

Crowd simulation datasets. Format description

January 30, 2018

1 Overview

Two data sets are provided for crowd-evacuation simulations in emergency situations. The first dataset (`tunnel.h5`) is the result of a simulation in a tunnel with two exits. The second dataset (`office.h5`) is the result of a simulation on a particular floor of an office building.

2 Structure of the data

The data sets are stored in HDF5 files, a hierarchical data format for numerical data. Each file contains the PNG image of the environment and locations of cameras and fires (under `/scene`, as well as a snapshot of the simulation every Δt seconds (under `/i`, where `i` represents an integer).

The coordinate system follows mathematical conventions. Coordinates $(0,0)$ corresponds to bottom-left, angles are in radians and are measured counter-clockwise from the x -axis.

3 Scene information

The dimensions of the scene are stored in the dataset under `/scene/image`. The PNG image that represents the environment is stored as a base64 encoded string in the attribute `environment`. Green areas indicate exits, black areas indicate obstacles, and grey areas indicate accessible, but preferably avoided areas. The camera settings are stored in `/scene/cameras`. Each row contains three entries (x, y, θ) and represents one camera with position (x, y) and angle θ . The cameras are placed a bit above the agents, perhaps at the ceiling.

4 Snapshots

Each snapshot contains positions and velocities of all the particle (as $n \times 1$ or $n \times 2$ datasets marked `micro`), continuum quantities (as N_x, N_y datasets marked `macro`), and the time passed since the last snapshots (as attribute `dt`). Here, n represents the maximum number of agents the simulation handles simultaneously, and $N_x \times N_y$ represents the resolution of the continuum data.

4.1 Agents

Important to note is that the simulation does not always contain n agents, even though the arrays always have size $n \times 2$. When an agent exits the simulation, his counter will be removed from the data set `map`. However, since the simulation framework works with fixed-length arrays, the

length of the arrays (and the entries) in `positions` and `velocities` is not adjusted. Similarly, if a new agent enters the simulation, he will be placed in an unused entry, his counter is registered in `map`, and from then on, the datasets `positions` and `velocities` will reflect his evacuation trajectory.

In short, in order to correctly keep track which trajectories belong to which agents, read out `map`. This is a map from the dataset indices (rows) to the (unique) agent counters. Every agent that was present during the snapshot is registered in the map, and if an index points to -1 , it is not used. *The data set `active_entries` is auxiliary and contains the same information as covered by `map`.*

4.2 Continuum

The continuum quantities are independent on the number of agents present in the scene. They represent the same information, but interpolated to a continuum. These are important to keep track of the flow of the crowd and could be visualized with colormaps and vector fields (or perhaps something even better).