Formula Student Driverless System Architectures: Literature Review

Wei Hai Cui

Division of Engineering Science
UNIVERSITY OF TORONTO

UTFR
University of Toronto Formula Racing

Overview

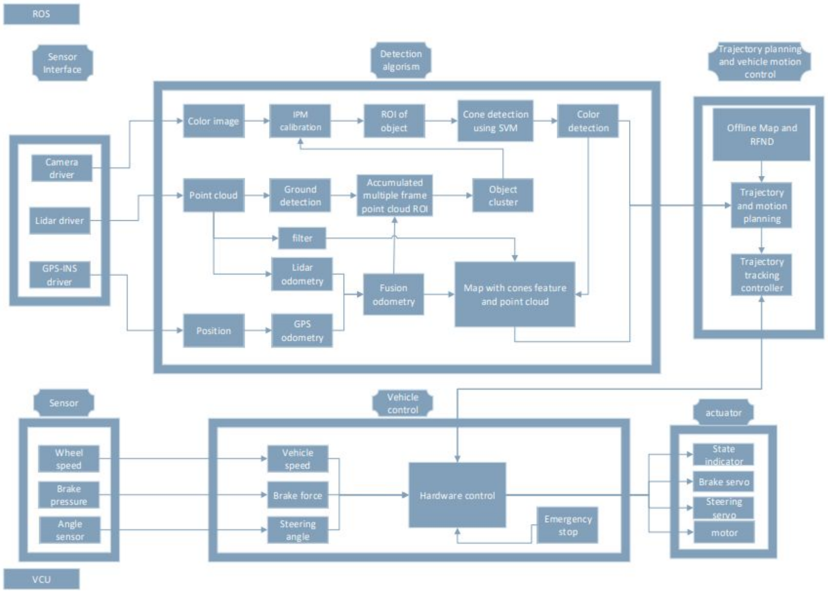
- Assumptions simplify problem at hand:
  - Pose can be represented with x, y and yaw variables rather than full 3d space
  - Cones are known to be a set color & minimal distance apart
  - Acceleration & Skidpad track shapes are known prior
- Perception typically relies on Cameras and/or Lidar Pipelines to obtain cones
  - Perception often struggles with occlusion and weather effects
- SLAM is typically used for AutoX events, using cones as landmarks
  - Odometry is limited due to wheel slip
- Control systems need to work in real time & account for complex plant model

Team Architectures Reviewed:

- BIT-FSA (2016)
- TUW (2017)
- KIT (2019)
- AMZ (2018)

BIT-FSA (2016)

HanQing Tian, Jun Ni, Jibin Hu
Beijing Institute of Technology



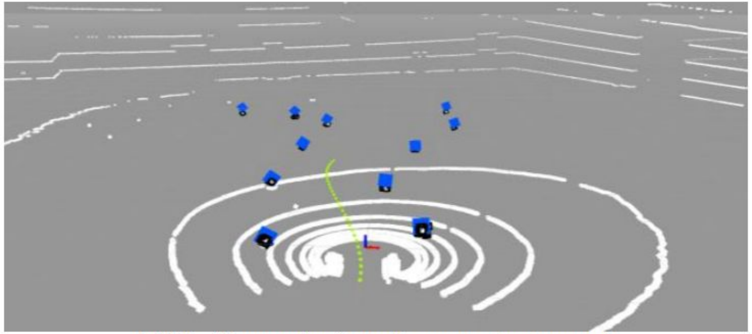
- Two controllers:
  - **Upper level** runs ROS on x86 computer
    - Runs Perception, Planning & Controls
  - **Lower level**
    - 32-bit Micro programmed with Simulink
    - Controls actuators, interfaces with sensors, runs emergency stop system

https://arxiv.org/ftp/arxiv/papers/1809/1809.07636.pdf

BIT-FSA (2016)

Perception

- **Lidar:**
  - Ground points are removed assuming locally planar ground, leaving only cones left
  - Sliding window of measurements are taken to obtain dense point clusters for each cone
  - Clusters are grouped based on euclidean distance from which cone centroids are obtained
- **Monocular Camera:**
  - Cone detection is done with SVM Classifier
  - Color detection is based on HSV Heuristics
  - Matched with lidar detection to obtain cone type



(a) The blue marks show the cones cluster result

(b)The SVM classification and color detection result

Fig. 10. The cones detection result

Division of Engineering Science
UNIVERSITY OF TORONTO

UTFR
University of Toronto Formula Racing

BIT-FSA (2016)

Localization & Navigation

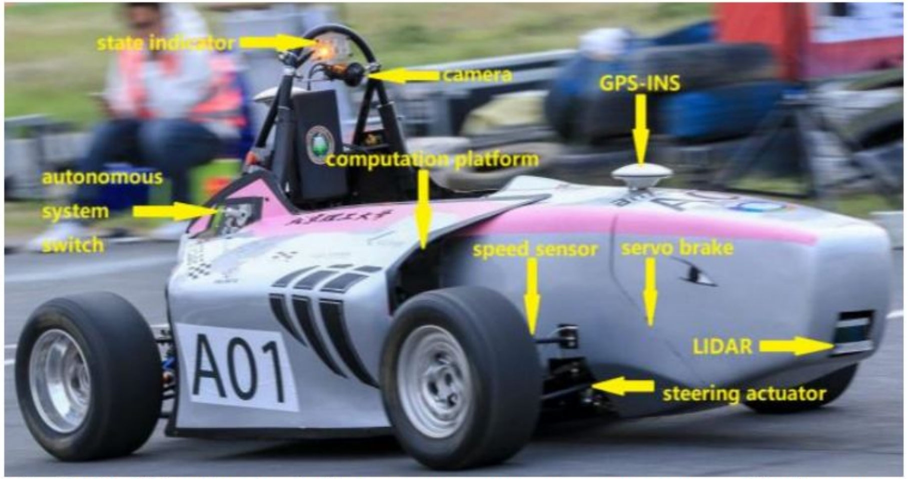


Fig. 2. Modifications including sensors, computer, actuators and indicator

- Odometry is obtained by fusing:
    - **GPS-INS**
        - RTK provides high accuracy outdoors
        - INS provides acceleration measurements
    - **Lidar**
        - Cones are assumed to be non-moving

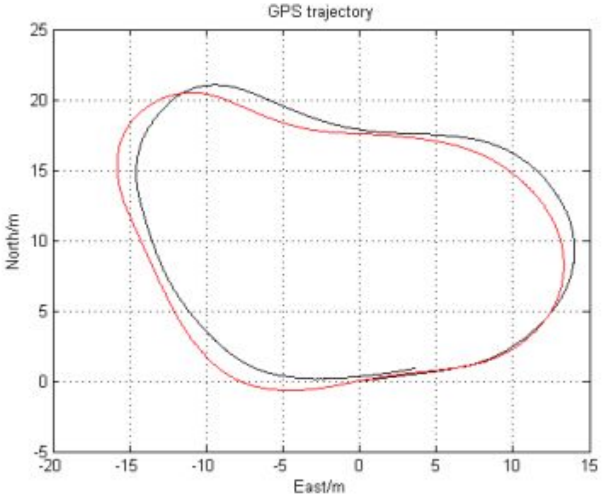
- Perception is relied upon to build offline map of environment
- Once loop-closure is detected, GPS Odometry is relied upon over perception

BIT-FSA (2016)

Summary

- Reliance on heuristic allows for lower computation requirements compared to deep methods
- **Winner of Formula Student Autonomous China 2016**

- *Expensive hardware requirements (GPS, Lidar)*
- *Lack of redundant sensors*
  - *Object detection only from Lidar - susceptible to false detections from weather/debris*
  - *Camera based on heuristics may miss-identify occluded cones or cones in different lighting conditions*
  - *GPS +- 2m accuracy*



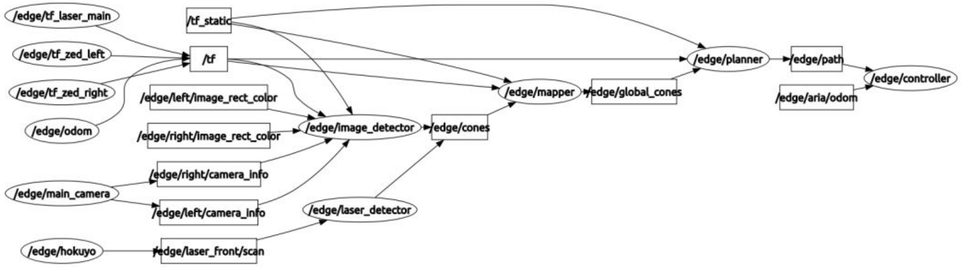
Division of Engineering Science
UNIVERSITY OF TORONTO

University of Toronto Formula Racing

TUW (2017)

Marcel Zeilinger, Raphael Hauk, Markus Bader and Alexander Hofmann
Technical University Wien

- Uses ROS as a backbone for Machine vision, SLAM, Path Extraction and Motion Control



TUW (2017)

Drive By Wire

- Brake System:
  - Actuators mechanically move the pedal due to existing brake balance configuration
  - Can also decelerate through reverse operation of motors
  - Able to stop the car within 10m
- Steering:
  - Additional steering motor is mounting to the steering strut
  - Can apply 25Nm of force to turn the wheels with a driver in the car

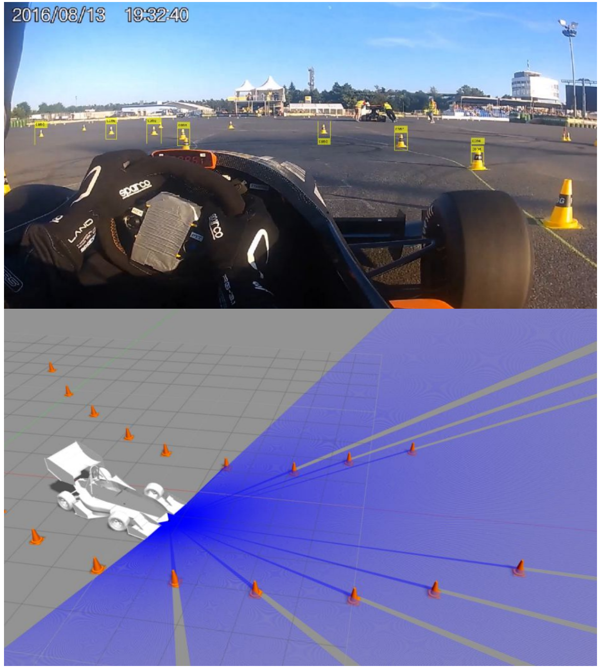
Division of Engineering Science
UNIVERSITY OF TORONTO

University of Toronto Formula Racing

TUW (2017)

Perception

- Stereo Camera (**ZED Stereo Camera**)
  - Camera intrinsics are calculated using visual markers on chassis
  - Cones are detected in both images separately through a cascade classifier
  - Disparity, z-depth and 3D position can then be calculated based on the centroids of the same object compared in the left and right frames
  - Color is sampled at the centroid - if a black/white stripe is detected, the sample point moves down until color match is made
- Lidar (**Hokuyo 20LX Planar Scanner**)
  - Simple algorithm to search for clusters of distances that match the radius of the cone at lidar height
- Results of both are fed into an Extended Kalman Filter



Division of Engineering Science
UNIVERSITY OF TORONTO

University of Toronto Formula Racing

TUW (2017)

Localization & Navigation

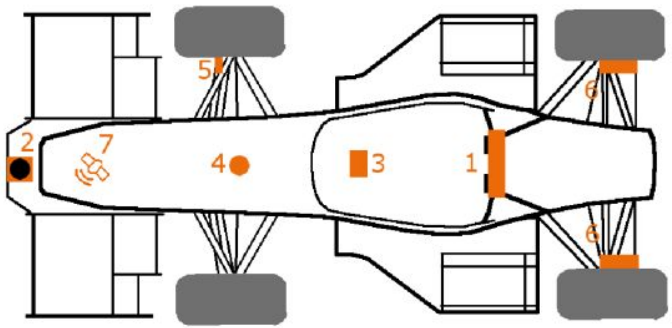


Fig. 2. Sensor placement on the vehicle: 1 stereo camera, 2 laser scanner, 3 IMU, 4 steering angle encoder, 5 wheel speed encoder, 6 rotor position encoder, 7 GPS

- GPS (Piksi Multi GNSS) and IMU are used to provide position and acceleration measurements
- Odometry is provided with inductive wheel speed sensors at the front wheels, and rotor position encoders for the rear motors
- EKF-SLAM uses traffic cone perception combined with above localization to build track map
- Heuristics allow cones to be filled in upon detection failures:
  - Cones are known to be at 5m apart
  - Left and Right cones are different colors
  - Cones are known to be ground truth - adjustment to right lane cones requires the same adjustment to left lane cones
- Center Trajectory path is generated using the center of each pair of left and right cones.

TUW (2017)

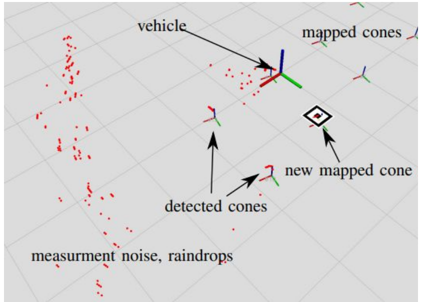
Controls

- Model Predictive Control is used to calculate optimal local trajectory
  - Optimization is based on vehicle motion model
  - Takes in mechanical constraints:
    - Maximum acceleration and deceleration
    - Actuator & Pipeline delays
- Feedback for steering angle is accomplished with a rotary position encoder connected to steering shaft

TUW (2017)

Summary

- Lightweight detection & planning algorithms reduce computational complexity
- Redundant Perception allows for detection in the case of failure in one modality, across different distances
- Multiple Localization sensors allows for detection of wheel-slip and improvement in odometry accuracy.



- *Classical Lidar detection algorithm is not robust to weather effects*
- *Classical Camera conde detection is unable to deal with changing light conditions*
- *Model Predictive Control requires good modelling of vehicle dynamics, and is computationally expensive compared to simple PID*

KIT (2019)

Sherif Nekkah , Josua Janus , Mario Boxheimer , Lars Ohnemus , Stefan Hirsch, Benjamin Schmidt, Yuchen Liu, David Borbély, Florian Keck, Katharina Bachmann, Lukasz Bleszynsk

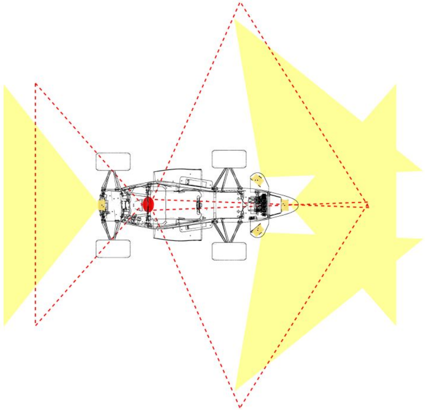
Karlsrushe Institute of Technology

- Autonomous Control Unit runs ROS on consumer computer hardware
    - I7-9700k, 32GB Ram
- Main Control Unit takes control commands from ACU and EBS to send to motor, brake and steering controllers.

KIT (2019)

Perception

- Four Lidars are used to obtain Cone Landmark Proposals
  - Points clouds are first transformed into the same frame
  - Empirical evidence from density, variance and spacing is used to filter the cones
- Three cameras are used for Cone Landmark Validation
  - Landmark proposals are projected to the image plane to form a bounding box
    - Bounding box size is adjusted based on distance from camera, and centroid is moved to the center of cone color mass
  - CNN is used to classify the type of cone in each bounding box, or if the proposal is a false detection
- Synchronization ECU sends synchronization pulse to all perception sensors to allow for asynchronous merging in ACU

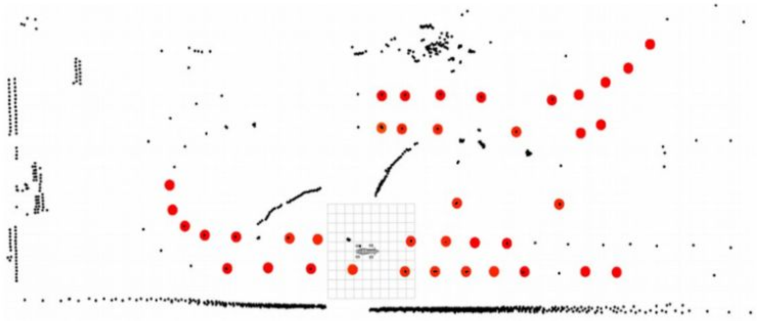


Division of Engineering Science
UNIVERSITY OF TORONTO

UTFR
University of Toronto Formula Racing

KIT (2019)

Localization & Navigation

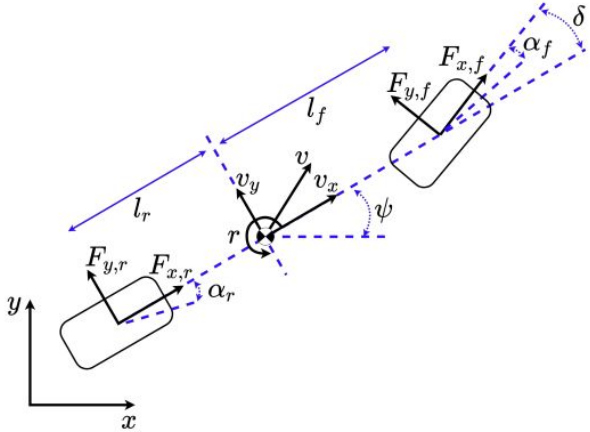


- Extended Kalman Filter uses an IMU & Rear Wheel Odometry for the prediction step
  - Removes the need for a vehicle model based on inputs
- Correction step is done using fixed landmarks from perception
  - Fixed landmarks are only used if it's been tracked more than once in the same color
- Joint Compatibility Branch and Bound algorithm is used to associate detected landmarks and mapped landmarks
  - Individual Compatibility Nearest Neighbour created false associations due to lack of minimum distance between cones

KIT (2019)

Controls

- Geometric Path optimization yielded small performance benefits compared to Vehicle Control optimization
    - Trajectory is simply set to centerline of path
- Model Predictive Control is used for lateral control
    - Current velocity is assumed to be constant throughout prediction horizon
    - Dynamic Bicycle model is used to model vehicle behaviour
    - Lateral tracking becomes a parameter optimization problem
- Feedforward PI is used for longitudinal velocity control
    - Target velocities are using lap-time simulation methods
    - Combined lateral & longitudinal acceleration limits at different speeds are computed offline, providing a map that can be read online
    - Limits are capped so that tires remain within linear operation regions



KIT (2019)

Summary

- Rear facing perception allows for better mapping in Skidpad & AutoX
- Splitting longitudinal & lateral control reduces computational load
- JCBB improves recognition over commonly used ICNN
- Winner of Skidpad & AutoX, FSG19

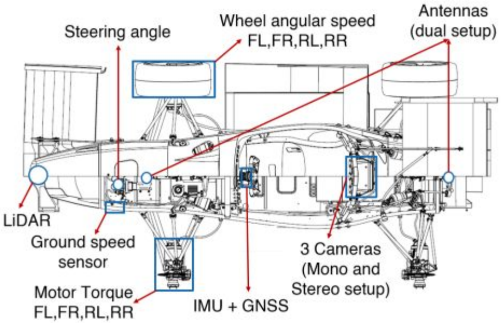
- *Increased complexity with perception (Sync-ECU, 7 sensors total)*
- *Perception pipeline not robust to missed lidar proposals*
    - *Classical lidar detection not robust to noise & weather*
- *Odometry is not robust to wheel-slip*

AMZ (2018)

Juraj Kabzan, Miguel I. Valls, Victor J.F. Reijgwart, Hubertus F.C. Hendrikx, Claas Ehmke, Manish Prajapat , Andreas Buhler , Nikhil Gosala , Mehak Gupta, Ramya Sivanesan, Ankit Dhall, Eugenio Chisari, Napat Karnchanachari, Sonja Brits, Manuel Dangel, Inkyu Sa, Renaud Dube, Abel Gawel, Mark Pfeiffer, Alexander Liniger, John Lygeros and Roland Siegwart

ETH Zurich

- Features fully independent Lidar and Camera perception pipelines
- Multiple failure detection algorithms for sensors and detections



Division of Engineering Science
UNIVERSITY OF TORONTO

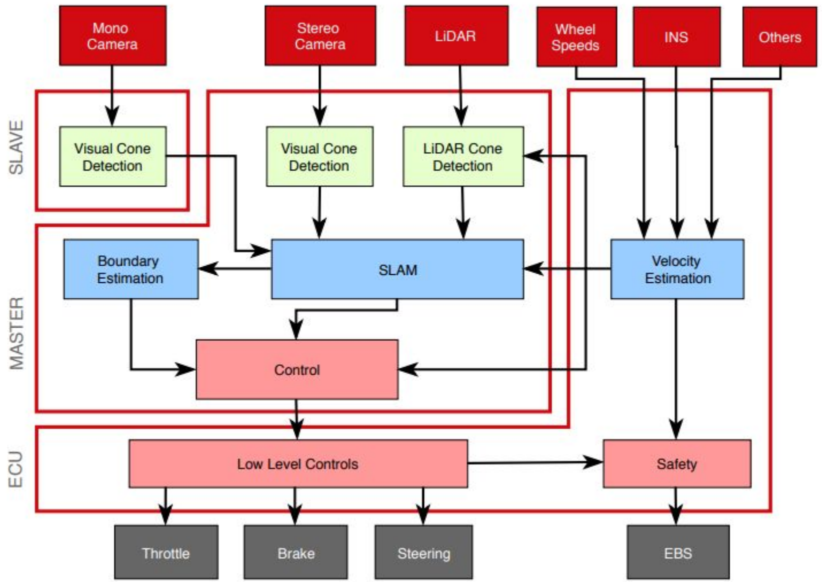
https://arxiv.org/pdf/1905.05150.pdf

University of Toronto Formula Racing

AMZ (2018)

Architecture Overview

- Uses two computing units to carry out Perception Navigation and Controls:
  - Follower Jetson TX2
  - Leader Pip39 (i7, 16Gb Ram) and GTX 1050Ti
- Leader/Follower architecture allows system to race despite failure in follower
- ECU is the only system that interacts with actuators

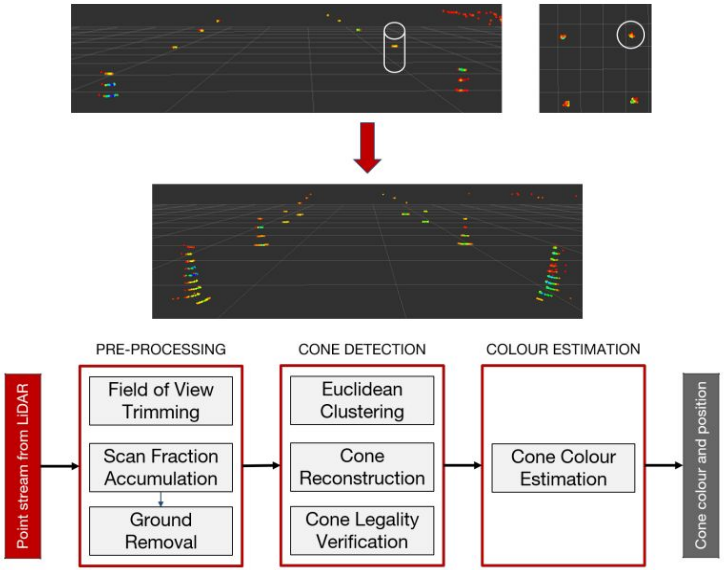


Division of Engineering Science
UNIVERSITY OF TORONTO

University of Toronto Formula Racing

AMZ (2018)

Perception - Lidar



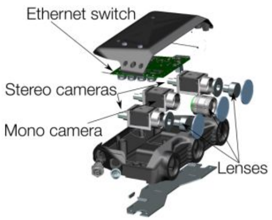
- Lidar point clouds are first Pre-Processed:
    - Adaptive ground removal algorithm splits point clouds into radial bins, then creates ground planes based on lowest points in each segment.
- Remaining points are and points in a cylindrical area around cluster are reconstructed
    - If reconstructed cone matches expected points for a cone at that distance, it passes the rule-based filter
- CNN is used to estimate cone colour based on point cloud intensities

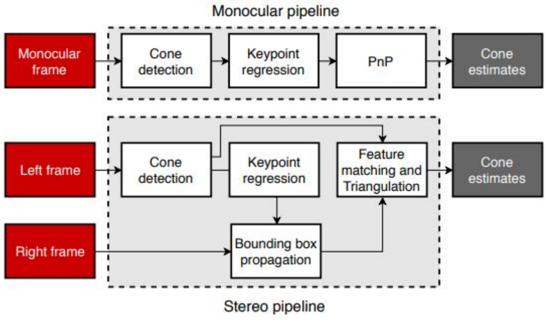
Division of Engineering Science
UNIVERSITY OF TORONTO

University of Toronto Formula Racing

AMZ (2018)

Perception - Camera



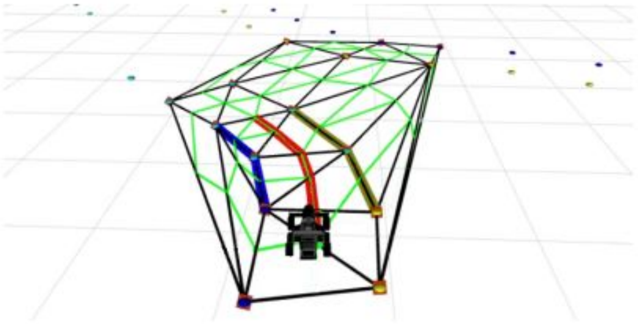


- YoloV2 is used for detection and colour classification
- Keypoints of each cone are found using a custom CNN based on ResNet.
  - Prior known geometric information about the cones can be leveraged with keypoints
- Monocular pipeline uses RANSAC Perspective and Point (PnP) to estimate 3D Pose
- Stereo pipeline projects detections from Left frame to Right Frame, then feature matches with SIFT feature extractor
  - 3D Pose is then estimated based on triangulation of matching features between frames

AMZ (2018)

Motion Estimation and Mapping

- Extended Kalman filter fuses Ground Speed, Inertial Measurement, GPS and Wheel Speed Sensors to estimate velocity and position
  - Process model is a constant acceleration model, using the motor torque and steering commands as inputs
  - Wheel slip ratio is calculated for each wheel and incorporated into calculations to offset wheel speed sensor information
  - Measurement model has failure detection to counter sensor outlier (chi-square test) and drift (variance)
- SLAM assigns recall counters for each cone
  - If cone counter doesn't increase despite being within sensor FOV, cone position is assumed to be dubious
- Once loop closure is detected, Localization calculates center line through Delaunay Triangulation of each cone



Division of Engineering Science
UNIVERSITY OF TORONTO

University of Toronto Formula Racing

AMZ (2018)

Control

- Model Predictive Control is used to follow trajectory
- Vehicle Model is a combined dynamic and kinematic model:
    - Bicycle model with non-linear tire force law for high speed
        - Combined slip is ignored by constraining lateral and longitudinal load on tires
    - Simple kinematic model for low speed and tight turns assuming tires are well within grip
    - Allows high performance at all velocities and solvable in real time by MPC
- Additional constraints are placed so that vehicle does not exceed track boundaries
- Nonlinear optimization problem is solved using ForcesPro NLP with a receding horizon

AMZ (2018)

Summary

- Multiple redundancy sensors for every section to prevent failures while running
- Error tracking for localized cones, as well as sensor measurements
- MPC vehicle model accurately defines motion but can be solved in real time
- Easily the most advanced FSAE DV architecture, winner of FSG for multiple years in a row

- *YOLOv2 is outdated by today's standards*
- *Lack of Fusion detection methods fails to combine strengths of both lidar and camera modalities*
- *360 Lidar adds additional mass compared to solid state lidar*
- *Computationally expensive architecture requires multiple powerful control units*
- *High cost + complexity architecture*

Key Takeaways:

- Perception:
  - Tradeoff between robust detection methods and computational power
  - Camera only pipelines are possible, at the cost of accuracy
  - Learning methods require large datasets -> FSOCO most common
- Localization:
  - EKFs commonly used to fuse IMU and Odometry (sometimes GPS and GNSS)
  - Wheel Slip must be accounted for for accurate results
- Navigation:
  - Simple PID Feedforward is enough for longitudinal velocity control
  - Accurate tire data is required to ensure control scheme falls within linear modelling + no combined slip
  - MPC commonly used for trajectory optimization in AutoX event, but may not be necessary for Acceleration & Skidpad

- **A simple architecture that finishes the event slowly is better than a complex one that doesn't finish at all**