Hi! I Know Where You are Going

Tong Bai
tongb@andrew.cmu.edu

Sai Nihar Tadichetty
ntadiche@andrew.cmu.edu

1 Motivation

Detection and segmentation approaches in computer vision are all about sensing the world and interpreting the details captured in a scene. While these are great at understanding the scene, tracking algorithms are what add life to it; they not only tell you what is in the scene but also allow you to predict their states based on previous motion patterns. This has interesting applications in land and aerial locomotive industries as tracking can allow machines to have a true sense of what is happening around them. A version of this is being used in autonomous driving cars and now we would like to extend this to aerial vehicle safety.

2 Approach

Multiple Object Tracking is a classic but challenging task in the field of computer vision. Traditionally, researchers used optical flow and Kalmann filter methods to encode motion information and predict the next state using previous motion patterns. With the advent of deep learning methods that revolutionized the world of computer vision, many researchers proposed ways for tracking which are highly efficient compared to their traditional counter parts. Kim et.al's Multi-hypothesis Tracking (MHT) framework [1] , Alahi's "pure machine learning" approach also called Social LSTM [2] are some works in this field which produced promising results. However, neither of these are fast and therefore cannot be applied to real-time tracking.

In this project, we are planning to experiment and modify existing motion tracking and prediction methods to produce a fast tracking system that can be used to address motion planning problems in autonomous vehicle systems.

Right now we are looking at branching off from Mask RCNN to another LSTM network that will allow us to implement our tracking algorithm on in. We are also looking into "Learning to Track at 100 FPS with Deep Regression Networks" by David Held [3]. We haven't zeroed down on what method we will be implementing, but this is the approach we are looking at currently.

3 Work designation

Tong has experience about object tracking before, who will go through previous work of motion prediction and tracking, and set up experiment environment. Nihar works with robotics, who will collect and process the real-world dataset and also doing data processing on the exist bench marks. Both of us will discuss and doing experiments of new algorithm to solve our particular problem.

References

[1] C. Kim, F. Li, A. Ciptadi, and J. M. Rehg, "Multiple hypothesis tracking revisited," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 4696–4704.

[2] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social lstm: Human trajectory prediction in crowded spaces," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[3] D. Held, S. Thrun, and S. Savarese, "Learning to track at 100 fps with deep regression networks," in *European Conference on Computer Vision*. Springer, 2016, pp. 749–765.