Analysis of the Available Data Sets

Halit Anil Donmez

Cyprus Data

Academic article it was used:
http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8659.2010.01808.x/full

Link to data: https://graphics.cs.ucy.ac.cy/research/downloads/crowd-data
Crowd samples are from crowds in uncontrolled environment (not from an experiment). All scenarios are the same where pedestrians are walking straight and towards each other. There are no obstacles (objects) in the path of pedestrians.

Figure 1: Scenario of the crowd data. In all datasets, pedestrians are either walking towards opposite or same direction

Overview

Data is given as splines in the form of a vsp file. Each spline starts with the total number of control points. Then the coordinates are given in the form:

x, y , frame number, gaze direction

X and y coordinates are positions in the pixel space ((0,0) is the center of the frame). Frame number is the time when the location was tracked. Gaze direction is given in degrees (0 degrees means person is looking upwards.).

They provided a software for visualizing the tracked data for the given video **(figure 1)**. I downloaded and used it. It is an analysis tool and used to create the spline data for the videos. However, data sets they provided as samples were also usable. Splines for each pedestrian was visible.

Putting the data points would not be enough for analyzing the data. A spline structure should be created in order to create a crowd from the given data set (bezier curve or another spline structure).

For example, in a data set, first position at frame 0 is x=70, y = -35. Next given position is x= 136 y = -23.000000 at frame 37. We need to get the curved path along this frames to fully simulate the path.
Another problem is with the frames. Frame numbers should be converted to time. This way we can get a better calculation.

BIWI Dataset

Link to data set: http://www.vision.ee.ethz.ch/en/datasets/
Link to paper it was used (pdf):
http://vision.cse.psu.edu/courses/Tracking/vlpr12/PellegriniNeverWalkAlone.pdf
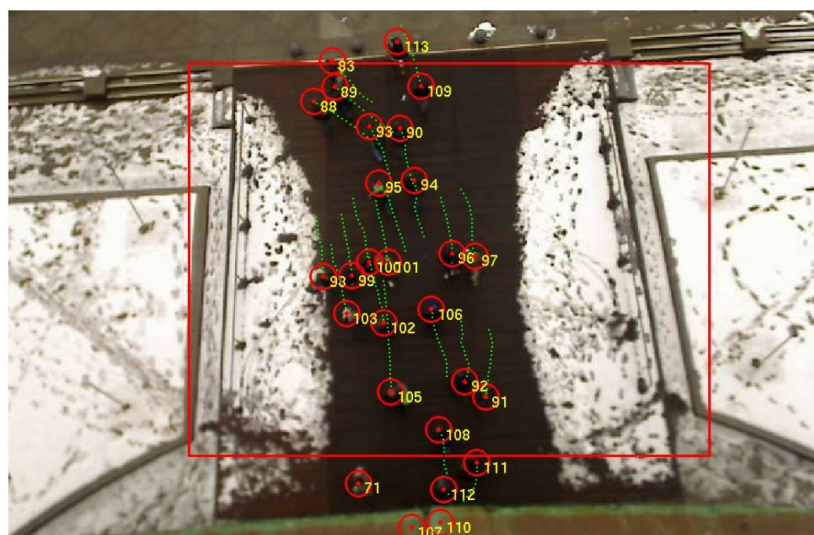

Figure 2: Scenario for the dataset. Pedestrians are walking through a corridor.

Overview

Data is taken from the pedestrian video **(figure 2)**. There are no obstacles in the path of the pedestrians. Data does not contain information about the pedestrians that were at the edge of the scene.

There are several information about the camera calibration. The data is stored in the form
[frame_number pedestrian_ID pos_x pos_z pos_y v_x v_z v_y]

pos_x, pos_y, pos_z, velocity_x, velocity_y and velocity_z are obtained in meters. There are more data available on the obstacles in the scene and the goals for each pedestrian.

These data was recorded at 2.5 fps and a time step of 0.4 seconds.

Obstacle information is given as a png image and destinations are given in the form of x and y coordinates.

Information about the groups are also available. ID's of the pedestrians who are walking in groups are stored in that file. For example, if a line has "5 4", this means that pedestrian 5 and pedestrian 4 are walking as a group.

Velocity and x,y coordinates are available. Yet they are available in meters. Timestep and frame number is given along with velocities so the position can be calculated. Only problem is converting the meters into unity coordinates. This may require setting a specific screen resolution and assign meter to a pixel accordingly.

I wrote an email to the contact person for the data about how did they handled the tracking and asked how should I handle this data. Still waiting for a response.

To summarize, we have the velocity and location in x and y plane is given along with time step. This seems sufficient enough to create a simulation on Unity.

BUW University Dataset (controlled experiments)

Link to dataset: http://www.asim.uni-wuppertal.de/datenbank/own-experiments.html
Link to paper referenced: (for the unidirectional dataset)
http://iopscience.iop.org/article/10.1088/1742-5468/2011/06/P06004/meta;jsessionid=6F72F7FE548A05159E87DF05020EBED6.c5.iopscience.cld.iop.org

(pdf) for the bidirectional corridor experiment: https://arxiv.org/pdf/1107.5246.pdf

Figure 3.1: Bi-Directional dataset. Pedestrians are walking towards the same direction

Figure 3.2: Uni-Directional dataset. People in red and people in black walk towards opposite direction.

Figure 3.3: Bottleneck dataset scenario

Overview

The dataset in these experiments were gathered from controlled environments. Meaning, they have made the experiments and told the pedestrians where to go.

For the bidirectional corridor experiment, 22 runs were made and they were collected in separate folders. For the unidirectional corridor experiment 28 runs were made

Experiments are available in the form of Corridor **(figures 3.1 and 3.2)** and bottleneck **(figure 3.3)** (another experiment with bikes are also available). For corridor data, bidirectional **(figure 3.1)** and unidirectional **(figure 3.2)** data is available. I took both of the data.

Data is recorded with two cameras and with 16 fps. Data is given in the following format:
ID,frame, x-coordinate, y-coordinate, z-coordinate

X,y and z coordinates are given in centimeters. The data files are sorted by the ID's. Starting points and how long the pedestrians

They made experiments with Bottleneck scenario. Data available there is given in the same way with the corridor scenes.

TU Berlin

Link to data and paper:
http://www.math.tu-berlin.de/projekte/smdpc/v_menue/simulation_of_multi_destination_pedestrian_crowds/project_overview/parameter/en/

Data is given in the .dat files. Although it is not clear what the files indicate. I sent an email to the author about this.

Directions of two pedestrian groups are perpendicular to each other **(figure 4)**.

They used Lucas Kanade algorithm to track the pedestrians and obtained B-splines. They merged the data from separate cameras with Kuhn-Munkres algorithm. The data files does not make much sense in their raw form.
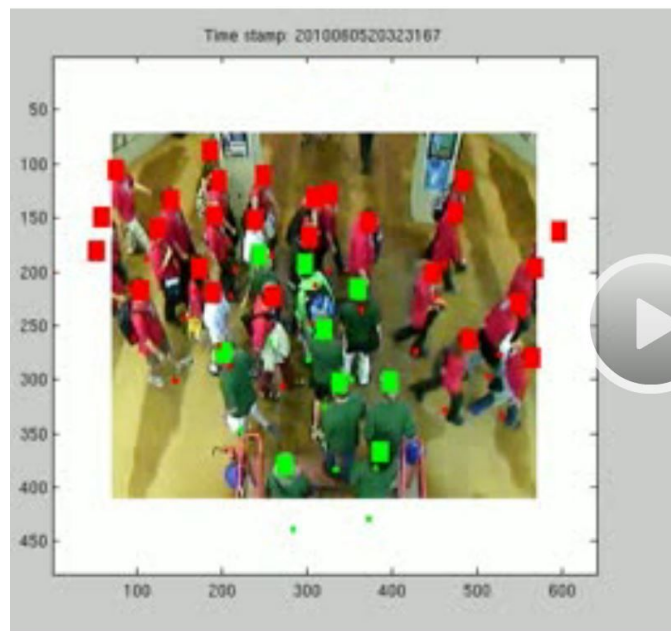

Figure 4: Scenario for 90 degree intersection dataset

New York Train Station dataset

Link: http://www.ee.cuhk.edu.hk/~xgwang/grandcentral.html
Paper: http://people.csail.mit.edu/bzhou/project/cvpr2012/zhoucvpr2012.pdf

Figure 5: Dataset scenario for corpora. There are no defined directions. Pedestrians are either walking towards or opposite of each other

Overview

This dataset was recorded with 25 fps with 50010 frames. Data is available as a Matlab data structure.

In the MATLAB file, it contains the extracted trajectories from the video of the train station. They used Kanade-Lucas-Tomasi tracking algorithm to extract the trajectories.

In the MATLAB file, there is x, y and t variables. Label meanings are not indicated in the website of the data set.

SNU Movement Research Lab Dataset

Link:
http://mrl.snu.ac.kr/research/ProjectMorphableCrowds/open_crowd_data/CrowdData.htm
Paper: http://mrl.snu.ac.kr/research/ProjectVideoCrowd/Group_Behavior_SCA2007.pdf

Figure 4.1: Aggressive   **Figure 4.2:** One way

Overview

Real crowd data was captured from 7 scenarios. Yet the captured data comes from experiments and not from natural crowds.

Data is given in the following form:

<div align="center">

40 // number of agents
0 // start frame of first agent
561 // end frame of first agent
44.5058 0 -40.7135 // first frame position of first agent
44.5242 0 -41.3203 // second frame position of first agent
44.6053 0 -44.115 // third frame position of first agent

</div>

I have looked into the file. There are some inconsistencies. For example, in some agents, start frame is greater than end frame which should not be the case. Also the given data format is not the same with the data structure in the website.

Extraction of the data will require the same effort with BIWI and Cyprus data set since the way data sources are represented are the same.

Name of the dataset	Criteria of the data set				
	Frame number	Coordinates (in some form of x,y,z)	Data gathered from real crowds (not experiment)	Velocity/Direction	Data given in single file
Cyprus	X	X	X	X	X
BIWI	X	X	X	X	X
BUW University	X	X		X	
TU Berlin	X			X	
New York Train station	X		X		X
SNU Movement Research	X	X			X