Comparison of Different Methods to Determine Pedestrian Shoulder Orientation from Stereo Recordings

Deniz Kılıç*and Maik Boltes

Institute for Advanced Simulation 7: Civil Safety Research, Forschungszentrum Jülich, 52428 Jülich, Germany

Abstract This work compares imaging based methods for pedestrian orientation measurements. They are tested on overhead stereo recordings from corner-flow and bottleneck experiments with different densities. Classical computer vision based methods do not outperform using the movement direction and get worse with higher densities. A machine learning method outperforms the classical methods and does not degrade on higher densities.

Keywords Orientation, Shoulder, stereo computer vision, machine learning

Introduction

The orientation of the upper body of pedestrians can give insights into areas such as space requirements, collision avoidance and clogging or be used in further steps to e.g. determine social groups. To study the orientation of pedestrians, accurate measurements are needed. Different methods have been used in the literature such as IMUs [4], shoulder markers [1, 3], classical computer vision on depth data [2] or machine learning (ML) on depth data [5]. For large crowds, instrumentation of each individual is not feasible, so this contribution focuses on a comparison of the imaging based methods. Additionally, we compare to using the walking direction to determine the pedestrians orientation.

Method

All methods work with depth data from a stereo camera, except the movement direction based method. A stereo camera takes images from two cameras and computes the disparity between these images to determine the distance of objects from the camera. (1) The method from [2] is based on a PCA on points from the shoulder region. Said region is extracted from 3D imaging data using a segmentation algorithm based on growing clusters. Cluster growing starts with the highest points, making use of the fact that pedestrians are better separated at head height. It has been re-implemented with small changes to the segmentation method that have been observed to improve results on high density crowds by the authors. (2) Additionally, a self-developed method for the measurement of pedestrian orientation based on (1), working with constrained ellipsoid fitting instead of PCA was implemented. (3) To infer the orientation from the movement direction, the head trajectories were smoothed with a central moving average with a window size of 0.8 s. The orientation is then assumed to be orthogonal to the moving direction. (4) Finally, a neural network (NN) based approach was implemented by using the movement direction to automatically label data, albeit with large noise, similar to [5]. The NN uses patches of the depth image as input, centred around the point 32 cm below the head to counteract perspective distortion. These are generated from head trajectories. All of these methods were then applied on three runs of a corner flow experiment, with varying density ($0.54 \pm 0.14\,\mathrm{ped/m^2}$, $0.83 \pm 0.4\,\mathrm{ped/m^2}$, $1.84 \pm 0.66\,\mathrm{ped/m^2}$) and a bottleneck experiment ($1.82 \pm 1.31\,\mathrm{ped/m^2}$). Some data was manually annotated to estimate the error of the different methods.

Preliminary Results

Preliminary results for the corner flow can be seen in Figure 1. We see that method (3) using the movement direction is not outperformed by neither (1) nor (2). We also see that higher density leads

*d.kilic@fz-juelich.de

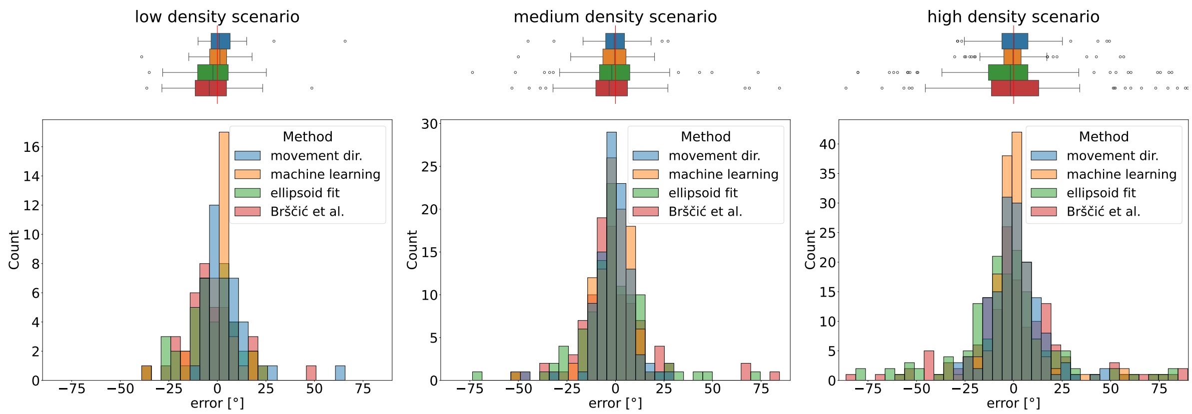

Figure 1: Error of different methods compared to manual annotation. Both the boxplot and the histogram use the same scale and colour scheme.

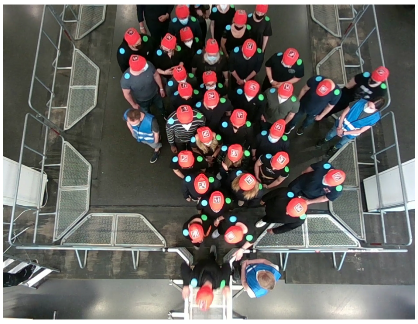

Figure 2: Frame from a RGB-D camera recording of bottleneck experiments from 10.34735/ped.2022.5 with shoulder markers.

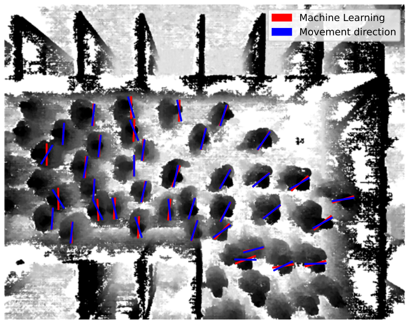

Figure 3: Comparison of ML method (4) and movement direction (3) on an example frame of the high density corner flow.

to worse results on the methods (1) and (2) and to lesser extent on method (3). The ML method (4) does not meaningfully degrade in performance with higher density and outperforms the methods (1) and (2) in all scenarios. However, the average error only gets lower than (3) in the high density scenario. It should be noted that the higher density scenario takes up a larger part of the training dataset, due to the higher number of pedestrians involved and the way of automatically creating the dataset. It should also be noted that we used laboratory experiments with an order of magnitude less training data than [5].

The bottleneck scenario a RGB-D recording (see Figure 2) including shoulder markers is available, which enables a comparison of all methods, including shoulder markers [1]. The high density is a challenge for the classical CV methods and the shoulder markers, while the low average speed mandates a different method for determining the walking direction from trajectory data.

References

[1] A. K. Boomers and M. Boltes. Shoulder Rotation Measurement in Camera and 3D Motion Capturing Data. In K. R. Rao, A. Seyfried, and A. Schadschneider, editors, *Traffic and Granular Flow '22*, pages 35–42, Singapore, 2024. Springer Nature.

[2] D. Bršćić, T. Kanda, T. Ikeda, and T. Miyashita. Person Tracking in Large Public Spaces Using 3-D Range Sensors. *IEEE Transactions on Human-Machine Systems*, 43(6):522–534, Nov. 2013.

[3] L. Luo, G. Yang, C. Chen, Z. Yuan, and Z. Fu. Empirical study on pedestrian rotation mechanisms through bottlenecks. *Physical Review E*, 111(1):014103, Jan. 2025.

[4] J. Schumann. *Utilizing Inertial Sensors as an Extension of a Camera Tracking System for Gathering Movement Data in Dense Crowds*. PhD thesis, Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag, 2022.

[5] J. Willems, A. Corbetta, V. Menkovski, and F. Toschi. Pedestrian orientation dynamics from high-fidelity measurements. *Scientific Reports*, 10(1):11653, July 2020.