

# Driverless Formula Student Problem Statement

## CS461 Spring 2018

Lucas Frey

10/10/18

Driverless vehicles require robust internal representations of the world around them in order to make correct driving decisions. The vision system of a driverless car needs to take sensor input from lidar, RGB camera, accelerometer, and GPS in order to construct a virtual representation of itself and of the world around it. The various sensor readings are used to both map the environment around the car as well as localize the car within that environment. Due to the inherent error within the sensors and as well as low sampling rates, high-speed motion can cause data about the environment to be sparse. Our team is tasked with constructing the vision system for a driverless formula student vehicle. Our objective is to take the sensor readings from the cameras, lidar, accelerometer, GPS and other sensors and turn it into a virtual representation of the world. The driverless formula student vehicle is expected to compete in August 2019.

# 1 The Problem

The driverless vehicle must take advantage of sensor readings from the cameras, lidar, accelerometer, GPS and other sensors and turn it into a virtual representation of the world. The virtual track that is constructed will be used to compute the optimal path and actuator (motor and break) velocities for the vehicle to complete laps around a race track in the shortest amount of time. Blue and yellow traffic cones make up the left and right boundaries of the track and are set up differently depending on the type of competition. The driverless vehicle will compete in 4 types of competitions, figure-eight, straight-drive, autocross, and trackdrive. In figure-eight, the vehicle must complete several laps around a figure-eight style track. The exact specifications of the track are known before the race, so constructing a map is not necessary. Only the location of the vehicle within the track is needed. In straight-drive, the car must accelerate in a straight line as fast as possible. In autocross, the vehicle must complete one lap around a custom track as quickly as possible. And in trackdrive, the vehicle must complete 10 laps around a custom track as fast as possible. In both autocross and track drive, the vehicle will have no previous knowledge of the layout of the track before it begins the loop, so it must utilize lidar and camera data in order to generate a map of the world as quickly as possible. However, sensors and actuators are prone to error. Error must be accounted for and the system must be robust to incorrect sensor readings as well as sensor failure.

# 2 The Solution

Our vision system must utilize some form of the Simultaneous Localization and Mapping (SLAM) algorithm in order to keep track of how the world changes around the vehicle as it moves. To do this, information from lidar, RGB camera, accelerometer, GPS, and other sensors must be used to construct a robust representation of the world and establish where the vehicle is in that world. Faster-RCNN (Region Convolutional Neural Network) and YOLOv3 (You Only Look Once version 3) are powerful neural networks for detecting objects. One of these networks can be used to find and classify the blue and yellow cones. Lidar can then be used to determine how far those cones are from the. When the cones in the image have been localized, information from GPS and the accelerometer can be used to compute the distance those cones are from the car as well as how the cones are moving around the vehicle. Once a lap around the track has been completed, the vehicle should have a good understanding of where all the cones are on the map. The vehicle can then compute the optimal path around the entire track and follow that path as closely as possible.

# 3 Evaluation

Simulations of our vision system will be run and the location and count of traffic cones in our virtual reconstruction will be compared to the count and location of the traffic cones in the simulation in order to validate the vision systems accuracy. Detection speed is also an incredibly important factor. The vehicle will be moving as fast as 60 miles per hour or 88 feet per second. Our mapping and localization should ideally run at a minimum of 88 frames per second, or 1 frame per foot traveled in order to have enough time to correct for errors in the trajectory. Additionally, the false positive and false negative cone detection should be no greater than 5% of all cones detected during the operation of the vehicle.