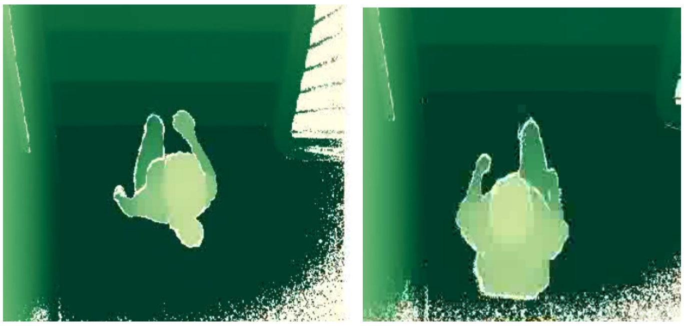
As the identification test occurred several months after the videos were collected, some volunteers had left the place, or some didn't have the availability to perform the test (going through the 24 videos of duration about 5s each, takes between 2 and 3 min for each quality level). Besides, we chose to perform the test with some volunteers with rather neutral appearance. Indeed, some characteristics like carrying a cup of tea (see 4a), wearing a bun (again 4a), or carrying a backpack (4b) would indeed have biased the recognition process by making it much easier. Still these specific cases (6 among 24) were kept in the set of presented videos. As a result, only 6 members of the laboratory performed the identification test.

3 Results analysis

Surprisingly, and at odds with our prior hypothesis, it turns out that the degree of pixelization was not a relevant factor.

Instead, we found that there were mostly two types of testers: Four of them could identify themselves even with a high degree of pixelization (3 with $q = 8$, 1 with $q = 6$), while 2 others could never identify them, whatever the quality of the image (the pixelization levels that they explored ranged from $q = 8$ to 1).

The two persons who could not identify themselves were convinced from the beginning that they would not manage ("It's impossible", "Impossible to recognize oneself, there are no distinctive details").



(a) (b)

Figure 4 Examples of specific features: (a) person with a bun and carrying a cup; (b) person with a backpack.

More surprisingly, the same doubt was expressed by those who identified themselves successfully. For example, one of them first claimed : "Impossible to recognize oneself". Once this person had nevertheless succeeded, they explained that they had proceeded by elimination of people who had distinctive characteristics, and then their final choice was led, they thought, by the speed of walk as main criterion.

Another one mentioned some motion of the arm without which they would not have recognized themself.

A third one claimed that the recognition was rather instinctive, and that they could not explain it. When we asked them to make nevertheless some hypothesis, they mentioned the rather elongated silhouette and the swaying of the head.

The last one mentioned that the most striking feature was "the way people walk".

4 Conclusion

The initial aim of this study was to determine the degree of pixelization that would ensure anonymity in a future experiment. This aim was implicitly implying that recognition would be more difficult with an increased pixelization level.

Our experimental results, though quite preliminary, suggest that this hypothesis is wrong, at least for the pixelization levels we explored, and that pixelization is not so much a relevant factor. By contrast, it seems that it is rather the characteristics of the *dynamics* of the walk that led to recognition, in the absence of distinctive features.

Actually, in the field of human-sensing [10], gait's characteristics was already identified as a powerful signature allowing to identify people [11]. Gait can be detected for example through Dopler [12, 13] or just sound [14]. Kinect depth-field cameras were also used, but usually from a side view [15–18]. Indeed all these approaches aim at maximizing identification, at odds with our approach.

Besides, this litterature focuses on automatic recognition. By contrast, we only tested recognition by human beings. Though we do not have enough statistics to conclude, our results seem to indicate that, while some people would spontaneously use gait dynamics, some others are unable to do it. This raises questions from the point of view of the cognitive process at work, that it would be interesting to explore further.

Our results underline also that there was a high degree of uncertainty associated with the recognition process, even when the latter was successful. Here the fact that volunteers knew they were present at the test was certainly an important information. If we had not shown all videos, and had asked "Are you present among these people?", it is possible that the results would have been quite different. The fact that participants knew a large fraction of the other participants, even if they were not aware of who was participating, could certainly also help them in the recognition process.

It would also be interesting to do again the test while asking to recognize other people than oneself. A few informal attempts seem to indicate that again, pixelization would be less relevant than the dynamical characteristics of the walk, but more systematic trials would be necessary. Comparing recognition from videos or still images could also bring interesting informations.

Another point that we didn't explore further and that may deserve more attention is the following. On fixed images, the less pixelization, the more information we get from the image. But on videos, one can see some scintillation at the boundaries of objects, in particular around the head of the pedestrians, that could hinder the interpretation of the visual signal. Though it is less clear on fixed images, it can partially be guessed on the images of Fig. 2. This sparkling is suppressed by the local averaging, and one may wonder whether videos are not even more easy to analyze for $q = 4$ than for $q = 1$.

As a conclusion, anonymity is not easy to preserve, in particular when it is connected with extra information (as here, the a priori knowledge that the person was indeed included in the set). Besides, the dynamical features mentioned above could be accessible by AI. This questions the possibility to ensure a high level of anonymity, even with depth cameras. Computing trajectories on the fly would allow not to store the images - but then no a posteriori check can be performed.

The python code of the graphical interface allowing the test of anonymity is available at [9].

Acknowledgements We thank Alexandre Nicolas for introducing us to the use of Kinect Cameras.

Ethics Statement This study was performed in the frame of a larger project implying pedestrian tracking on the campus of our university, project for which we had the official agreement of the university.

The test reported here was a preliminary study necessary to assess the degree of anonymity that we could ensure, requested by the authorities in charge of data protection, in order to determine the type of administrative declaration that was requested by our project of tracking in natural settings.

Volunteers were members of our laboratory. They signed an agreement before performing the walk under the Kinect camera.

Author Contributions S. Habet and C. Appert-Rolland realized the experimental data collection. C. Appert-Rolland developed the software and conducted the survey.

References

[1] Haghani, M., Sarvi, M.: Crowd behaviour and motion: Empirical methods. Transp. Res. Part B: Method. **107**, 253–294 (2018). doi:10.1016/j.trb.2017.06.017

[2] Shi, X., Ye, Z., Shiwakoti, N., Grembek, O.: State-of-the-art review on empirical data collection for external governed pedestrians complex movement. Journal of Advanced Transportation **2018**, 1063043 (2018). doi:10.1155/2018/1063043

[3] Feng, Y., Duives, D., Daamen, W., Hoogendoorn, S.: Data collection methods for studying pedestrian behaviour: A systematic review. Building and Environment **187**, 107329 (2021). doi:10.1016/j.buildenv.2020.107329

[4] Seer, S., Braendle, N., Ratti, C.: Kinects and human kinetics: A new approach for studying pedestrian behavior. Transp. Res. Part C-Emerging Technologies **48**, 212–228 (2014). doi:10.1016/j.trc.2014.08.012

[5] Corbetta, A., Bruno, L., Muntean, A., Toschi, F.: High statistics measurements of pedestrian dynamics. Transp. Res. Procedia **2**, 96–104 (2014). doi:10.1016/j.trpro.2014.09.013

[6] Corbetta, A., Meeusen, J., Lee, C.M., Toschi, F.: Continuous measurements of real-life bidirectional pedestrian flows on a wide walkway. In: Song, W., Ma, J., Fu, L. (eds.) Proceedings of Pedestrian and Evacuation Dynamics 2016, pp. p.18–24. Collective Dynamics A11 (2016). doi:10.17815/CD.2016.11

[7] Kroneman, W., Corbetta, A., Toschi, F.: Accurate pedestrian localization in overhead depth images via height-augmented HOG. Collective Dynamics **5**, 33–40 (2020). doi:10.17815/CD.2020.30

[8] Birnstill, P., Ren, D., Beyerer, J.: A user study on anonymization techniques for smart video surveillance. In: 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1–6 (2015). doi:10.1109/AVSS.2015.7301805

[9] Appert-Rolland, C.: sharingsomecode/test-anonymity: testanonymity.py (2024). doi:10.5281/zenodo.11068943

[10] Teixeira, T., Dublon, G., Savvides, A.: A survey of human-sensing: Methods for detecting presence, count, location, track, and identity. ENALAB Technical Report 09-2010 (2010). URL https://thiagot.com/papers/teixeira_techrep10_survey_of_human_sensing.pdf

[11] Wan, C., Wang, L., Phoha, V.: A survey on gait recognition. ACM Comput. Surveys **51**, article 89 (2018). doi:10.1145/3230633

[12] Geisheimer, J., Greneker, E., Marshall, W.: High-resolution Doppler model of the human gait. In: Proceedings, Radar Sensor Technology and Data Visualization, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 4744, pp. 8–18 (2002). doi:10.1117/12.488286

[13] Kalgaonkar, K., Raj, B.: Acoustic doppler sonar for gait recogination. In: Proceedings of the 2007 IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 27–32. IEEE Computer Society Washington, DC, USA (2007)

[14] Geiger, J., Kneißl, M., Schuller, B., Rigoll, G.: Acoustic gait-based person identification using hidden markov models. In: Proceedings of the 2014 Workshop on Mapping Personality Traits Challenge and Workshop, MAPTRAITS '14, p. 25–30. Association for Computing Machinery, New York, NY, USA (2014). doi:10.1145/2668024.2668027

[15] Staranowicz, A., Brown, G., Mariottini, G.L.: Evaluating the accuracy of a mobile Kinect-based gait-monitoring system for fall prediction. In: Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments, PETRA '13, p. article number 57. Association for Computing Machinery,New York, NY, USA (2013). doi:10.1145/2504335.2504396

[16] Schmitz, A., Ye, M., Shapiro, R., Yang, R., Noehren, B.: Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system. J. Biomech. **47**, 587–591 (2014). doi:10.1016/j.jbiomech.2013.11.031

[17] Dikovski, B., Madjarov, G., Gjorgjevikj, D.: Evaluation of different feature sets for gait recognition using skeletal data from kinect. In: Proceedings of the 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 1304–1308 (2014). doi:10.1109/MIPRO.2014.6859769

[18] Roy, G., Bhuiya, A., Mukherjee, A., Bhaumik, S.: Kinect camera based gait data recording and analysis for assistive robotics-an alternative to goniometer based measurement technique. In: Proceedings of International Conference on Robotics and Smart Manufacturing (RoSMa2018), *Procedia Computer Science*, vol. 133, pp. 763–771 (2018). doi:10.1016/j.procs.2018.07.121