

Methodology for Generating Individualized Trajectories from Experiments

Wolfgang Mehner and Maik Boltes and Armin Seyfried

Abstract Traffic research has reached a point where trajectories are available for microscopic analysis. The next step will be trajectories which are connected to human factors, i.e information about the agent. The first step in pedestrian dynamics has been done using video recordings to generate precise trajectories. We go one step further and present two experiments for which ID markers are used to produce individualized trajectories: a large-scale experiment on pedestrian dynamics and an experiment on single-file bicycle traffic. The camera set-up has to be carefully chosen when using ID markers. It has to facilitate reading out the markers, while at the same time being able to capture the whole experiment. We propose two set-ups to address this problem and report on experiments conducted with these set-ups.

1 Introduction

Laboratory experiments are a valuable tool when conducting research into traffic or pedestrian dynamics. The data can be analysed to uncover effects which are important for modelling, and it can be used to validate simulations.

When conducting such experiments, one always has at least two conflicting requirements. On the one hand, the observed subjects have to be detected reliably, which requires a large focal length, i.e. a larger zoom factor, of the used camera. On the other hand, one wants to observe as much of the experiment as possible, which requires a small focal length. When using ID markers, this situation gets worse, since the required resolution is most likely higher than with other markers,

Wolfgang Mehner

Visual Computing Institute, RWTH Aachen University, e-mail: mehner@vision.rwth-aachen.de

Maik Boltes

Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, e-mail: m.boltes@fz-juelich.de

Armin Seyfried

Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, e-mail: a.seyfried@fz-juelich.de

e.g. coloured caps. Employing multiple cameras may be the only viable solution when using ID markers, which will make it more difficult to calibrate the cameras. As a result, camera set-ups for such experiments have to be carefully engineered.

In the following, we present two experiments using ID markers and the solutions we have chosen in order to address the above mentioned problems. In detail, those consist of: (1) A large-scale experiment with pedestrians, where the markers are read out everywhere, which however requires a complex camera system. (Section 3) (2) An experiment on single file bicycle traffic, where the markers are only read out in one camera view. Using overlapping fields of view, the IDs can still be transferred to trajectories generated from detections in other views. (Section 4)

2 Related Work

Characteristics of pedestrians have been the subject of experiments before. A motion capture system was used by Jelić et al. [5], while otherwise video recordings seem to be a common choice for data acquisition. Boltes et al. [1] use colour-coded markers to associate detected trajectories to the pedestrians' heights, which is important for a precise localisation on the ground plane. Daamen et al. [3] investigate the influence of different classes of pedestrians (children, elderly, disabled) on the capacity of a door, and to that end equipped participants with caps of different colour, indicating their class. Individualized trajectories are used by Bukáček et al. [2] to link the behaviour of the participants across different runs of an experiment. Stuart et al. [11] use ID markers in experiments with individuals with disabilities, to investigate their impact on pedestrian dynamics. However, the ID markers used in both these experiments have the undesirable property that they protrude over the participants' heads, which makes experimentation at high densities more difficult.

Bicycle traffic has been investigated before by Navin [9], who used video to analyse single-file bicycle traffic on an oval track. Experiments similar to ours were performed by Rui et al. [10], who to the best of your knowledge do not produce individualized trajectories. Zhang et al. [15] used video recordings to obtain fundamental diagrams, and also focus on properties of bicycles and electric bicycles. The dynamics during bicycle races were investigated by Trenchard et al. [13]. Single-lane experiments for cars, in the same spirit as our bicycle experiments, were performed by Nakayama et al. [8] and Tadaki et al. [12].

3 Pedestrian Experiments

The pedestrian experiments were performed as part of the project BaSiGo in Düsseldorf, Germany, in 2013. The aim was to investigate pedestrian flows in various geometries, such as corridors and different intersections, especially at very high densities. For example, corridors of 4 and 5 m width and a length of over 10 m were

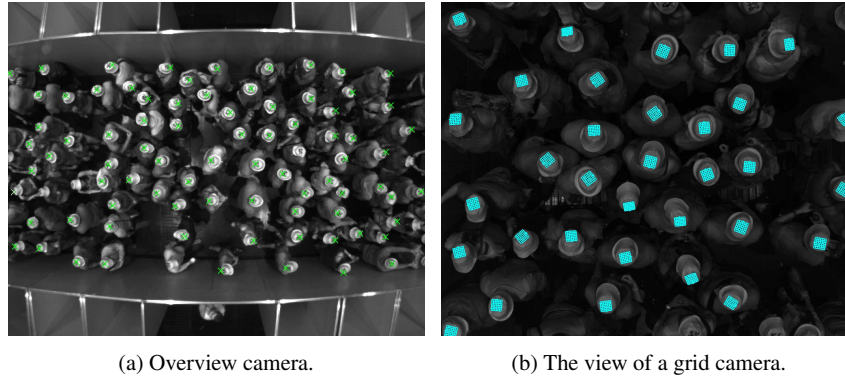


Fig. 1: Camera views of the pedestrian experiment with detected markers.

set up (see Figure 1a). Therefore a huge number of participants is required to reach a steady state at high densities in such large geometries.

This requires a very large number of ID markers and powerful detection and tracking algorithms, already presented in [7]. However, to guarantee success a carefully engineered camera set-up and calibration procedure is equally important. The cameras used to record the experiments have to be able to read out the markers, while at the same time covering the $10m \times 10m$ area used for the experiments. The ID markers also provide information on the head rotation, therefore the markers are read out in the entry area, to get the full information (position and rotation) everywhere. While the universal availability of the IDs simplifies the task of linking the trajectories across the different camera views, this set-up makes the calibration of the camera grid more difficult.

3.1 Camera Grid

The experiments were recorded with a grid of 6×4 overhead cameras (see Figure 1b), mounted $7.5m$ above the floor. Each camera covered an area of a little over $2.5m$ ($= \frac{10m}{4}$) \times $1.67m$ ($= \frac{10m}{6}$) measured at $2m$ above the floor, the maximum presumed height of the participants. With this set-up, the markers can be read out everywhere in the observed area. A little overlap between the views of neighbouring cameras makes it possible to “hand over” trajectories without losing them in blind spots. Given the resulting small opening angles of the cameras, an image resolution of 1280×1024 pixels turned out to be sufficient to read out the markers. The markers had a size of $8.5cm \times 8.5cm$ and fitted comfortably on a normal hat, making them usable at high densities of up to 10 persons per square meter.

Since the set-up uses monocular cameras, the height of a person is needed to compute the location on the ground plane from the marker detected on the head

(compare the discussion in [1]). The camera grid included an additional camera for computing the height of the participant. This camera's field of view overlapped with other cameras entirely. If a marker is detected in this camera and an overlapping view, then the height can be computed via triangulation. The computed heights can be saved in a list and retrieved via the marker ID, and can thus be used everywhere in the grid to compute the exact locations of the participants.

The intrinsic calibration of the cameras (focal length, lens distortion, ...) is done using a standard technique [16]. Bundle adjustment [4] is used to obtain the external calibration (positions and rotations of the cameras) of the grid. To estimate these parameters, a minimisation problem is set up. The positions of identifiable objects in the scene are projected into the view of each camera which can see the objects. For each object, this yields the image position of the projection, which is compared with the known image position for this object, which must be obtained by a different method, e.g. manual annotation. The difference in these two positions is called the reprojection error. The average reprojection error for all objects and cameras is minimized. This yields the camera calibration parameters, because these parameters influence the projection process.

One challenge with this approach based on reprojection errors is that the estimation of the parameters is unreliable in some regards. For example, the camera can be moved along the optical axis (which in our set-up corresponds to the height of the overhead cameras above ground) without much change to the reprojection error. This is a bad property, given that measurement errors (imprecise annotations of the image positions) also figure into the minimisation problem. Other properties of the set-up, the small opening angles of the cameras and the small overlaps, further complicate the calibration.

3.2 Discussion

The described set-up is able to deliver detailed and precise individualized trajectories, including the rotation of the heads, provided for the entire area used for experimentation. While this is ideal with regards to the type of data one would want for investigating pedestrian dynamics, this set-up requires a lot of effort to design and use correctly. This effort stems from the amount and the processing of the produced image data, as well as the calibration and synchronisation of the large number of cameras. All these problems have to be addressed in the design of the recording system, which requires expertise in a number of computer vision topics. On the other hand, given the size of the experiments, over 2000 participants in four days of experimentation, this effort seems justified. The only limitation of a design like this is that it can probably not be extended to cover a larger area at reasonable cost.

In conclusion, this set-up puts its focus on the utility of the produced data without much regard for the resulting effort. For experiments with a lower number of participants, easier and faster solutions would be preferable.

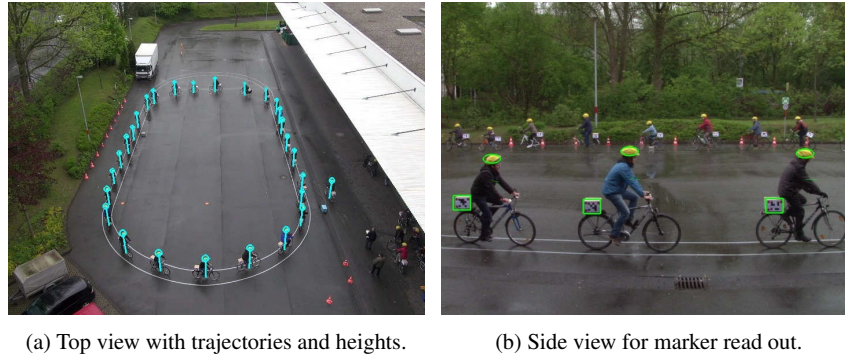


Fig. 2: Camera views of the bicycle experiment with detected helmets and markers.

4 Bicycle Experiments

The bicycle experiments were performed in Wuppertal, Germany, in 2012. The single-file set-up allows to compare the behaviour of the cyclists to pedestrians and cars [14], for which such experiments have been performed before.

One camera (“top view”, see Figure 2a) was set up to observe the entire experiment from above. However a top-down perspective could not be achieved easily given the surroundings. The bicycles were equipped with markers, which could however not be read in the top view. Therefore, an additional camera (“side view”, see Figure 2b) was set up to identify the drivers. They additionally wore yellow helmets, which can be detected in both views, and thus be used to associate the trajectories generated from both views. Additionally, the drivers are visible in the second row of the side view, which could be used for further measurements.

4.1 Detection and Tracking

The helmets in the top view are detected by finding local extrema in the Laplacian scale-space, followed by suppressing low-texture areas (compare [6]), and colour filtering. The helmets in the front of the side view are found by thresholding the HSV colour-value of each pixel and then performing connected component labelling. The marker boxes are found by detecting the red dots using the same technique as for the helmets in the top view. Three red dots can be associated to yield the position of the marker grid, which is then binarized and read out. The tracking procedure uses Kalman filtering to deal with missing detections, but is otherwise very simple because of the relatively reliable detections.

It is not possible to read out the markers in the top view, or in the back of the side view. Therefore, only the trajectories in the front of the side view have IDs attached to them, while the others are “anonymous”.

4.2 Associating Anonymous Trajectories

In the pedestrian experiment, we could associate trajectories across different views using their IDs. Here, this is not possible anymore, since only the trajectories in the front of the side view have IDs attached to them. We need a more advanced technique for associating the remaining “anonymous” trajectories with the ones in the front view. This will then allow us to transfer the IDs to them, and thereby other information, such as the height of the helmets above the ground.

We make use of the fundamental matrix [4]. Given points P_i in world coordinates and their projection into two different camera views, $p_i^{(1)} = (x_i^{(1)}, y_i^{(1)}) = \text{project}^{(1)}(P_i)$ and $p_i^{(2)} = (x_i^{(2)}, y_i^{(2)}) = \text{project}^{(2)}(P_i)$, we can find a fundamental matrix F , such that:

$$(x_i^{(1)}, y_i^{(1)}, 1) \cdot F \cdot (x_i^{(2)}, y_i^{(2)}, 1)^T = 0, \text{ for all } i \quad (1)$$

Given eight pairs of corresponding points, this matrix is unique up to scale. More points can be used to estimate it in a least-squares fashion. For more details see [4]. Note that we only need corresponding pixel coordinates in both camera views. The 3D geometry of the scene can thus be bypassed.

Given a point in the first camera view, with unknown world coordinates, we are still able to constrain its position in the second view, given F :

$$(x_i^{(1)}, y_i^{(1)}, 1) \cdot F = (a, b, c) \text{ and } (a, b, c) \cdot (x_i^{(2)}, y_i^{(2)}, 1)^T = 0 \quad (2)$$

It can be seen that a point in the first view parametrizes a line in the second view in implicit notation, via the fundamental matrix.

The relationship given by the fundamental matrix can be used to associate trajectories tracking the same physical object in two views. Given trajectories in two views $P^{(v)}(t) = (x^{(v)}(t), y^{(v)}(t))$ and a number of overlapping frames T , we can compute an average distance between the lines $(a(t), b(t), c(t)) = (x^{(1)}(t), y^{(1)}(t), 1) \cdot F$ and the points $(x^{(2)}(t), y^{(2)}(t))$. In each step, the line parametrisation has to be normalized in order to compute meaningful distances across different frames:

$$(a_i(t), b_i(t), c_i(t)) = (x_i^{(1)}(t), y_i^{(1)}(t), 1) \cdot F \quad (3)$$

$$D_{i,j} = \frac{1}{|T|} \sum_{t \in T} \left| (a_i(t)^2 + b_i(t)^2)^{-\frac{1}{2}} \cdot (a_i(t), b_i(t), c_i(t)) \cdot (x_j^{(2)}(t), y_j^{(2)}(t), 1)^T \right| \quad (4)$$

This yields a distance $D_{i,j}$ between every pair $P_i^{(1)}(t)$ and $P_j^{(2)}(t)$ of trajectories. We only associate them if:

$$j = \operatorname{argmin}_h D_{i,h} \text{ and } i = \operatorname{argmin}_h D_{h,j} \quad (5)$$

This is preferable to a greedy strategy of simply picking the closest match for one trajectory and then moving on to the next. Our strategy leads to a unique solution, and helps to suppress associations with meaningless trajectories resulting from

false-positive detections. More advanced approaches, such as the Hungarian algorithm, are possible as well.

4.3 Evaluation

We provide some numbers to give an impression of the performance of the automated trajectory generation for the bicycle experiments. We manually cleaned up the trajectories in the top view for one of the experiments and compared them to the trajectories reported by the detection and tracking algorithm. There were 33 participants in the experiment. In 9400 frames, 228888 positions needed to be reported. The algorithm only failed to report 254 (false negatives), but missed no trajectory completely. There were 2082 additional detections reported, belonging to 19 trajectories made up of false positive detections. While these numbers show that the system does not produce perfect results, the performance is high enough that a manual clean up of the results can be done with very little effort.

4.4 Discussion

The linking of the cameras is much easier to do with this approach, since the estimation of the fundamental matrix takes less effort than the full 3D calibration in the pedestrian experiments. Additionally, the required precision is easier to achieve. The external calibration of the camera which actually reports the measurements is still required, of course. Overall, this set-up, with one camera to view the entire experiment and one camera to read the ID markers, seems better suited for smaller experiments. A few prerequisites have to be met, however. For example all participants have to pass in front of the camera reading out the IDs at least once per experimental run. The density in the field of view of this camera should also not be too high, or the association of the trajectories will get more difficult.

5 Conclusion

We have presented two experiments we conducted, and discussed the requirements and the resulting solutions for the data capturing. Both experiments included information that was obtained from one camera in the grid and then transferred to the trajectories of the entire experiment. We showed two techniques to accomplish this, once using the marker IDs themselves, once by exploiting the interaction of trajectories and the camera geometry.

Acknowledgements This study was performed within the project BaSiGo (Bausteine für die Sicherheit von Großveranstaltungen, Safety and Security Modules for Large Public Events) funded by the Federal Ministry of Education and Research (BMBF) Program on “Research for Civil Security – Protecting and Saving Human Life”. Wolfgang Mehner was partly funded by the ERC Starting Grant Project CV-SUPER (ERC-2012-StG-307432).

References

1. Boltes, M., Seyfried, A.: Collecting pedestrian trajectories. *Neurocomputing* **100** (2013)
2. Bukáček, M., Hrabák, P., Krbálek, M.: Experimental study of phase transition in pedestrian flow. *Transportation Research Procedia* **2** (2014)
3. Daamen, W., Hoogendoorn, S.: Capacity of doors during evacuation conditions. *Procedia Engineering* **3** (2010)
4. Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*. Cambridge University Press (2003)
5. Jelić, A., Appert-Rolland, C., Lemercier, S., Pettré, J.: Properties of pedestrians walking in line: Fundamental diagrams. *Physical Review E* **85**(3) (2012)
6. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International journal of computer vision* **60**(2) (2004)
7. Mehner, W., Boltes, M., Mathias, M., Leibe, B.: Robust marker-based tracking for measuring crowd dynamics. In: *Proceedings of the 10th International Conference on Computer Vision Systems* (2015)
8. Nakayama, A., Fukui, M., Kikuchi, M., Hasebe, K., Nishinari, K., Sugiyama, Y., Tadaki, S., Yukawa, S.: Metastability in the formation of an experimental traffic jam. *New Journal of Physics* **11**(8) (2009)
9. Navin, F.P.: Bicycle traffic flow characteristics: experimental results and comparisons. *ITE journal* **64**(3) (1994)
10. Rui, J., Mao-Bin, H., Qing-Song, W., Song, S.W.G.: Experimental feature of bicycle flow and its modeling. arxiv.org/abs/1411.1136 (2014)
11. Stuart, D., Christensen, K., Chen, A., Kim, Y., Chen, Y.: Utilizing augmented reality technology for crowd pedestrian analysis involving individuals with disabilities. In: *Proceedings of the ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (2013)
12. Tadaki, S., Kikuchi, M., Fukui, M., Nakayama, A., Nishinari, K., Shibata, A., Sugiyama, Y., Yosida, T., Yukawa, S.: Phase transition in traffic jam experiment on a circuit. *New Journal of Physics* **15**(10) (2013)
13. Trenchard, H., Richardson, A., Ratamero, E., Perc, M.: Collective behavior and the identification of phases in bicycle pelotons. *Physica A: Statistical Mechanics and its Applications* **405** (2014)
14. Zhang, J., Mehner, W., Holl, S., Boltes, M., Andresen, E., Schadschneider, A., Seyfried, A.: Universal flow-density relation of single-file bicycle, pedestrian and car motion. *Physics Letters A* **378**(44) (2014)
15. Zhang, S., Ren, G., Yang, R.: Simulation model of speed–density characteristics for mixed bicycle flow – comparison between cellular automata model and gas dynamics model. *Physica A: Statistical Mechanics and its Applications* **392**(20) (2013)
16. Zhang, Z.: A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **22**(11) (2000)