PKU Omni Smart Sensing (POSS)

Towards an Omni-directional Sensing of a Large Dynamic Environment

Key Lab of Machine Perception (MOE), School of EECS, Peking Univ.
Huijing Zhao zhaohj@cis.pku.edu.cn +86-10-62757458

Introduction

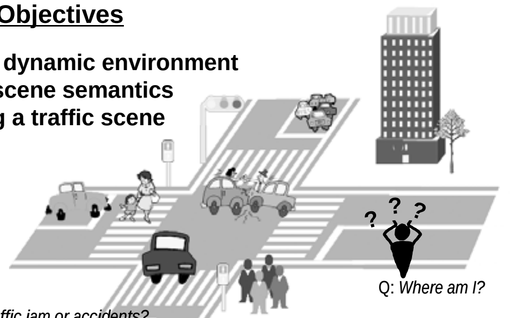
http://www.poss.pku.edu.cn

This research focus on the perception and reasoning techniques using either an intelligent vehicle or a network sensing system or a collaboration of them.

We have a goal of developing an intelligent vehicle of Omni-directional eyes monitoring both static and dynamic objects when it cruises around a large dynamic environment; and a network sensing system that monitoring constantly the dynamic objects in a cluttered environment. Based on such platforms, we study methodologies on perception, modeling and reasoning of a dynamic procedure and a dynamic traffic scene, where the fundamental issues such as multi-modal sensor fusing, calibration, SLAM (simultaneous localization and mapping), moving object detection and tracking, behavior modeling and situation awareness, abnormal detection, semantic mapping, as well as potential applications are focused.

Research Objectives

- **Sensing a dynamic environment**
- **Learning scene semantics**
- **Simulating a traffic scene**



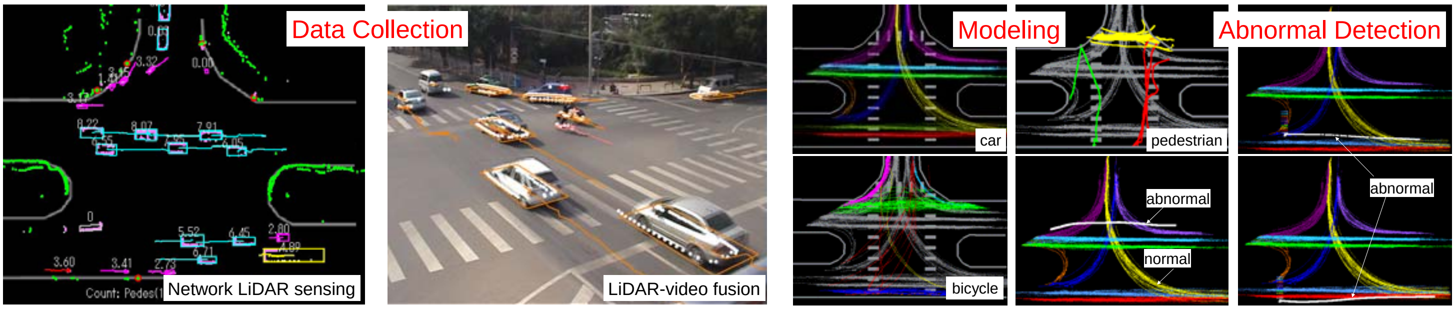
Q: Why always traffic jam or accidents?

A: We need to sense the data of both motion (e.g. people, cars) and motionless (e.g. building, tree, road, traffic sign) objects at the environment, analyze their relations, locate the problem and find a possible solution.

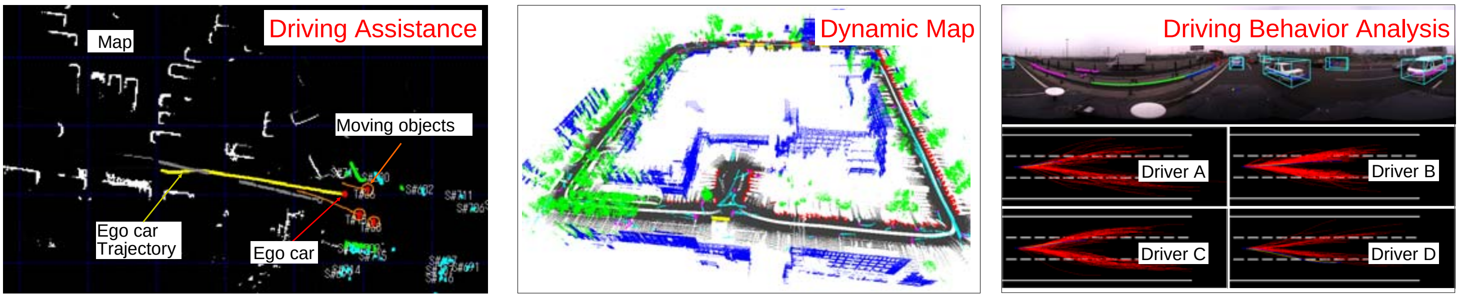
Q: Where am I?

A: Locate "I" to the virtual space.

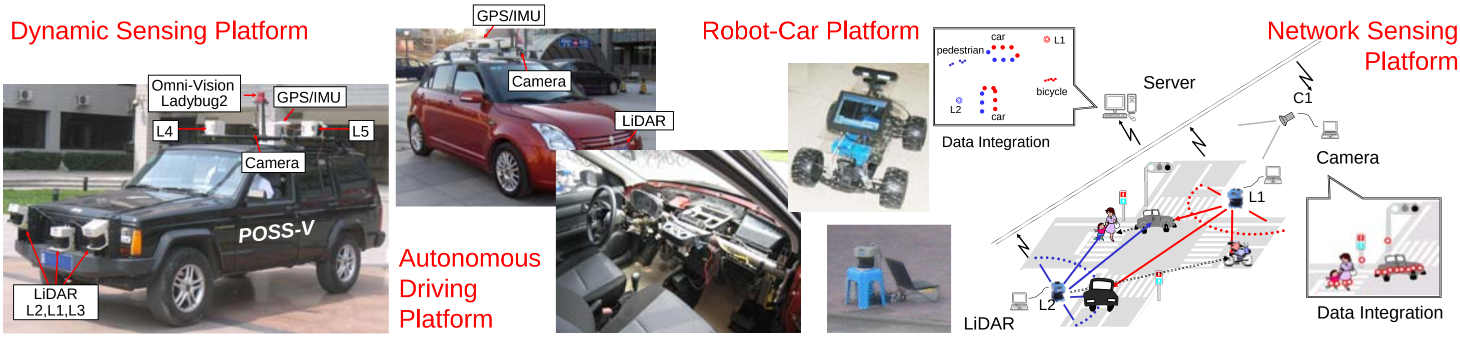
Monitoring a traffic scene through network sensing



Monitoring a dynamic procedure and scene using an intelligent vehicle



Intelligent vehicle and network sensing platforms



Omni-directional detection and tracking of on-road vehicles using multiple horizontal laser scanners

Huijing Zhao*, Chao Wang*, Wen Yao*, Franck Davoine** , Jinshi Cui*, Hongbin Zha*

Introduction

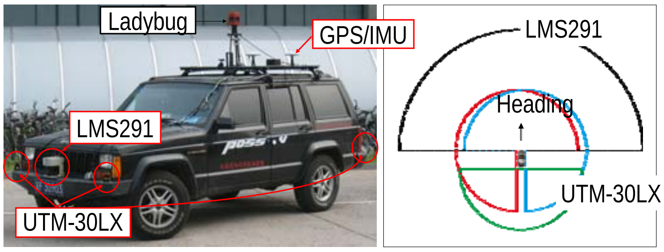
This research is motivated by two potential applications:

(1) Detecting and tracking the moving objects at the ego-vehicle's 360° local surroundings, so as to assist for safety driving;

(2) Obtaining the motion trajectories of other traffic participants, so as to support for driving behavior modeling and reasoning.

This research studies at the on-road traffic environment, such as freeway, which is free of intersections, traffic signals, and pedestrians. So that the algorithm focuses on the detection and tracking of vehicles only.

Sensor Setting



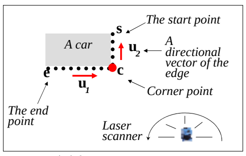
In this paper, the Ladybug is for visualization only.

Contributions

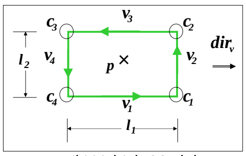
In this research, a system of detecting and tracking on-road vehicles using multiple horizontal laser scanners on a vehicle platform is developed.

Algorithms are developed with focuses on solving data association of simultaneous measurements to single objects, and state estimation in case of partial observations in dense traffic situations.

Vehicle Model



(a) Laser measurement



(b) Vehicle Model

1) **u** is distinct in each side of a vehicle;

2) **u** is equal in the measurements of different laser scanners.

Grouping overlapped measurements: by matching **u** with **v**,

1) the overlapped measurements to the same vehicle side can be associated;

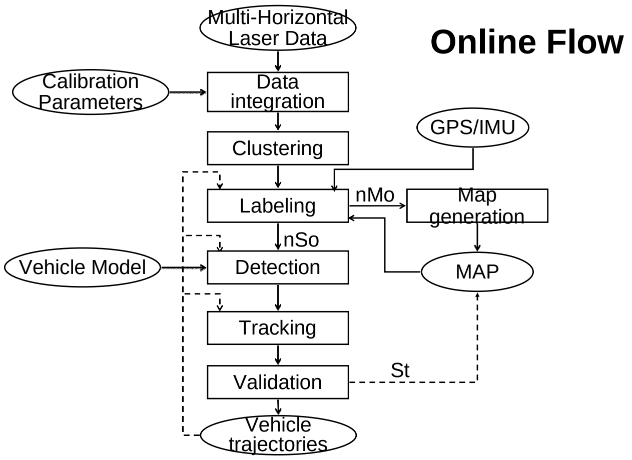
2) mis-grouping the data measurements can be avoided.

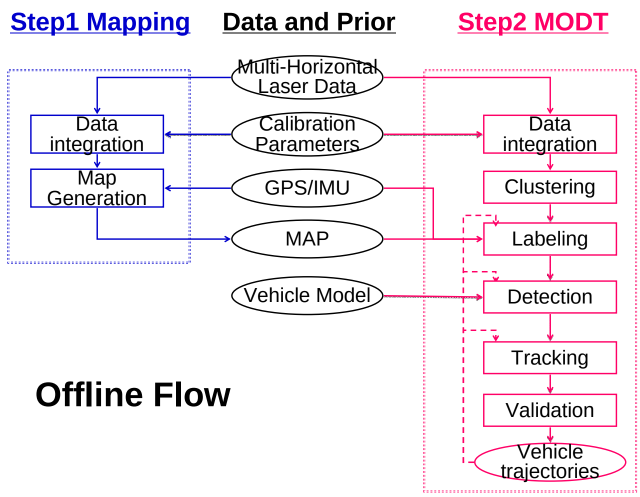
Accounting partial observations:

1) Reliability items are defined to denote whether the corresponding features are estimated on direct observations or inferred through the assumption on vehicle model.

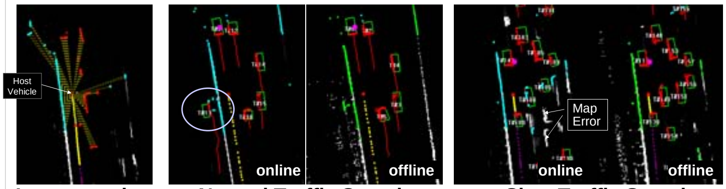
2) Reliability items are accounted in data association and vehicle track estimation.

Approaches

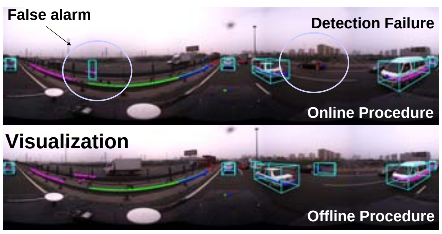


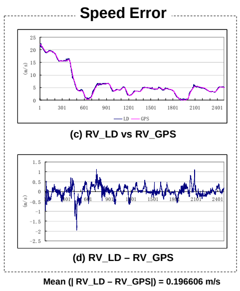


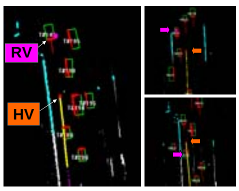
Experimental Results



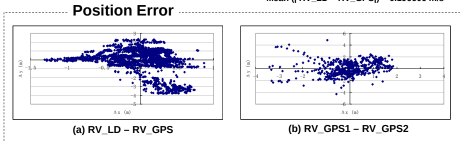
Laser sensing　Normal Traffic Speed　Slow Traffic Speed







Accuracy Examination



Contact : Huijing Zhao, zhaohj@cis.pku.edu.cn

*Key Lab of Machine Perception, Peking University, Beijing, China
**CNRS / LIAMA Sino French Laboratory, Beijing, China

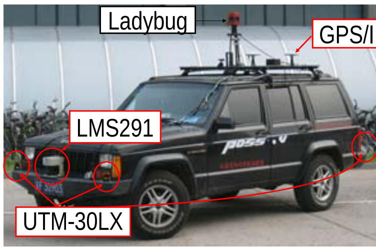
Omni-directional detection and tracking of on-road vehicles using multiple horizontal laser scanners

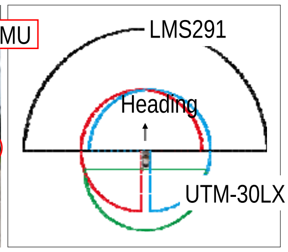
Huijing Zhao*, Chao Wang*, Wen Yao*, Franck Davoine** , Jinshi Cui*, Hongbin Zha*

Introduction

Motivations: (1) Detecting and tracking the moving objects at the ego-vehicle's 360° local surroundings, so as to assist for safety driving; (2) Obtaining the motion trajectories of other traffic participants, so as to support for driving behavior modeling and reasoning.

Scenario: on-road traffic environment, such as freeway, which is free of intersections, traffic signals, and pedestrians.

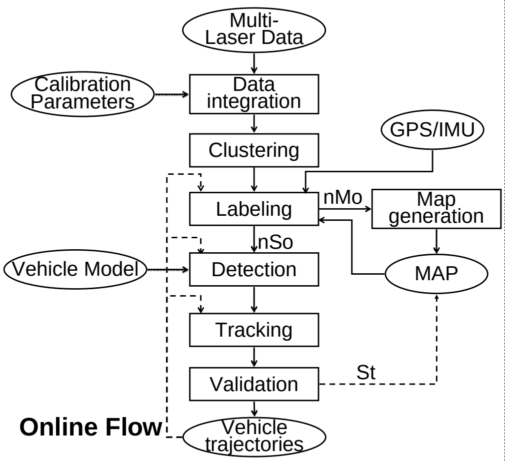


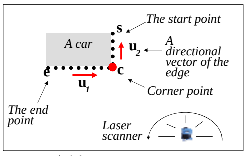


In this paper, the Ladybug is for visualization only.

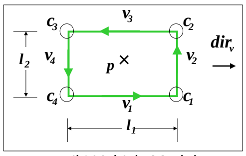
Contributions: In this research, a system of detecting and tracking on-road vehicles using multiple horizontal laser scanners on a vehicle platform is developed. Algorithms are developed with focuses on solving data association of simultaneous measurements to single objects, and state estimation in case of partial observations in dense traffic situations.

Approach



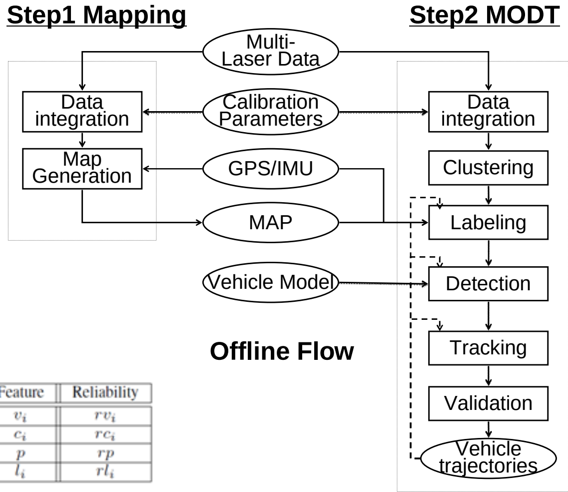


(a) Laser measurement



(b) Vehicle Model

Item	Feature	Reliability
directional vectors ($i = 1,..,4$)	v_i	rv_i
corner points ($i = 1,..,4$)	c_i	rc_i
a center point	p	rp
lengths on two vertical edges ($i = 1,2$)	l_i	rl_i

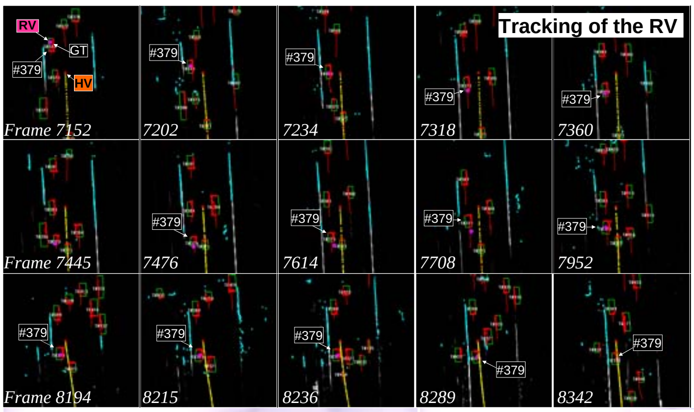


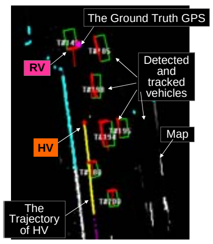
1) **u** is distinct in each side of a vehicle;
2) **u** is equal in the measurements of different laser scanners.

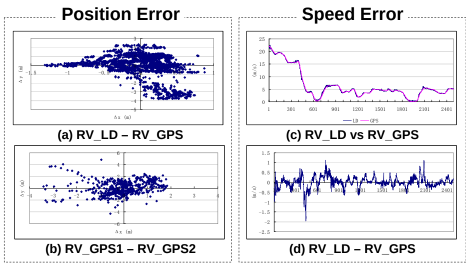
Grouping overlapped measurements:
by matching **u** with **v**, the overlapped measurements to the same vehicle side can be associated.

Addressing partial observations:
1) Reliability items are defined to denote whether the corresponding features are estimated on direct observations or inferred through the assumption on vehicle model.
2) Reliability items are accounted in data association and vehicle track estimation.

Experimental Results

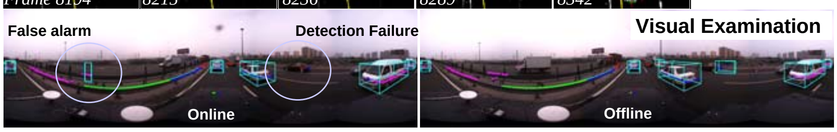




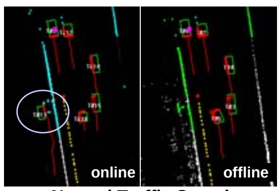


Accuracy Examination

Mean (| RV_LD – RV_GPS|) = 0.196606 m/s



Visual Examination



Normal Traffic Speed

Slow Traffic Speed

Online vs Offline

Contact : Huijing Zhao, zhaohj@cis.pku.edu.cn

*Key Lab of Machine Perception, Peking University, Beijing, China
**CNRS / LIAMA Sino French Laboratory, Beijing, China

Lane Change Trajectory Prediction by Using Recorded Human Driving Data

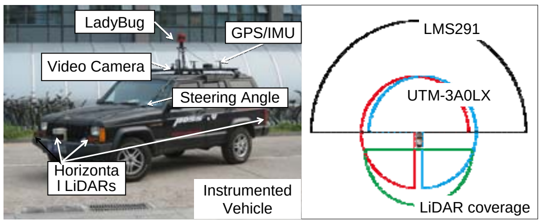
Wen Yao*, Huijing Zhao*, Philippe Bonnifait** , Hongbin Zha*

http://www.poss.pku.edu.cn

Introduction

Motivations: 1) Efficiently collect human driving data using an instrumented vehicle to build a human lane change database; 2) Predicting human driver's potential lane change trajectory by referring to the database for risk assessment which helps the driver to decide whether to change lane.

Scenario: urban traffic scene, such as freeway. We mainly focused on lane change behavior on straight road currently, which is very common in daily driving.



Contributions: 1) A human driving database is generated through on-road data collection using an instrumented vehicle, which consists synchronized trajectories of ego vehicle and surrounding traffics during lane change maneuvers; 2) A lane change trajectory prediction is developed by referring to the human driving data of similar situations.

Work Flow

Offline data processing

- **1.1 On-road Driving Data Acquisition** — *GPS/IMU Vehicle Parameters/ Lidar/ Video*
- **1.2. Lane Change Data Extraction** — *Ego vehicle/ Traffic participants/ Road/ Driver/*

1. Human Driving Database Generation
human driving data acquisition using an instrumented vehicle

Synchronized trajectories of ego and environmental vehicles in lane change

Human Lane Change Database

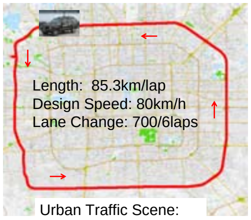
Online risk assessment

- **2.1 Situation-based Trajectory Search** — *Find lane change samples in similar initial situation*
- **2.2 Lane Change Trajectory Prediction** — *Trajectory generation on predicted end state*

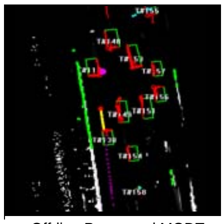
2. Lane Change Trajectory Prediction
lane change intention is detected as the trigger of this model

Database Generation

1)Synchronized motion trajectory collection of the ego and surrounding vehicles (Zhao, IV09)



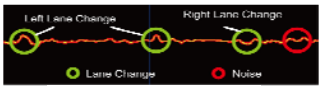
Urban Traffic Scene: 4th Ring Road of Peking



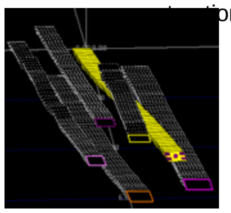
Off-line Processed MODT

2)Lane change samples

a) Manually record lane change start/end time
b) Steering wheel data
c) Video check



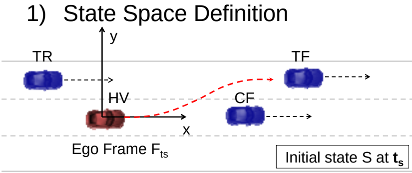
Steering wheel data

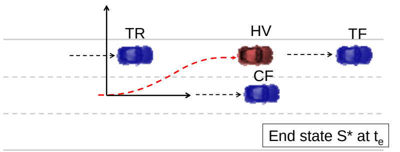


Lane change data sample

Trajectory Prediction

1) State Space Definition

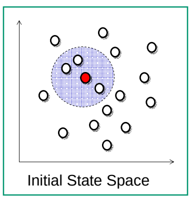




State Definition S, S* = (s_{HV}, s_{CF}, s_{TF}, s_{TR})
$s_{HV,CF,TF,TR}$ = (*Pos, Spe., Acc.*) at $F_{ts,te}$

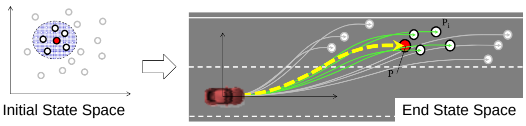
2) Trajectory Prediction Approach

Step1.
Situation-based Neighborhood Trajectory Search



- Samples in the human driving database
- The current state when a lane change intention is detected in online processing
- **Neighborhood of the current state with a Distance Measure** $D = \|S - S_{ego}\|$

Step2.
End State Interpolation from End States of Selected Samples

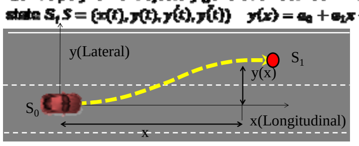


End state P interpolated from selected samples $\{P_i\}$ with similar input states:

$P = \sum \omega(D_i) P_i$, P_i – end states of trajectory samples, $\omega(D_i) = \frac{1}{D_i^2} / \sum \frac{1}{D_i^2}$.

Step3.
Trajectory Generation Using Polynomial Model (Werling, ICRA 10')

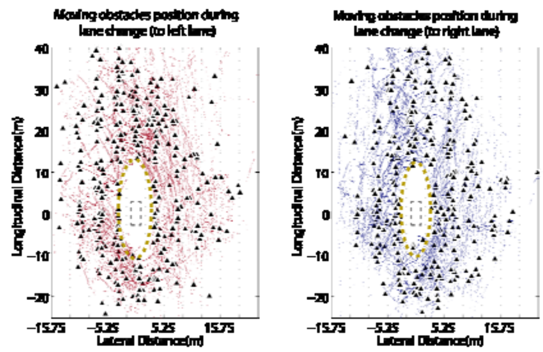
Quintic polynomial trajectory generation between initial S_0 and predicted end state S_1, $S = (x(t), y(t), \dot{y}(t), \ddot{y}(t))$ $\quad y(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5$



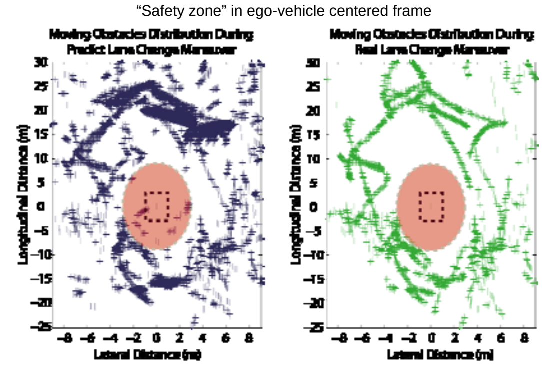
solving

$$y(x_1) = y_1,\ y(0) = 0$$
$$y'(0) = y'(x_1) = 0$$
$$y''(0) = y''(x_1) = 0$$

Experimental Results



"Safety zone" in ego-vehicle centered frame



Comparison of predicted and real trajectories in ego-frame

Contact：
Wen YAO, yaowen@cis.pku.edu.cn

*Key Lab of Machine Perception, Peking University
**the Universit´e de Technologie de Compiegne, Heudiasyc CNRS UMR 7253

A Probabilistic Framework for Occluded Vehicle Detection

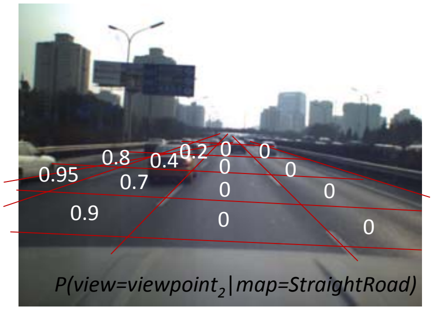
Chao Wang[1], Huijing Zhao[1], Chunzhao Guo[2] , Seiichi Mita[3] , Hongbin Zha[1]

Abstract

In this research we propose a novel probabilistic framework for occluded vehicle detection on road scene. We build probabilistic vehicle's viewpoint map on image with the road structure priority from GPS and road map. With different viewpoints, part-based vehicle detectors are trained to find vehicle part candidates, and a probabilistic model for parts' locations relative to vehicle's center is learnt to infer the vehicle's viewpoint and occluded parts. The framework is evaluated on data from Nagoya downtown and Beijing freeway.

Probabilistic Viewpoint Map

Determine the viewpoint of vehicle in different place according to the map priority. The probability can be obtained by traffic rules and large real data statistic analysis.
We divide the multi-view vehicles into 4 classes according to viewpoints. For each viewpoint class, probabilistic map is generated by considering the road structure. *e.g.* $P(view=viewpoint_2|map=StraightRoad)$

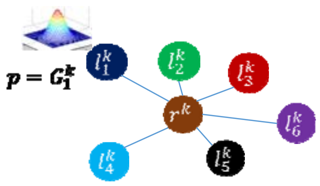


Probabilistic Parts Location Model

Parts location distribution model across different viewpoints. In viewpoint k, for each part l_i^k, the distribution model is

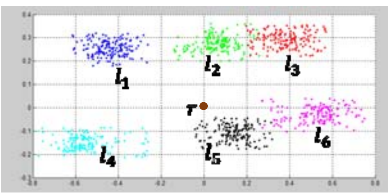
$$p(l_i^k|r^k) = G(\mu_{i1}^k, \mu_{i2}^k, \sigma_{i1}^k, \sigma_{i2}^k, \rho_i^k)$$

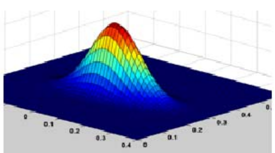
Part samples



Parts location distribution model

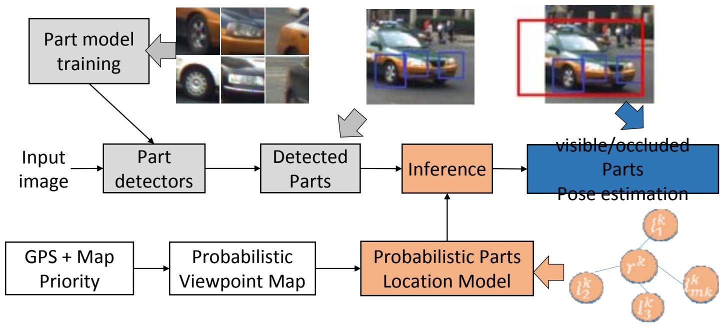
Parts locations





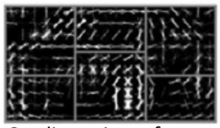
$$p(l_2^k) = G_2^k$$

Framework



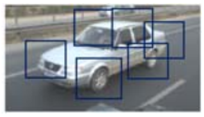
Part Detectors

Train part detectors based on HoG features. Using unsupervised method, automatically select most representative parts for detectors, and record each part's position for parts location distribution model learning. The part detectors are training in each viewpoint class.



Gradient view of a car, and representative parts

Examples of selected parts for training



Parts detection result

Inference

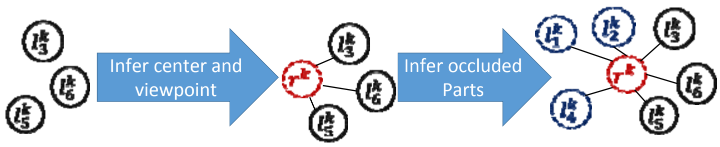
Problem definition: infer the vehicle's center and viewpoint r^k,

$$\widehat{r^k} = arg\max_{r^k} p(r^k|L^k, M_g^k, M_r),$$ k is viewpoint class.

Solution:

$$p(r^k|L^k, M_g^k, M_r) \propto \prod_i^{vis} p(l_i^k|r^k, M_g^k) \prod_j^{occ} p(l_j^k|r^k, M_g^k)\, p(r^k|M_r)$$

Then infer the occluded parts, $p(occ) = p(l_{occ-j}^k|r^k)$

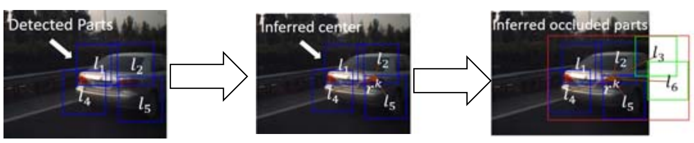


Detected arts

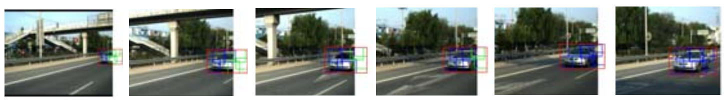
Inferred center and viewpoint

Inferred occluded parts

Experimental Results



Detection procedure



Detections on car image sequence

Key Lab of Machine Perception, Peking University[1]
Toyota Central R&D Labs., Inc. [2]
Research Center for Smart Vehicles, Toyota Technological Institute [3]

Contact :
Chao Wang, chao.wang@pku.edu.cn

Monocular Visual Localization using Road Structural Features

Yufeng Yu*, Huijing Zhao*, Franck Davoine+, Jinshi Cui*, Hongbin Zha*

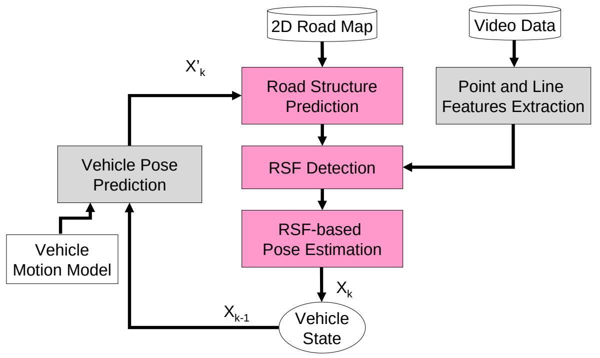
Introduction

Motivations: 1) Precise localization is an essential issue for autonomous driving applications, while low-cost GPS can not meet such requirements ; 2) General visual-based localization algorithms use point feature, which is highly affected by the moving objects in a normal traffic scenario; 3) Line segments with three perpendicular directions, which are road's longitude, latitude and vertical directions, are stable in rotation estimation and those on environmental cars provide additional supports.

Scenario: Major roads in downtown Beijing, which are structured and with intense dynamic traffic.

Contributions: 1) We designed a novel feature named Road Structural Feature (RSF), which is robust dealing with moving objects, for on-road localization. 2) We proposed a real-time precise visual localization method based on RSF, and the experiment showed good performance.

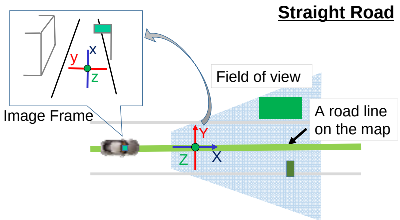
Work Flow

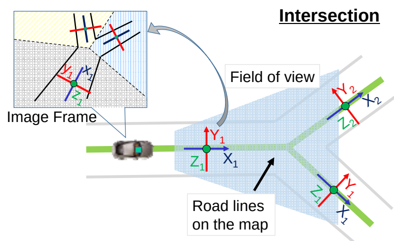


Road Structure Prediction

Given: A rough 2D road map with roads' directions and intersections' positions

Solution: 1) Get the camera's Field of view on the map
2) Project the road's perpendicular axes onto the image plane





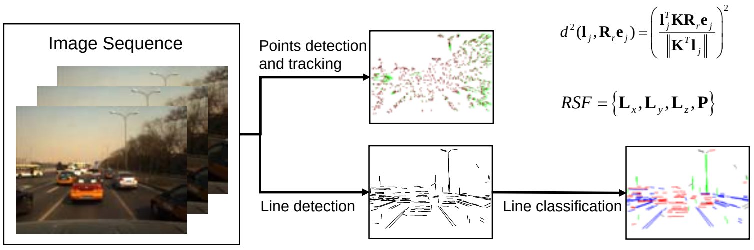
RSF Detection

Given: video data, predicted road structure

Solution: 1) Detect and track general feature points
2) Detect general line segments, then label them with a distance measure

$$d^2(\mathbf{l}_j, \mathbf{R}_r \mathbf{e}_j) = \left(\frac{\mathbf{l}_j^T \mathbf{K} \mathbf{R}_r \mathbf{e}_j}{\left\| \mathbf{K}^T \mathbf{l}_j \right\|} \right)^2$$

$$RSF = \{\mathbf{L}_x, \mathbf{L}_y, \mathbf{L}_z, \mathbf{P}\}$$



RSF-based Pose Estimation

Given: road structure feature, vehicle speed

Solution:

1) Sample RSF candidate with $C_{RSF} = \{\mathbf{l}_1, \mathbf{l}_2 \in \mathbf{L}_u, \mathbf{l}_3 \in \mathbf{L}_v, u \neq v, \mathbf{p}_1, \mathbf{p}_2 \in \mathbf{P}\}$

2) Calculate x_t with given C_{RSF}

$$\left.\begin{array}{l} \mathbf{l}_1 \times \mathbf{l}_2 = \mathbf{v}_1 = \mathbf{KRd}_1 \\ \mathbf{R} = (\mathbf{I}-\mathbf{S})^{-1}(\mathbf{I}+\mathbf{S}) \\ \mathbf{S} = \begin{bmatrix} 0 & -c & -b \\ c & 0 & -a \\ b & a & 0 \end{bmatrix} \\ \mathbf{l}_3^T \mathbf{KRd}_3 = 0 \end{array}\right\} \Rightarrow \mathbf{R} \qquad \left.\begin{array}{c} \mathbf{R}_{rel} = \mathbf{R}\mathbf{R}'^T \\ \mathbf{F} = \mathbf{K}^{-T} [\mathbf{t}_{rel}]_\times \mathbf{R}_{rel} \mathbf{K}^{-1} \\ p_1^T \mathbf{F} p_1' = 0 \\ p_2^T \mathbf{F} p_2' = 0 \end{array}\right\} \Rightarrow \mathbf{t}_{rel}$$

$$\mathbf{t}_k = \mathbf{t}_{rel} - \mathbf{R}_{rel}^T \mathbf{t}_{k-1}$$

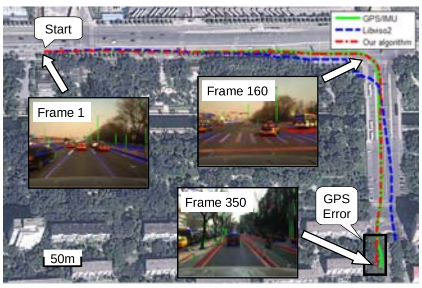
$\|t_{rel}\|$ is calculated by the speed data

3) Evaluate x_t using an observation error measure

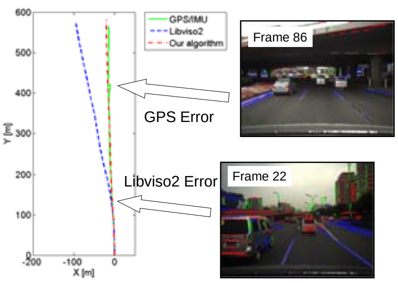
$$E = E_l + \lambda E_p = \sum_j length^2(\mathbf{l}_j) \cdot \left(\frac{\mathbf{l}_j^T \mathbf{KR}\hat{\mathbf{d}}_j}{\left\| \mathbf{K}^T \mathbf{l}_j \right\|} \right)^2 + \lambda \sum_i \left(d^2(\mathbf{p}_i, \mathbf{F}\mathbf{p}_i') + d^2(\mathbf{p}_i', \mathbf{F}^T \mathbf{p}_i) \right)$$

Experimental Results

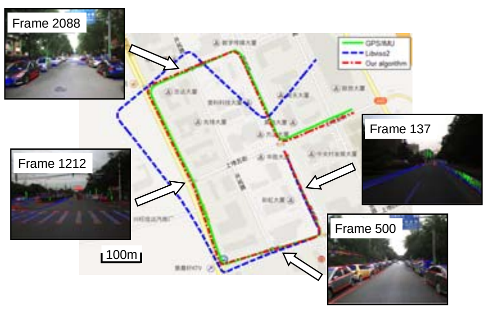
(a) Normal Traffic

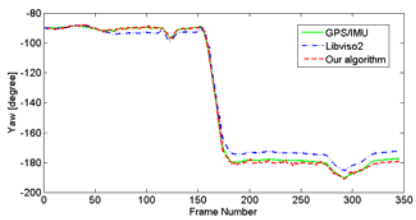


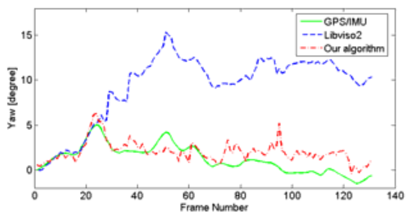
(b) Straight Road Situation

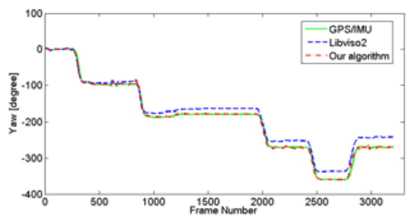


(c) Complex Situation









Contact:
Yufeng YU, yuyufeng@pku.edu.cn

*Key Lab of Machine Perception, Peking University, Beijing, China
+CNRS, LIAMA Sino-French Laboratory, Beijing, China

Scene Understanding in a Large Dynamic Environment through a Laser-based Sensing

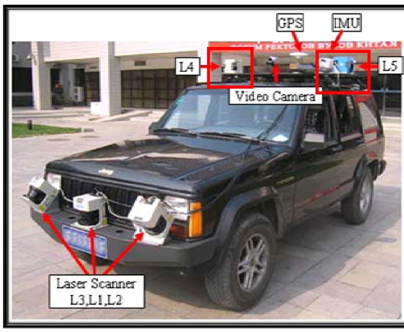
Huijing Zhao, Yiming Liu, Xiaolong Zhu, Yipu Zhao, Hongbin Zha
Key Lab of Machine Perception (MOE), Peking University, Email: zhaohj@cis.pku.edu.cn

Introduction

Objective

We aim to provide a **map** with **high-level** representations.

This map enables a robot to have **semantic knowledge** of the environment which is large and dynamic, such as objects, their types and their spatial relationships.

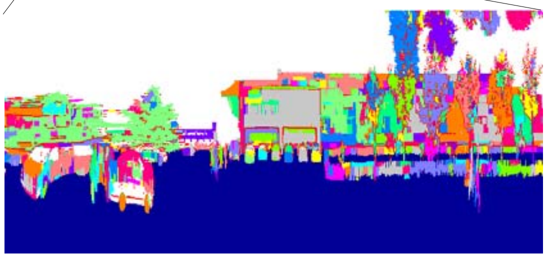


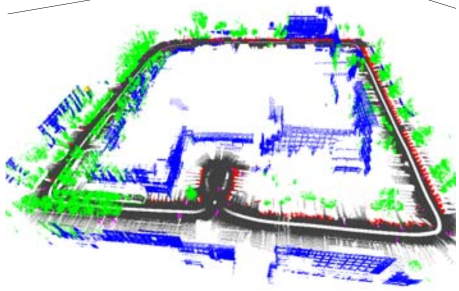
Laser #	Scan Mode	Major Purpose
L1	Horizontal	SLAM with MODT
L2	Downward	Road geometry
L3	Upward	Objects above the road
L4	Vertical Right	Objects to the right of the car
L5	Vertical Left	Objects to the left of the car

Data Acquisition

System Flowchart

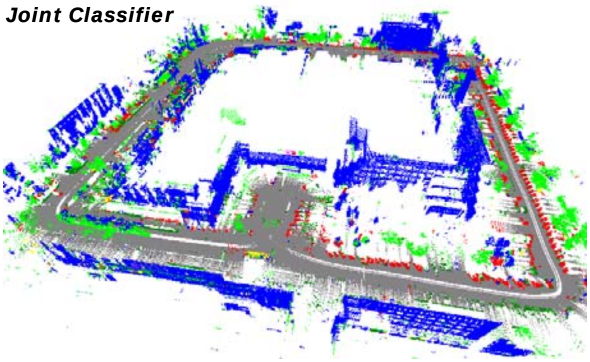
Range Image → Over-segmentation → Joint Merge with Classification → Objects with Semantic Label





Experimental Results

Joint Classifier

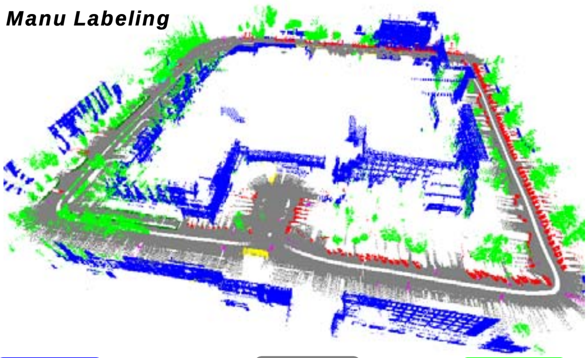


Line Classifier

Point Classifier

Joint Classifier

Manu Labeling



Manu Labeling

Segmentation

Super Segmentation

Building Road Tree Car People Bush Bus

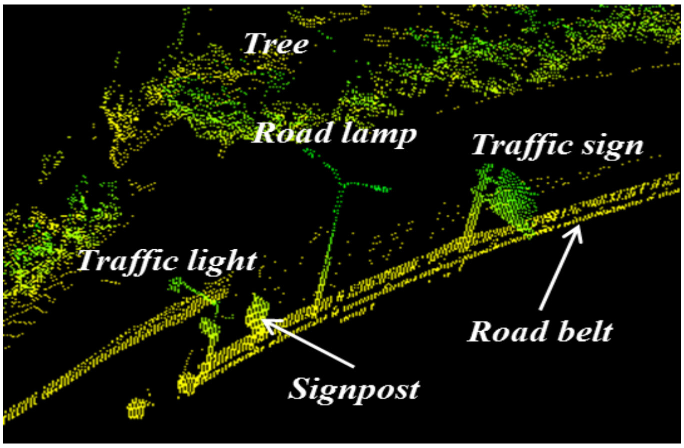
Object Detection Using 3D Street Data

Key Lab of Machine Perception (MOE), School of EECS, Peking Univ.
Yipu Zhao zhaoyp@cis.pku.edu.cn

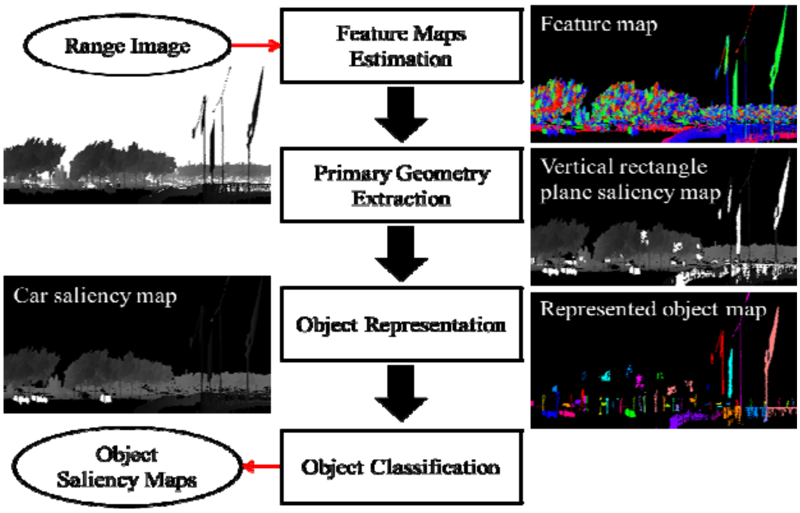
Introduction

http://www.poss.pku.edu.cn/people/zhaoyp

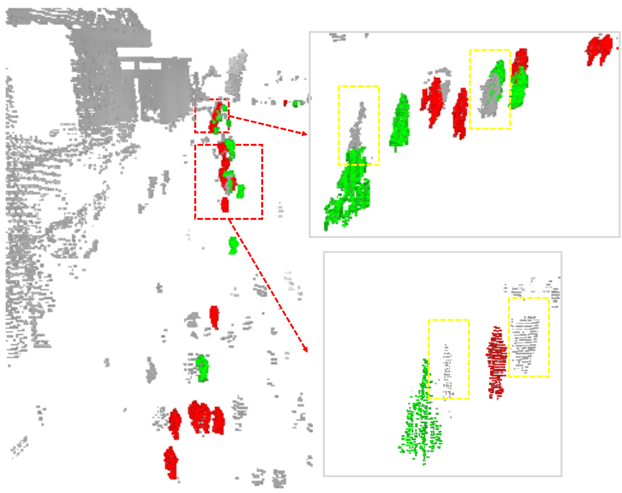
It becomes a well-known technology that a low level map of complex environment containing 3D laser points can be generated using a robot with laser scanners. Given a cloud of 3D laser points of an urban scene, this paper proposes a method for locating the objects of interest, e.g. traffic signs or road lamps, by computing object-based saliency. Our major contributions are: 1) a method for extracting simple geometric features from laser data is developed, where both range images and 3D laser points are analyzed; 2) an object is modeled as a graph used to describe the composition of geometric features; 3) a graph matching based method is developed to locate the objects of interest on laser data. Experimental results on real laser data depicting urban scenes are presented; efficiency as well as limitations of the method are discussed.



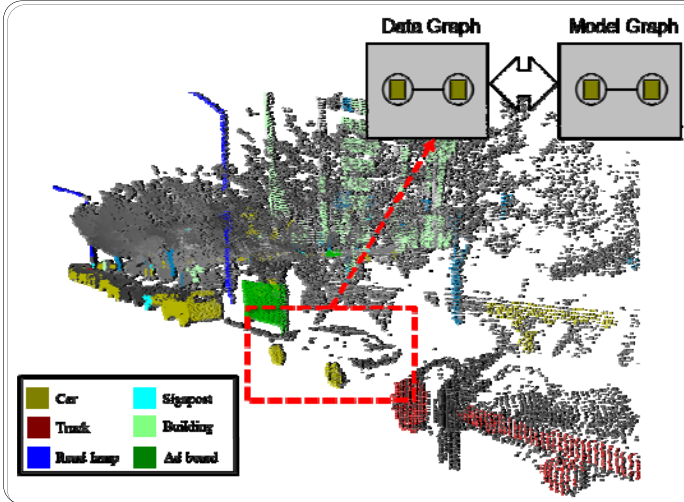
Approach

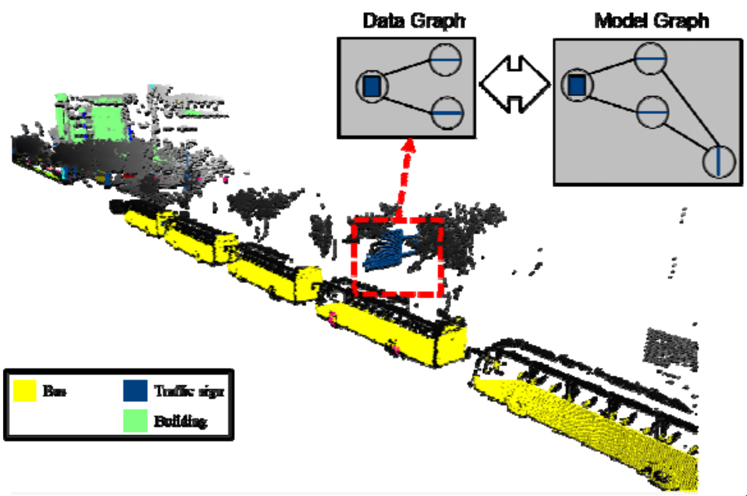


Learning-based Object Detection



Graph-based Object Detection





* Y. Zhao, et al., “Computing Object-based Saliency in Urban Scenes Using Laser Sensing”, *2012 IEEE International Conference on Robotics and Automation*

2D-LIDARs' Calibration Using Multi-Type Geometry Features in Urban Outdoor Scene

Mengwen He alexanderhmw@gmail.com
Huijing Zhao zhaohj@cis.pku.edu.cn

Introduction

For a multi-LIDAR system (Fig.1), calibration is essential for collaborative use of LIDAR data. In our research, we developed a 2D-LIDAR calibration method using *3D point-clouds alignment*. In order to enable the calibration method work in *urban outdoor scene*, we proposed a alignment method based on *multi-type geometry features* extracted from 3D point-clouds of urban outdoor scene (Fig.2).

Figure 1: POSS-V Multi-LIDAR System

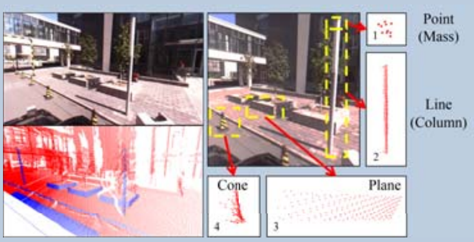


Figure 2: Multi-Type Geometry Features

Method Overview

Suppose a LIDAR, called as the *reference LIDAR*, has been calibrated with mobile platform and would produce *undistorted 3D point-cloud*. Whenever a new LIDAR, called as the *target LIDAR*, is introduced to the on-board system, it could be calibrated by alignment of 3D point-clouds from reference and target LIDARs to find its geometric transformation to the mobile platform. [1, 2]
This research aims at an *online self-calibration*, therefore, an approach is proposed by using multi-types of geometric features (Eq.1) in urban outdoor scene to cope with the challenges such as *small overlapped area, different viewpoints, occlusion and scene dynamics.*

$$G = \{P, \{\phi, f^{\phi}\}, \{\mu, \sigma^2\}\} \quad (1)$$

P : 3D point-set $\{\mathbf{p}_i\}$
ϕ : Geometry type descriptor
f^{ϕ} : Parametric equation
μ : Confidence degree, $E(f^{\phi}(\mathbf{p}_i)^2)$
σ^2 : Noise level, $D(f^{\phi}(\mathbf{p}_i)^2)$

Experiments

Experiments are conducted using the data sets of an intelligent vehicle platform POSS-V (Fig.1) through a driving in the campus of Peking University (Fig.3).

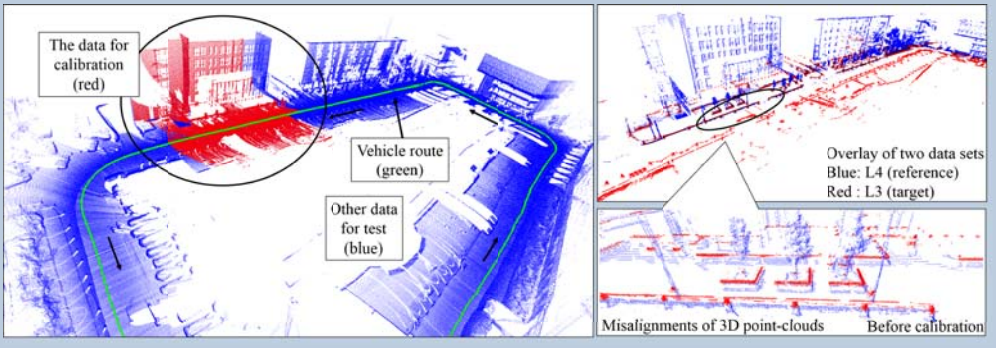


Figure 3: Outdoor Environment in PKU (Left); 3D Point-Clouds Misalignment before Calibration (Right)

We semi-automatically extracted 3 types of matched geometry features (Fig. 4 right) by region grow algorithm from the 3D point-clouds of selected calibration area (Fig. 4 left). Then we used the optimization method defined in [1] to align matched geometry features.

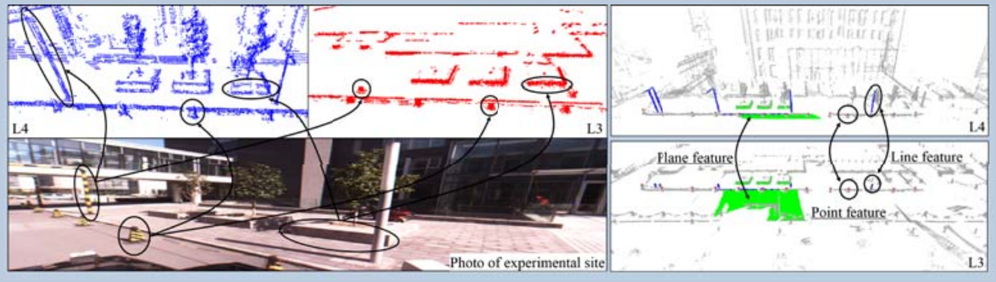
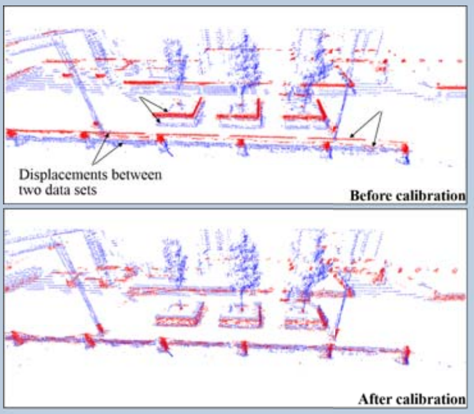


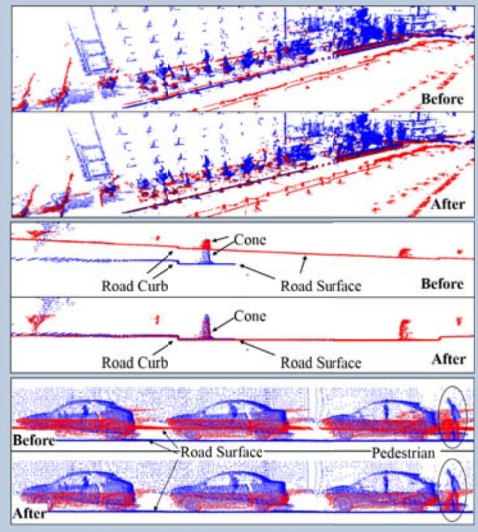
Figure 4: Selected Calibration Area (Left); Extracted Matched Multi-Type Geometry Features (Right)

Alignment in Calib. Area



After alignment of geometry features, we aligned corresponding 3D point-clouds, and also calibrated the target LIDAR L3 in theory [1, 2].

Alignment in Test Area



Calibration Verification (Accuracy and Robustness)

Feature Types	Geometric Features	e_1 without noisy features	Accuracy Evaluation	e_2 with noisy features	Error Ratio $= e_2/e_1$
Mono	Point (P)	0.03232	Bad	0.04615	1.427
	Line (L)	0.07797	Bad	0.15548	1.994
	Plane (PL)	0.01939	Good	0.05878	3.031
Dual	P + L	0.01363	Good	0.01429	1.048
	P + PL	0.00980	Excellent	0.00993	1.013
	L + PL	0.02471	Good	0.05510	2.229
Triple	P + L + PL	0.00981	Excellent	0.00989	1.008

References

[1] M. He, H. Zhao, F. Davoine, J. Cui, and H. Zha, "Pairwise LIDAR Calibration Using Multi-Type 3D Geometric Features in Natural Scene," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2013.

[2] M. He, H. Zhao, J. Cui, and H. Zha, "Calibration Method for Multiple 2D LIDARs System," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2014.