Sam Park

Summer Reading 2: Social attention: Modeling attention in human crowds

Social Robot Navigation Project @ Bot Intelligence Group

Paper:
Vemula, Anirudh, Katharina Muelling, and Jean Oh. "Social attention: Modeling attention in human crowds." 2018 IEEE international Conference on Robotics and Automation (ICRA). IEEE, 2018. https://ieeexplore.ieee.org/abstract/document/8460504

Sam Park

Summary:

Abstract

Early works on human trajectory prediction focused on proximity between humans. The authors' model focused on social attention instead of proximity. Social attention is a new trajectory prediction model that captures the relative importance of each person in the crowd. The authors compared the performance between their method (interactions between humans as a function of social attention) and previous approaches (a function of proximity) using 2 publicly available crowd datasets, ETH and UCY.

Introduction

Early works used individual human motion patterns in crowds to predict human trajectories. However, independent modeling does not capture the complex and subtle interactions between humans in the crowd which resulted in a suboptimal path for the robot.

Recent works modeled a joint distribution to capture dependencies between trajectories of interacting humans, assuming that human-human interaction is the only affecting factor. However, it is not true and velocity, time-to-collision, acceleration, heading, etc. need to be counted.

The authors used Recurrent Neural Network (RNN) to model trajectories of humans in the crowd.

Problem Definition

The authors obtain spatial coordinates (x_i^t, y_i^t) of all agents/people (i for each person) from a scene at different time steps t from time t = 1 to t = T_observe, and then predict their future locations for the next time steps from time t = T_observe+1 to t = T_predict.

Related Work

1. Social Force - Attractive forces & Repulsive forces based on relative distances. However, this can't model complex crowd behavior (e.g. cooperation).
2. Structural RNN (S-RNN) - Can be jointly trained to model dynamics in spatio-temporal tasks. The authors used a variant of S-RNN to human motion modeling, driver maneuver anticipation. etc.

Approach

The authors aimed to model the influence of all surrounding agents in the crowd (spatially local neighbors and other factors such as velocity, acceleration, heading, etc.).

The authors used a jointly trained RNN model to capture interactions between humans.

The authors used Spatio-Temporal Graph (ST-Graph) where the nodes in the graph represent humans in the crowd, the spatial edges connect 2 different humans at the same time-step in the scene, and the temporal edges connect the same human at adjacent time steps for the same scene.

- Spatial edges: To capture the dynamics of human-human interactions (relative orientation and distance) between two humans
- Temporal edges: To capture the dynamics of the human's motion (own trajectory)

Evaluation

The authors computed the prediction error using 2 different metrics:

1. Average Displacement Error to compute mean euclidean distance over all estimated points at each time step in the predicted and true trajectories
2. Final Displacement Error to compute the mean euclidean distance between the final predicted location and the final true location

Conclusion

The authors modeled an attention-based human trajectory prediction, called Social Attention. Social Attention learns the relative influence of each human in the crowd and predicts their future trajectories. Social Attention outperforms other novel approaches in prediction errors on 2 publicly available datasets, ETH and UCY.

Future works are:

1. Investigate further about the reason for prediction failures of the Social Attention model.
2. Extend the Social Attention model to count static obstacles in the environment.
3. Implement the Social Attention model into a real robot placed in a human crowd to verify and validate the model.

Sam Park

Glossary: