Shoulder rotation measurement in camera and 3D motion capturing data

Ann Katrin Boomers and Maik Boltes

Abstract The individual movement of pedestrians and their body parts, as for example shoulders, is of great interest to understand body movement and interactions and thus to improve pedestrian models. Nearly all laboratory experiments in pedestrian dynamics use camera data to obtain trajectories. A perpendicular top view of the camera does not only allow to extract the head position but also data of upper body segments. The detection is more reliable if shoulders are tagged with markers and for low densities of people. In this study a head-shoulder model is used to assign coloured shoulder markers to a person. The location of a marker is predicted by taking head position, basic body dimensions, movement direction and camera angle into account. It is implemented as a new feature in the software PeTrack. This paper shows a comparison of shoulder rotation measurements obtained from 3D motion capturing systems (Xsens) with those from camera data using the newly introduced model and detection technique. Detection rates and limits of the camera-based rotation measurement are shown and implications are given for the future application at high densities in crowds.

Keywords: Shoulder detection, Rotation measurement, laboratory experiments

1 Introduction

Understanding pedestrian dynamics is crucial for developing and improving models and safety concepts that help to design proper functioning and safe facilities. Well calibrated and reproducible controlled experiments serve to develop model ideas as well as provide high quality data to calibrate and validate a model. Most pedestrian experiments use camera recordings for

Ann Katrin Boomers · Maik Boltes
Institute for Advanced Simulation 7: Civil Safety Research, Forschungszentrum Jülich, 52428 Jülich, Germany, e-mail: a.boomers@fz-juelich.de , m.boltes@fz- juelich.de

K. R. Rao et al. (eds.), *Traffic and Granular Flow '22* , Lecture Notes in Civil Engineering 443, https://doi.org/10.1007/978-981-99-7976-9_5

documentation and qualitative analysis as well as for extraction of head trajectories. Different detection methods with different degrees of accuracy are commonly used. All methods have in common, that the level of extractable data is limited to what information is visible for the camera. Which is based on the fact that the closer people stand to each other the less information can be obtained from camera data due to occlusion or non distinguishable contours.

Therefore experimental data of rotation, that is becoming of larger interest for improving the description of interactions and space requirements, are rare and often (but not always) limited to people being equipped with inertial sensors. A big advantage of inertial sensors compared to camera data is, that they allow to obtain information of people in a crowd that is not visible for the camera and that they can easily be strapped to body parts of interest, e.g shoulders, hip or head. However, while showing high accuracy in angular and acceleration data, positional data are prone to cumulative error due to the relative measurements of inertial measuring units [10]. Some 3D-motion-capturing (mocap) systems combine multiple inertial sensors attached to independently movable segments of a body and hence provide position for several body parts in time [9]. A hybrid system is described in [3] yielding high accuracy motion data of equipped persons in a crowd. Nevertheless monetary resources most likely limit the equipment of multiple pedestrians to small groups and leave the provision of rotational data for large groups for model validation purposes as a remaining issue.

In this paper we present a newly implemented method for detecting marked shoulders in laboratory experiments from camera data with the software PeTrack [2, 4]. The method is based on color blob detection and an underlying simple shoulder-head model, described in Section 2. An application of the method is presented in Section 3.

2 Methods

Two data sets of the same experiment are presented and compared in this paper. Data set one consists of position data for head and shoulders from hybrid mocap systems. Data set two consists of head and shoulder points extracted from camera data with a newly introduced shoulder-head model in the software PeTrack. The hybrid system is shortly resumed hereafter whereas the shoulder-head model is explained in more detail in the following paragraph.

The mocap system Xsens [9] is used in this study. The system itself is a relative system without reference in global space. Head trajectory data from the camera system show a maximum positional error of 1.40 cm (assuming fixed vertical head-to-camera distance) for the experiment described in Section 3 in global space. A combination of both methods, by mapping the top of head trajectory from the motion capturing system to the head trajectory of the camera system, yields dynamic data of the whole body with an accuracy of 0.86 cm with respect to the camera trajectory. The mechanism is described in more detail in [3]. As the hybrid system has a very small positional error, it is taken as the ground truth to compare and assess the value of the newly introduced shoulder detection method from camera data.

Shoulder-head model There are two types of head positions in PeTrack, one being a head labeled as being 'recognized' based on chosen recognition methods (e.g. code marker or color blob) and the other being labeled as 'tracked', where the head could not be recognized but is tracked based on the structure after a previous recognition. The shoulder detection is only possible when a head is recognized and is implemented in [1]:

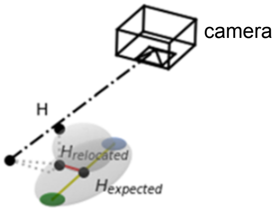

Fig. 1 Scetch showing shoulder-head model for a tilted head. Point H, $H_{relocated}$ and $H_{expected}$ as well as camera angle are visualized. Blue and green patches denote right and left shoulder markers.

Step 1: Whenever a head is recognized, the position of the head H is known in real-world coordinates (cf. Fig. 1). Point H is then lowered to shoulder height, taking the angle between head and camera into account, assuming a height difference between top of head and shoulder of 32 cm [7]. $H_{relocated}$ is the position of the head relocated to shoulder height.

Step 2: A square shaped search region with $H_{relocated}$ as the center is created for each recognized head in each frame (with side length s_{max} = 55 cm referring to the maximum expected shoulder width from [7] plus additional buffer, s_{min} = 25 cm resp.). Within the search region a color blob detection is performed storing possible candidates for left and right shoulder distinguished by color.

Step 3: Between the candidates each 'left shoulder candidate' is paired with each 'right shoulder candidate' and two parameters are calculated: First, the distance between the two candidates (ref. actual shoulder width $s_{candidate}$) and second, the expected head position at shoulder height ($H_{expected}$, Fig. 1). The later is assumed to be at the middle between both shoulder candidates. The calculated parameters for each pair are checked for the following criteria:

Criterion 1: $s_{min} < s_{candidate} < s_{max}$.

Criterion 2: $||H_{relocated} - H_{expected}|| < 21$ cm (one head diameter [7], red line in Fig. 1).

Criterion 3: If criterion one and two apply to more than one candidate pair, the pair where criterion 2 is smallest is chosen.

If the criteria are not met, no shoulder is assigned. The implementation does not predict the location of the second shoulder, if only one shoulder is detected. If the person is keeping the head straight up the deviation between 'expected head position at shoulder height' and 'the position of the head relocated to shoulder height' is expected to be small. If the person lowers its head, the deviation is expected to be bigger, but never larger that one head diameter.

Shoulder rotation Shoulder rotation γ is the angle between the shoulderline and the x-axis, where shoulderline is the vector pointing from the right to the left shoulder. The shoulder rotation is calculated for both methods (camera and mocap) separately (γ_{camera}, γ_{mocap}). Rates of change of the shoulder rotation ($d\gamma/dt$) are calculated as the differences between two consecutive elements of γ.

3 Application

Rotational data of the upper body is of interest as it alters the space requirement of people and therefor is likely to have an effect on dynamics inside crowds. Possibly, a better knowledge of the turning behavior in reality as an additional collision avoidance strategy, could reduce the occurrence of deadlocks in simulations. However, as it is very expensive to equip hundreds of people with full body mocap systems, to study crowds dynamics, other approaches are valuable. Therefore we developed a new shoulder detection method from camera data. To evaluate the performance of the method two parameters have been chosen: the difference of the angle between shoulderline and x-axis and the amount of data points with non-plausible rates of change of the shoulder rotation.

Experiment The experiments presented here were performed at Forschungszentrum Jülich, Germany in 2020. The number of participant was 13 and consisted of staff members. A bottleneck of 0.2 m length and a variable width w was installed. The bottleneck width varied between 0.4 m and 1.0 m with an increment of 0.2 m. The participants walked all by oneself from seven angles starting 4 m in front of the bottleneck. The angles included straight walking 0°, ±30°, ±60° and ±90°. One camera had been installed overhead to obtain the trajectories of the participants (cf. Fig. 2). All participants were wearing mocap systems from Xsens, orange caps with an individual Aruco Code [8] on top and colored markers on the shoulders (left shoulder: green; right shoulder: light blue; both ∅ 7 cm).

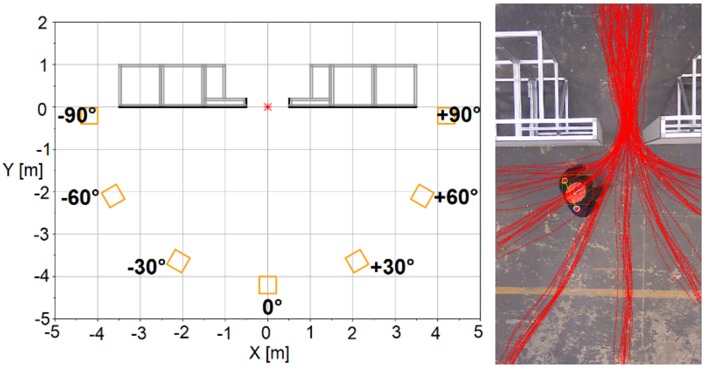

Fig. 2 (left) Schematic topview of the experimental setup for a bottleneck width of w=0.8 m. Orange squares show starting positions of pedestrians. (right) Overhead camera view of the experiment (w=0.4 m) with people wearing orange caps, code marker at the head and color markers at the shoulders. Exemplary trajectories are shown in red. Yellow line shows detected shoulders, marked with red dots.

Table 1 The table shows four experimental runs. Column 2 refers to the number of frames where a head was reported by the software PeTrack, column 3 to the number (percentage) of frames where a head was recognized by code and column 4 to the number (percentage) of frames where shoulders were assigned with respect to the number of frames with recognized heads.

width	no. of data points	no. of recognized heads	no. of assigned shoulders
0.4 m	8562	6233 (72.8 %)	4180 (67.1 %)
0.6 m	7242	4894 (67.6 %)	4004 (81.8 %)
0.8 m	6724	4118 (61.2 %)	3641 (88.4 %)
1.0 m	6532	4100 (62.8 %)	3041 (74.2 %)

Results The shoulder detection based on the simple shoulder-head model described in Section 2 is performed for four runs comprising more than 350 trajectories. Detection rates are given in Table 1 and show that shoulders are detected in 70 %-80 % of the cases were a head has been recognized. For the run with a bottleneck width of 0.4 m detection rates are lower as the participants rotated their shoulders in a way, that the head covered the frontal shoulder. As the shoulder position is only assigned if two markers are detected, none is assigned in cases of partly covered shoulders.

Deviation in rotation In this study the angle γ is calculated individually for each system as described in Section 2 and compared by their deviation. The results are shown in Figure 3, on the left side separated for different starting positions of the pedestrians and on the right side for all starting positions combined. The median for $\gamma_{camera} - \gamma_{mocap}$ is 0.06° (σ =7.09°) with a maximum error from error propagation of positional data of 4.1°. For most data points the deviation for angle calculation between both systems can be

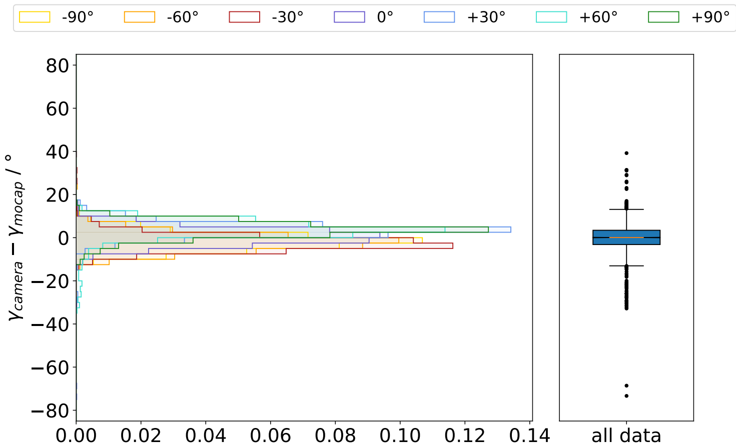

Fig. 3 Deviation of rotation angle between method from camera based detection and mocap system (left) as normalized histogram seperated for different angles of the starting point and (right) as a boxplot for all runs. Orange line denotes the median, box limits refer to inter-quartile range (IQR) of Q1 and Q3, whiskers show Q1-1.5IQR and Q3+1.5IQR respectively. Outliers are shown as black dots.

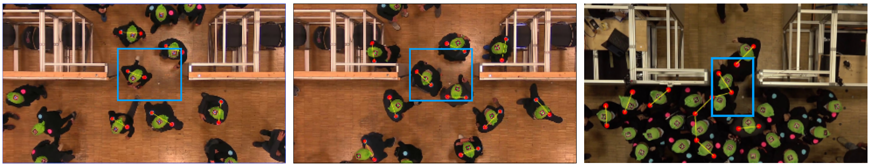

Fig. 4 Snapshots of experiment runs [5] with a mean density of (left) $1\,\mathrm{P/m^2}$, (middle) $2\,\mathrm{P/m^2}$ and (right) $6\,\mathrm{P/m^2}$. Densities are measured during steady state within the measurement area (blue rectangle) with the Voronoi method. Yellow lines, connecting red dots indicate assigned shoulders.

rated as very good. Only 0.76% of data points show a deviation of more than 13°, which refer to the whiskers is Figure 3. This means that the newly introduced method of detecting shoulders from video material is consistent with the ground truth mocap system and valuable information can be gathered within the limits of the method.

Outlook for multiple person experiments Multiple person or high density experimental data where participants are equipped with marked shoulders was gathered only recently within the CroMa project. Three exemplary runs ([5] runs: D_1_3_20-20_norm100_2, D1_1_d_4_11, 4D180_w070_l021_h1_interrupt, Fig.4) with approximate densities of ρ =1, 2 and 6 were chosen and shoulders detected with PeTrack for rates of change of the shoulder rotation to be compared between the previous experiments with single and multiple persons. Detection rates are computed as it can be assumed that outliers can

Table 2 Comparison of the mean of the single person runs with three runs with multiple persons. Column 2 refers to the percentage of frames where a head was recognized by code and column 3 where shoulders were assigned with respect to the number of frames with recognized heads. Column 4 refers to the number (percentage) of data point that exhibit the threshold of 108° per second.

experiment	recognized heads	assigned shoulders	data points with $d\gamma/dt$ > 108 and single values
single persons	66.6%	76.8%	50832/267416 (0.19%)
$\rho = 1\,P/m^2$	93.7%	66.6%	76/4823 (1.6%)
$\rho = 2\,P/m^2$	84.7%	74.6%	571/36666 (1.6%)
$\rho = 6\,P/m^2$	95.2%	31.0%	3518/29255 (12.0%)

be identified as potentially erroneous shoulder assignments. Using the mocap data from the individual experiments, the threshold was determined (3σ of the rate of change) below which the alterations in rotation occur naturally. The threshold was determined to be 1.8° per frame or 108° per second (for comparison of different frame rates). Individually occurring rotation values are added to the points exceeding the threshold due to the lack of comparability with surrounding data points. Table 2 shows, runs with a density of up to 2 P/m^2 show only very little potentially erroneous data points, that could be corrected by manually editing or interpolation without much effort. The run at ρ=6 P/m^2 shows low shoulder assignment rates combined with a high rate of potentially erroneous data points. It can therefore be summarized, that the newly introduced methods yields good results for densities up to ρ=2 P/m^2 and potentially a little higher but needs further improvements for high densities.

4 Discussion and conclusion

This paper introduced a newly implemented method for detecting the position of shoulders from camera data under laboratory conditions. The method is based on a simple head-shoulder model and implemented in the software PeTrack. The percentage of data points that the algorithm assigned shoulders for is about 70 % or more, relative to the number of data points were a head is recognized. For the future it might be worth to think about changing the code to not only detect shoulders for detected but also tracked heads. This would increase the number of data points for shoulders and possible insights into dynamics due to better statistics. One could also think of adapting the shoulder-head model for the use with stereo cameras and therefore a markerless detection. We furthermore presented exemplary applications of the new detection method by applying the method to bottleneck experiments performed in June 2020 in Jülich. We compared the results from the camera based detection method with data from the 3D motion capturing system Xsens based on the rotation angle. The shoulder rotation based on the detection of shoulder position from camera data evaluates to be really good when compared to the mocap system. The detection operates well for the multiple person test cases up to a medium density of 2 P/m^2. For high densities of >6 P/m^2 further improvements of the method are necessary. In which range between two and six persons per square meter, the limit for a good usability lies, no statement can be made yet. Future enhancements include e.g. a shoulder assignment for tracked data points and the prediction of a second shoulder based on movement direction and surrounding persons. The authors think that the new detection method will yield valuable data for shoulder rotation in crowds if keeping the limits of the method and margins of error in mind.

Acknowledgements We want to thank all colleagues of our institute for taking part in the experiment. Special thanks go to Juliane Adrian for maintaining the 3D motion capturing systems during the experiments and processing of the data after.
`Data Availability Statement:` Supplementary data of mentioned experiments will be published after finishing the content-related data analysis at the data archiv [6].
`Funding:` This research was supported by the German Federal Ministry of Education and Research (BMBF) within the project "CroMa - Crowd-Management in traffic infrastructures"; grant numbers 13N13950 and 13N14533.
`Informed Consent Statement:` Informed consent was obtained from all subjects involved in the study.

References

1. Boltes et al.: PeTrack (Version TGF2022) (2022) doi: 10.5281/zenodo.7426553
2. Boltes et al.: PeTrack (2022) doi: 10.5281/zenodo.6320753
3. Boltes et al.: A hybrid tracking system of full-body motion inside crowds. Sensors (2021) doi: 10.3390/s21062108
4. Boltes, M. and Seyfried, A.: Collecting pedestrian trajectories. Neurocomputing (2013) doi: 10.1016/j.neucom.2012.01.036
5. Boomers et al.: Pedestrian Crowd Management Experiments: A Data Guidance Paper. Collective Dynamics. (submitted), doi: 10.48550/arXiv.2303.02319
6. Forschungszentrum Jülich, Institute for Advanced Simulation: Data Archive of Experiments on Pedestrian Dynamics. http://ped.fz-juelich.de/da
7. Jurgens et al.: International Anthropometric Data for Wok-Place and Machinary Design. In: Arbeitswissenschaftliche Erkenntnisse, Forschungsergebnisse für die Praxis, Vol. 108, pp 1-12. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund, Germany (1998)
8. Romero-Ramirez et al.: Speeded up detection of squared fiducial markers. Image and Vision Computing (2018) doi: 10.1016/j.imavis.2018.05.004
9. Schepers et al.: Xsens MVN : consistent tracking of human motion using inertial sensing, Xsens Technologies Technical Report. Xsens Technologies (2018) https://www.researchgate.net/publication/324007368 . Cited 15 Mar 2022
10. Woodman, O.J.: An introduction to inertial navigation. University of Cambridge, Computer Laboratory (2007) doi: 10.48456/tr-696