

On-Road Vehicle Trajectory Collection and Scene-Based Lane Change Analysis: Part I

Huijing Zhao, *Member, IEEE*, Chao Wang, Yuping Lin, Franck Guillemard, Stéphane Geronimi, and François Aioun

Abstract—This two-part paper aims to study lane change behaviors at the tactical level from an on-road perspective, with a special focus on analyzing the interactions between an ego and surrounding vehicles during the procedure. Part I addresses vehicle trajectory collection, whereas Part II addresses lane change extraction and scene-based behavioral analysis. Different from the general technique of moving object detection and tracking, trajectory collection for tactical driving behavior study is required to have the properties of consistency, completeness, continuity, and accuracy. This paper proposes a system of on-road vehicle trajectory collection, where an instrumented vehicle is developed with multiple horizontal 2-D lidars that have 360° coverage. The software is developed by fitting the laser points of all lidars on a vehicle model using a coupled estimation of features and reliability along frames to achieve accurate state estimations of occluded data and robust data association in multiviewpoint sensing. The performance is investigated extensively, and a large trajectory set is developed through on-road driving at the Fourth Ring Road in Beijing for a total distance of 64 km, with more than 5700 environmental trajectories with a total length of over 19 h. The performance is demonstrated to be of high quality in terms of the required properties. To the authors' knowledge, this is the first system that is able to automatically collect all-around vehicle trajectories during on-road driving and to demonstrate good performance in providing a high-quality database for driving behavior studies from an on-road perspective that addresses vehicle interactions in real-world traffic at the trajectory level.

Index Terms—Lidar processing, on-road, tactical driving behavior, trajectory collection, vehicle.

I. INTRODUCTION

RECENTLY, amazing progress has been made in terms of the development of autonomous driving and active safety systems. Currently, autonomous driving systems have demonstrated safety and autonomy. To achieve a higher level of intelligence such as having close interactions with other human-driven cars when they need to share the same roads, a more intensive understanding of the complex driving behaviors at

Manuscript received August 21, 2015; revised February 23, 2016 and April 9, 2016; accepted May 19, 2016. This work was supported in part by the PSA's OpenLab program, by the National Natural Science Foundation of China under Grant 61573027, and by the Hi-Tech Research and Development Program of China under Grant 2012AA011801. The Associate Editor for this paper was F.-Y. Wang.

H. Zhao, C. Wang, and Y. Lin are with the Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China (e-mail: zhaohj@cis.pku.edu.cn).

F. Guillemard, S. Geronimi, and F. Aioun are with Groupe PSA (Peugeot Société Anonyme) formerly PSA Peugeot Citroën, 78943 Vélizy, France.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TITS.2016.2571726

a tactical level is required, e.g., how a driver/vehicle governs safe interactions with other traffic participants during on-road driving.

This two-part work aims to study the lane change behaviors at a tactical level from an on-road perspective. Motor way scenes are addressed, with a special focus on analyzing the interactions between an ego and the surrounding vehicles during the procedure. The research framework is shown in Fig. 1 and consists of studies at four levels. This work addresses the signal and trajectory levels, and a system of all-around vehicle trajectory collection through on-road driving in real-world traffic is proposed. The resulting trajectories based on long-term on-road driving at changing traffic conditions and across large areas are demonstrated to have high quality for tactical driving behavior studies. Based on this result, Part II addresses the semantic and behavioral levels, and an automated method for lane change extraction from a continuous driving sequence is proposed. An on-road database is subsequently developed, in which each data entry contains the trajectories of both the ego and environmental vehicles during the entire procedure of a lane change, and they are analyzed along the axes of both the ego's lane change trajectories and the interactions with scene vehicles.

Data collection aids in establishing a basis for driving behavior studies. Many studies are conducted using the data from driving simulators. However, for complex behavior in which the interactions with surrounding objects are crucial, the data from a state-of-the-art driving simulator are far from accurate. Recent research studies have invested great efforts to collect real-world data, among which the NGSIM (The Next Generation Simulation Program) [1] and NDS (Naturalistic Driving Study) [2]–[6] are the most notable programs. NGSIM aimed to support traffic simulation with a primary focus of microscopic driver behavior modeling. An important outcome of the program is the set of high-quality trajectory data, which was developed using video processing of the data from airborne or infrastructure-based cameras [7]. Although each data set contains the vehicle trajectories across a small road section for a short period, it provides a good basis for validation and calibration of the microscopic traffic models [8]–[11]. The 100-Car Naturalistic Driving Study [2] was the first large-scale naturalistic driving data collection project, and 100 instrumented vehicles were driven during participants' daily travel over a year. Recently, much larger efforts have been made, such as the SHRP2 [3], ANDS [4], UDRIVE [5] and an international NDS collaborative program [6]. The data have been used to analyze a driver's risk assessment near crash scenes [12], [13], detect events and assess driving quality [14]–[16]. In these NDS programs, data collection is addressed at the signal/object level, and trajectory

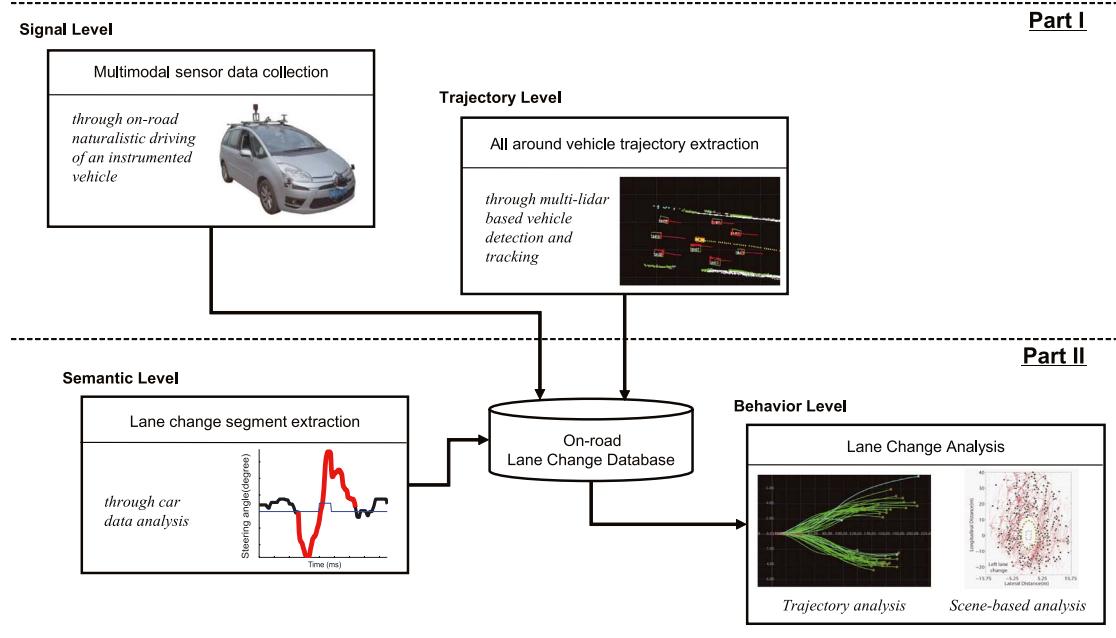


Fig. 1. Research framework.

data are not collected from the vehicles at the scene. Collecting vehicle trajectories during an on-road procedure still remains to be a challenging problem. No such dataset is publicly available, and to the authors' knowledge, such a technique has not been reported in the literature.

A trajectory is a time series of a vehicle's position across a continuous period. Trajectory collection in on-road procedures is crucial for quantitatively analyzing the interactions among different traffic components across extensive areas. This research aims to develop such a system, where during the driving of an instrumented vehicle (i.e., the ego) on roads, trajectories are collected of the ego as a subject and all the surrounding vehicles alongside. For tactical driving behavior studies, the trajectories are required to have the following properties:

- **Consistency.** All trajectories are synchronized and georeferenced to a common spatial-temporal frame.
- **Completeness.** Trajectories are collected simultaneously for all the vehicles that affect the subject's decision and/or maneuvers.
- **Continuity.** Trajectories are long enough to cover the entire procedure of complex behaviors such as lane-change or overtaking.
- **Accuracy.** Trajectories describe the time series of vehicles' positions accurately at an ego or world coordinate system.

This research proposes a system of on-road vehicle trajectory collection by using multiple 2D lidars that compose an ego-centered 360° horizontal coverage. To achieve accurate state estimation of severely occluded data and robust data association in the case of discontinuous observations that occur in multi-viewpoint sensing, software is developed by fitting the laser points of all lidars on a vehicle model using a coupled estimation of features and reliability along frames. The system performance is investigated extensively, and a large trajectory set is developed based on on-road driving at the 4th Ring Road

in Beijing over a total distance of 64 km that contains more than 5700 environmental trajectories with a total length of over 19 hr that is demonstrated to be of high quality in terms of consistency, completeness, continuity and accuracy, using an on-board omni-vision system, a reference vehicle for on-road experiments and maps. The system is used in Part II for lane change studies at the trajectory level, and a database is developed using ten days of data on on-road driving for a total distance of 1232 km. This work has general meaning as it can be used for studies on other driving and traffic behaviors, and a general knowledge of the system's limitation and data quality is crucial for designing experiments as well as the subsequent processing algorithms. To the authors' knowledge, this is the first system that is able to automatically collect all-around vehicle trajectories during on-road driving and is demonstrated to have a good performance in developing a high-quality database across large areas in real-world traffic for tactical driving behavior studies. A data is published at <http://www.poss.pku.edu.cn/download>.

This paper is structured as follows. A literature review is provided in Section II. An outline of the on-board system is described in Section III for both the sensing platform and multi-lidar data. The algorithm for vehicle trajectory extraction is described in Section IV. We present the experimental setting for an extensive performance study in Section V, followed by evaluations of the required properties as well as a description of the resultant trajectory data set in Section VI. Conclusion and future work are discussed in Section VII.

II. BACKGROUND

We provide an overview of related works, focusing on real-world data generation in support of microscopic driving behavior studies and state-of-the-art sensing techniques for on-road vehicle-trajectory acquisition.

A. Data for Microscopic Driver Behavior Studies

The data usually utilized in modeling driving maneuvers include the vehicle's proprioceptive data, such as braking, steering, gear shifts, and acceleration, the driver's gaze/head motion through in-vehicle sensing, and parameters available through exteroceptive sensing, such as lane, geographical position, and headway distance to the front vehicle [17]–[19]. In recent NDS studies [15], [16], the ego's drive is analyzed by correlations with the lane and surrounding vehicles to detect lane change events and characterize the drive quality in terms of risks and violations. These works generate models based on the current or short periods of historical data. However, the real driving behavior is much more complex: actions are planned based on anticipation of future conditions during a period of several tens of seconds, and the evolution of the ego vehicles position and its negotiations with surrounding vehicles play a crucial role. For such behavioral studies, trajectory data sets that represent vehicle interactions across the entire period of the procedures are necessary.

The earliest trajectory data sets supporting microscopic driving behavior studies can be traced to [20], which were collected at six freeway sections of typical geometries using airborne imaging systems. For example, the data collected in 1983 over a section of I-395 covered a four-lane highway section that was 997 m in length. Each data set corresponded to one hour with a frame rate of 1 Hz, and each frame contained the observations of the position, lane and dimensions of every vehicle within the section. Since then, considerable efforts and resources have been spent on trajectory data collection using the data from airborne [21] or infrastructure-based cameras, e.g., mounted on poles or tall buildings nearby [7], [22], [23]. Such real-world data provide a good basis for the validation and calibration of microscopic traffic models [8]–[11], whereas the data sets in the literature are collected at only a few roadway sites, and each covers only a small roadway section for a short period. A major reason for this limitation is the high cost of and numerous restrictions on data collection. Video processing is always time-consuming and labor-intensive, and site selection and sensor setting are not trivial. Normally, the site is required to have a tall building or poles nearby so that cameras can be positioned with good vantage points for measuring vehicle trajectories as well as reducing manual operations in video processing: the cameras views of the roadway should not be obstructed by trees, overpasses or buildings; the image blocks of different vehicles should be separated for automated estimation of vehicle positions; and the weather conditions should be conducive for video processing.

On the other hand, [24] gives an extensive review of traffic behavioral studies by analyzing vehicular trajectories. In this body of research, infrastructure videos are the major data resources, and few have been addressed from on-road perspectives, except [25], where the trajectories in a triangular zone in front of the ego are analyzed. However for applications such as ADAS or autonomous driving systems, modeling and reasoning methods from on-road perspectives are essential. The shortage of data and necessary techniques for collecting high-quality on-road trajectories constitutes a major research bottleneck.

B. On-Road Vehicle Trajectory Acquisition

Many studies on on-road traffic analysis have been conducted using the vehicle trajectories of probe vehicles that are instrumented with GPS receivers [26], [27]. These data are useful for providing a qualitative description of the roadway operation conditions; however, they contain no information on the scene. Surrounding vehicles can be measured from an ego frame using various sensors, such as cameras, lidar and radar, where a key issue is moving-object detection and tracking. The visual-based approach has been widely used as a major technique for developing active safety systems, which constitutes a large body of research [28]–[30]. The state-of-the-art methods were evaluated within the PASCAL Visual Object Classes Challenge [31] and recently in the KITTI's Vision Benchmark Suite [32] with challenging real-world datasets and benchmarks. Vision data provide rich cues for object detection, while great challenges are faced in collecting high-quality trajectories. In addition to the challenges from illumination changes and occlusions, vehicles have various shapes, sizes and colors, and they may be viewed perspective from different directions on multi-lane roads, yielding great difficulties in accurate pose estimations in a world coordinate system.

Range-based approaches using lidar are popular for detecting and tracking moving objects along a horizontal plane [33]–[35], and fusion-based approaches have been developed to improve efficiency and compensate for the limitations of mono-modal sensors [36]–[39]. In order to maximize sensing coverage, [35]–[38] use multiple lidars, radars and/or cameras. However, these systems are mainly used for timely awareness of scene objects nearby for applications such as collision warning or motion planning of an autonomous driving system, and the resultant trajectories of the tracking modules are not assessed based on their quality or used in tactical behavior studies. Recently, more studies have addressed using a powerful 3D lidar, Velodyne High-Definition LiDAR (HDL) [40], [41], which generates a 3D view of omni-directional surroundings with a frame rate of up to 15 Hz, so that the motions of all of the interacting moving objects can be clearly captured. However, the sensor has to be mounted extruding from the roof to reduce occlusion from the car body. This makes hardware maintenance difficult, which is not trivial for long-term operation in real-world traffic conditions. In addition, the sensor is very expensive, a big restriction for many potential applications.

Different from the general techniques of moving-object detection and tracking, trajectory collection for tactical driving behavior studies has additional challenges as it is required to have accuracy in vehicle state estimation, continuity across the period of complex maneuvers, omni-direction coverage of all interacting vehicles and spatial-temporal consistency. Such properties have never been demonstrated by any detection or tracking technique in the literature. A vision-based system is developed in [25], where vehicle trajectories are collected by combining active learning for monocular vehicle detection [42] and stereo vision for ranging and tracking. However, the system covers only a triangle zone in front of the ego. On the other hand, an omni-video-based approach is proposed in [43] to generate a dynamic panoramic map, and a visualization tool

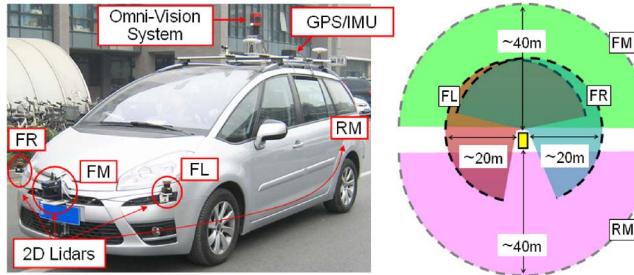


Fig. 2. On-board sensing system. (Left) Example of sensor setting. (Right) Lidar sensing coverage at a horizontal plane.

is developed in [44] with an iconic representation of the ego vehicle's vicinity. However, the demonstrated trajectories are fragmentary, providing only short periods of historical data.

In contrast with previous works in literature, this study develops a system of on-road vehicle trajectory collection in real-world traffic, and an extensive experiment is conducted to examine the trajectory quality on the aspects of consistency, completeness, continuity and accuracy for tactical driving behavioral studies.

III. ON-BOARD SENSING SYSTEM

A. Sensor Setting

The sensor setting of the on-board system is illustrated in Fig. 2. A GPS/IMU suite (XW-ADU7600 with FOG sensors) is used for determining the ego's location and orientation, through which the resultant trajectories are geo-referenced to a global coordinate system. However, due to GPS error and gyro drift, the localization accuracy is sometimes degraded. The ego position is also estimated by dead reckoning using the yaw rate and wheel speed from the vehicle's CAN. Although dead reckoning lacks global accuracy during long-term operation, it provides a smoother ego trajectory, which is more important for studying the driving behaviors at a tactical level that last for several tens of seconds. Both approaches are used in this research and are compared in the experimental study.

For measuring the trajectories of all-around vehicles, 2D lidars are mounted at the bumper level and scan horizontally along different directions. The major advantages of such a setting are that 1) ego-centered cross-sections with 360° coverage at that level can be easily generated by transforming all of the lidar data at the same time instances into a vehicle frame according to the sensors' calibration parameters and that 2) sensors could potentially be embedded into the vehicle body in the future, which is favorable from the viewpoint of maintenance for long-term operation. In this study, four HOKUYO lidars are used: two UTM-30LXs (short range model) on the left and right front bumper (FL and FR) and two UXM-30LXs (long range model) on the middle of the front and rear (FM and RM). All of the lidar profiles lie nearly along the same horizontal plane. An illustration of the lidars' coverage is given in Fig. 2. Based on experience, the normal measurement range of UTM-30LX and UXM-30LX for on-road vehicles is approximately 20 m and 40 m, respectively. However, it can be shorter for dark objects

or longer for bright objects. In addition to the above sensors, an omni-vision system, the PGR Ladybug 3, is mounted on the platform as well for visualization and result validation.

B. Multi-Lidar Data

Fig. 3 presents a single frame of multi-lidar scans as an example. (a–d) show the laser points of the FL, FM, FR and RM sensors in green, red, blue and pink, respectively. The laser beams of valid returns are visualized using yellow dashed lines, and the data from other lidars are shown in white for reference. (e) is an integrated frame of all lidar scans, which is visualized in the omni-image of the nearest time stamp in (g). (e–h) are annotated to illustrate the same vehicles for different data views. The results of vehicle detection and tracking using the proposed system are shown in (f) and (h) for reference, and each vehicle has an ID that can be used to correlate the same vehicle at different views. There are seven vehicles around the ego (v1–v7) in the scene. It can be found that after integrating the multi-lidar data, all these vehicles are measured simultaneously with more complete information than those from a single lidar. Such a sensor configuration has the advantage of achieving omnidirectional horizontal coverage, while introduces challenges too in data processing: 1) The occlusion is severe. As shown in (e), all the vehicles are observed in one or two facades, and some facades are only partially measured. 2) Data association is challenging. In case a vehicle shifts from the viewpoint of one lidar to another, a discontinuous change occurs in the data appearance.

IV. VEHICLE TRAJECTORY EXTRACTION

With the constraint that the cross-section of a vehicle at a horizontal plane can be approximated using a rectangular shape of reasonable size, a vehicle model is defined that contains observable and hidden state parameters, time-variant and invariant ones, and each is associated with a reliability item to tackle the occlusion problems. An algorithm is designed to allow for inference of *unreliable* features from *reliable* ones and for inheriting *reliable* time-invariant features along frames while updating only *unreliable* ones so as to achieve accurate state estimation and data association. Currently, only motor way scenes are addressed in this work. Such a condition greatly reduces the challenges in object recognition, as vehicles are the only moving objects.

A. Problem Statement

A trajectory T_i is a sequence of the state parameters of a vehicle i along the frames $k = 1, \dots, n$ from its detection to the end of tracking.

$$T_i = \{(t, p, \dots)_k | k = 1, \dots, n\}_i \quad (1)$$

where t_k is the time of frame k and p_k is the center point of vehicle i at the frame. The notation \dots , denotes that other parameters such as speed and direction can be exported in addition to the two basic ones according to the application needs.

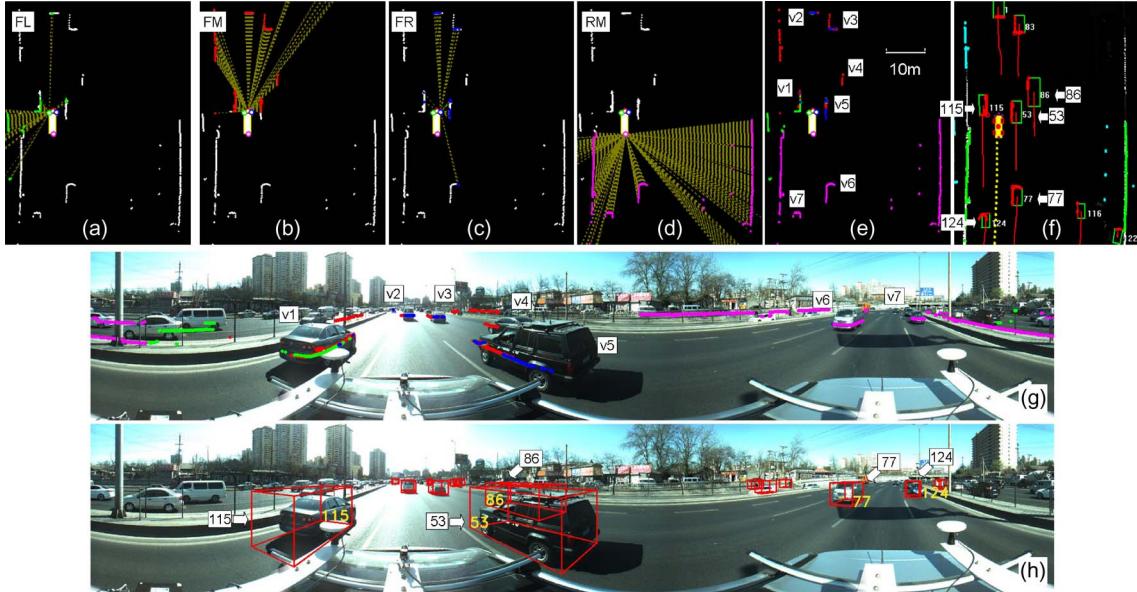


Fig. 3. Multilidar data integration. (a)–(d) Laser points of the FL, FM, FR, and RM sensors. (e) Integrated frame of all of the lidar scans. (f) Vehicle detection and tracking results using the proposed method. (g) Visualization of the laser points on the synchronized omni-image. (h) Visualization of the detection and tracking results on the omni-image.

However, a vehicle body is measured only partially, with its center point p , speed and directions being hidden states. According to the vehicle's pose relative to the ego, it could be measured on its front, rear and/or side parts, and as its relative location changes, e.g., from the ego's rear to the front, the measured body parts could have dramatic change. Especially in a multiple lidar system, a vehicle could be measured by multiple lidars simultaneously or could change from the viewpoint of one single lidar to another at successive frames, yielding discontinuous changes in the measured body parts as well as data appearance. The above discussion suggests that a lidar observation could show a center location biased with the vehicle's true center, and the data appearance could exhibit dramatic and discontinuous changes along frames. An improper estimation could yield a sequence of vehicle center points displaying irregular motion, and more importantly, it could lead to inaccurate state prediction and data association, thereby causing tracking failure and fragmented trajectories. Estimating the vehicle state parameters is key in achieving high-quality vehicle trajectory collection.

B. Vehicle State Estimation

We design a method by inferring the vehicle's hidden states on the other observable ones (called dependant features) with the assumption of a vehicle model. By extending the work in [46], a vehicle model is defined as illustrated in Fig. 4, where the horizontal contour of a vehicle is represented by using a rectangle with a sequence of variables. In this research, the features are defined as listed in Table I, which are broadly divided into two groups describing a vehicle's observable and hidden states at the frame. A full set of observable parameters is estimated at each frame with their reliability evaluated, where the more direct an observation is with less occlusion and uncertainty, the more *reliable* the estimation is. Reliability of a feature is defined as below.

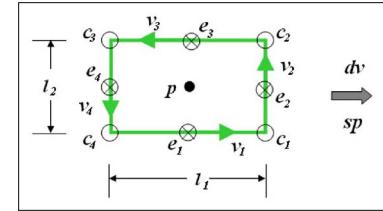


Fig. 4. Definition of a vehicle model.

TABLE I
PARAMETERS IN A VEHICLE MODEL

Item	Feature	Reliability
<i>Observable parameters</i>		
– Time-variant		
edge vectors ($i = 1, \dots, 4$)	v_i	rv_i
edge points ($i = 1, \dots, 4$)	e_i	re_i
corner points ($i = 1, \dots, 4$)	c_i	rc_i
– Time-invariant		
lengths on two vertical edges ($i = 1, 2$)	l_i	rl_i
<i>Hidden states</i>		
center point	p	rp
speed	sp	rsp
direction	dv	rdv

Reliability Definition: A feature is *reliable* if its value is estimated using a direct observation or inferred using the vehicle model with all *reliable* dependent features; it is *unreliable* if it is inferred using the vehicle model with some *unreliable* dependent features.

These parameters are used to provide a prediction to a vehicle's contour location, so as to associate the lidar measurements to solve the problem of discontinuous observation due to multi-viewpoints sensing, and infer the vehicle's hidden states with all or more *reliable* dependent features, so as to improve reliability in case of occlusion. Problems still remain: the features of horizontal and vertical edge lengths are indispensable in predicting a vehicle's contour location and

inferring the vehicle's hidden state such as the center point, while they are usually not fully measured at one frame, and as environmental vehicles may have various sizes, their values cannot be known beforehand. However different from the other parameters, the edge lengths are time invariant and possess the following properties:

Time-Invariant Property: Although the edge lengths of different environmental vehicles could be variant, the same vehicle has invariant edge lengths along the frames; an edge can be fully or partially observed or totally occluded, while without considering processing error, an edge should have a length no less than what has been observed.

Therefore, the observable parameters listed in Table I are further divided into two groups: time-variant ones, such as edge vectors and edge and corner points, and time-invariant ones, such as edge lengths. The following measures are taken to estimate the edge length by addressing time-invariant properties to improve the accuracy and reliability in the case of occlusion and discontinuous changes in the measurement data:

Time-Invariant Estimation: Once an edge length is reliably estimated, e.g., the corner points at both sides of an edge are reliable (i.e., directly measured), the value can be succeeded in the following frames, even when the edge disappears from the sensors' viewpoint; before an edge length is reliably estimated, it should only be updated when a new measurement has a larger value to gradually refine the estimation while preventing disturbances from further occlusions.

Estimation of the full set of parameters as well as their reliability is subsequently developed as detailed in List 1.

List1: Coupled Estimation of Feature and Reliability

- Edge vectors:

$v_{i=1,\dots,4}$ correspond to the four orthogonal sides of a vehicle. The v_i that is associated with the longest data cluster is first updated, with the others estimated on v_i accordingly to compose a counterclockwise loop. An edge vector is reliable, i.e., $rv_i = \text{true}$, if it is associated with a data cluster.

- Edge points:

$e_{i=1,\dots,4}$ are estimated as the middle points of edges. An edge point is reliable, i.e., $re_i = \text{true}$, if the corresponding edge is reliable.

- Corner points:

$c_{i=1,\dots,4}$ are the intersection points of neighboring edges. If both edge vectors are reliable, the c_i at their intersection is reliable, i.e., $rc_i = \text{true}$.

- Edge length:

l_1 and l_2 are the lengths on the directions of $v_{1,3}$ and $v_{2,4}$ respectively. If an edge length is already reliable, the feature and reliability values are succeeded; otherwise they are updated: $l_i = \max(l_i, \bar{l}_i)$, where \bar{l}_i is the currently length observation; if the pair of edge vectors or corner points on both ends of the subject edge are reliable, the estimated edge length is reliable, i.e., $rl_i = \text{true}$.

- Center point:

p is a hidden state, estimated on the vehicle's observation parameters. If all dependant parameters in estimation are reliable, p is reliable, i.e., $rp = \text{true}$.

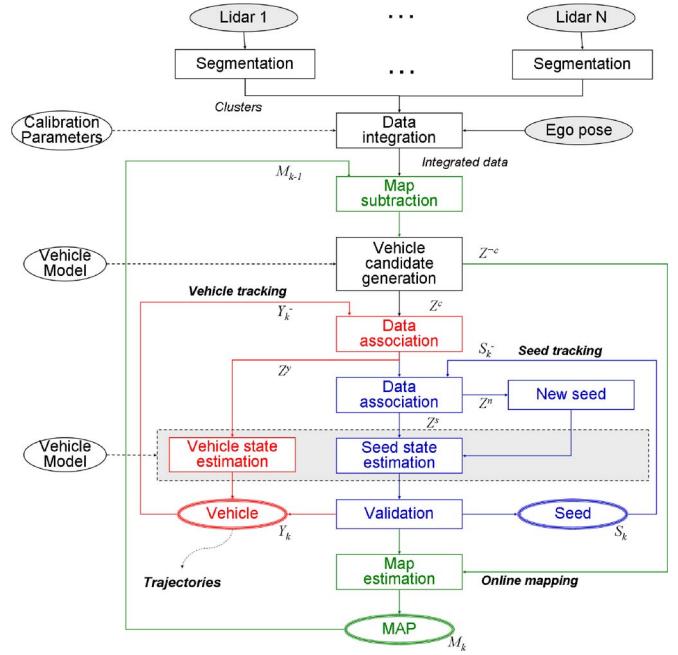


Fig. 5. Processing flow of vehicle trajectory extraction.

- Speed:

sp is inter-frame speed. If there exists a pair of corresponding points (center, corner or edge points with decreasing priority) of the successive frames that are both reliable, sp is estimated on them with $rsp = \text{true}$. Otherwise, sp is estimated on the pair of center points with $rsp = \text{false}$.

- Direction:

dv is the inter-frame directional vector. It is initialized with $rdv = \text{true}$, if a travel distance during the past frames is detected larger than a threshold. In case $rdv = \text{true}$, the edge vector v_i on the same direction with dv is set to v_1 , with the other parameters aligned accordingly.

C. Processing Flow

Fig. 5 illustrates the processing flow at each frame. An early version was reported in [45], which has been refined with more details. Each lidar scan of the latest measurement is first segmented according to range consistency and linear constraints to generate clusters. Each cluster consists of a set of laser points that could be fitted onto a line, which is an instantaneous observation of a side of an environmental vehicle or other objects. The clusters from all lidars are transformed into a vehicle frame according to their calibration parameters and subsequently into a global coordinate system according to the ego's pose at the moment, which is estimated using a GPS/IMU suite or dead reckoning.

Given a sequence of integrated data $Z_{0:k}$, where each Z_k contains the lidar clusters from all scans at frame k , the objective is to estimate a set of vehicles Y_k , seeds S_k and a map M_k . The processing is conducted in a framework of simultaneous mapping with vehicle detection and tracking as shown in Fig. 5.

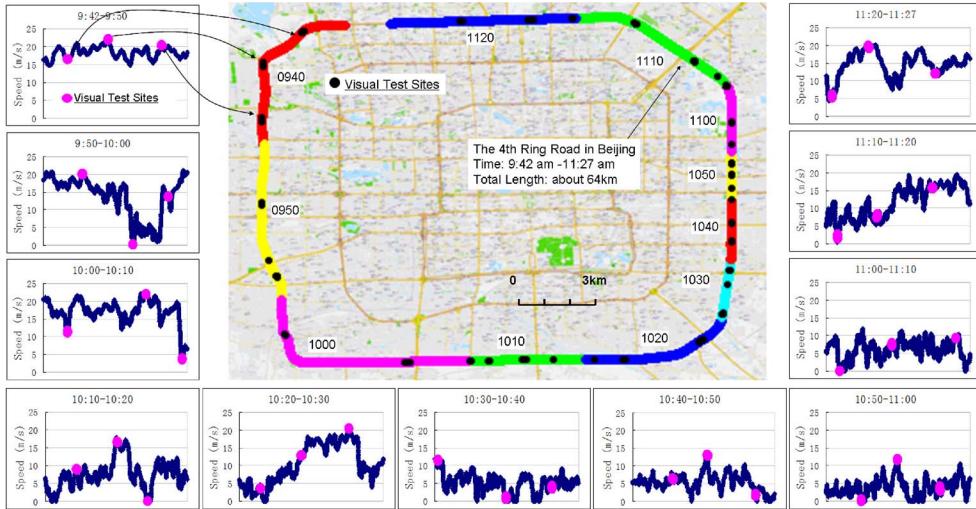


Fig. 6. Experimental setting, the ego's driving route, and speed.

A seed is an object that could potentially be a vehicle, e.g., its data can be fitted on a vehicle model within a reasonable size, but it has not been surely validated as either a vehicle or a static object due to insufficient observations. Seeds are tracked for detection, in order to solve the occlusion problems in on-road traffic, i.e., as the seed object is being tracked, more data are collected and more cues could be used to make a reliable inference. The seeds data set is conjugated with those of the vehicles and the map in a validation module. All seeds are validated at each frame with the criteria below, where the thresholds are predefined based on experiences.

- **Motion criteria**

- 1) Obvious motion: the trajectory during the past several frames is longer than the threshold.
- 2) Regular motion: the sequence of inter-frame speeds and directional vectors during the past frames has a variance lower than the threshold.

- **Shape criteria**

- 1) Perpendicular edges: a *reliable* edge vector is detected perpendicular to the motion direction, or one *reliable* edge vector has been continuously detected during frames that are longer than the threshold.
- 2) Reasonable size: both longitudinal and lateral edge lengths lie within thresholds, representing reasonable sizes for a normal vehicle.

If both the motion and shape criteria are met, the seed is validated as a vehicle and upgraded to the vehicle set. On the other hand, if no obvious motion is detected during a long frame or the reasonable size criteria are violated, the seed is considered as a static one, and its data are forwarded for map updating. If the regular motion criteria are violated, the seed is discarded as an irregular one. The remaining seeds are retained in the seed set as a decision that could not be made with certainty.

Data association and state updating is performed sequentially as depicted in Fig. 5. Given an integrated data set that comprises

the clusters from all lidar scans of the frame, map subtraction is first conducted with the previous M_{k-1} to extract foreground clusters from the integrated data. These data are then fitted onto the vehicle model so as to generate a set of candidates Z_k^c that have rectangular shapes within reasonable sizes. The rest set of the data Z_k^{-c} is forwarded to a map estimation module that exploits a general occupancy grid mapping (OGM) method [47] to conduct a frame-by-frame update. Meanwhile, the detected vehicle candidates Z_k^c are matched with the predicted rectangular contours of the vehicles Y_k^- then seeds S_k^- , and the associated observations Z_k^y and Z_k^s are used to update object states to the current frame as described in List 1. The rest sets $Z_k^n = Z_k^c - Z_k^y - Z_k^s$ are treated as a newly appeared potential vehicle and registered as new seeds.

V. EXPERIMENTAL SETTING

The proposed system has been used to collect trajectory data during naturalistic driving of an instrumented vehicle on the ring roads in Beijing, which are bi-directional motor ways with 4–5 lanes on each side that have no signalized intersections. In these experiments, raw sensor signals were first collected during on-road driving and processed afterwards using a simulated online mode for trajectory extraction. One set of experimental data was used to examine the trajectory quality as depicted in Fig. 6, where the instrumented vehicle (the ego) was driven entering the 4th ring road at 9:42 A.M. and leaving at 11:27 A.M. The data captured during its journey were studied. The raw sensor data were saved into different files (called segments) every 10 minutes with each named according to its starting time. The route of each segment is visualized in a different color, accompanied with a graph of the ego speed over the period. As the ego was driven naturally and continuously, its speed reflects the on-road traffic conditions, i.e., low for jammed while high for smooth traffic. The ego speed changed greatly during the course of the experiment, and the data were acquired across an extended area. With such a

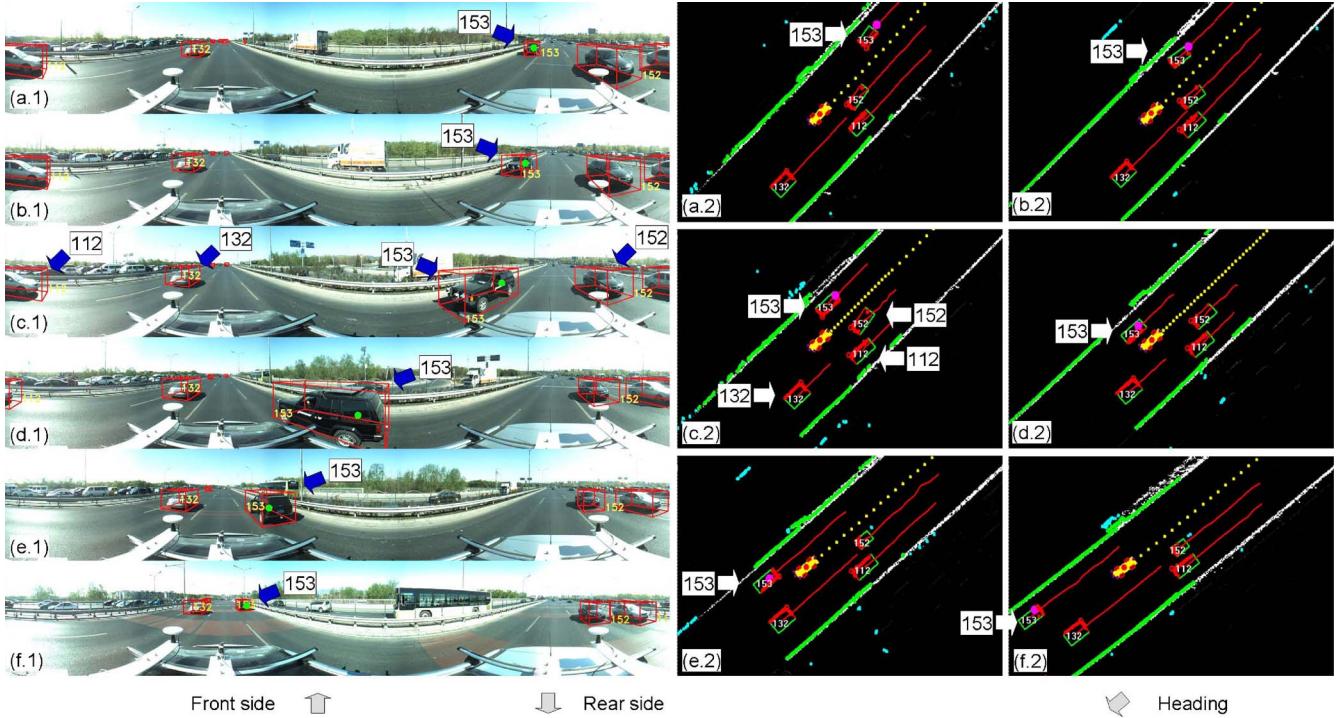


Fig. 7. Visual-based examination of trajectory extraction results in smooth traffic (successful results). (a.1–f.1) Visualization on the synchronized omni-images (image views). (a.2–f.2) Snapshots of multilidar processing (lidar views).

data set, the resultant trajectories were evaluated in a three-fold configuration.

1) *Completeness and Accuracy, Visual-Based Examination:* The examination determines to what extent the trajectories of all-around vehicles are collected. The results of multi-lidar processing are visualized in the synchronized omni-images. Statistics, such as *recall* and *precision*, are determined by manually counting *true* and *false* using the visually observable vehicles in omni-images as the ground truth, and “all around” is defined at two levels, within the direct neighbor lanes and the whole road. Statistics are reported at the test sites and depicted as pink and black dots in Fig. 6.

2) *Continuity and Accuracy, Reference Vehicle-Based Examination:* The examination determines how accurate and continuous the resultant trajectories are. Another vehicle (as the reference) participates in the on-road experiment and has a stand-alone GPS to record its position. It acts as an environmental vehicle that has a known trajectory for reference. The two vehicles were not tightly coordinated, but were to keep a close distance if the traffic allowed. Continuity and accuracy were examined by comparing the remotely sensed trajectories with the directly measured GPS one.

3) *Consistency and Accuracy, Map-Based Examination:* The examination was conducted by referring to a map-based representation of the road space. A lidar map was generated by processing the homologous lidar data, which provides a description of the road space, i.e., on-road, boundary and outside. Trajectories were subsequently labeled into positive and false, using which the accuracy was examined. Consistency was examined by overlapping the trajectories on the lidar map as well as a commercial map.

VI. RESULTS

We present results and discussions for each of the above addressed configurations, followed by a description of the resultant trajectory data set. A video of the results can be found at <http://www.poss.pku.edu.cn/video>.

A. Visual-Based Examination

A sequence of results for smooth traffic is shown in Fig. 7. The right two columns are snapshots of the lidar processing results (lidar view). They are projected to the synchronized omni-images (image view) for a more intuitive visualization. Each vehicle is detected and tracked using multi-lidar data in a horizontal plane and represented in the lidar view using a rectangle at the current position, an ID and a line for its historical trajectory. However, when projecting to perspective image views, the 2D results nearly appear as line segments. Therefore, a uniform height value (e.g., 2.0 m) is assigned to each 2D result to compose a cubic frame and then projected onto the omni-image, where the IDs of nearby vehicles (e.g., within 20 m from the ego) are shown, which can be used to correspond to the same detections in both views. There are four nearby vehicles in the scene. The sequence depicts a procedure where a vehicle (#153) drove from the rear to the front of the ego, during which all of the vehicles were detected and tracked continuously, i.e., with consistent IDs.

Based on the image views, the completeness and precision of vehicle detection under different traffic conditions are examined. As the ego was driven naturally, its speed, to some extent, reflects the traffic density. Three sites are picked up

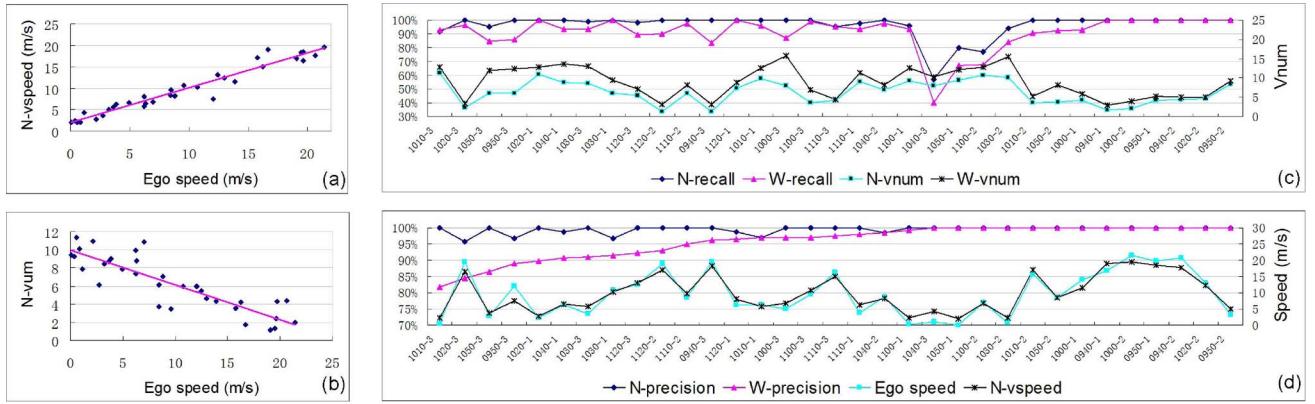


Fig. 8. Cross correlation of the statistics on the visual-based examination.

in each segment on the graphs of the ego speed at its low, middle and high levels. The distribution of the 33 test sites is shown in Fig. 6. At each site, the statistics are recorded for 10 continuous frames at 1 Hz using an operator for the vehicles that are visually observable in the omni-images as the ground truth. The vehicles concerned are those that have their front bumper enter and before their rear bumper leave, and the lines denoting a distance of 25 m from the ego center are marked specially for manual counting. The numbers of true positives (N_{TP}), false positives (N_{FP}) and false negatives (N_{FN}) are counted at two levels, i.e., within the direct neighboring lanes and the whole road, and are denoted using the prefixes of “N-” and “W-” respectively. Four measures are defined as below and estimated at both levels of all segments, where *recall* represents the completeness of detection, *precision* is the percentage of correctness, *vnum* is the actual number of vehicles in the scene, which is an average over the frames of the site with a frame number $N_{frame} = 10$, and *vspeed* reflects the traffic speed, which is an average over the vehicle speeds ($tj.sp$) of all true positives (TP)

$$recall = \frac{N_{TP}}{(N_{TP} + N_{FN})} \quad (2)$$

$$precision = \frac{N_{TP}}{(N_{TP} + N_{FP})} \quad (3)$$

$$vnum = \frac{(N_{TP} + N_{FP} + N_{FN})}{N_{frame}} \quad (4)$$

$$vspeed = \frac{\sum_{tj \in TP} tj.sp}{N_{TP}} \quad (5)$$

The cross-correlation of these measures and the ego speed are presented in Fig. 8. To exclude the effects of GPS error, these results are developed using the ego position with dead reckoning. (a) and (b) correlate the ego speed with *vnum* and *vspeed*, where, in general, the ego speed is inversely proportional to *vnum* and directly proportional to *vspeed*; the pink lines depict their tendencies. (c) correlates *recall* with *vnum*, compares the measurements at both the “N-” and “W-” levels, and sorts

N-recall in increasing order. Except for the lowest three, all of the *N-recall* values are above 0.92, with an average of 0.989, which means that the completeness of the trajectory collection is good within the direct neighboring lanes, a factor that is the most important for analyzing vehicle interactions. *W-recall* is generally lower than *N-recall*, while except for the lowest three, all of the *W-recalls* are above 0.83, with an average of 0.94. Low *recall* occurs when *N-vnum* is large, e.g., ≥ 7 , implying that many vehicles are missed in detection. (d) correlates *precision* with *vspeed* and is sorted based on *W-precision* in increasing order. All of the *N-precisions* are above 0.95, with an average of 0.994, implying a high rate of correctness in detecting the vehicles within direct neighboring lanes. The *W-precisions* are lower, with an average of 0.958, and five values are less than 0.9. A major reason for this decrease is that the road has 4–5 lanes in each direction, and false detections occur for the static objects on the roadside, which are beyond the neighboring lanes and thus have more effect on *W-precision*. Such false detections could be removed using offline map-based processing, which is addressed in Section VI-C.

Typical negative cases are analyzed in Fig. 9. Occlusion is a major reason as indicated by “A,” which is heavily occluded during frames. However, due to the asynchronous motion of the front vehicle, the observable part changed frequently. The data failed in being recognized as a vehicle, but were forwarded to map estimation, resulting in incorrect map pixels, as indicated by “B.” Sensing uncertainty is another reason, as indicated by “C.” The valid sensing distance can be different based on the vehicle’s material, color and shape. The body of vehicle “C” had its headlights measured only at the distance, resulting in two detached point clouds that failed in vehicle model fitting. Similar results were obtained for dark objects. A big limitation of the proposed algorithm has also been indicated by “D.” An obvious motion is a criterion for upgrading a seed to a vehicle, and in the case of stopped or slow motion, the vehicle detection performance is poor. These negative cases reveal the major weak points in lidar-based processing, which could be solved by incorporating visual cues that infer on vehicle appearance, e.g., [30]. A lidar-vision-fusion-based method will be developed in future work to improve the performance in slow or jammed traffic with heavy occlusions.

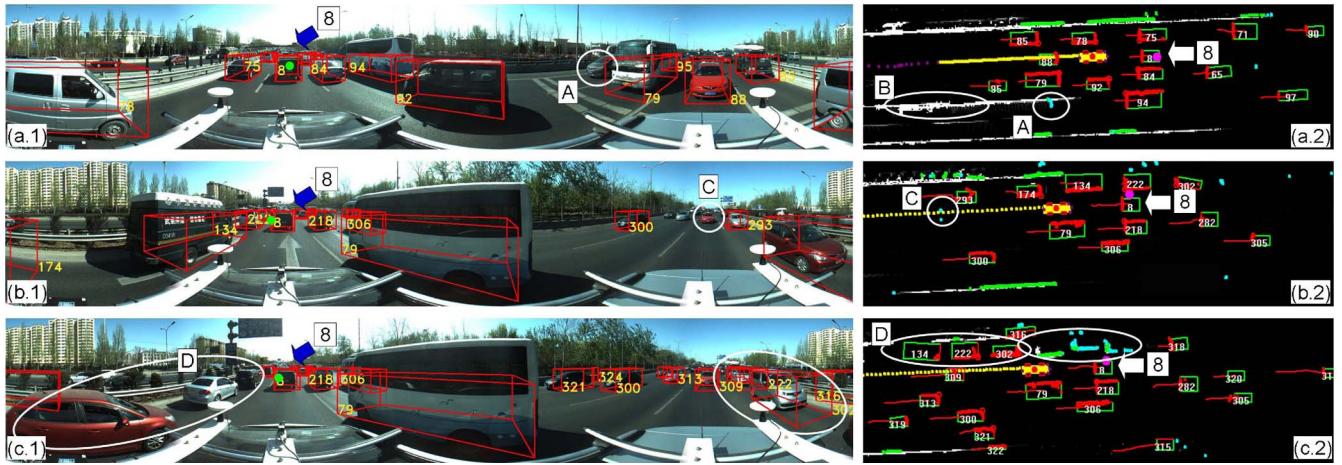


Fig. 9. Visual-based examination of trajectory extraction results in dense traffic (the challenges). (a.1–c.1) Visualization on the synchronized omni-images (image views). (a.2–c.2) Snapshots of multilidar processing (lidar views).

TABLE II
 STATISTICS OF THE RESULTS THROUGH REFERENCE VEHICLE-BASED EXAMINATION

Reference Vehicle Data			Succeeded Tracking						Failure Analysis			
Data segment	TOT(s)	TOD(m)	TTN	TT(s)	TD(m)	TTR	TDR	TSPD (m/s)	1st Reason	BR1	2nd Reason	BR2
0940	430	7817.89	7	388	7129.05	0.90	0.91	18.36	too far	0.08	-	-
0950	600	7973.77	12	526	6980.45	0.88	0.88	13.27	too far	0.06	occlusion	0.03
1000	600	10186.96	3	590	10049.96	0.98	0.99	17.02	too far	0.02	-	-
1010	600	4525.73	7	554	4401.32	0.92	0.97	7.94	algorithm failure	0.04	occlusion	0.03
1020	600	6290.10	5	488	4585.30	0.81	0.73	9.38	occlusion	0.14	too far	0.05
1030	600	3047.85	7	389	2203.85	0.65	0.72	5.66	occlusion	0.35	-	-
1040	600	3335.90	6	558	3041.64	0.93	0.91	5.45	occlusion	0.06	-	-
1050	600	2629.33	7	437	1770.91	0.73	0.67	4.05	GPS error	0.26	-	-
1100	600	3606.49	4	559	3328.98	0.93	0.92	5.95	GPS error	0.06	-	-
1110	600	7308.68	5	581	7042.50	0.97	0.96	12.11	GPS error	0.03	-	-
1120	504	7195.68	5	407	5794.06	0.81	0.81	14.22	too far	0.11	GPS error	0.08

TOT: total time; TOD: total distance; TTN: the number of tracked trajectories; TD: tracked distance; TT: tracked time;

TDR: the ratio of tracked distance (=TD/TOD); TTR: the ratio of tracked time (=TT/TOT);

TSPD: average speed of tracked trajectories;

1st Reason: the 1st reason for tracking failure; BR1: time ratio of tracking failure on 1st Reason;

2nd Reason: the 2nd reason for tracking failure; BR2: time ratio of tracking failure on 2nd Reason

B. Reference Vehicle-Based Examination

A jeep with a stand-alone GPS (Ref-GPS) was used as the reference in the experiment. In Figs. 7 and 9, the Ref-GPS data are visualized as green and pink dots on the image and lidar views respectively, so that the estimated trajectories (Lidar-Traj), which are #153 in Fig. 7 and #8 in Fig. 9, can be retrieved.

During the 64 km of driving, the reference is tracked by the ego for a total of 67 trajectories, with the longest one lasting for 330 s and 4906 m. There are a number of reasons for broken trajectories, among which 10 are caused due to file division. Table II lists some statistics of the tracking results at the level of each segment. The results of segment 0950 are analyzed below as an example. With the reference of the Ref-GPS trajectory that lasts for 7973.77 m (TOD) and 600 s (TOT), 12 Lidar-Trajs (TTN) are extracted by the ego and have a total distance of 6980.45 m (TD), a total time of 526 s (TT), and an average speed of 13.27 m/s (TSPD). Thus, the ratios of valid tracking time and distance are TT/TOD = 0.88 (TTR) and TD/TOD = 0.88 (TDR), respectively. One of the reasons for the $1 - TTR = 0.12$ failure in tracking time is that the reference is at a position too far to be measured by the lidars on the ego, which possess 0.06 (BR1); another is due to occlusion.

by other traffic, which possesses 0.03 (BR2). GPS error is also a major reason for tracking failure, which is severe in segment 1050. Tracking failures also occur due to improper data processing, where hijacking of the trajectories by different vehicles occurred twice. On the other hand, given the Ref-GPS as the ground truth, a statistic is recorded as shown in Fig. 11 to find the ratio of tracked time (TTR) when the reference vehicle is within a certain distance of the ego. Because the reference vehicle could be continuously blocked by others, the TTRs are low in some data segments such as 1030 and 1050. However, a general tendency of increasing TTRs is found as the distance decreases, and on average, TTRs within distances from 15 m to 40 m are in a range of [88%, 93%], which provides a general knowledge of the percentage of a vehicle being tracked in real-world traffic if it is within a certain distance.

A number of Lidar-Traj are compared with Ref-GPS, both under good GPS conditions. In Fig. 10 the coordinates of Ref-GPS are transformed to the synchronized ego frame, with the vertical (y-) axes corresponding to the direction of vehicle heading. Their residuals are evaluated quantitatively as shown in Fig. 12. For each Lidar-Traj point, dx and dy are estimated by considering the residual with the synchronized Ref-GPS

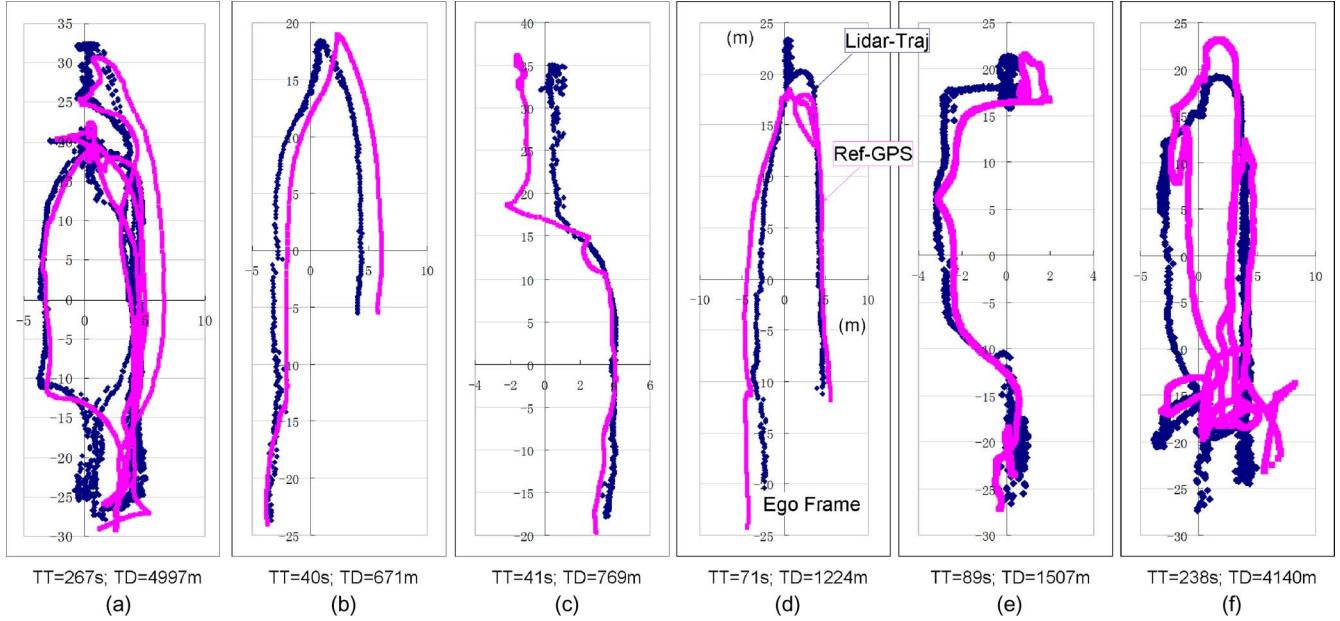


Fig. 10. Comparison of the reference vehicle's trajectories through multilidar processing (lidar-traj) and by a stand-alone GPS (ref-GPS).

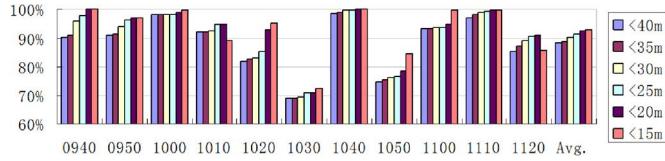


Fig. 11. TTR (the ratio of tracked time) of the reference vehicle with respective to its distance with the ego.

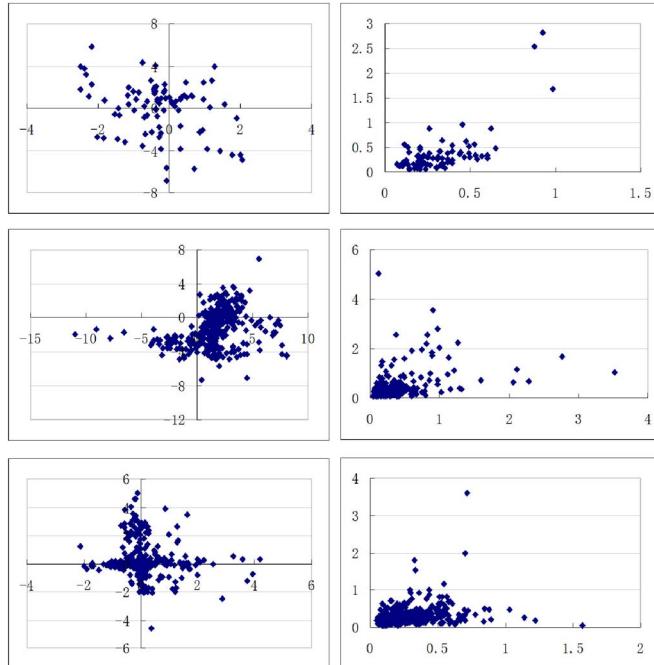


Fig. 12. Trajectory accuracy examination by comparing with the stand-alone GPS data of the reference vehicle. (Left) Mean error points. (Right) Standard deviation points. (Top) At the ego frame. (Middle) At a global frame. (Bottom) At a global frame after time compensation. Unit: meter.

one. The mean and standard deviation are estimated for the sequences of dx and dy every 10 seconds at both the ego and global frames. Because in this experiment, the ego and the reference are measured using different GPS modules, i.e., a GPS/IMU suite (XW-ADU7600) with FOG sensors and a stand-alone GPS (u-blox), a time delay is found in the two systems. Time compensation is conducted by looking for the point with the shortest Euclidean distance within the previous 500 ms, and a result is shown in the bottom of Fig. 12. After time compensation, more than 58% have mean errors less than 1 m, and more than 60% have standard deviations less than 0.235 m. It can be concluded that the accuracy of the remotely sensed trajectories reaches the level of direct measurement, when the GPS is under good conditions.

C. Map-Based Examination

A map-based labeling tool is developed to examine the trajectory accuracy, and the processing flow is illustrated using a sequence of results for a typical road segment in Fig. 13. A traditional OGM approach is used in map generation such that each grid pixel is estimated with a probability that it is occupied by an obstacle. As shown in (a), the static environment is represented as black for occupied and white for the spaces that are free of objects. The white strip zone corresponds to on-road space, and the black pixels at its edge are the barriers on the road sides. Road boundary lines are extracted as below. Initiated from the ego trajectory on a sequence of equally tessellated points, edge points on the left and right are detected by searching laterally to the nearest occupied pixels, shown as the green dots in (b). Because there could be errors in estimations on wrong map pixels, manual corrections are performed in (c). Finally, boundary lines are developed on continuous edge points within a reasonable distance, and red lines denote the pairs of successive edge points that are far from



Fig. 13. Map-based examination results at a typical road segment. (a) OGM. (b) Automated estimation of the edge points on road boundary. (c) RBD (road boundary) generation with manual correction. (d) Map space labeling. (e) Trajectory labeling. (f) Visualization of the trajectory set on Google Earth.

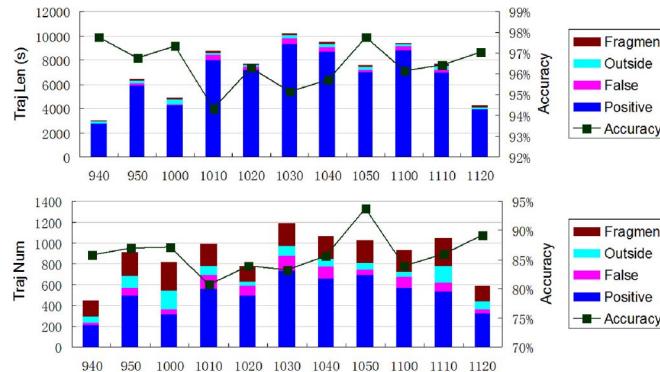


Fig. 14. Trajectory accuracy examination through map-based labeling.

each other, indicating an exit or entrance. The mapped space is thus divided into three parts as shown in (d) after removing the short trajectories that are less than 3 s and are labeled as *fragment*, the rest are labeled as *positive*, *false* and *outside* and are within on-road, boundary and outside spaces and shown in pink, green and blue in (e), respectively, where four-lane traffic is clearly depicted using the trajectories. In the case of GPS-based ego localization, the trajectories can be analyzed using a GIS map, and an overlapped visualization on Google Earth is presented in (f). Based on the above labeling result, the trajectory accuracy is examined using the following measures.

$$\text{Accuracy}(\text{TrajNum}) = \frac{T_{tjn}}{(T_{tjn} + F_{tjn})} \quad (6)$$

$$\text{Accuracy}(\text{TrajLen}) = \frac{T_{tjl}}{(T_{tjl} + F_{tjl})} \quad (7)$$

Here, respectively, T_{tjn} and T_{tjl} are the number and length (in seconds) of the *positive* trajectories and F_{tjn} and F_{tjl} are those of the *false* ones. The accuracy measures do not concern fragmentary or outside trajectories because their correctness is difficult to prove.

The trajectory accuracy is shown in Fig. 14. At each segment, the composition of the trajectory set is analyzed on the aspects of both TrajNum and TrajLen, and the estimated accuracy is plotted. The accuracy ranges in [80.7%, 93.8%] for TrajNum,

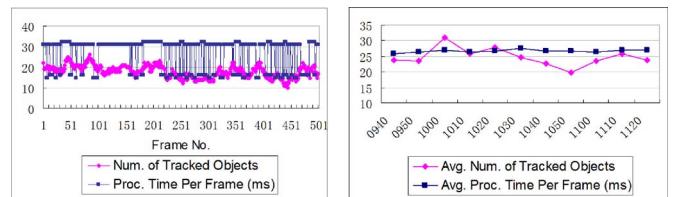


Fig. 15. Computation time with respect to the number of tracked objects (vehicles and seeds). (Left) Result in processing a sequence of frames. (Right) Average time per frame in processing each data segment.

and is largely increased to [94.3%, 97.8%] for TrajLen. A reason is that *false* trajectories are much shorter than *positive* ones. Within the automatically collected trajectories, the *positive*, *false*, *outside* and *fragment* trajectories occupy 91.4%, 3.6%, 3.1% and 1.9% of their lengths, respectively.

D. Computation Cost

The computational cost of each frame is shown in Fig. 15, where (a) shows the results of a sequence of frames and (b) shows the average values for processing each data segment. Such a cost does not include that of sensor data collection. The processing time per frame lies mainly within a range of [15, 31] ms, and the average values per data segment are nearly constant, around 26–27 ms. The experiment is conducted on a ThinkPad X230 with a 2.60 GHz CPU. Although the accuracy in time estimation is limited to within the error range of Windows APIs, compared to the usually required frame rate of trajectory estimation, which is 10 Hz (100 ms per frame), such a computational cost allows for on-line processing. In addition, the number of tracked objects, containing both vehicles and seeds, are shown in the figure for reference. Although the number of tracked objects varies in frames, the processing time does not show correlations with them. One reason is that the system contains other processing modules such as clustering, vehicle candidate generation, map estimation and subtraction. Especially for map estimation, a lower number of tracked objects means more data for map updating. With the same lidar sensor configuration, the computational cost could be kept nearly constant on average.

TABLE III
TRAJECTORY DATA SET

	GPS-based Ego Localization				DR-based Ego Localization			
	Ego Traj		Env Traj		Ego Traj		Env Traj	
	ED	EL	TC	TL	ED	EL	TC	TL
0940	(m)	(s)	(s)	(s)	(m)	(s)	(s)	(s)
7818	7818	430	209	2747	7647	430	205	2750
0950	7974	600	509	5830	7794	600	498	5875
1000	10187	600	309	4181	10088	600	318	4204
1010	4526	600	581	7984	4442	600	565	7960
1020	6290	600	492	7034	6018	600	494	7165
1030	3048	600	759	9153	2949	600	729	9293
1040	3336	600	694	7579	3202	600	661	8661
1050	2629	600	706	6734	2179	600	694	7034
1100	3606	600	610	8788	3608	600	571	8808
1110	7309	600	522	6639	6978	600	531	6903
1120	7196	504	336	3926	7038	504	325	3952
Total	(km)	(hr)	(hr)	(hr)	(km)	(hr)	(hr)	(hr)
	63.9	1.76	5727	19.6	61.9	1.76	5591	20.2

ED: ego traj dis; EL: ego traj len; TC: the num of env vehicle traj; TL: total len of env vehicle traj.

E. The Trajectory Data Set

The experiment resulted in a data set that is broadly divided into three groups at the sensor level, i.e., multi-modal sensor data such as lidar and videos; object level, i.e., the maps; and trajectory level. The data at the trajectory level are listed in Table III, and according to the method of ego localization, i.e., GPS or dead reckoning (DR), they are further divided into those in a global coordinate system, which can be associated with a GIS map, or in a relative coordinate system that has good local consistency for behavior analysis. All of the data frames are stamped with synchronized time logs. For the total of all segments, during the ego's drive in 1.76 hr, an ego trajectory lasting for 63.9 km and 5727 environmental trajectories with a total length over 19.6 hr are extracted in the global coordinate system, while an ego trajectory lasting for 61.9 km and 5591 environmental ones with a total length over 20.2 hr are extracted in a relative frame.

VII. CONCLUSION AND FUTURE WORKS

This article has addressed on-road vehicle trajectory collection on motor ways as the first part of a two-part work for lane change behavior studies at the tactical level. The trajectories for such a study are required to have the properties of consistency, completeness, continuity and accuracy. To this end, an onboard system of all-around vehicle trajectory collection is developed using multiple horizontal lidars that compose an ego-centered 360° coverage. Software is developed to achieve accurate vehicle state estimation of occluded data and robust data association with multi-viewpoint sensing. The system performance is investigated extensively using a large-scale trajectory collection experiment, and the results demonstrate high quality in terms of the above addressed properties. The results of this work serve as necessary prerequisites, while not limited, for Part II for lane change behavior studies from an on-road perspective that addresses vehicle interactions in real-world traffic at the trajectory level. Future work will address a lidar-vision-fusion-based method to achieve continuous tracking in slow or jammed traffic with heavy occlusions and improve the accuracy of object recognition at downtown streets and intersection scenes, where collecting the trajectories of all kinds of traffic participants is crucial.

REFERENCES

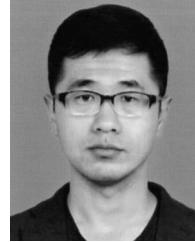
- [1] The Next Generation Simulation Program (NGSIM). [Online]. Available: <http://ngsim-community.org/>
- [2] V. L. Neale, T. A. Dingus, S. G. Klauer, J. Sudweeks, and M. Goodman, "An overview of the 100-car naturalistic driving study and findings," in *Proc. 19th Int. Tech. Conf. Enhanc. Safety Veh.*, 2005, pp. 1–10.
- [3] K. L. Campbell, "The SHRP 2 naturalistic driving study: Addressing driver performance and behavior in traffic safety," *TR News*, vol. 282, pp. 30–35, 2012.
- [4] M. A. Regan *et al.*, "The Australian 400-car naturalistic driving study: Innovation in road safety research and policy," in *Proc. Australasian Road Safety Res., Policing Educ. Conf.*, 2013, pp. 1–13.
- [5] R. Eeninka, Y. Barnardb, M. Baumannc, X. Augrosd, and F. Uteschc, "UDRIVE: The European naturalistic driving study," in *Proc. Transp. Res. Arena*, 2014, pp. 1–10.
- [6] K. Takeda, J. Hansen, P. Boyraz, C. Miyajima, and H. Abut, "International large-scale vehicle corpora for research on driver behavior on the road," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1609–1623, Dec. 2011.
- [7] V. Kovvali, V. Alexiadis, and L. Zhang, "Video-based vehicle trajectory data collection," in *Proc. Annu. Meet. Transp. Res. Rec.*, 2007, pp. 1–18.
- [8] H. N. Koutsopoulos and H. Farah, "Latent class model for car-following behavior," *Transp. Res. B, Methodol.*, vol. 46, no. 5, pp. 563–578, Jun. 2012.
- [9] M. Montanino and V. Punzo, "Making NGSIM data usable for studies on traffic flow theory: Multistep method for vehicle trajectory reconstruction," *J. Transp. Res. Board*, no. 2390, pp. 99–111, 2013.
- [10] J. Zheng, K. Suzuki, and M. Fujita, "Predicting drivers lane-changing decisions using a neural network model," in *Proc. IEEE Intell. Veh. Symp.*, 2014, pp. 108–114.
- [11] S. Moridpour, E. Mazloumi, and M. Mesbah, "Impact of heavy vehicles on surrounding traffic characteristics," *J. Adv. Transp.*, vol. 49, no. 4, pp. 535–552, 2014.
- [12] F. Guo and Y. Fang, "Individual driver risk assessment using naturalistic driving data," *Accident Anal. Prevent.*, vol. 61, pp. 3–9, 2013.
- [13] J. K. Jonasson and H. Rootzen, "Internal validation of near-crashes in naturalistic driving studies: A continuous and multivariate approach," *Accident Anal. Prevent.*, no. 62, pp. 102–109, 2014.
- [14] R. K. Satzoda and M. M. Trivedi, "Overtaking & receding vehicle detection for driver assistance and naturalistic driving studies," in *Proc. IEEE ITSC*, 2014, pp. 697–702.
- [15] R. K. Satzoda, P. Gunaratne, and M. M. Trivedi, "Drive quality analysis of lane change maneuvers for naturalistic driving studies," in *Proc. IEEE Intell. Veh. Symp.*, 2015, pp. 654–659.
- [16] R. K. Satzoda and M. M. Trivedi, "Drive analysis using vehicle dynamics and vision-based lane semantics," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 9–18, Feb. 2015.
- [17] P. Kumar, M. Perrollaz, S. Lefevre, and C. Laugier, "Learning-based approach for online lane change intention prediction," in *Proc. IEEE Intell. Veh. Symp.*, 2013, pp. 797–802.
- [18] M. Mori, C. Miyajima, T. Hirayama, and N. Kitaoka, "Integrated modeling of driver gaze and vehicle operation behavior to estimate risk level during lane changes," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, 2013, pp. 2020–2025.
- [19] S. Martin and M. M. Trivedi, "Continuous head movement estimator for driver assistance: Issues, algorithms, and on-road evaluations," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 2, pp. 818–830, Apr. 2014.
- [20] S. Smith, "Freeway data collection for studying vehicle interactions," Federal Highway Admin., US Dept. Transp., McLean, VA, USA, Tech. Rep. FHWA/RD-85/108, 1985.
- [21] M. Schreuder, S. P. Hoogendoorn, H. J. Van Zuylen, B. Gorte, and G. Vosselman, "Traffic data collection from aerial imagery," in *Proc. IEEE Conf. Intell. Transp. Syst.*, 2003, pp. 779–784.
- [22] R. Ervin *et al.*, *System for Assessment of the Vehicle Motion Environment (SAVME)*, UMTRI-2000-21-1, 2000.
- [23] A. Skabardonis, "Estimating and validating models of driver behavior with video data," California PATH Program, Inst. Transp. Stud. Richmond Field Station, Richmond, CA, USA, PATH Res. Rep. UCB-ITS PRR-2005-14, 2005.
- [24] B. T. Morris and M. M. Trivedi, "Understanding vehicular traffic behavior from video: A survey of unsupervised approaches," *J. Electron. Imaging*, vol. 22, no. 4, 2013, Art. no. 041113.
- [25] S. Sivaraman, B. Morris, and M. Trivedi, "Learning multi-lane trajectories using vehicle-based vision," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2011, pp. 2070–2076.
- [26] F. Dion, R. Robinson, and J. Oh, "Evaluation of usability of intellidrive probe vehicle data for transportation systems performance analysis," *J. Transp. Eng.*, vol. 137, no. 3, pp. 174–183, 2011.

- [27] Y. Zhu, Z. Li, H. Zhu, M. Li, and Q. Zhang, "A compressive sensing approach to urban traffic estimation with probe vehicles," *IEEE Trans. Mobile Comput.*, vol. 12, no. 11, pp. 2289–2302, Nov. 2013.
- [28] Z. Sun, G. Bebis, and R. Miller, "On-road vehicle detection: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 694–711, May 2006.
- [29] S. Sivaraman and M. M. Trivedi, "Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1773–1795, Dec. 2013.
- [30] C. Wang, Y. Fang, H. Zhao, C. Guo, S. Mita, and H. Zha, "Probabilistic inference for occluded and multiview on-road vehicle detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 215–229, Jan. 2016.
- [31] PASCAL Visual Object Classes Challenges. [Online]. Available: <http://pascallin.ecs.soton.ac.uk/challenges/VOC/>
- [32] The KITTI Vision Benchmark Suite. [Online]. Available: <http://www.cvlibs.net/datasets/kitti/>
- [33] C. C. Wang, C. Thorpe, S. Thrun, M. Hebert, and H. Durrant-Whyte, "Simultaneous localization, mapping and moving object tracking," *Int. J. Robot. Res.*, vol. 26, no. 9, pp. 889–916, 2007.
- [34] D. Wang, I. Posner, and P. Newman, "A new approach to model-free tracking with 2D lidar," in *Proc. Int. Symp. Robot. Res.*, 2013, pp. 557–573.
- [35] C. Mertz *et al.*, "Moving object detection with laser scanners," *J. Field Robot.*, vol. 30, no. 1, pp. 17–43, 2013.
- [36] J. Wei, J. M. Snider, K. Junsung, J. M. Dolan, R. Rajkumar, and B. Litkouhi, "Towards a viable autonomous driving research platform," in *Proc. IEEE Intell. Veh. Symp.*, 2013, pp. 763–770.
- [37] K. Schueler, T. Weiherer, E. Bouzouraa, and U. Hofmann, "360 degree multi sensor fusion for static and dynamic obstacles," in *Proc. IEEE Intell. Veh. Symp.*, 2012, pp. 692–697.
- [38] H. Cho, Y.-W. Seo, B. V. K. V. Kumar, and R. RajKumar, "A multi-sensor fusion system for moving object detection and tracking in urban driving environments," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2014, pp. 1836–1843.
- [39] T.-D. Vu, O. Aycard, and F. Tango, "Object perception for intelligent vehicle applications: A multi-sensor fusion approach," in *Proc. IEEE Intell. Veh. Symp.*, 2014, pp. 100–106.
- [40] A. Azim and O. Aycard, "Detection, classification and tracking of moving objects in a 3D environment," in *Proc. IEEE Intell. Veh. Symp.*, 2012, pp. 802–807.
- [41] F. Moosmann and C. Stiller, "Joint self-localization and tracking of generic objects in 3D range data," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2013, pp. 1138–1144.
- [42] S. Sivaraman and M. Trivedi, "A general active-learning framework for on-road vehicle recognition and tracking," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 2, pp. 267–276, Jun. 2010.
- [43] T. Gandhi and M. Trivedi, "Vehicle surround capture: Survey of techniques and a novel omni-video-based approach for dynamic panoramic surround maps," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 7, no. 3, pp. 293–308, Sep. 2006.
- [44] B. Morris and M. Trivedi, "Vehicle iconic surround observer: Visualization platform for intelligent driver support applications," in *Proc. IEEE Intell. Veh. Symp.*, 2010, pp. 168–173.
- [45] H. Zhao *et al.*, "Omni-directional detection and tracking of on-road vehicles using multiple horizontal laser scanners," in *Proc. IEEE Intell. Veh. Symp.*, 2012, pp. 57–62.
- [46] H. Zhao *et al.*, "Detection and tracking of moving objects at intersections using a network of laser scanners," in *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 655–670, Jun. 2012.
- [47] S. Thrun, "Robotic mapping: A survey," in G. Lakemeyer and B. Nebel, Eds., *Exploring Artificial Intelligence in the New Millennium*. San Mateo, CA, USA: Morgan Kaufmann, 2002.



Huijing Zhao received the B.S. degree in computer science from Peking University, Beijing, China, in 1991 and the M.E. and Ph.D. degrees in civil engineering from the University of Tokyo, Tokyo, Japan, in 1996 and 1999, respectively.

From 1991 to 1994, she was recruited by Peking University in a project of developing a geographic information system platform. In 2003, after several years of postdoctoral research with the University of Tokyo, she became a Visiting Associate Professor with the Center for Spatial Information Science. In 2007, she joined Peking University as an Associate Professor with the School of Electronics Engineering and Computer Science. Her research interests include intelligent vehicle, machine perception, and mobile robots.



Chao Wang received the B.S. degree in automation from Tsinghua University, Beijing, China, in 2010 and the Ph.D. degree in computer science (intelligent science and technology) from Peking University, Beijing, in 2015.

His research interests include intelligent vehicles, intelligent transportation systems, computer vision, and machine learning.



Yuping Lin received the B.S. degree in electrical engineering from Sun Yat-Sen University, Kaohsiung, Taiwan, in 1997, the M.S. degree in mechanical engineering from Tatung University, Taipei, Taiwan, in 2004, and the M.S. degree in computer science from the University of Detroit Mercy, Detroit, MI, USA, in 2007. He is currently working toward the Ph.D. degree in the Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China.

His research interests include computer vision

and machine perception.



Franck Guillemard was born in France on March 18, 1968. He received the Ph.D. degree in control engineering from the University of Lille, Lille, France, in 1996.

Currently, he is with the Scientific Department of Groupe PSA (Peugeot Société Anonyme), where he is in charge of advanced research concerning computing science, electronics, and photonics and control. He is also an expert in the field of modeling, design, and control of automotive mechatronic systems.



Stéphane Geronimi received the Ph.D. degree in physics and optronics from the University of Paris XI, Orsay, France, in 1996.

He joined the Groupe PSA (Peugeot Société Anonyme) in 2000. Since then, he has been involved in several research projects in the domain of driver assistance systems. He has been in charge of innovation projects dealing with advanced driver assistance systems (ADAS). He is currently an expert in ADAS and safety for ADAS and is working on the automation of driving functions.



François Aioun received the engineering degree in electronics, computer science, and automatic control from the École Supérieure d'Informatique, Électronique, Automatique High School, Paris, France, in 1988, the Postgraduate Diploma in automatic control and signal processing in 1989, and the Ph.D. degree in automatic control in 1993.

From 1989 to 1993, he was recruited by Électricité de France to study active vibration control of a structure. After the postdoctoral research at Ecole Normale Supérieure de Cachan, Cachan, France, and in different companies, he joined the Groupe PSA (Peugeot Société Anonyme) in 1997. His research interests in automatic control cover active vibration, power plant, powertrain, actuators, and more recently, autonomous and intelligent vehicles.