Vehicle Speed Estimation from 3D Object Detection for Surveillance Cameras

Manthan Chelenahalli Satish, MS in Robotics and Autonomous Systems (Artificial Intelligence)
Mentor: Dr. Yezhou Yang, Assistant Professor
School of Computing and Augmented Intelligence, Ira A. Fulton Schools of Engineering, ASU

Introduction

Traffic cameras have been used for many years to enforce traffic laws and catch speeding drivers. Today, traffic cameras are used for various purposes, including monitoring traffic congestion, detecting accidents, and enforcing speed limits.

Motivation

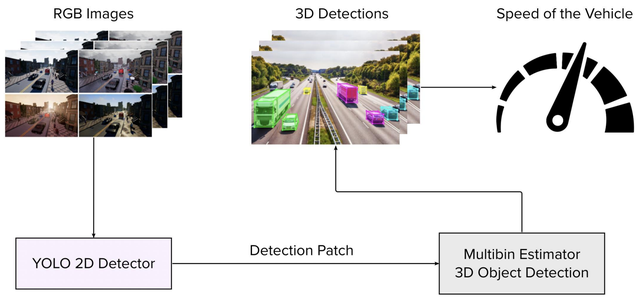
- Detecting vehicle speed in an accurate way and with fewer false positives.
- Need to account for factors such as the perspective of the camera, the varying speed of vehicles, and the different sizes of vehicles.

Contributions

- System on a traffic camera to calculate the speed of a vehicle by detecting vehicles in 3D for better efficiency.

Pipeline

The system uses a 3D object detector to detect vehicles in each image. It then estimates the speed of each vehicle by triangulating the position of the vehicle in the image over time.



Methodology

One of the major focus of the project is to detect vehicles in 3D from the perspective of a Traffic camera. Existing methods detect vehicles in 3D using datasets (ex: KITTI benchmark suite) that are recorded from the perspective of a car. To overcome this issue, we record a novel dataset that mimics the features of the KITTI benchmark dataset in the CARLA simulator but from the perspective of a Traffic camera.

The KITTI-CARLA Dataset

The system has two main parts, detecting vehicles in three-dimensional space and estimating the speed from the detections.

Monocular 3D Object Detection

- A two-step process: Detect objects in the image with YOLO
- Feed the 2D Detections to Multibin Estimator that computes:
  - Dimensions of 3D Bounding Box
  - Orientation
  - Confidence.

Speed Estimation

- The position of the vehicle known is from the 3D detections
- Speed is calculated by triangulating the positions of detected cars over time.
- A Kalman Filter smoothens the predicted velocity

Experiments

Dataset - The KITTI-CARLA dataset was recorded in 7 different scenes comprising of different landscape and traffic densities.

Metrics:

Use-case – Inference time
Evaluation metric - Mean Error
Per pose inference time PPIT (GTX 1080)

- KITTI – 0.112 s
- KITTI-CARLA – 0.114 s

Mean Error

- Transform 3D – 0.40 MPH
- KITTI-CARLA – 0.31 MPH

Future Work

- Record Data at different times of day and under different weather conditions.
- Implement the trained simulation model on real world traffic cameras and evaluate.
- Explore other 3D object detection frameworks for monocular RGB images

References

- V. Kocur, M. Ft´aˇcnik: “Detection of 3D Bounding Boxes of Vehicles Using Perspective Transformation for Accurate Speed Measurement”, arXiv preprint, arXiv:2003.13137 (2020)
- Lu, Duo, Varun C. Jammula, Steven Como, Jeffrey Wishart, Yan Chen, and Yezhou Yang. "CAROM--Vehicle Localization and Traffic Scene Reconstruction from Monocular Cameras on Road Infrastructures." arXiv preprint arXiv:2104.00893 (2021).

MORE
Masters Opportunity for Research in Engineering

Ira A. Fulton Schools of Engineering
Arizona State University