Crowdflow Dataset – diluted pedestrian dynamics in the Metaforum building of Eindhoven University of Technology

Alessandro Corbetta
Department of Applied Physics
Eindhoven University of Technology
a.corbetta@tue.nl

Federico Toschi
Department of Applied Physics and
Department of Mathematics and Computer Science
Eindhoven University of Technology

2017

1 Dataset DOI and official download link

https://doi.org/10.4121/uuid:25289586-4fda-4931-8904-d63efe4aa0b8

Reference publication:
[1] A. Corbetta, C. Lee, R. Benzi, A. Muntean, F. Toschi. **Fluctuations around mean walking behaviours in diluted pedestrian flows.** Phys. Rev. E. 95, 032316, 2017

Usage and basic scripts:
https://github.com/crowdflowTUe/MF_landing_data_analysis

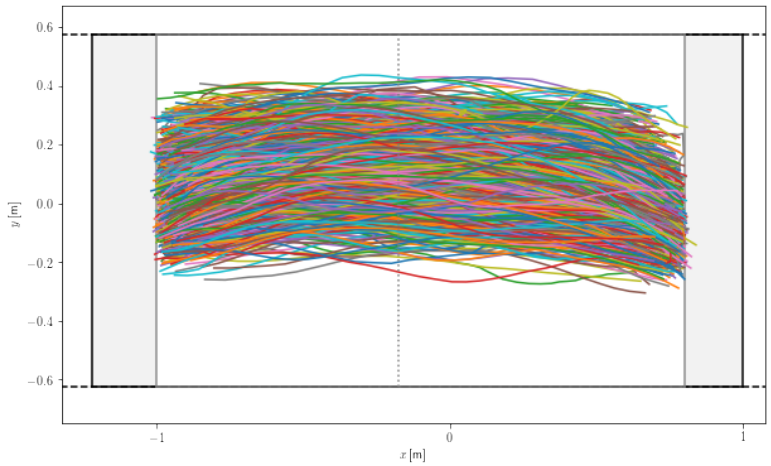

Figure 1: Random subset of about 2300 trajectories in the dataset.

2 Dataset description

This is a dataset of pedestrian trajectories recorded on a nearly 24/7 schedule in a landing in the Metaforum building at Eindhoven University of Technology. The data acquisition spanned over a year and, overall, about 250.000 trajectories have been collected. Depth imaging data has been first obtained via an overhead Microsoft Kinect sensor, then ad hoc localization algorithms and Particle Tracking Velocimetry-like tracking have been employed to estimate the trajectory of individual heads (cf. [1]). The current dataset includes 20.000 trajectories from pedestrians walking undisturbed i.e. in diluted conditions (individuals are walking alone in the facility, see Fig. 1). There are 10.000 trajectories of pedestrians crossing the landing entering from the left hand side (file: "left-to-right.ssv") and 10.000 trajectories of pedestrians entering in the opposite side (file: "right-to-left.ssv", right-left reference is given according to [1]). The purpose of the dataset is to enable ensemble analyses of diluted pedestrian motion.
The trajectories are in the following table format:

```
Pid  Rstep  X  Y  X_SG  Y_SG  U_SG  V_SG
```

where:

- `Pid`: unique identifier of a trajectory

- `Rstep`: identifier of the timestep (starts from zero, the first 5 and last 5 samples are eliminated as typically less precise)
- `X,Y`: position in Cartesian coordinates (in meters)
- `X_SG,Y_SG`: position in Cartesian coordinates after Savizky-Golay smoothing (in meters, cf. paper)
- `U_SG, V_SG`: velocity in Cartesian coordinates after Savizky-Golay smoothing (in meters per second, cf. paper).

To use the dataset please cite [1] as well as this dataset (DOI: `https://doi.org/10.4121/uuid:25289586-4fda-4931-8904-d63efe4aa0b8`).