Creating Autonomous Vehicle Systems

PERCEPTION INSIGHT INTELLIGENCE

Shaoshan Liu

shaoshan.liu@perceptin.io

Autonomous Driving

Sensing	Perception	Decision
GPS/IMU	Localization	Path Planning
LiDAR	Object Recognition	Action Prediction
Camera	Object Tracking	Obstacle Avoidance

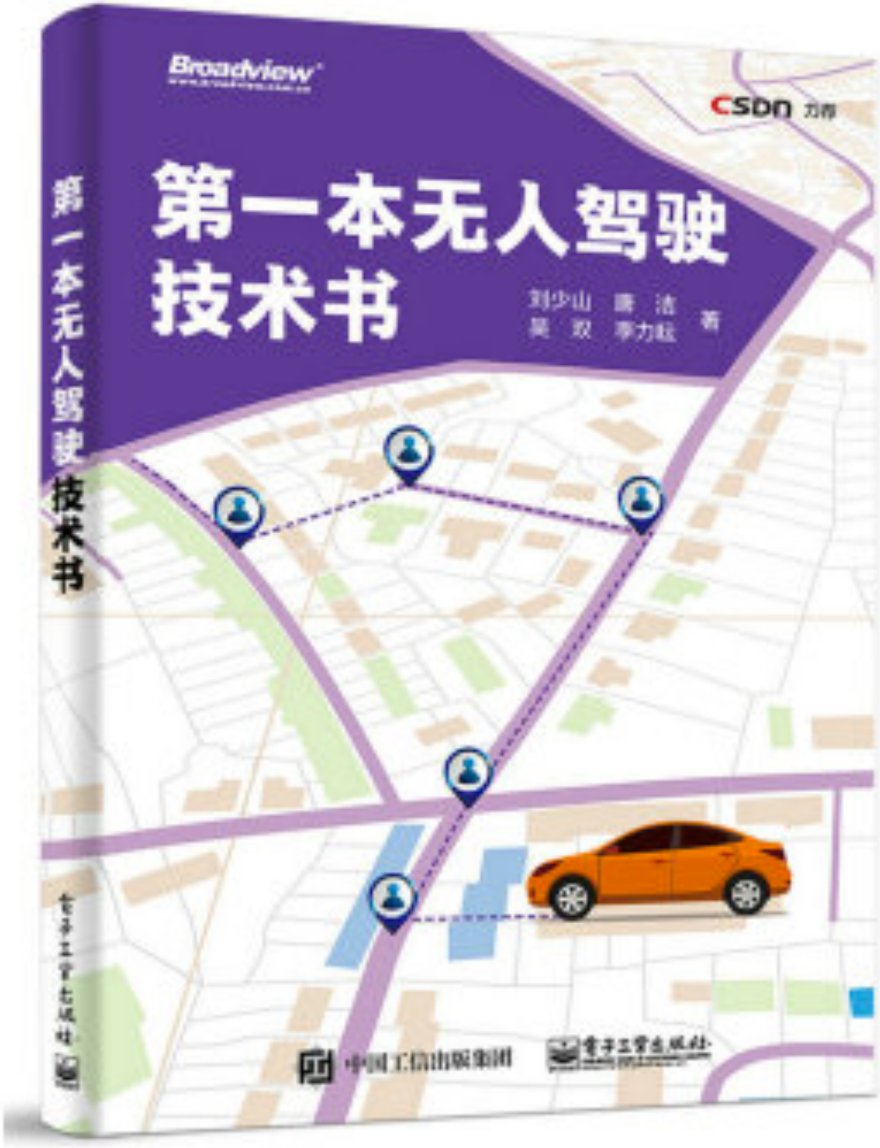
Operating System

Hardware Platform

Cloud Platform

- HD Map
- Model Training
- Simulation
- Data Storage

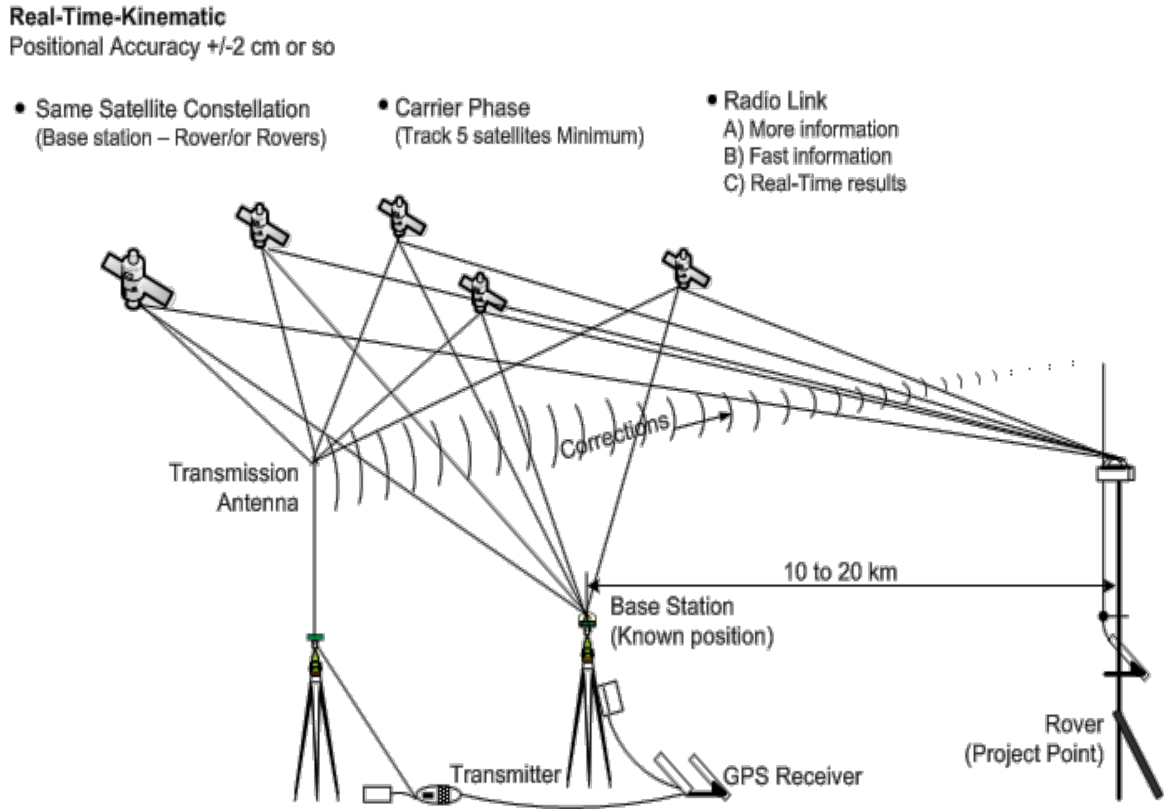
第一本无人驾驶技术书

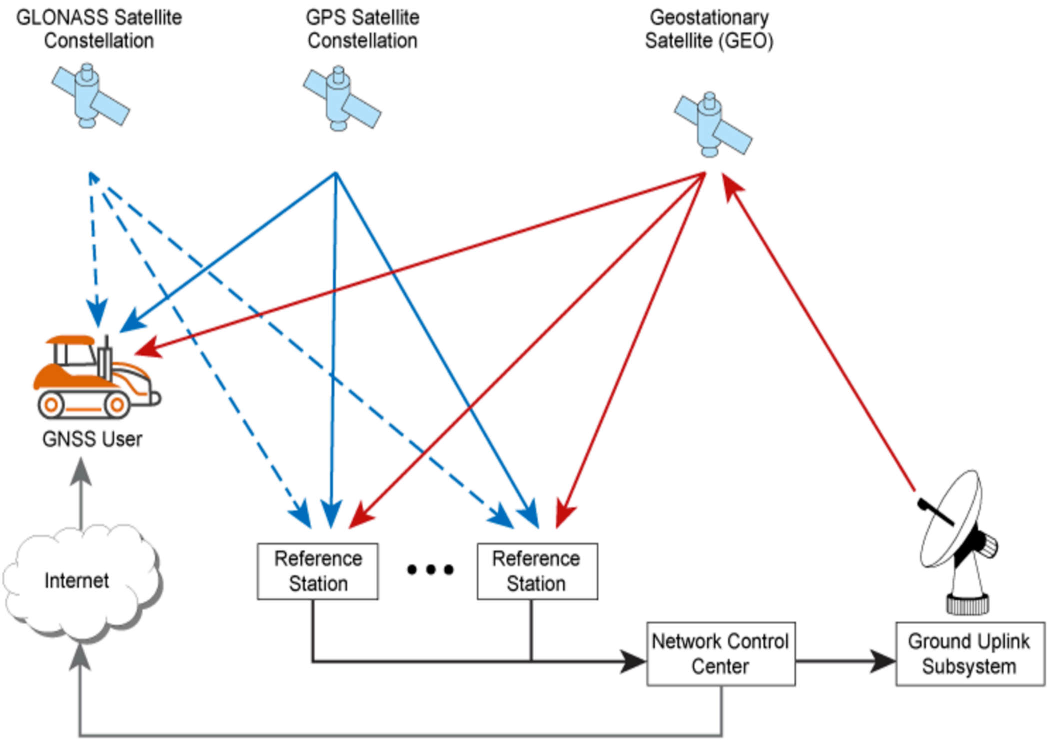


Module 1: Localization

Localization: GNSS/INS

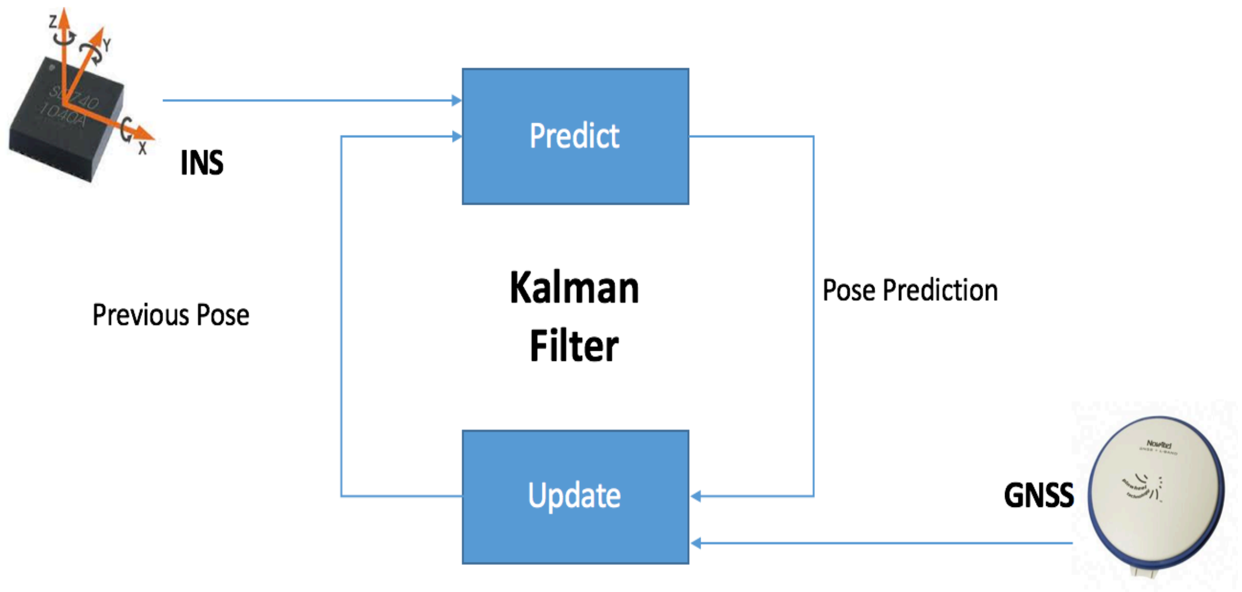
Contributing Source	Error Range
Satellite Clocks	±2 m
Orbit Errors	±2.5 m
Inospheric Delays	±5 m
Tropospheric Delays	±0.5 m
Receiver Noise	±0.3 m
Multipath	±1 m



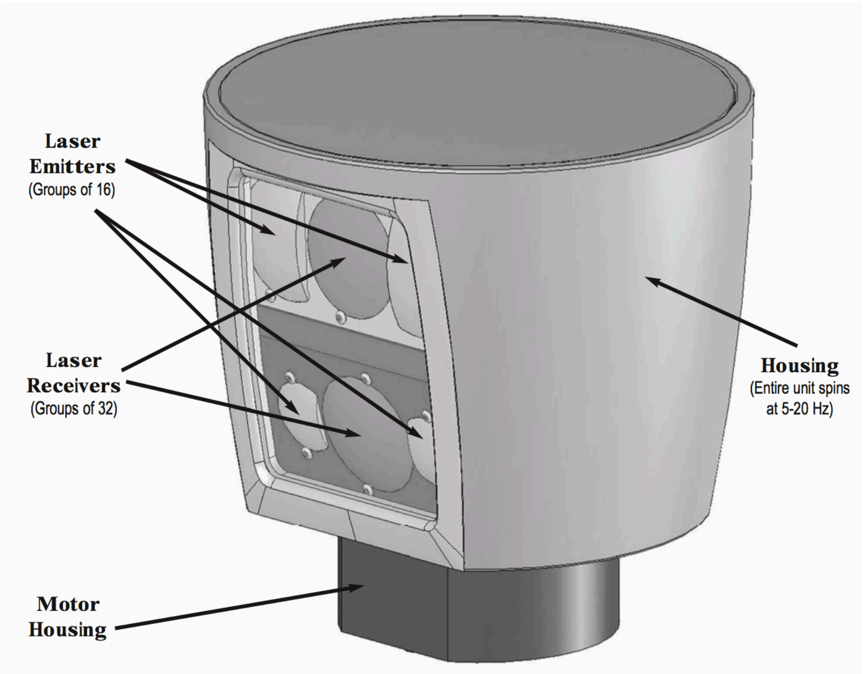


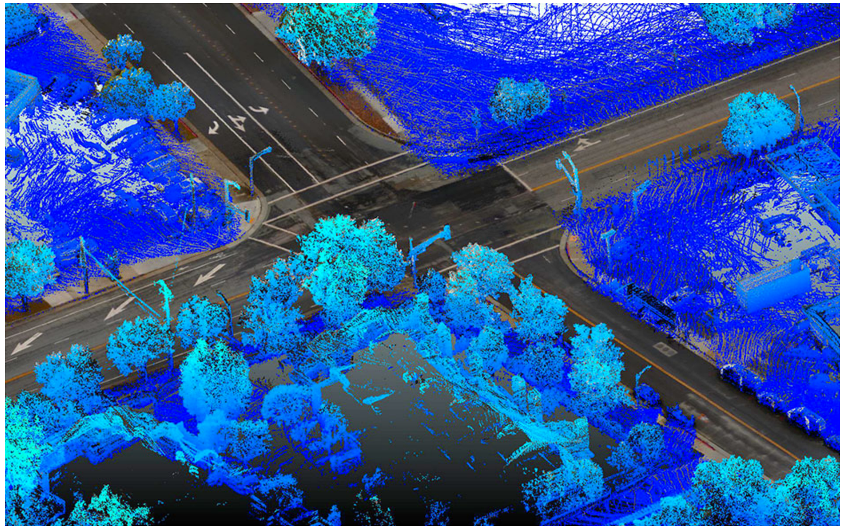
Localization: GNSS/INS

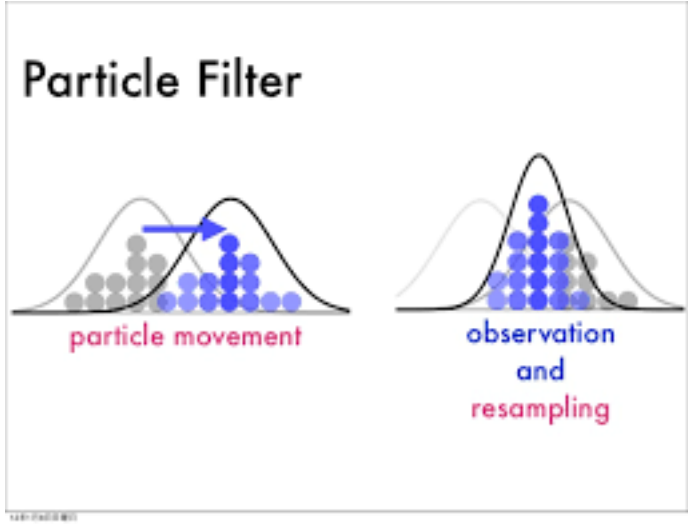
- **GNSS**：accurate but low frequency
- **INS**：high frequency but inaccurate
- **Kalman Filter:** get the best from both
- **Question:** is GPS/IMU enough?



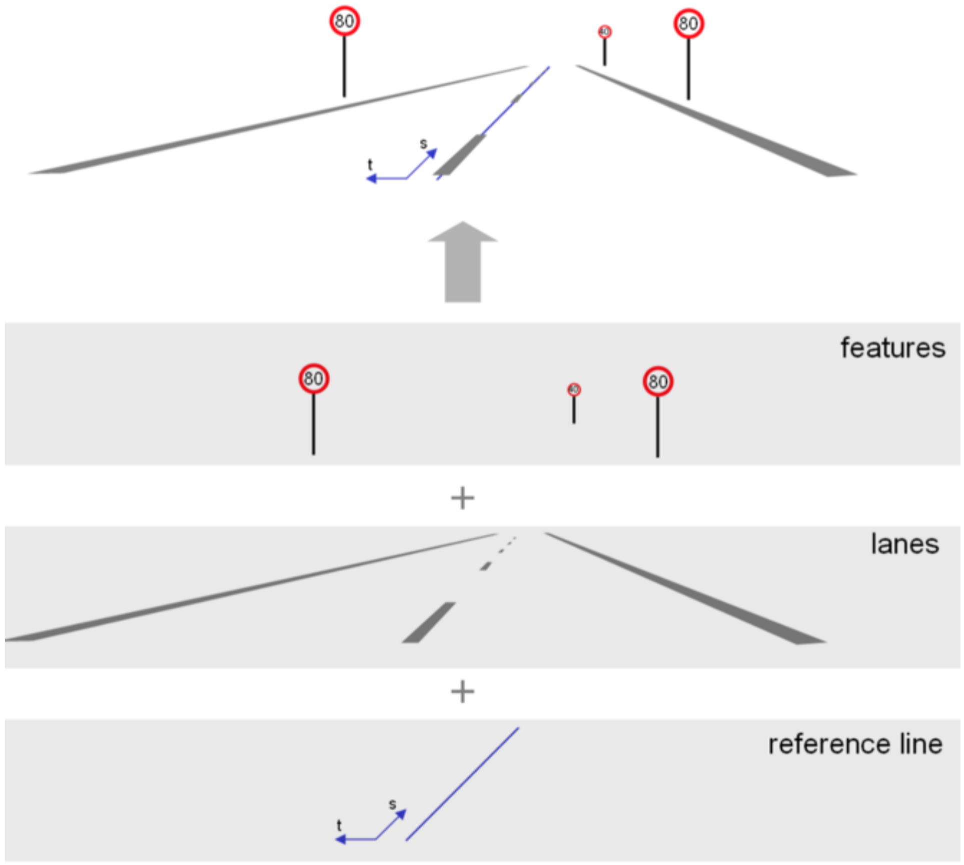
Localization: LiDAR and HD Map







Localization: LiDAR and HD Map



Layered Map:

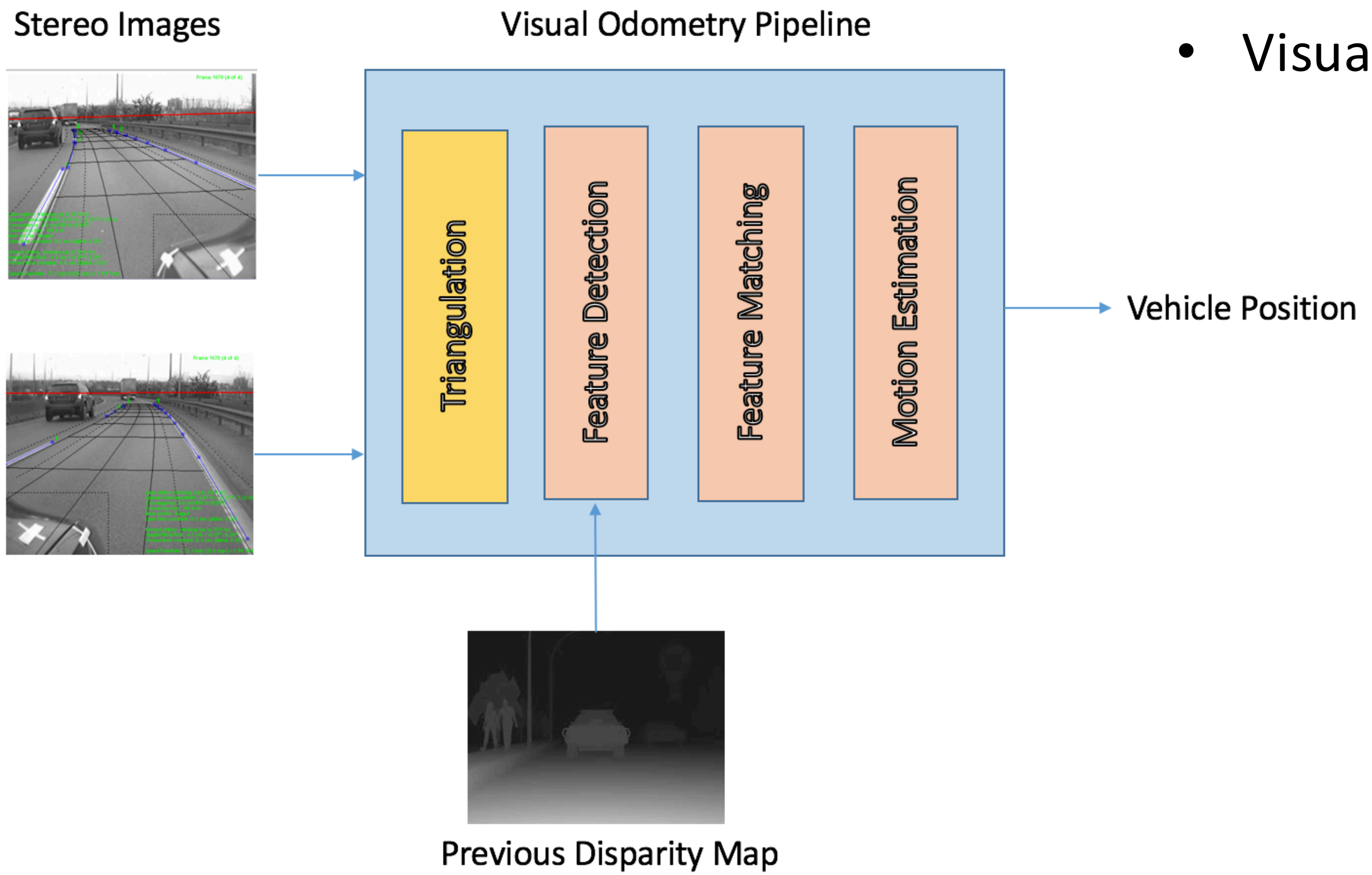
- Foundation layer is a 2D grid map with 5 cm by 5 cm resolution
- Road reference line is then added
- Next layer is the lane information
- Other semantic feature are then added

Requirements:

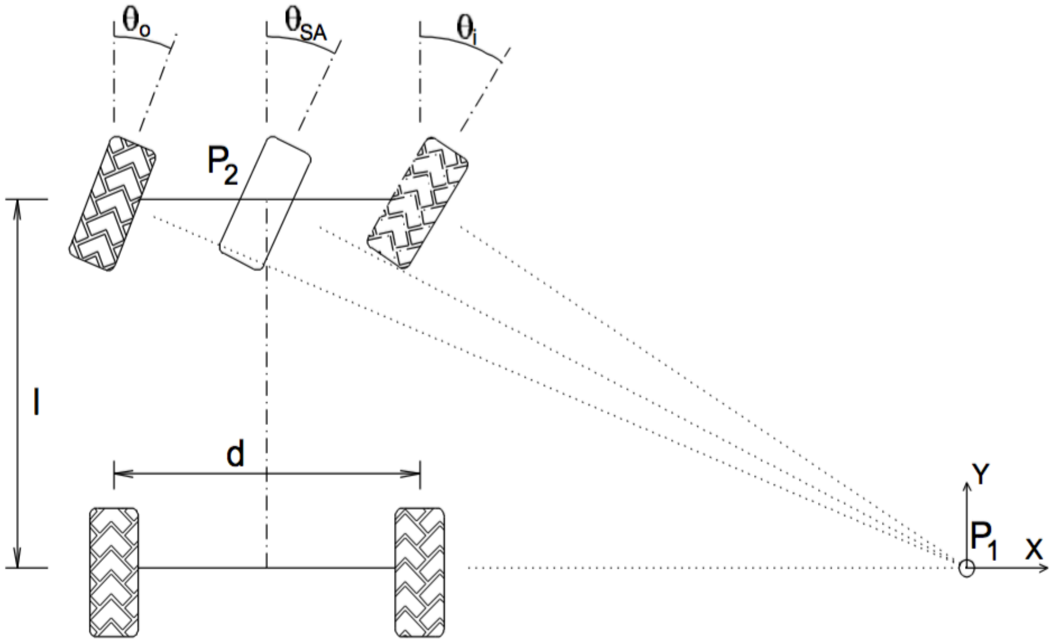
- Fresh
- Precise
- Integration with the client systems

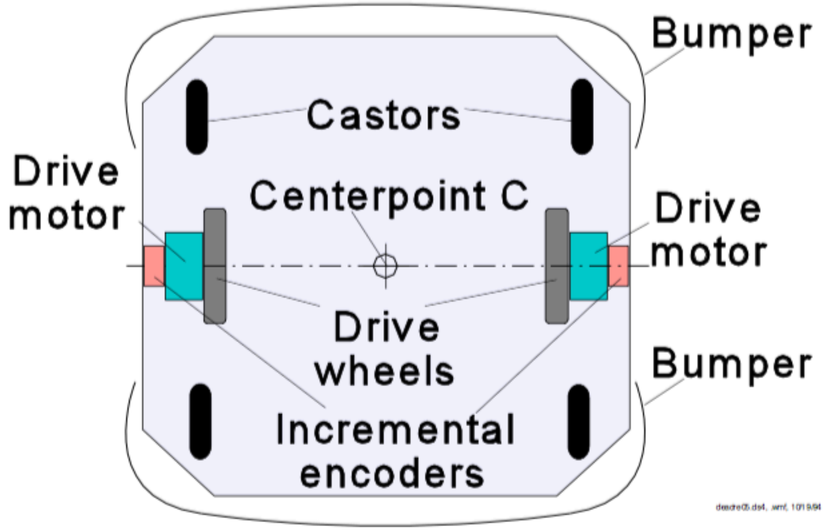
Localization: Visual Odometry

- Stereo visual odometry
- Mono visual odometry
- Visual inertial odometry

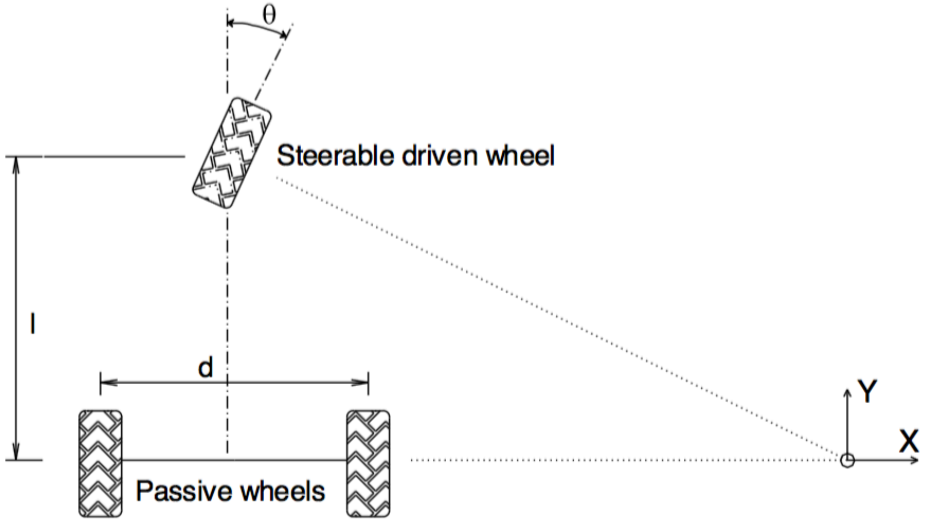


Localization: Wheel Odometry

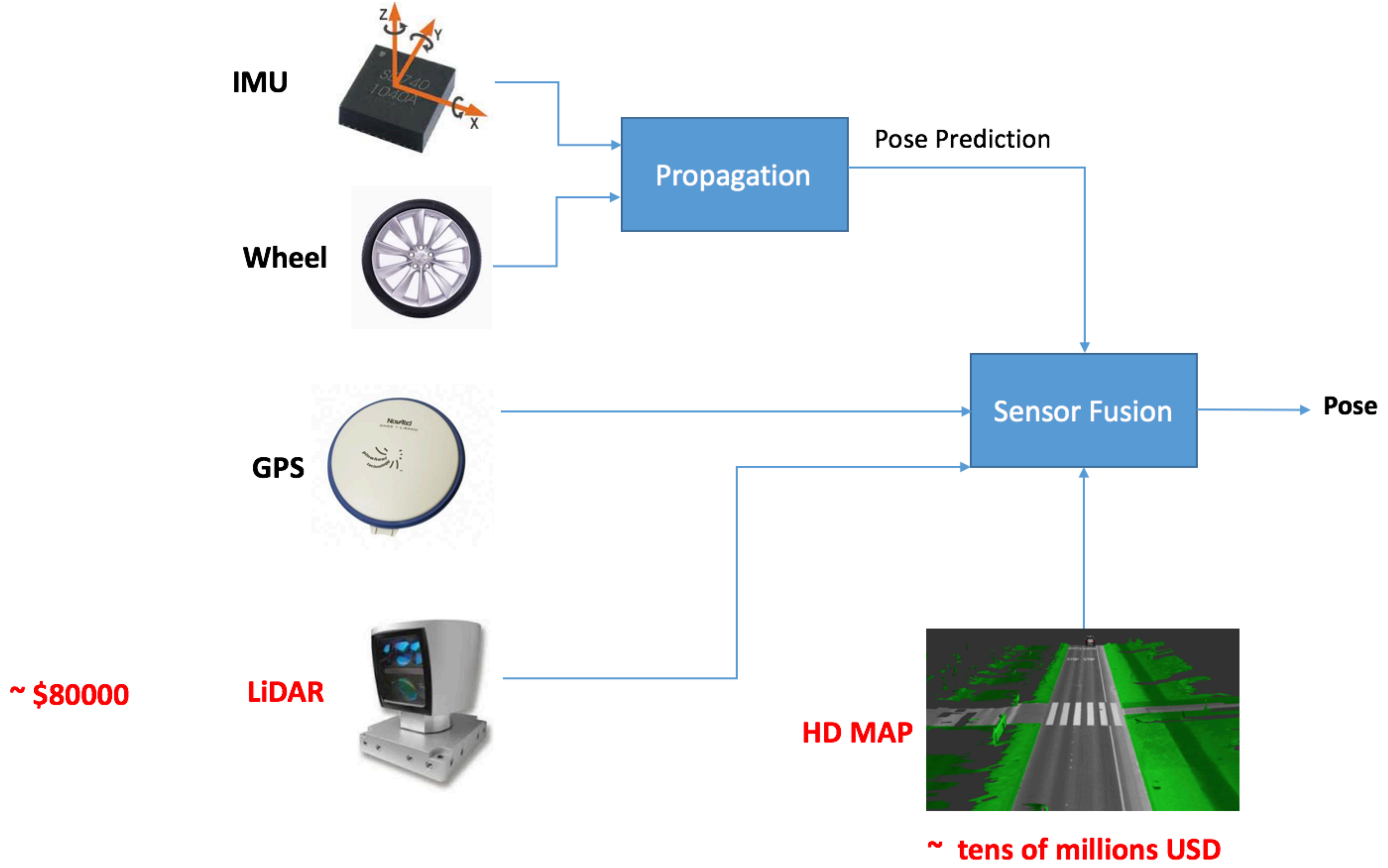




- Ackerman steering
- Differential drive
- Tricycle drive



Localization: Sensor Fusion I



Localization: Sensor Fusion

Module 2: Perception

Datasets

Geiger, Andreas ; Lenz, Philip ; Stiller, Christoph ; Urtasun, Raquel: Vision meets Robotics: The KITTI Dataset.
International Journal of Robotics Research (IJRR) 32 (2013), pp. 1229–1235

Vision meets Robotics: The KITTI Dataset

Andreas Geiger, Philip Lenz, Christoph Stiller and Raquel Urtasun

***Abstract*—We present a novel dataset captured from a VW station wagon for use in mobile robotics and autonomous driving research. In total, we recorded 6 hours of traffic scenarios at 10-100 Hz using a variety of sensor modalities such as high-resolution color and grayscale stereo cameras, a Velodyne 3D laser scanner and a high-precision GPS/IMU inertial navigation system. The scenarios are diverse, capturing real-world traffic situations and range from freeways over rural areas to inner-city scenes with many static and dynamic objects. Our data is calibrated, synchronized and timestamped, and we provide the rectified and raw image sequences. Our dataset also contains object labels in the form of 3D tracklets and we provide online benchmarks for stereo, optical flow, object detection and other tasks. This paper describes our recording platform, the data format and the utilities that we provide.**

***Index Terms*—dataset, autonomous driving, mobile robotics, field robotics, computer vision, cameras, laser, GPS, benchmarks, stereo, optical flow, SLAM, object detection, tracking, KITTI**

I. Introduction

The KITTI dataset has been recorded from a moving platform (Fig. 1) while driving in and around Karlsruhe, Germany (Fig. 2). It includes camera images, laser scans, high-precision GPS measurements and IMU accelerations from a combined

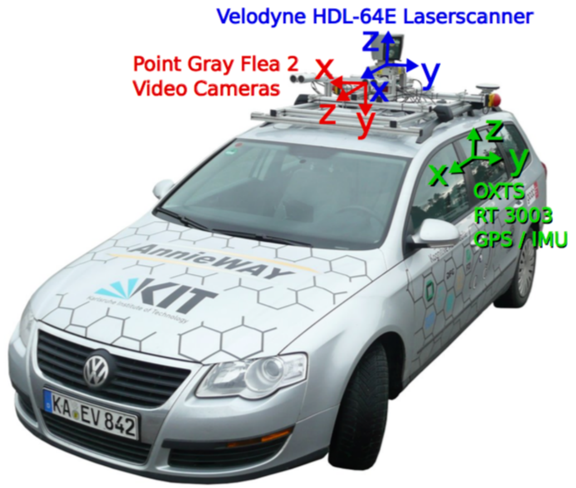
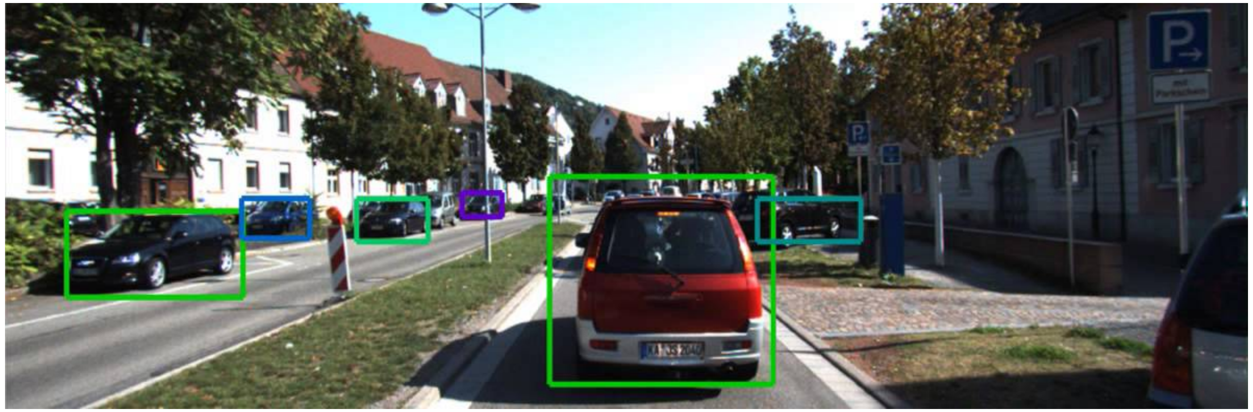
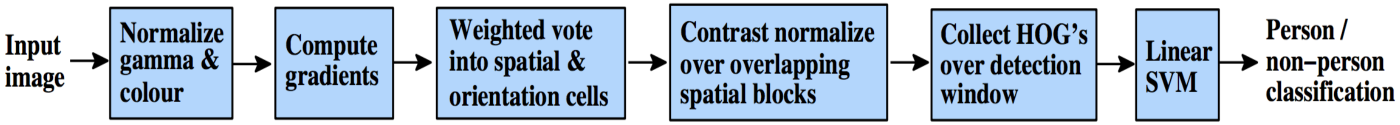


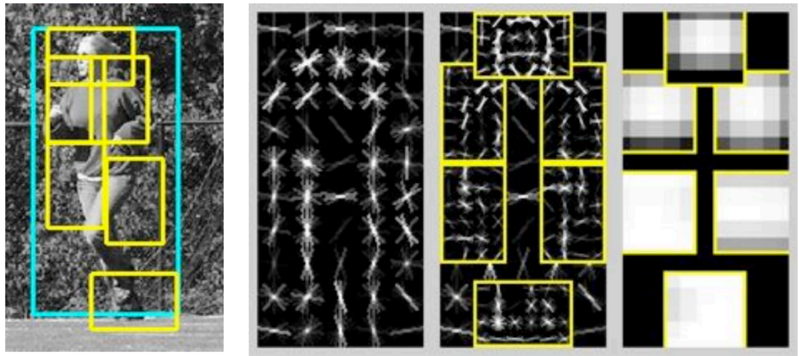
Fig. 1. **Recording Platform.** Our VW Passat station wagon is equipped with four video cameras (two color and two grayscale cameras), a rotating 3D laser scanner and a combined GPS/IMU inertial navigation system.

- 1 × OXTS RT3003 inertial and GPS navigation system, 6 axis, 100 Hz, L1/L2 RTK, resolution: 0.02m / 0.1°

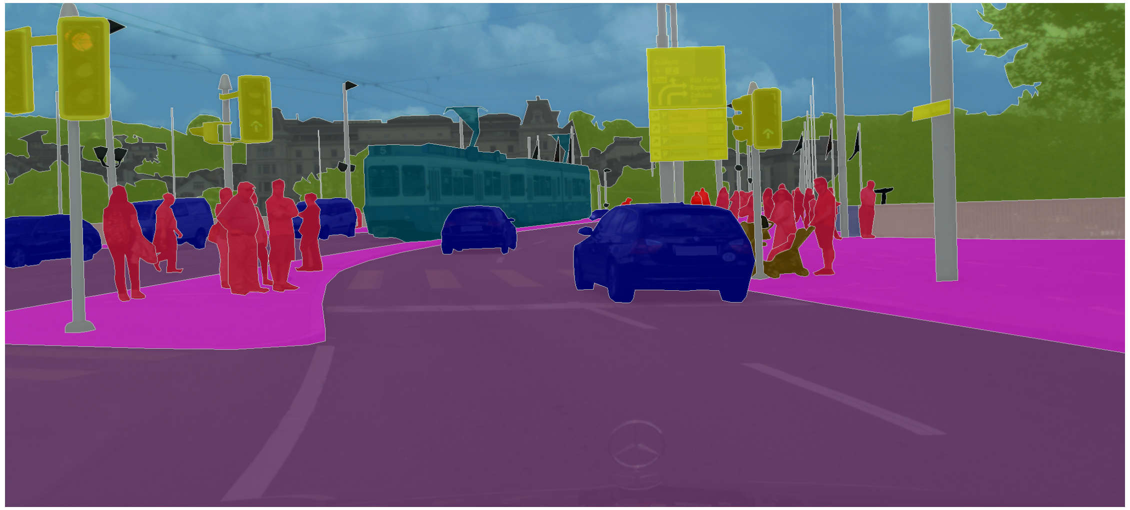
Object Detection

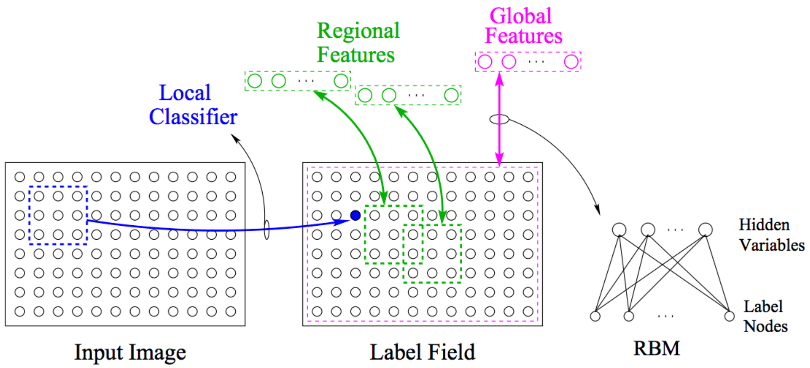




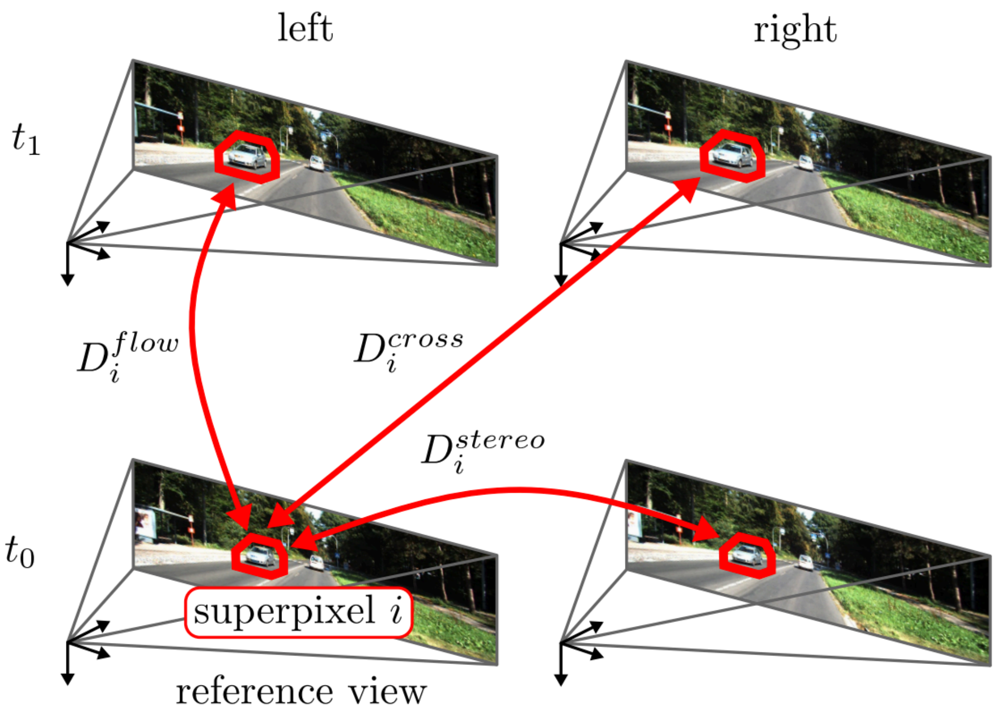


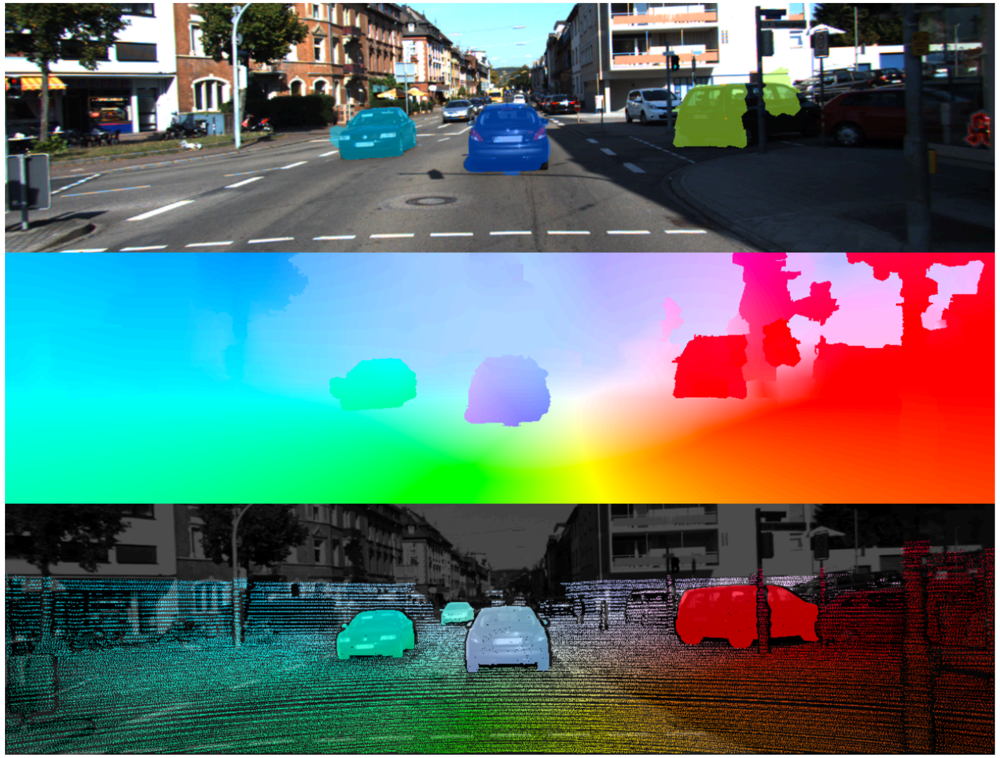
Segmentation



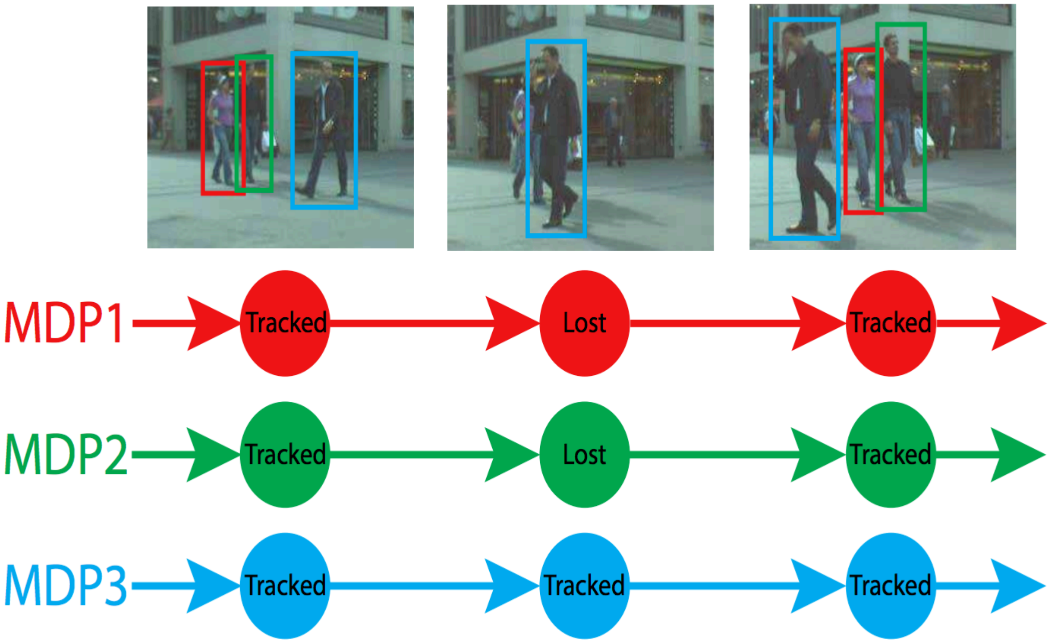


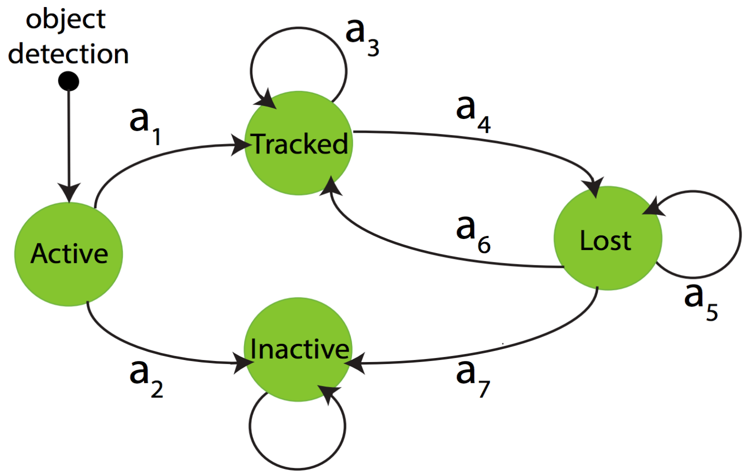
Stereo, Optical Flow, and Scene Flow



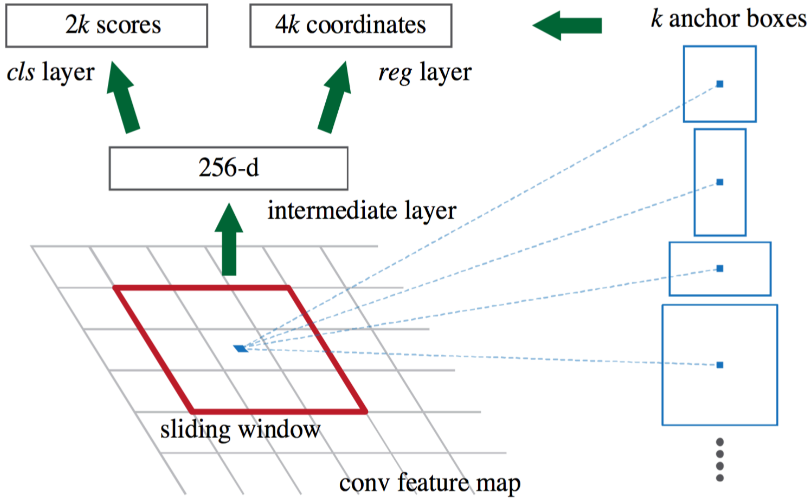


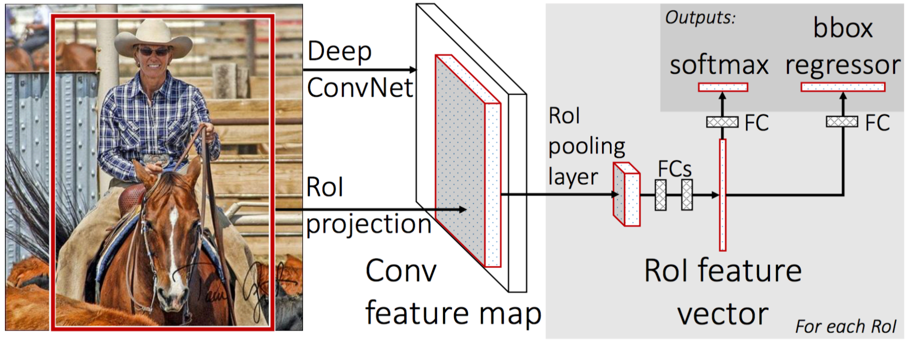
Object Tracking

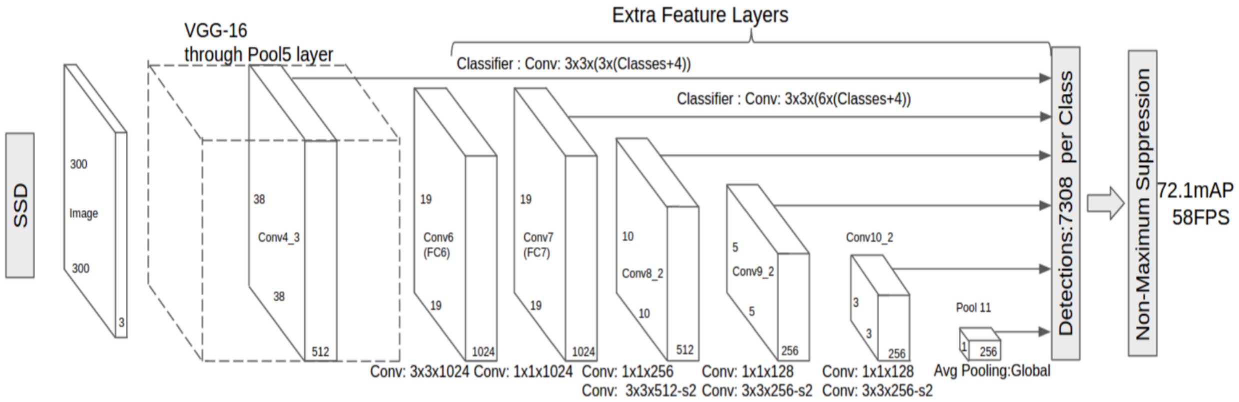




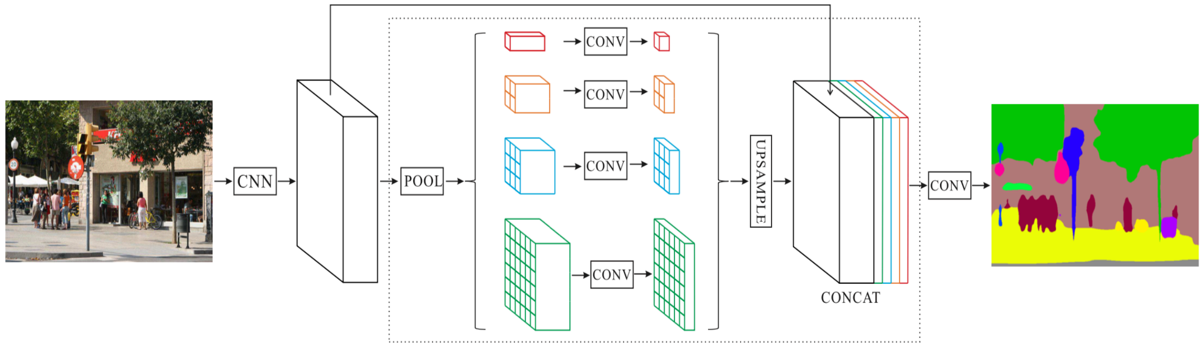
Detection: Faster R-CNN and SSD



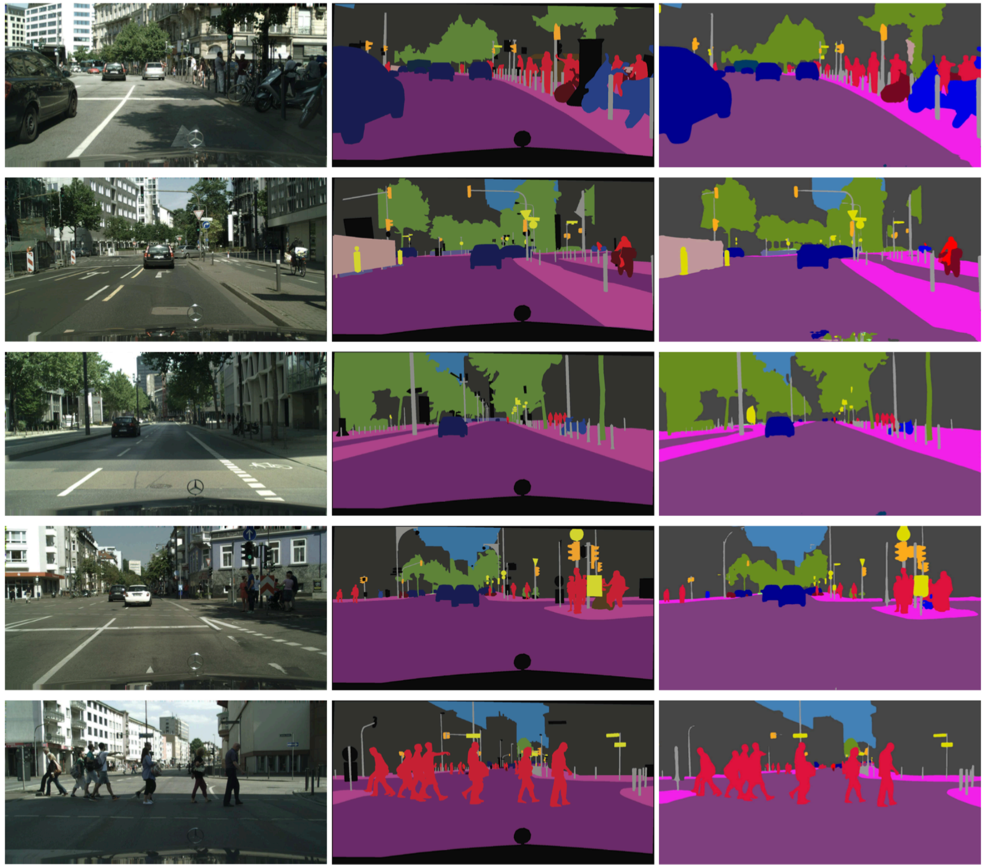




Semantic Segmentation

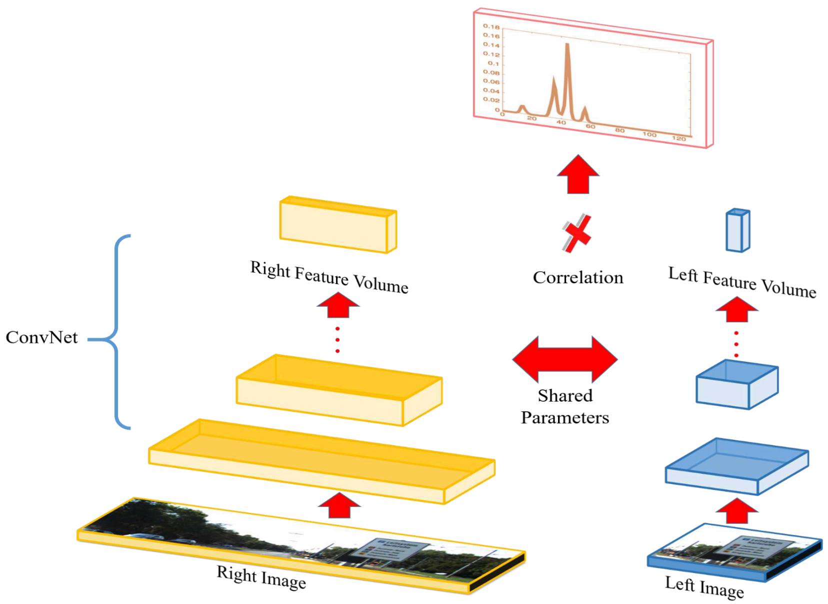


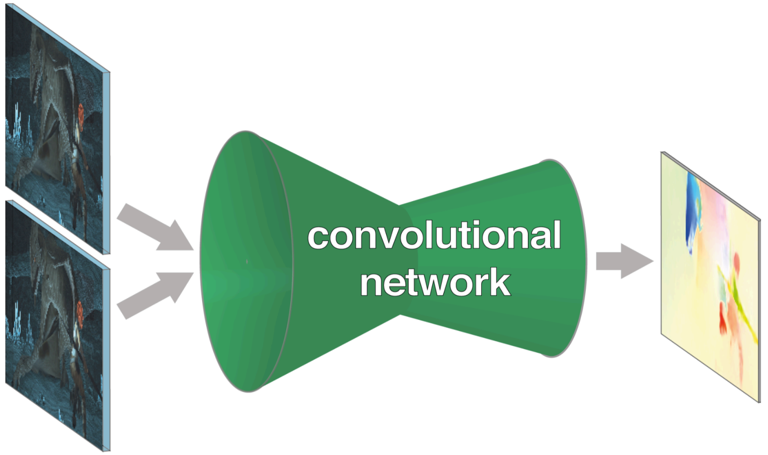
(a) Input Image (b) Feature Map (c) Pyramid Pooling Module (d) Final Prediction

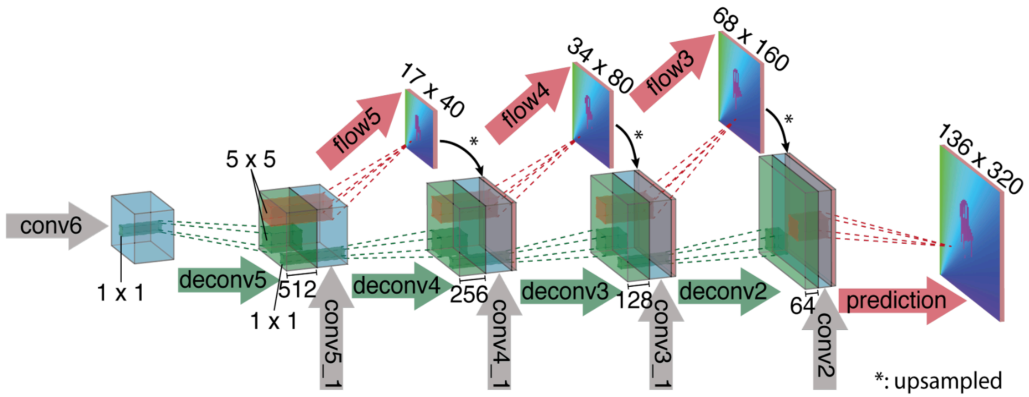


(a) Image (b) Ground Truth (c) PSPNet

Stereo and Optical Flow

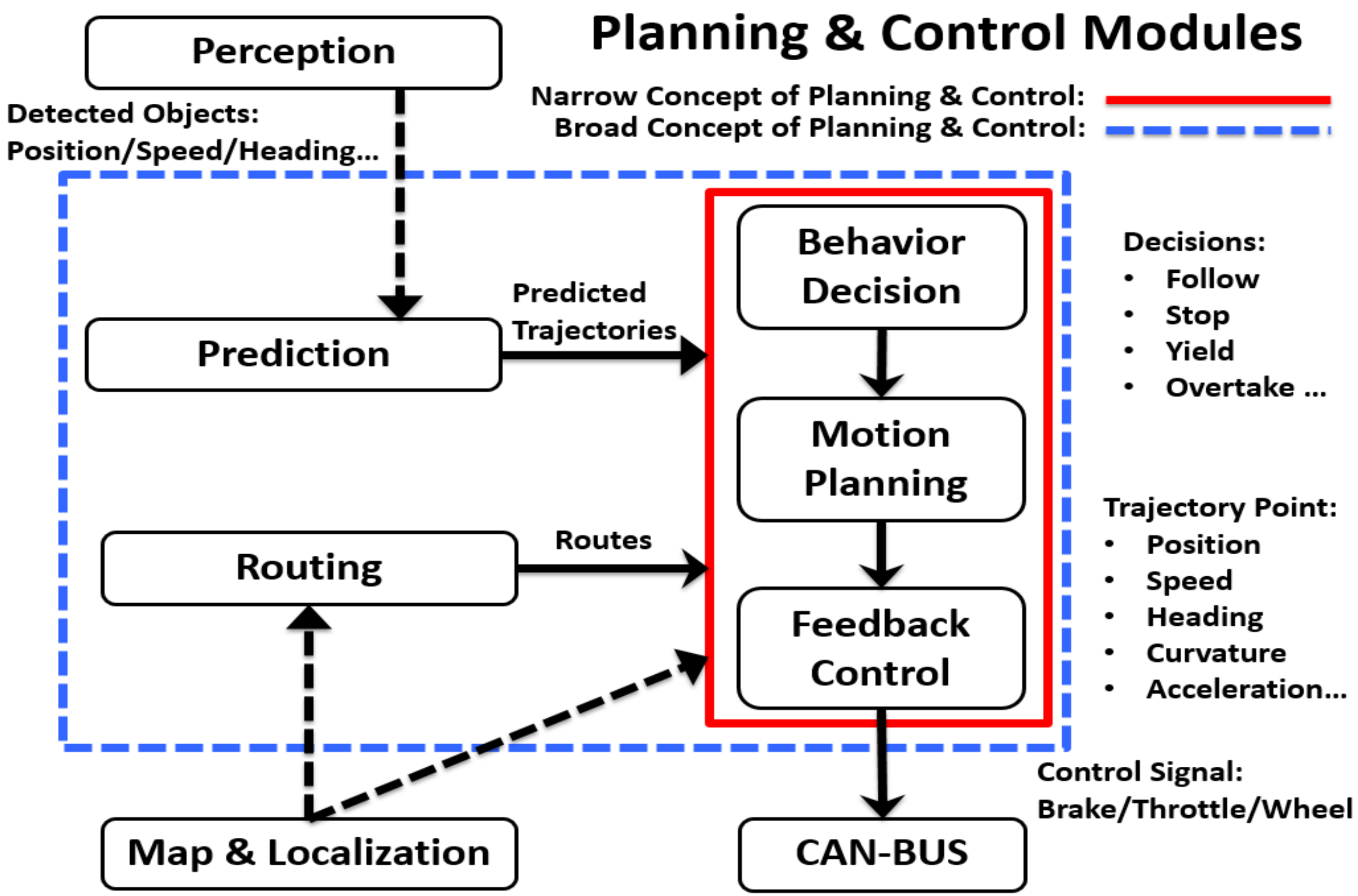






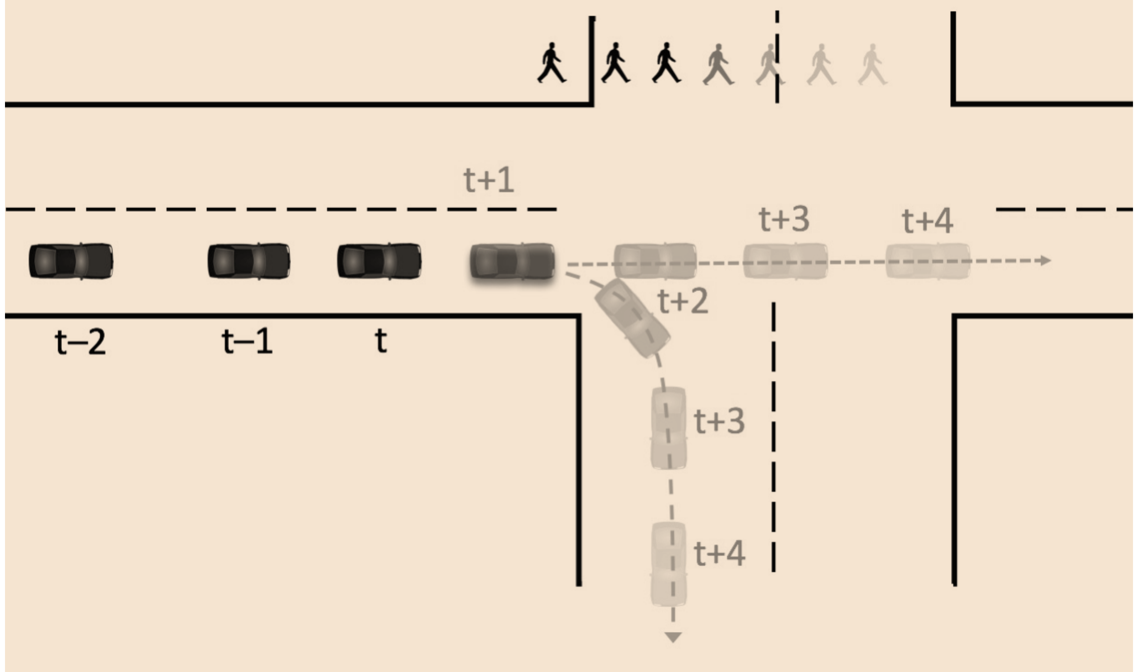
Module 3: Planning and Control

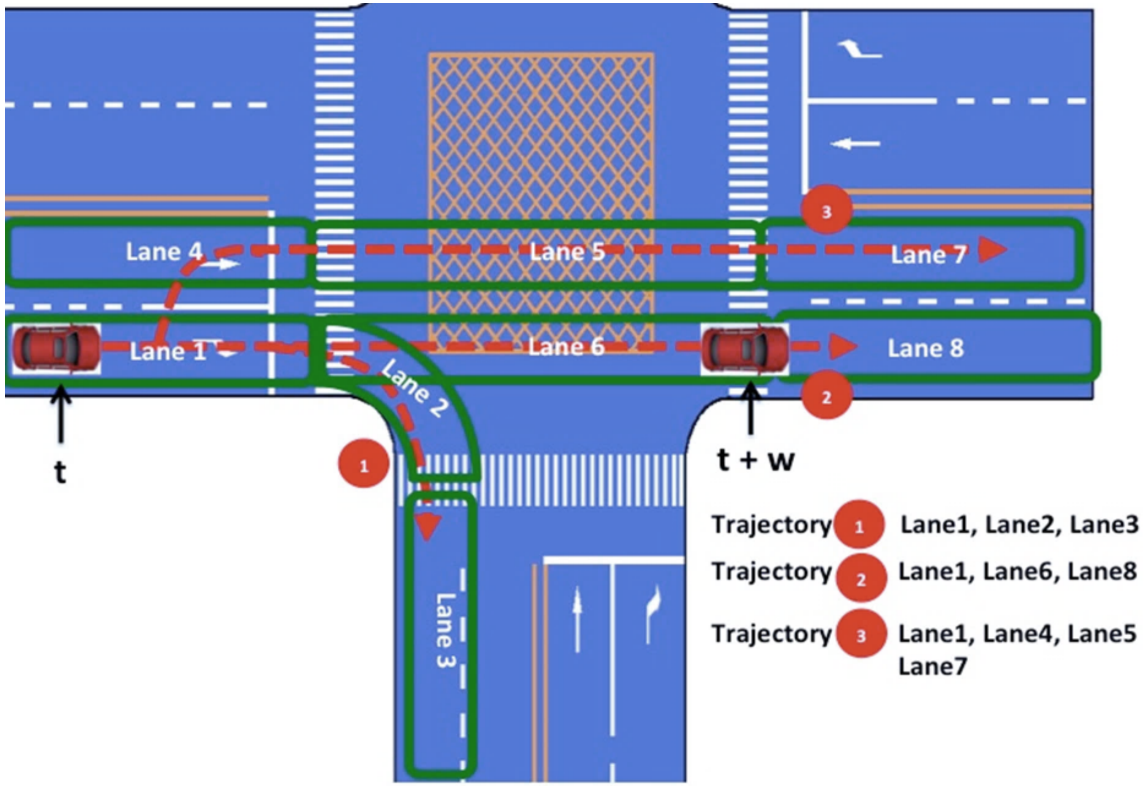
Architecture



Traffic Prediction

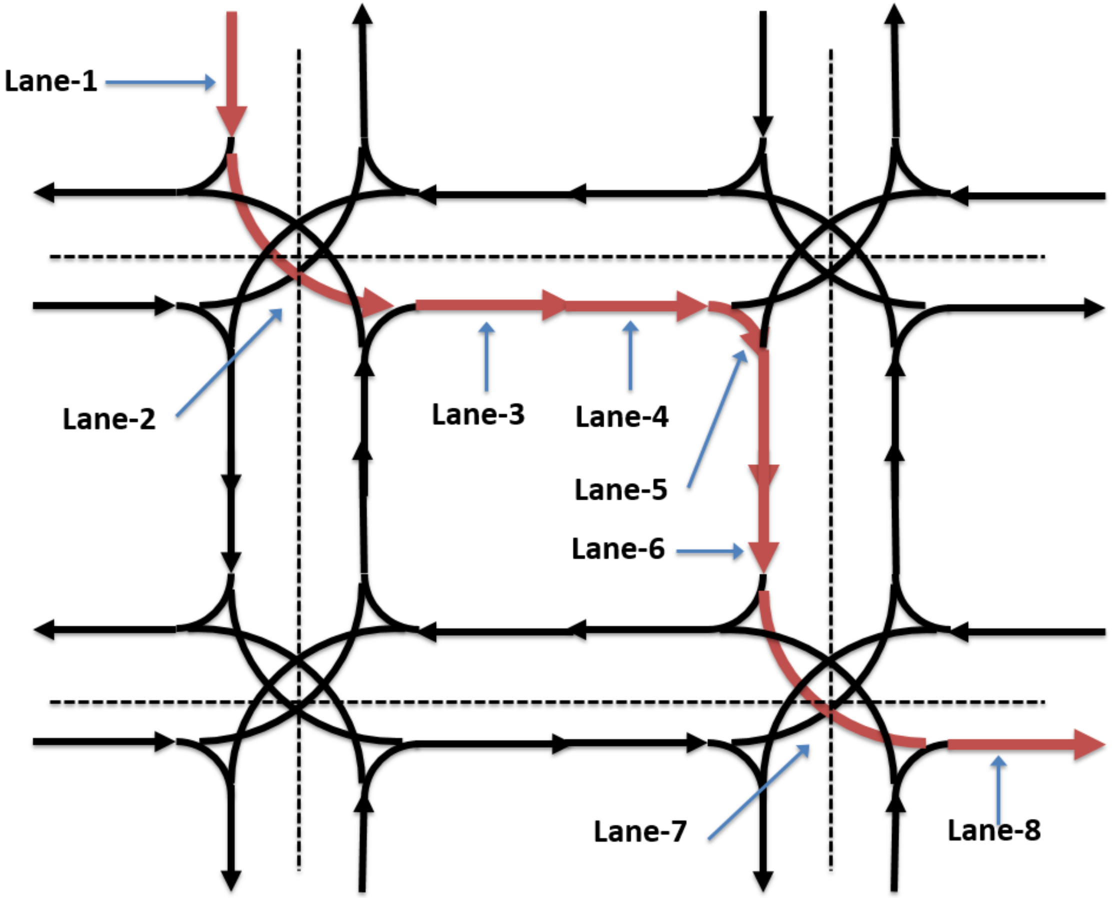
- *Classification problem* for categorical road object behaviors
- *Regression problem* for generating the predicted path with speed and time info

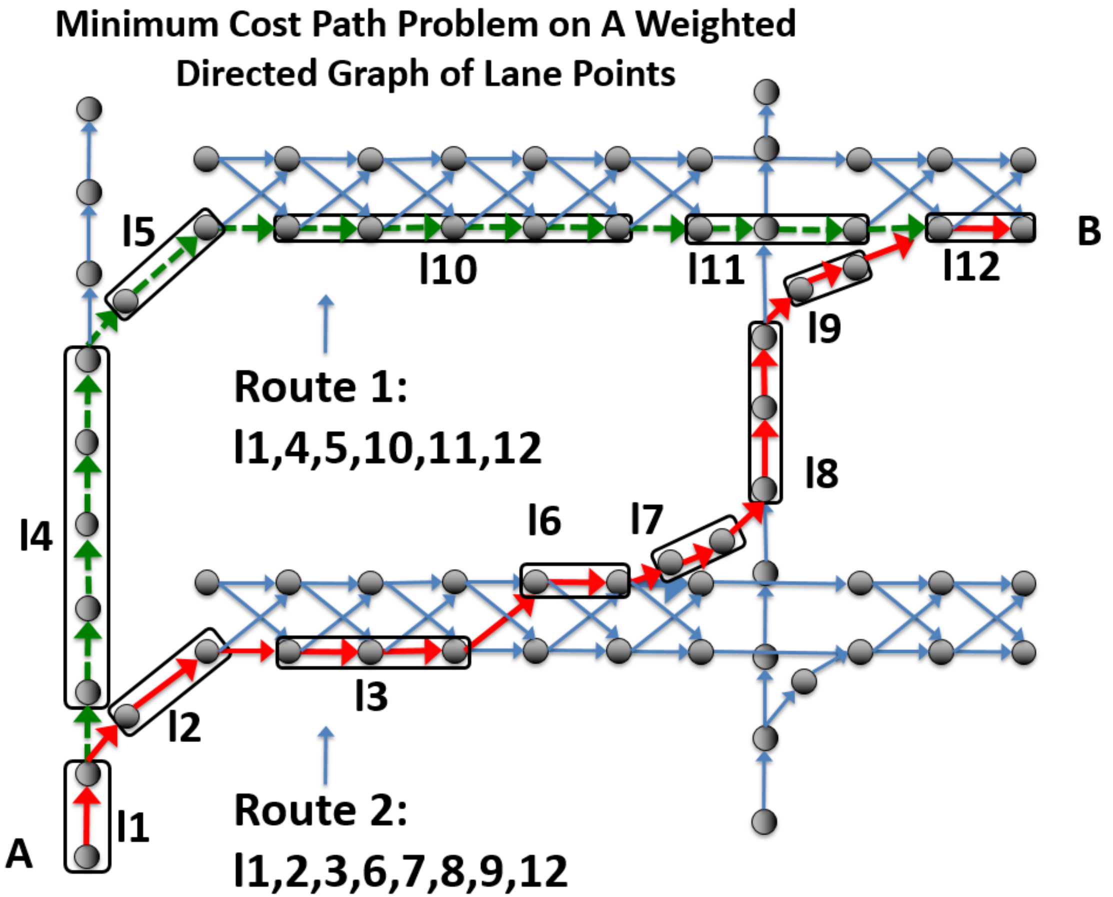




Lane-Level Routing

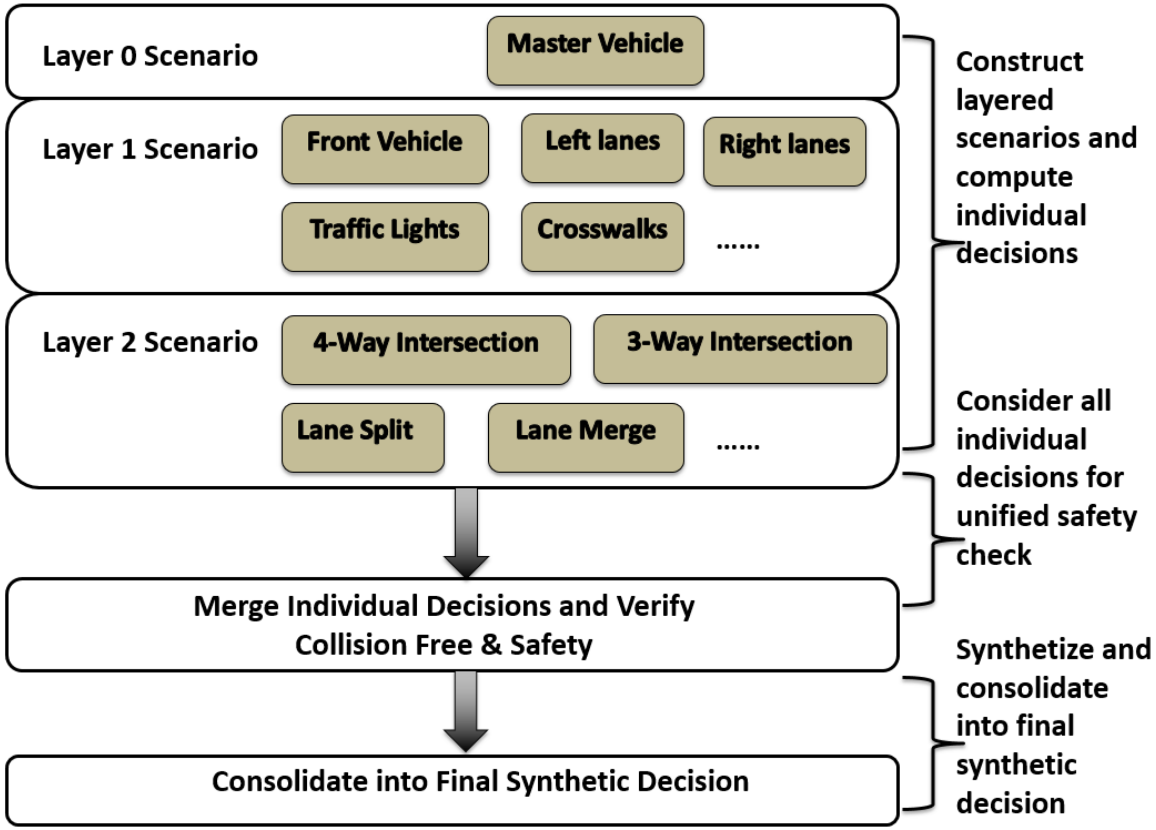
- Modeled as weighted directed graphs
- Shortest path problem: *Dijkstra* and *A**

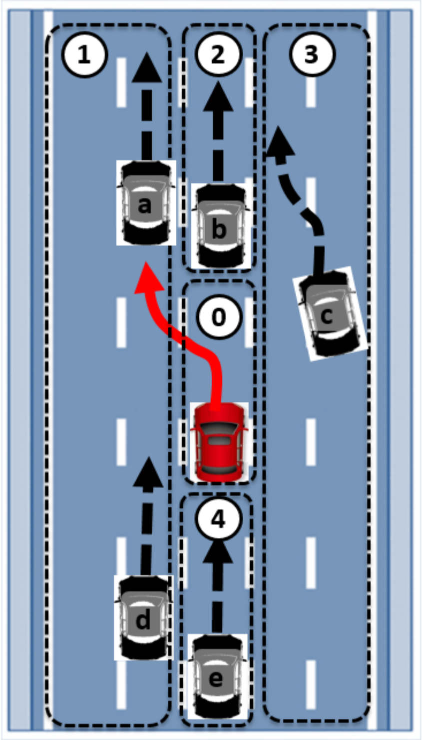




Behavioral Decisions

- Ruled-based “divide-and-conquer” approach: layered scenarios
- Markov Decision Process
- Synthetic decision and individual decisions



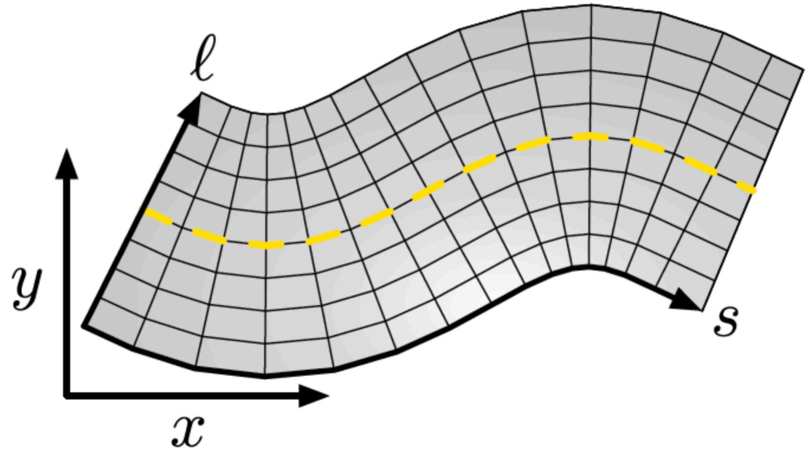


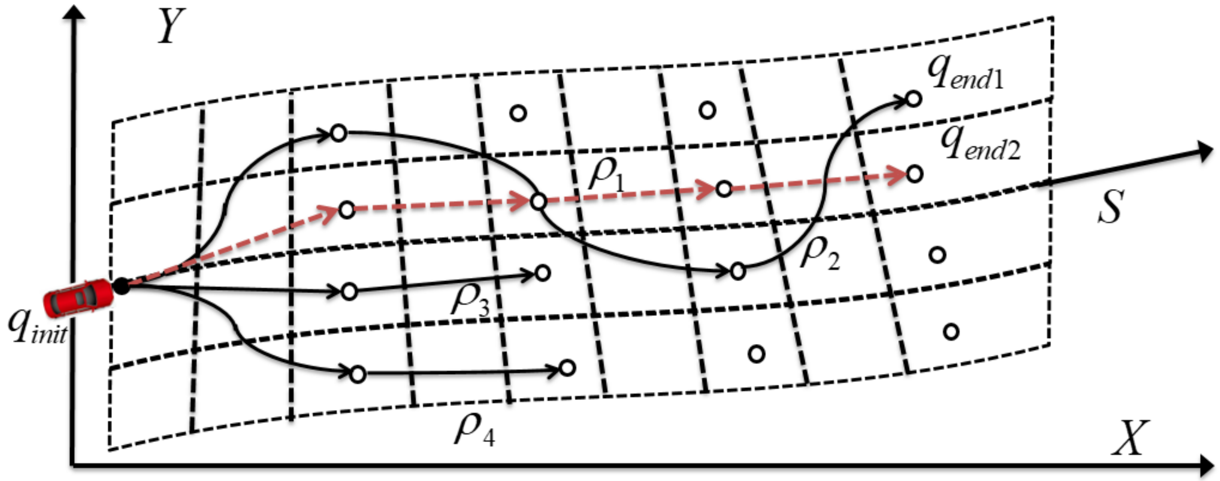
Synthetic Decision:
Switch lane from current lane to the left lane: yield vehicle a, overtake vehicle d, and attention to vehicle b at current lane

Scenarios and Individua Decisions:
0.Master Vehicle
1.Left Lane(s) : Overtake d and yield a
2.Front Vehicle(s): Attention b
3.Right Lane(s): Ignore c
4.Rear Vehicles(s): Ignore e

Motion Planning

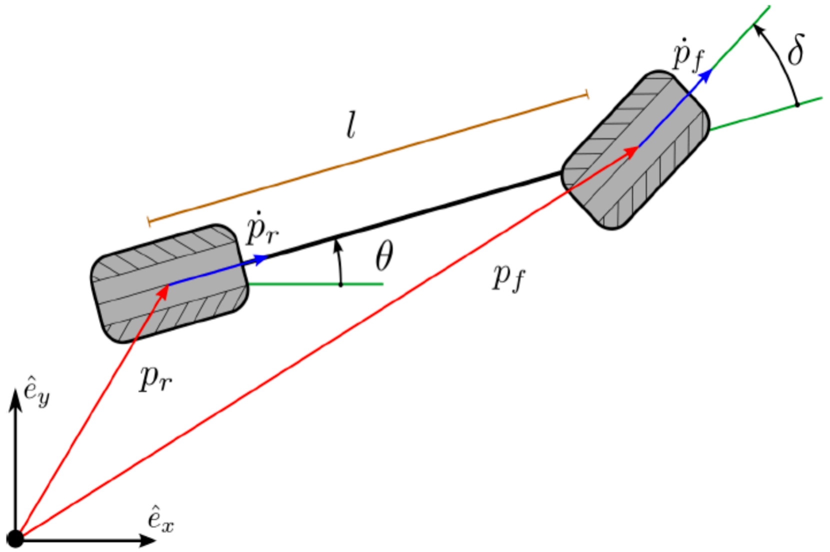
- Vehicle model, road model, and SL-coordination system
- Path planning and speed planning
  - Path planning: dynamic programming to minimize cost
  - Speed planning: ST-graph

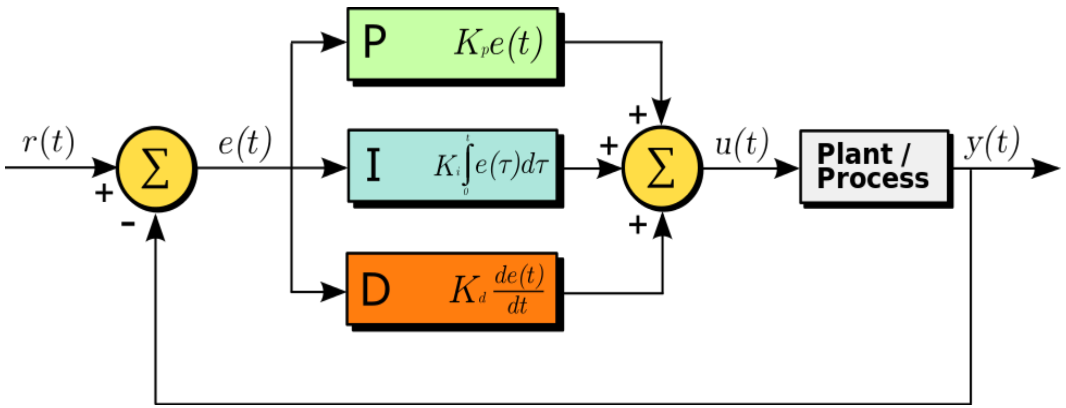




Feedback Control

- Bicycle model: model vehicles as rigid bodies with front and rear wheels
- PID control: proportional-integral-derivative controller



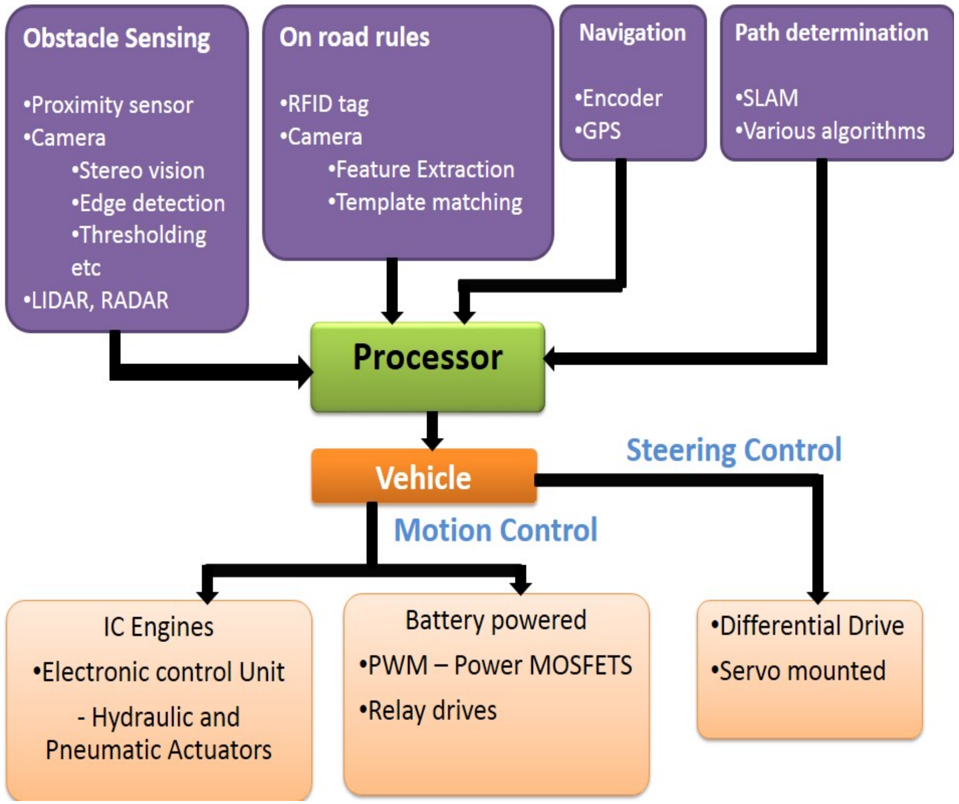


Module 4: Client Systems

Operating System

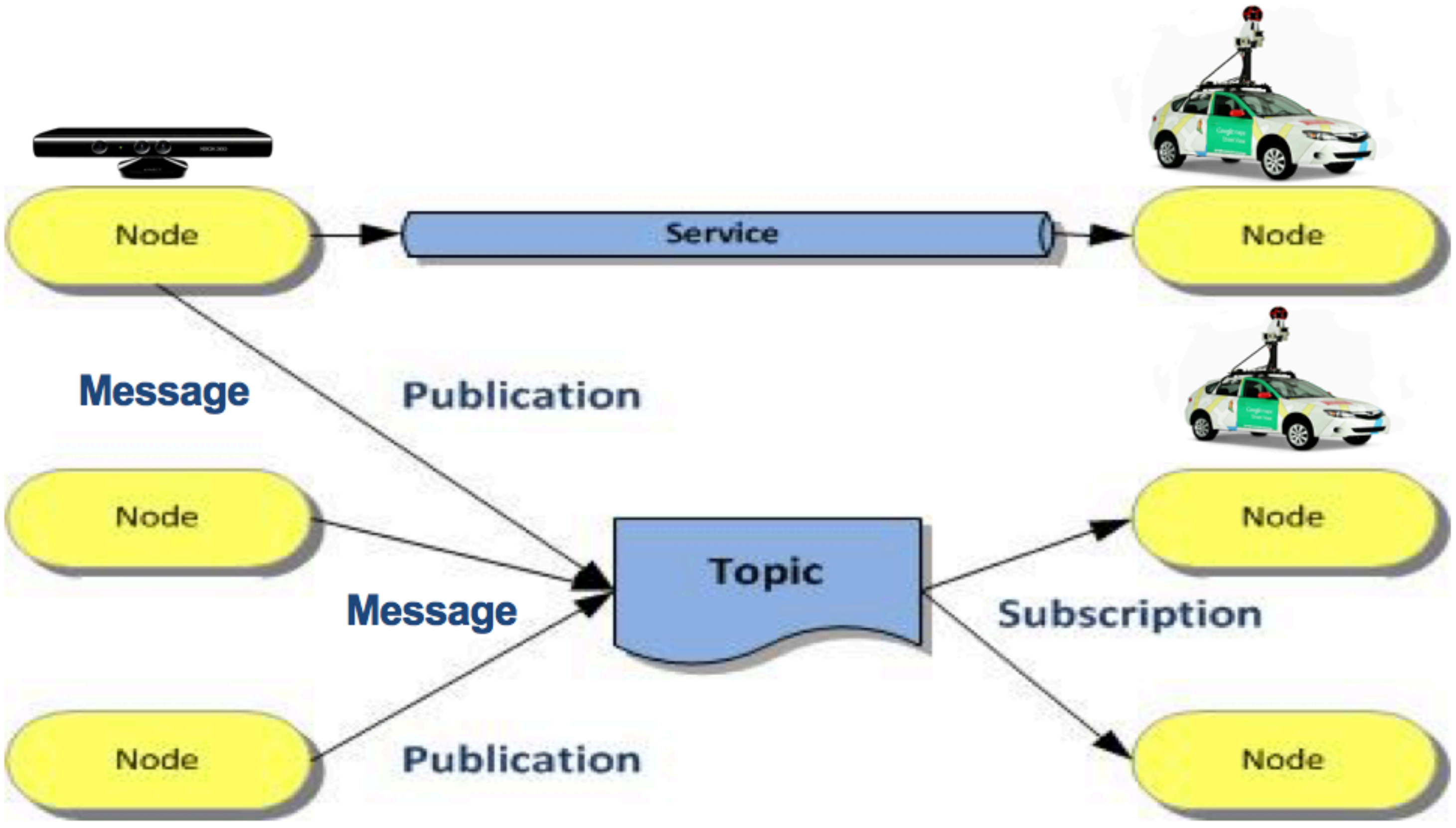
It takes a robust operating system to handle the complex interactions between different components

e.g. 60 FPS requirements for camera, leaving 16 ms for each frame, huge challenge when the amount of data becomes overwhelming



Operating System: ROS

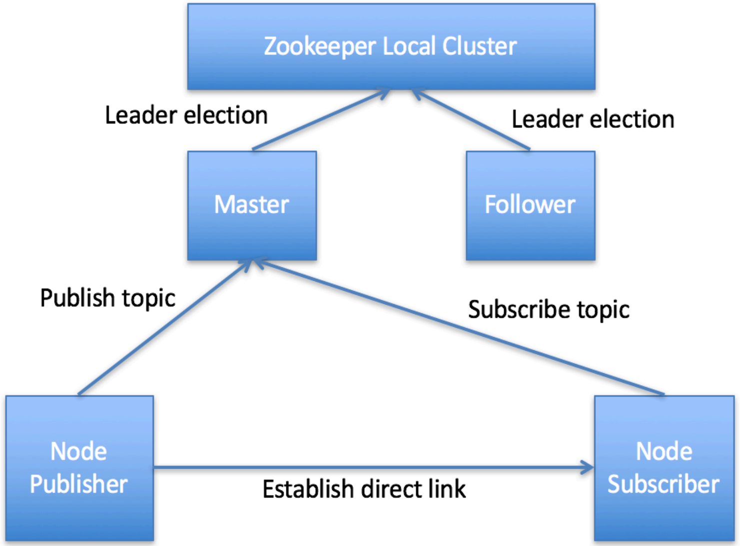
ROS provides a mechanism for the components to interact with each other

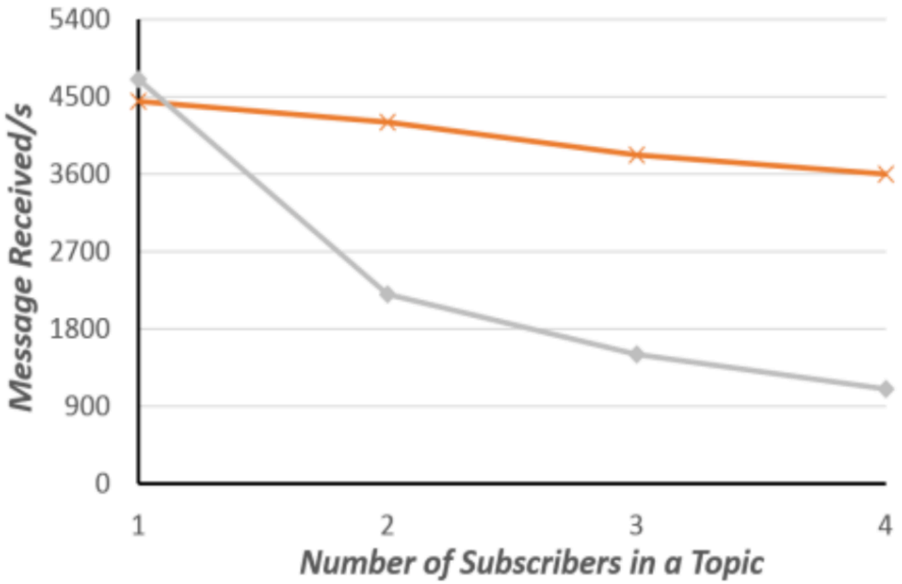


Operating System: ROS

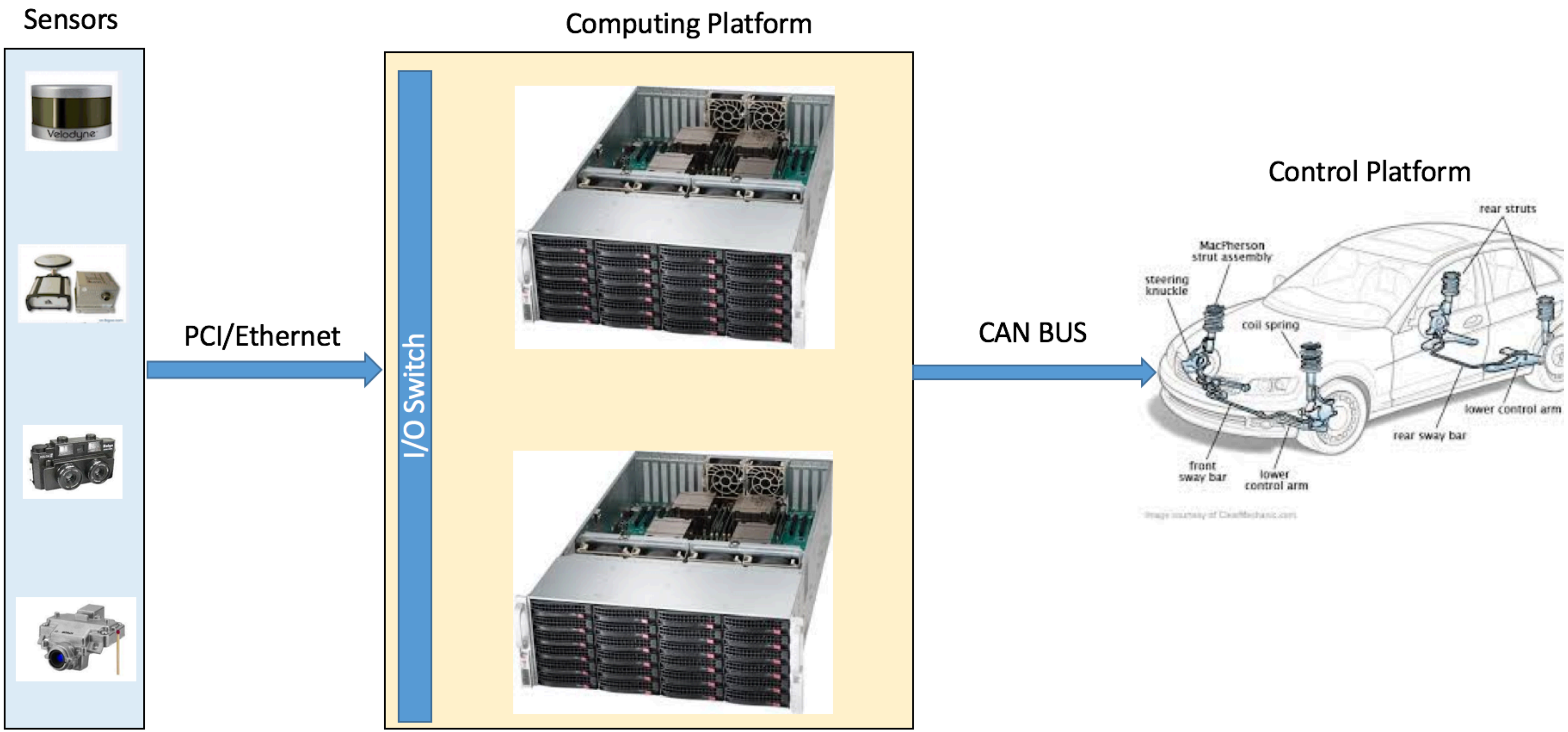
ROS Problems

- Easily crashed
- Communication becomes the bottleneck
- Not secure



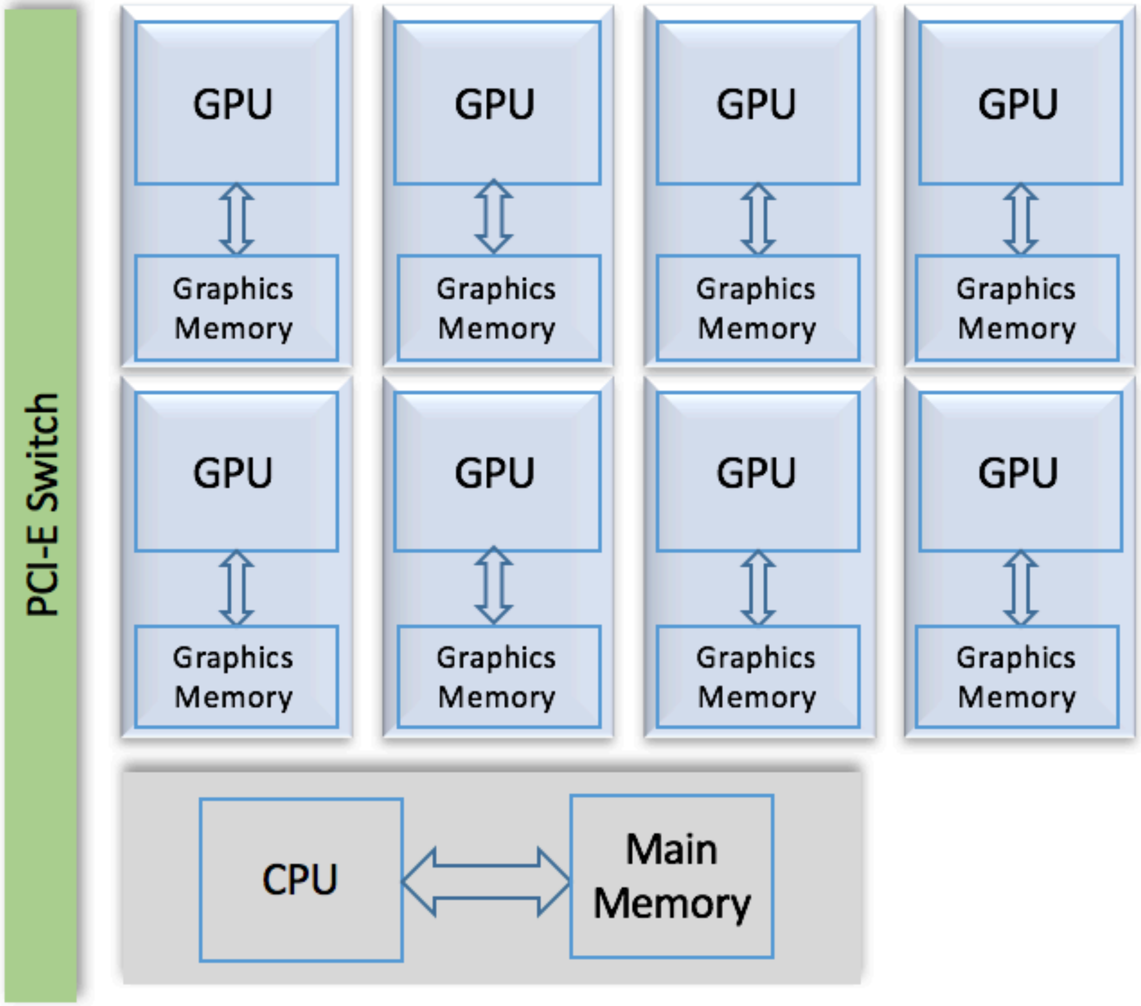


Hardware Platform



Hardware Platform

- High Performance
  - CPU + 8 ~ 16 GPUs
  - 60 TOPS/s
- High-Power Consumption
  - 3000 W at peak
- High Cost
  - $20000 ~ $30000
- Heat Dissipation
  - Special fan design needed



Autonomous Driving: on mobile SoC ?

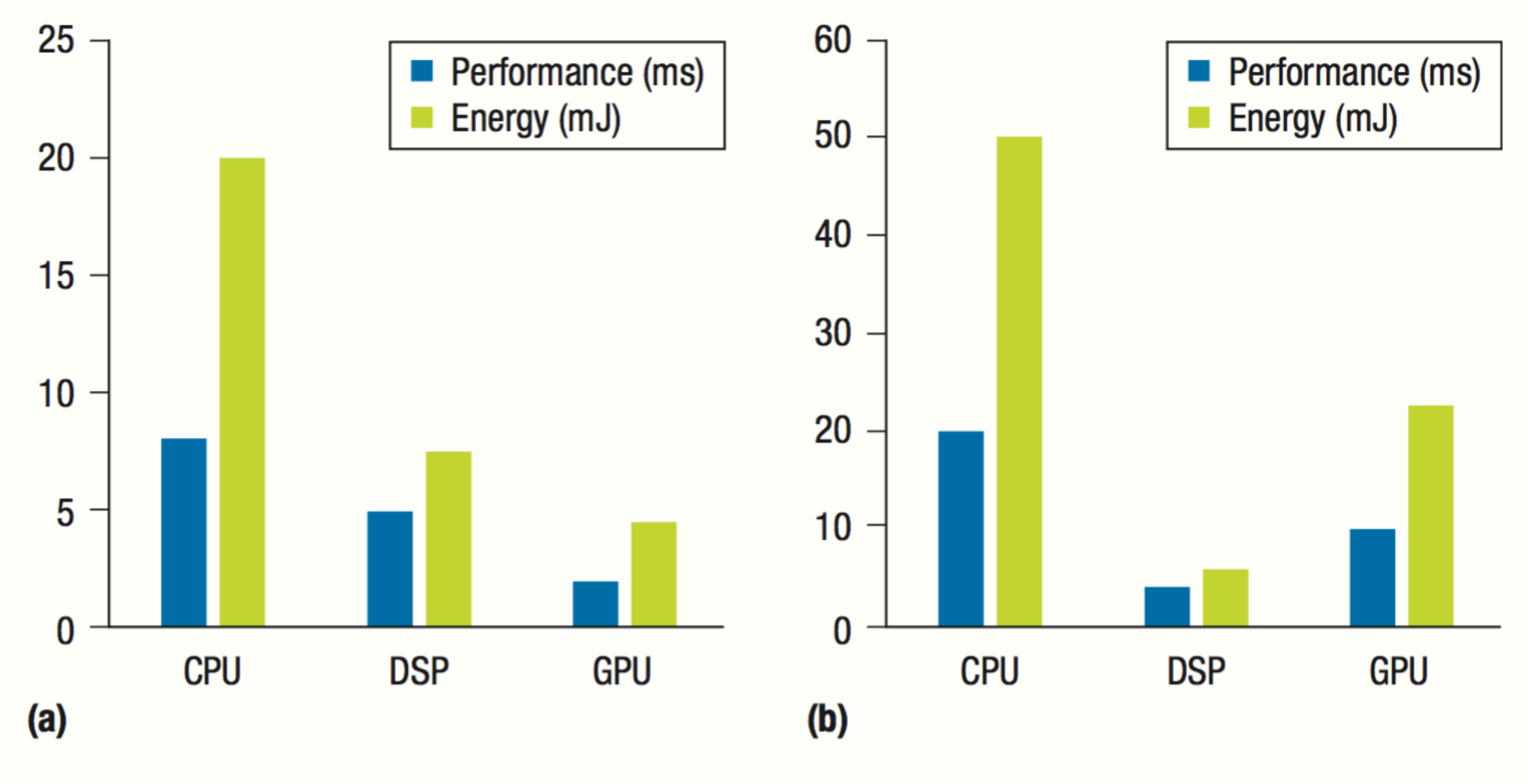
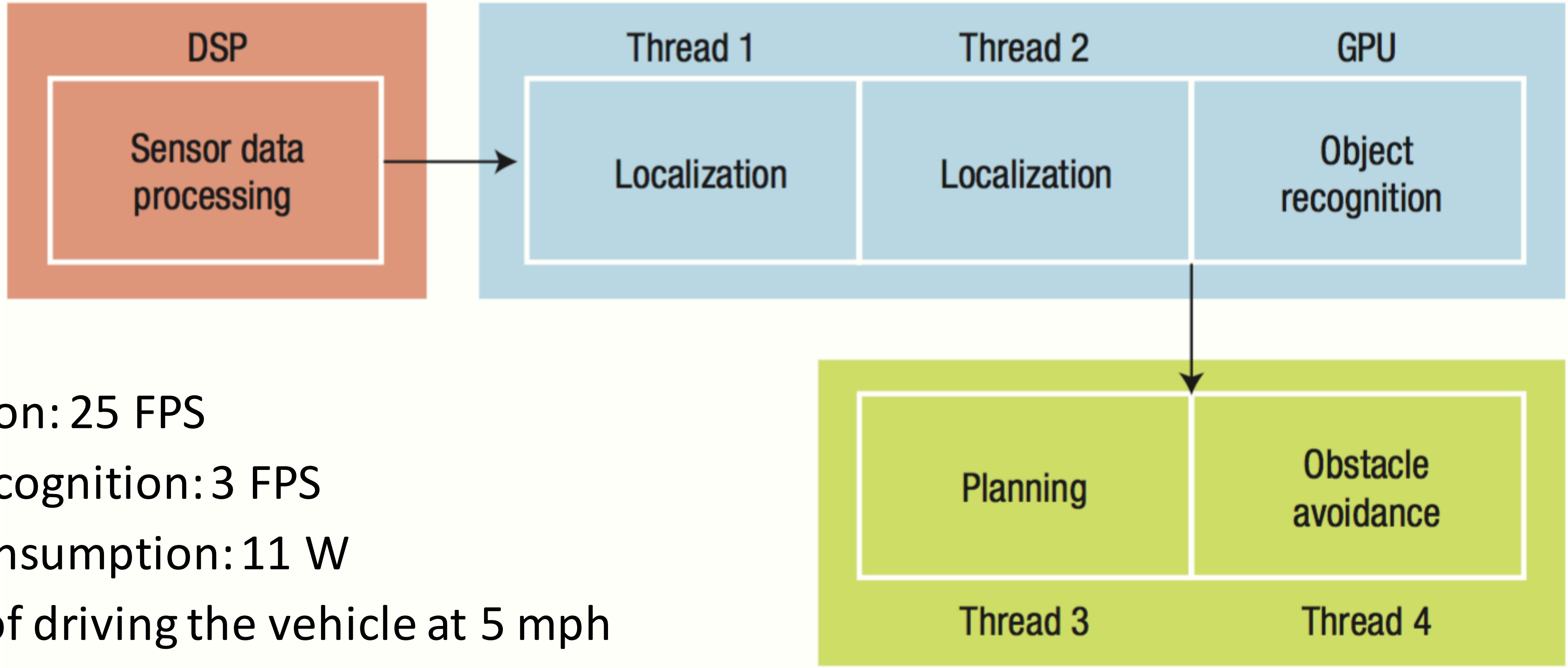


FIGURE 2. Performance and energy in (a) convolution and (b) feature-extraction tasks. In (a), the GPU takes only 2 ms and uses only 4.5 milliJoules (mJ) to complete convolution tasks. In (b), the digital signal processor (DSP) is the most efficient unit for feature extraction, taking 4 ms and consuming only 6 mJ to complete a task.

Mobile SoC:

- Quad-core CPU @ 2.2 GHz
- Hexagon 680 DSP
- Adreno 530 GPU
- Peak power ~ 15 W

Autonomous Driving: Heterogeneous Computing



- Localization: 25 FPS
- Object recognition: 3 FPS
- Power consumption: 11 W
- Capable of driving the vehicle at 5 mph

Module 5: Cloud Platforms

Cloud Infrastructure

Don't forget the cloud

arXiv.org > cs > arXiv:1704.02696

Search or Article ID inside arXiv | All papers | Broaden y

(Help | Advanced search)

Computer Science > Distributed, Parallel, and Cluster Computing

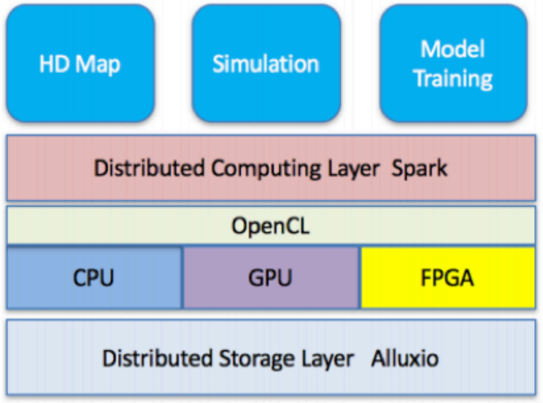
Implementing a Cloud Platform for Autonomous Driving

Shaoshan Liu, Jie Tang, Chao Wang, Quan Wang, Jean-Luc Gaudiot

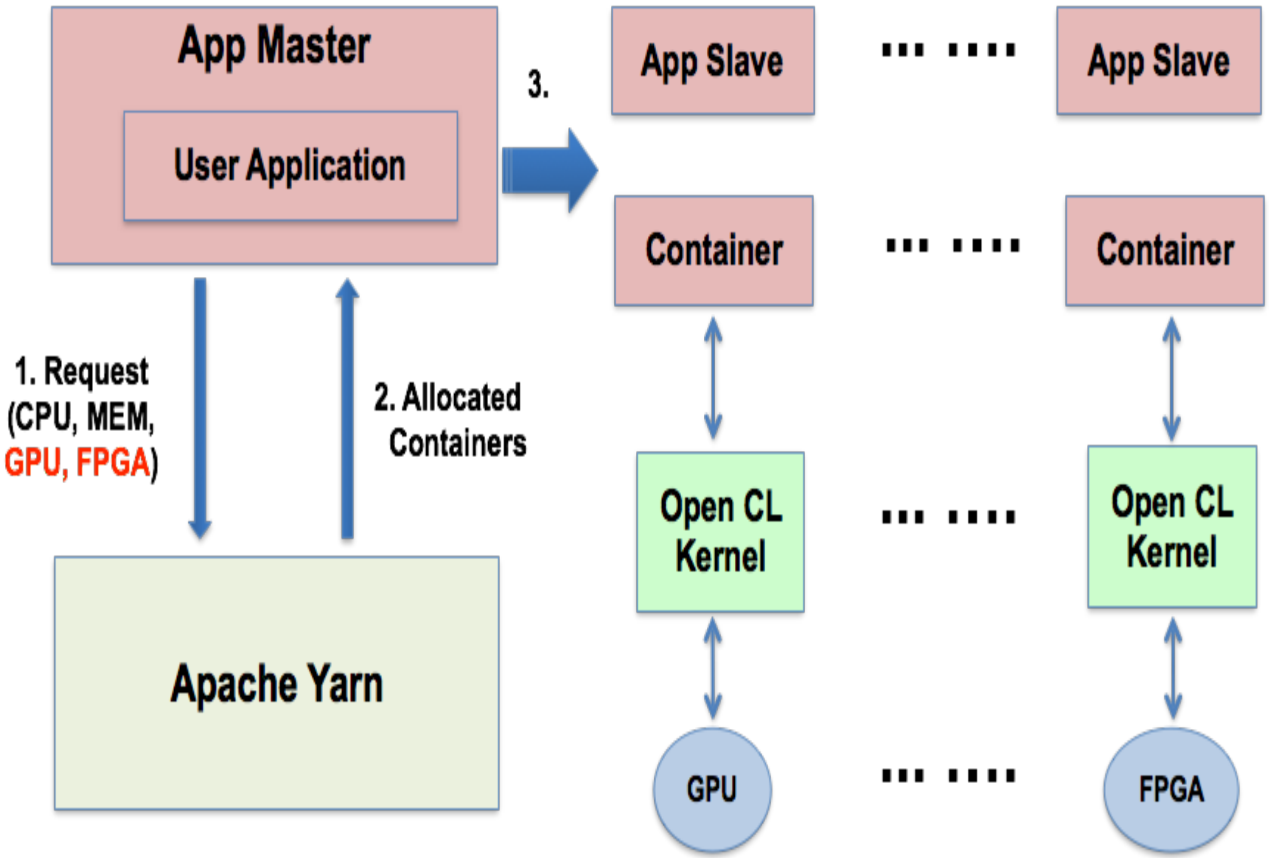
(Submitted on 10 Apr 2017)

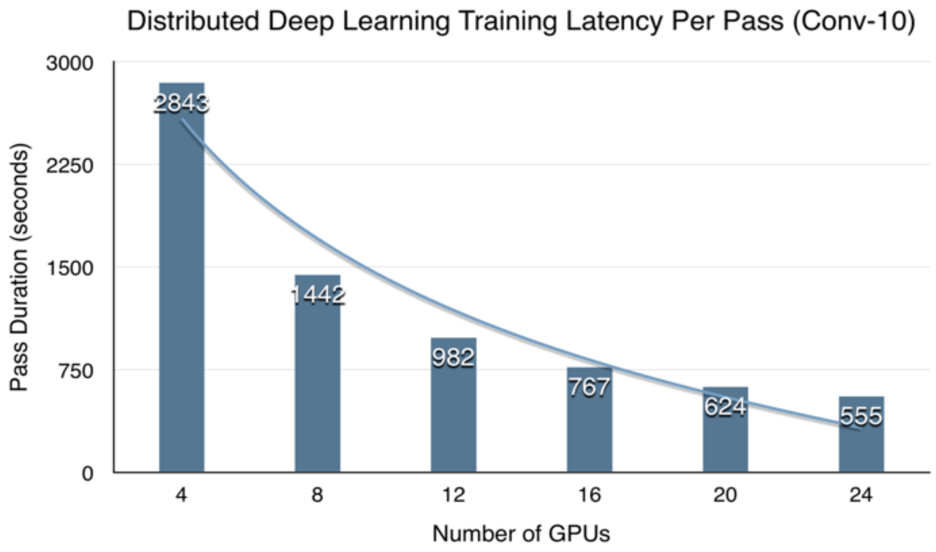
Autonomous driving clouds provide essential services to support autonomous vehicles. Today these services include but not limited to distributed simulation tests for new algorithm deployment, offline deep learning model training, and High-Definition (HD) map generation. These services require infrastructure support including distributed computing, distributed storage, as well as heterogeneous computing. In this paper, we present the details of how we implement a unified autonomous driving cloud infrastructure, and how we support these services on top of this infrastructure.

- Up to 2 GB / sec raw data generation
- Generated raw data used in different applications
- How do we make sense of the enormous amount of data?

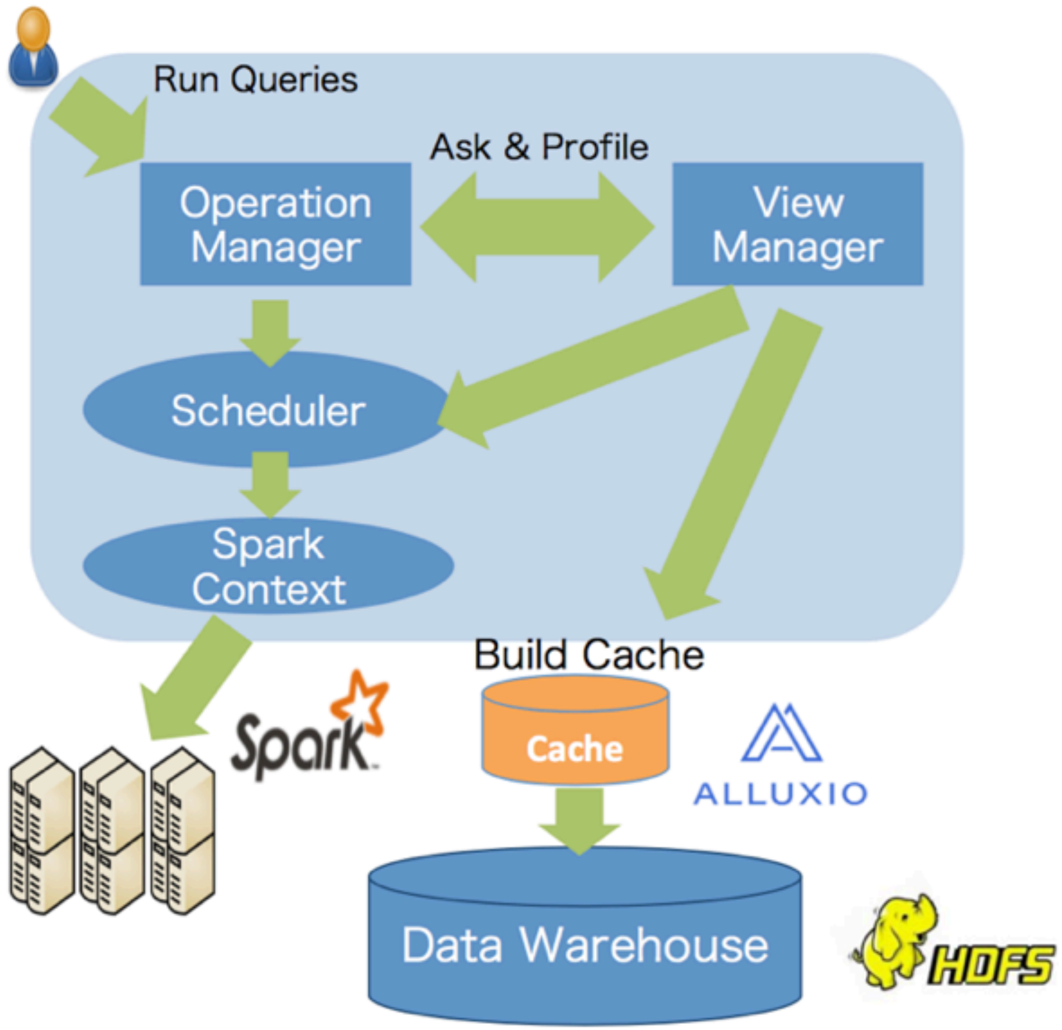


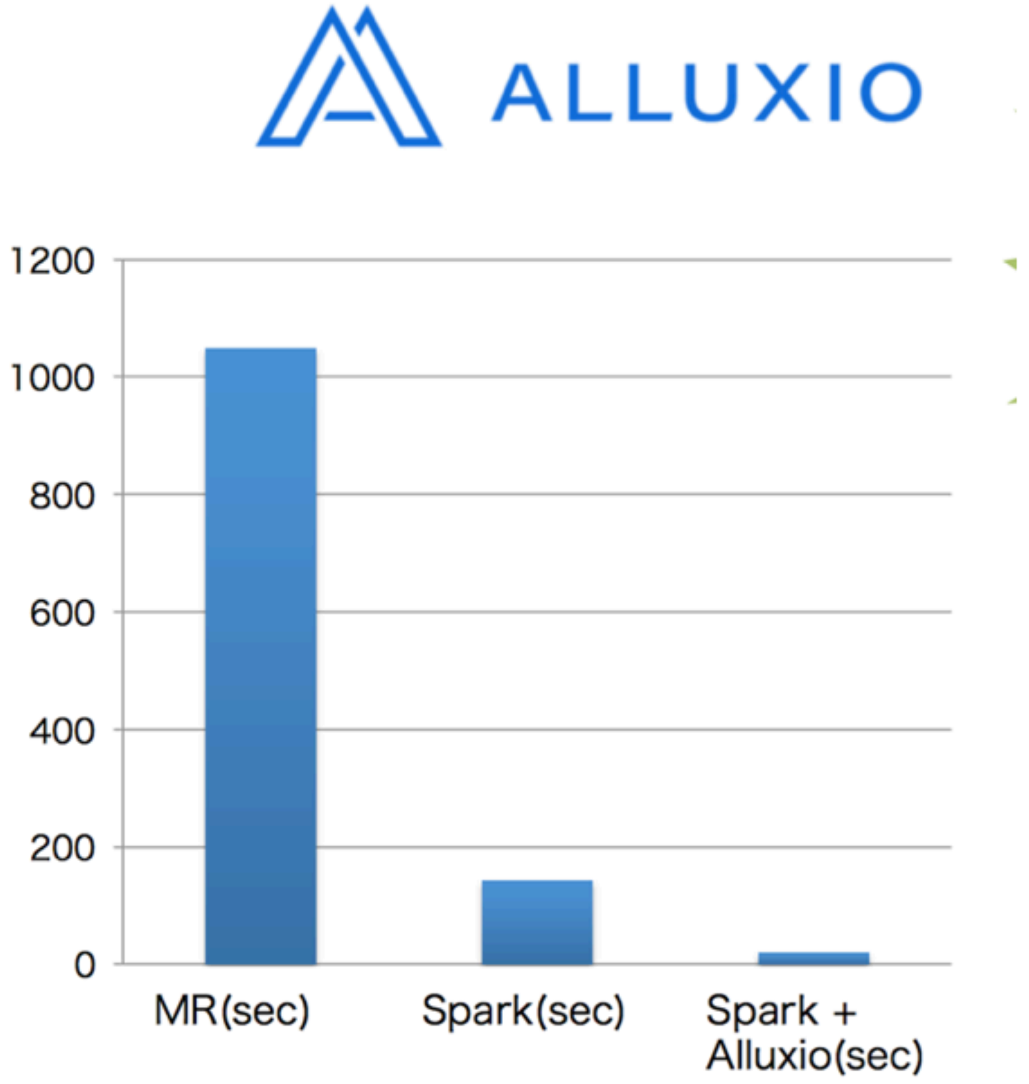
Cloud Infrastructure: Computing



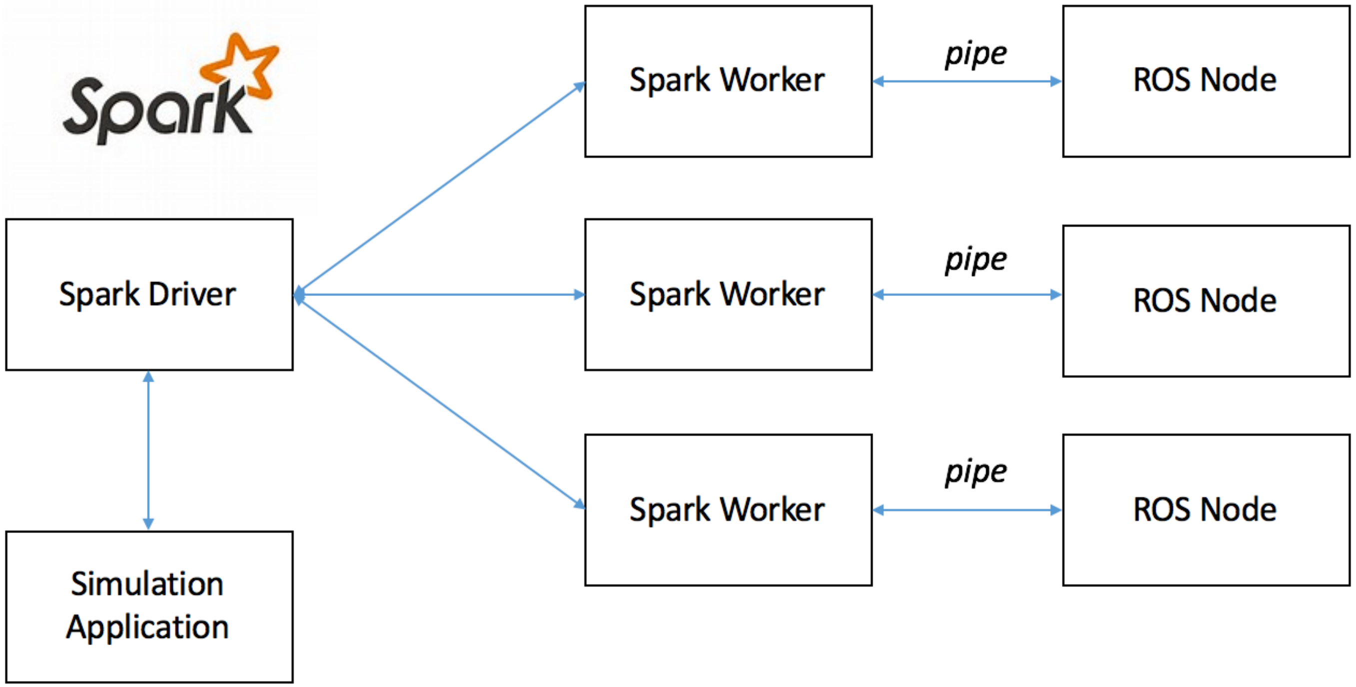


Cloud Infrastructure: Storage

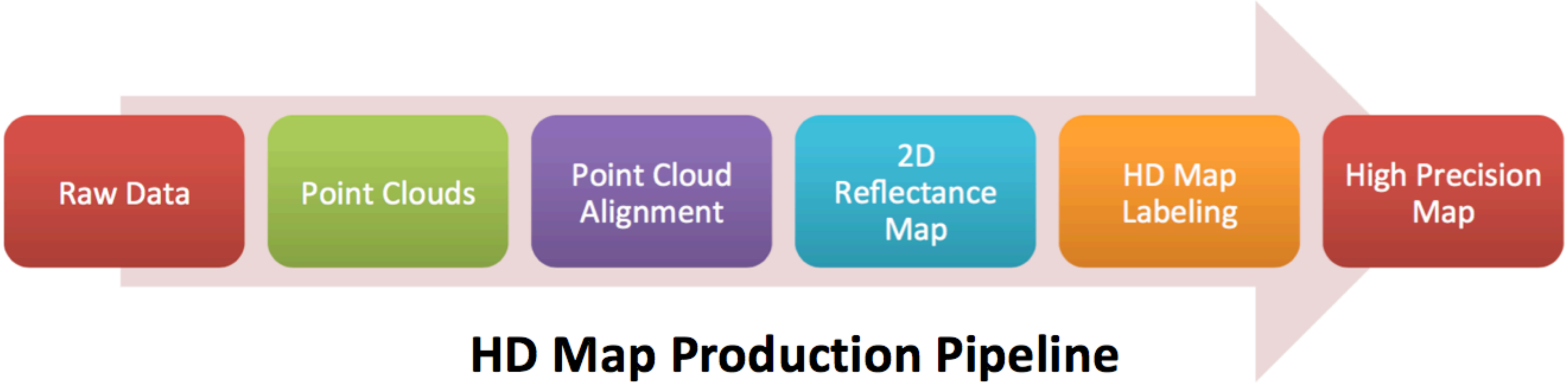




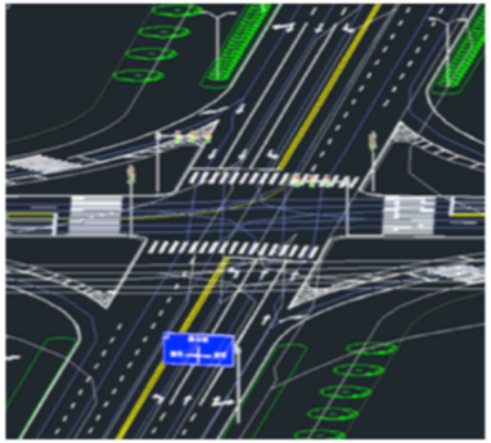
Cloud Infrastructure: Simulation



Cloud Infrastructure: HD Map



HD Map Production Pipeline



Cloud Infrastructure: Model Training