

Multivehicle Cooperative Driving Using Cooperative Perception: Design and Experimental Validation

Seong-Woo Kim, *Member, IEEE*, Baoxing Qin, Zhuang Jie Chong, Xiaotong Shen, Wei Liu, Marcelo H. Ang, Jr., Emilio Frazzoli, *Senior Member, IEEE*, and Daniela Rus, *Fellow, IEEE*

Abstract—In this paper, we present a multivehicle cooperative driving system architecture using cooperative perception along with experimental validation. For this goal, we first propose a multimodal cooperative perception system that provides see-through, lifted-seat, satellite and all-around views to drivers. Using the extended range information from the system, we then realize cooperative driving by a see-through forward collision warning, overtaking/lane-changing assistance, and automated hidden obstacle avoidance. We demonstrate the capabilities and features of our system through real-world experiments using four vehicles on the road.

Index Terms—Cooperative driving, cooperative perception, driving assistance, see-through system, vehicle communication.

I. INTRODUCTION

COOPERATIVE perception, via the exchange of sensor information between vehicles via wireless communications, is an emerging technology that is receiving attention from industry and academia [1]–[3]. The main advantage of cooperative perception can be summarized as an increase in situational awareness, without substantial additional costs. For example, cooperative perception can enable a driver to have a longer perception range—even beyond line-of-sight and field-of-view—because the sensing region can be extended to the union of the sensing regions of all connected vehicles. Since the prices of sensors and radio devices for cooperative perception are affordable compared with conventional long-range sensors, this results in a clear benefit for users.

Fig. 1 shows examples of this perception range extension. Fig. 1(a) and (b) are snapshots of maps of the ego vehicle running the proposed cooperative perception system, where the ego vehicle is represented as the box located at the bottom. The ego vehicle can see the preceding vehicles and an oncoming vehicle beyond line-of-sight in Fig. 1(a) and beyond field-

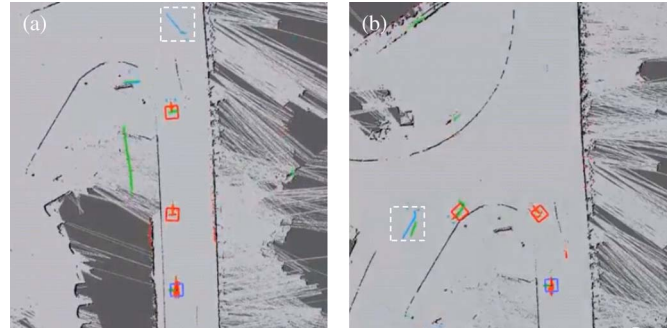


Fig. 1. Two examples of sensing range extension using cooperative perception: detecting an oncoming car (a) beyond line-of-sight, and (b) beyond field-of-view at sharp curb. The blue and red boxes correspond to the ego and preceding vehicle, respectively. The red, green, and blue dots are laser scan points of the ego, first and second preceding vehicle, respectively. The white dotted boxes indicate coming vehicles not detected by the ego vehicle.

view at the sharp corner in Fig. 1(b), where the oncoming vehicle is represented by the white dotted box. This aspect enables better driving decisions, such as overtaking, avoiding hidden or sudden obstacles, or early lane change against lane drop, and eventually could improve traffic safety and efficiency [4], [5]. Despite these advantages, however, there have been comparatively less research works on cooperative perception for Intelligent Transportation Systems (ITS)-related applications. This paper considers cooperative perception to enable cooperative driving.

In this paper, we explore a multivehicle cooperative driving system using cooperative perception along with experimental validation. The goals of this paper are to provide a comprehensive argument for such systems and to demonstrate the engineering feasibility. Building up to this, we first propose a cooperative perception system that can provide a far-sight see-through, lifted-seat, satellite or all-around view to a driver. The key purpose of this system is to provide as much visibility as if driver's seat would be lifted as much as express buses or trucks. Based on the augmented on-road sensing capability, we then present a cooperative driving system that can improve traffic safety and efficiency, specifically, by a see-through forward collision warning, overtaking/lane-changing assistance, and automated hidden obstacle avoidance. All the presented concepts are evaluated in experiments using four vehicles.

The contribution of this paper can be summarized as the following.

- We present our design of a multimodal cooperative perception system that can provide a far-sight see-through, lifted-seat, satellite or all-around view to a driver.

Manuscript received September 15, 2013; revised January 5, 2014, April 19, 2014, June 18, 2014, and June 24, 2014; accepted June 29, 2014. Date of publication July 28, 2014; date of current version March 27, 2015. This work was supported by the Future Urban Mobility project of the Singapore-MIT Alliance for Research and Technology (SMART) Center, with funding from Singapore's National Research Foundation. The Associate Editor for this paper was C. Olaverri-Monreal.

S.-W. Kim is with Singapore-MIT Alliance for Research and Technology, Singapore 138602 (e-mail: sungwoo@smart.mit.edu).

B. Qin, Z. J. Chong, X. Shen, W. Liu, and M. H. Ang Jr. are with the Department of Mechanical Engineering, National University of Singapore, Singapore 117575 (e-mail: baoxing.qin@nus.edu.sg; chongzj@nus.edu.sg; shen_xiaotong@nus.edu.sg; liu_wei@nus.edu.sg; mpeangh@nus.edu.sg).

E. Frazzoli and D. Rus are with Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: frazzoli@mit.edu; rus@csail.mit.edu).

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.2014.2337316

- We present a cooperative driving system using cooperative perception for traffic safety and efficiency along with experimental validation using four vehicles.
- We provide a comprehensive argument for the cooperative perception-based cooperative driving system from a viewpoint of engineering feasibility.

The remainder of this paper is organized as follows. Section II presents the related works. Section III introduces the system architecture of our proposal consisting of a multimodal cooperative system and perspective visualization for driver assistance, whose details are presented in Sections IV and V. Section VI presents how our system can support a see-through forward collision warning, overtaking assistance, and lane-changing assistance. Section VII provides scalability analysis. Section VIII provides experimental results performed with vehicles on urban roads. Section IX concludes this paper.

II. RELATED WORKS

A. Cooperative Perception on the Road

To enable cooperative perception, the distributed information from multiple vehicles should be properly fused. One of the key principles is to know the relative pose between spatial information from various sensors of different vehicles, which have been dealt with as a map merging problem [6] or relative localization [7]. The solving approaches include triangulation [8] and dead reckoning [9]. However, since rich sensing information has been recently available, sensing-information-based methods have been preferred due to their high accuracy. The principle is to find the best pose that maximizes the similarity of overlapping area between different maps using landmarks [10], topological maps [11], occupancy grid maps [12], or scan matching [13].

With the recent advancement of range sensor technology, scan matching has been one of the most essential technologies to enable map building, map merging, pose estimation [14], and visual odometry [15]. The principle of scan matching is to find a transformation between scan data from range sensors. The proposed solutions include iterative closest point or iterative correspondence point (ICP)-variant [16]–[18], adaptive random-walk [12], Hough transform [19], correlative scan matching (CSM) [20], and histogram approaches [21], [22].

Although these map merging and scan matching algorithms have been researched over the last two decades with hundreds of variants [18], [23], they still have several difficulties for on-line multivehicle map merging on the road. First, there is no guarantee that sensor configuration of each vehicle is identical. Second, each map has different scan points even in the overlapping area. For these reasons, scan points and features may not be obvious to match. Furthermore, bushes, trees, or moving pedestrians make the situation worse. Finally, the overlapping area may not be sufficient due to longer safety gaps for collision avoidance. In this paper, we devise, apply, and compare the scan matching method to map merging for driving on the road scenario using both a closed-form solution (ICP) and a probabilistic one (CSM).

Meanwhile, one of the primary goals of this paper is to establish a spatial map sufficient for driving control and assistance. For this purpose, there have been several research efforts using GPS. Tan and Huang proposed a Differential GPS (DGPS)-based cooperative collision warning system partially motivated by the inaccuracy of off-the-shelf GPS [24]. Wender and Dietmayer proposed an algorithm to deliver onboard sensing information via wireless communications, whose main motivation was that a position from DGPS is still less accurate than an onboard sensor, specifically a laser scanner [25].

Chong *et al.* experimentally demonstrated that onboard sensor-based localization can provide centimeter-level position accuracy sufficient for fully autonomous vehicle control without GPS or DGPS in [26]. In this context, we focus on sensing information, rather than GPS or DGPS, to obtain the accurate relative position of vehicles. One of the strongest aspects of this approach is no common coordinate system is assumed among vehicles, which makes the proposed perception system robust from signal measurement variation according to the change of weather, or instability of the central system.

B. Multimodal Perception and See-Through Systems

The sensors used for driving assistance or automation purposes can be largely classified into active sensors, (e.g., radar, laser, and ultrasonic sensors), or passive sensors, (e.g., monocular and stereo-vision sensors). Since both are complementary to each other, a multimodal approach fusing information from active and passive sensors has been actively researched [27], [28] for various purposes, such as vehicle detection and tracking [29], pedestrian detection and tracking [30], intersection safety [31], and see-through systems [32].

Recently, see-through systems have begun to be considered as one of the emerging driver assistance technologies to mitigate traffic accident caused by perception limitations. Olaverri-Monreal *et al.* [33] introduced a see-through system based on vehicular ad hoc networks for assisting overtaking maneuvers. The same group proposed a method that displays a see-through view to a human driver using augmented reality in [34]. Note that the see-through systems can be achieved through video streaming between vehicles [35], [36]. In contrast to simulation verification of these works, Li and Nashashibi provided experimental results using two vehicles on the road with the consideration of sensor multimodality in [32]. In the context of delivering additional sensor information for driver assistance, the see-through systems share a similar goal with this paper. In this paper, however, we will extend one step further toward cooperative driving.

C. Cooperative Driving Using Cooperative Perception

One of the most promising applications of cooperative perception on the road is cooperative driving for traffic safety and efficiency [24], [37], [38], because the results of map merging extend perception range beyond line-of-sight and sensing angles. Accordingly, they enable traffic flow prediction, early obstacle detection, and long-term perspective path planning. The work of Tsugawa *et al.* [39] included the architecture

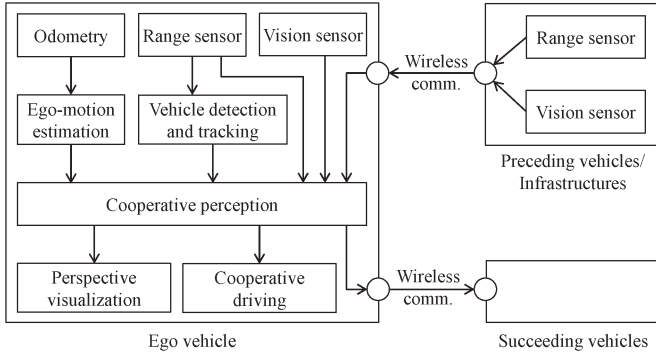


Fig. 2. Architecture of the proposed system.

for cooperative driving using cooperative perception. The agenda and experimental results of cooperative autonomous driving have been presented in [4], [40], and [41]. In particular, Kim *et al.* provided the motivation and experimental results of cooperative autonomous driving using cooperative perception in [3]. Liu *et al.* proposed the concept of motion planning using cooperative perception on the road [42]. Rebsamen *et al.* proposed infrastructure-support sensing for driving [43].

In addition, vehicle communication and vehicle identification problems are addressed in [44]–[48], respectively. In this paper, we focus on cooperative driving using cooperative perception in an engineering feasibility standpoint.

III. SYSTEM ARCHITECTURE

Fig. 2 shows the overall system architecture of the proposed system. The system consists of three subsystems: cooperative perception, perspective visualization, and cooperative driving.

For cooperative perception, each vehicle is equipped with range sensors, such as laser or radar scanners, vision sensors, and radios, as shown in the leftmost box of Fig. 2. An odometry system is utilized for ego-motion estimation such as moving direction and speed. The range sensors are used for vehicle detection and tracking. The vision sensors are used for classification and identification of vehicles and pedestrians and provide driver-friendly visual traffic information.

The vehicles exchange this local sensing information with other vehicles or infrastructure via wireless communications. In this paper, we define the exchanging of information as a *message*. Since there exists a tradeoff between communication performance and information quantity, the message profile, e.g., message size and transmission period, should be carefully chosen according to application requirements and driver preferences. In this paper, we investigate several message profiles, such as laser scan data only, raw image only, compressed image only, point clouds only, and both laser scan and point clouds, which are investigated in detail with the real measurement data on the road in Section VIII-B.

All local and remote sensing information is properly fused at the box of cooperative perception in Fig. 2. Compared with data fusion of on-board sensing information, data fusion of remote sensing information on the road includes a number of practical challenges [3], among which this paper focuses on the map merging problem and sensor multimodality.

Fig. 3. Concept of leader string where i is an ego vehicle. $i+1$ and $i+2$ are the first and second leader of i , respectively. $i-1$ is a following vehicle behind ego vehicle i .

After the information fusion procedure at cooperative perception, the fused information can be used for driving assistance for a human driver and motion planner for an autonomous driver. In addition, it can be delivered to the vehicles following behind or closest infrastructure. For the purpose of driving assistance, our proposed system represents the fused information in a perspective visualization manner, to provide intuitive guidance information, which is dealt with in Section V.

Finally, one of the primary goals of a cooperative driving subsystem is to let a driver know the moment when the driver should be careful, such as hidden obstacle detection or sudden braking of preceding vehicles beyond line-of-sight. The notification is performed by visual and sound alarms, which enable a driver to focus on driving until any dangerous situation is detected by our system. Moreover, the subsystem notifies a driver when there are any vehicles coming from behind or at blind spots, which can contribute to safe lane changing or overtaking. For self-driving vehicles, the notification triggers path re-planning for automated lane changing in Section VIII-D.

IV. COOPERATIVE PERCEPTION ON THE ROAD

On the road, the preceding vehicles highly affect the driving decision of the ego driver. In this sense, a leader is a preceding vehicle: 1) connected via wireless communications; and 2) observable by the ego vehicle through its local sensors or remote sensing information. Let $\mathcal{V}_i = \{i+1, \dots\}$ be a *leader string* of vehicle i , where, $j, j+1 \in \mathcal{V}_i$, j and $j+1$ are connected via wireless communications and $j+1$ is observable by local sensors of j . The concept of leader string is represented in Fig. 3.

In a broad sense, following vehicles of ego vehicle i can be included in the leader string \mathcal{V}_i , even though a vehicle $i-1$ is not a leader literally. The information of a vehicle $i-1$ can be useful for lane-changing assistance or blind spot detection, because the following vehicle $i-1$ can watch the ego vehicle in the third-person view from behind the ego vehicle.

In vehicle driving scenarios, sensing information is typically dealt with as a map. Let $\mathcal{M} = \{\dots, m, \dots\}$ be a map for navigation of vehicles, which consists of the set of points obtained and filtered from sensors, where $m \in \mathbb{R}^n$, $n = 2$, or 3. Given a position $p \in \mathcal{M}$ in a map, $\mathcal{M}[p] \rightarrow \mathbb{R}$ can be defined in several ways, such as the belief that the position p is obstacle-free, or the height of obstacle in case of $p \in \mathbb{R}^2$. \mathcal{M}_i denotes a map of a vehicle i .

Now, we can formulate cooperative perception as follows:

$$\mathcal{M}_i = \mathcal{M}_i \bigcup_{j \in \mathcal{V}_i} \mathcal{M}_j \quad (1)$$

where the operation \bigcup is called *map merging*. The map merging operation is merely a set union operation, if \mathcal{M}_i and $\mathcal{M}_{j \in \mathcal{V}_i}$

are mapped into a global coordinate frame, such as GPS coordinates. However, the observation from sensors is typically mapped into a local coordinate frame. In addition, there is no guarantee that the initial poses of vehicles are known. In this case, the relative pose between vehicles is necessary to merge different spatial information. The relative pose can be represented as $q = (\tau, \theta)$, where τ and θ correspond to translation and rotation, respectively. We define a transformation operator as $p \otimes q = R(\theta)p + \tau$. Finally, (1) can be rewritten in a more general form as follows:

$$\mathcal{M}_i = \mathcal{M}_i \bigcup_{j \in \mathcal{V}_i} \mathcal{M}_j \otimes q_{i,j} \quad (2)$$

where $q_{i,j}$ is a relative pose between a vehicle i and j . One of the key challenges to solve (2) is to obtain an accurate $q_{i,j}$.

A. Problem Formulation

1) *Peer-Vehicle Map Merging*: The primary problem of map merging can be formulated as follows:

$$q_{i,j}^* = \arg \max_{q_{i,j}} S(\mathcal{M}_i, \mathcal{M}_j, q_{i,j}) \quad (3)$$

where the similarity measure S is formulated as

$$\sum_p \mathcal{L}(\mathcal{M}_i[p], (\mathcal{M}_j \otimes q_{i,j})[p]) \quad (4)$$

where $\mathcal{L}(a_i, a_j)$ is a point-to-point similarity measure, which is positive if $a_i = a_j$ and 0 otherwise. Equations (3) and (4) attempt to find the relative pose that maximizes the overlapping area between two maps. Note that (3) and (4) are well-suited for a range-scanner-based approach. We present a method to merge vision-based data in Section IV-D.

2) *Multivehicle Map Merging*: Equations (3) and (4) can be extended to more than two vehicles. Let $Q_i = \{q_{i,i+1}, \dots, q_{i,i+N}\}$ be the set of relative poses with respect to (w.r.t) the ego vehicle, where N is the number of leader vehicles. The problem can be rewritten as follows:

$$Q_i^* = \arg \max_{Q_i} S(\mathcal{M}_i, \dots, \mathcal{M}_{i+N}, Q_i) \quad (5)$$

where the similarity measure S is formulated as

$$\sum_p \mathcal{L}(\mathcal{M}_i[p], \dots, (\mathcal{M}_j \otimes q_{i,j})[p], \dots) \quad (6)$$

where $\mathcal{L}(a_i, \dots, a_{i+N})$ is positive if $a_i = \dots = a_{i+N}$ and 0 otherwise.

B. Scan Matching Algorithm

To solve (3)–(6), we use a scan matching function $q_{i,j}^* = \text{ScanMatching}(\mathcal{M}_i, \mathcal{M}_j)$. Several approaches can be used for realizing the function, as mentioned in Section II. In this paper, we consider two kinds of approaches: One is ICP, a closed-form solution, and the other is CSM, a probabilistic solution.

1) *ICP Algorithm*: The ICP solution can be formally stated as follows:

$$q_{i,j}^* = \arg \min_{q_{i,j}} \sum_{m_j \in \mathcal{M}_j} \|C_{\mathcal{M}_i}(m_j) - m_j \otimes q_{i,j}\| \quad (7)$$

where $\|\cdot\|$ is an l_2 -norm, and $C_{\mathcal{M}_i}(m_j)$ is a point in \mathcal{M}_i corresponding to m_j . In the standard ICP, the closest point in \mathcal{M}_i is selected as the correspondence, which is defined as

$$C_{\mathcal{M}_i}(m_j) = \arg \min_{m_i \in \mathcal{M}_i} \|m_i - m_j\|. \quad (8)$$

ICP iterates (7) using $\mathcal{M}_j \leftarrow \mathcal{M}_j \otimes q_{i,j}^*$ until the solution has converged. $q_{i,j}^*$ can be obtained by using a closed-form solution that Arun *et al.* proposed in [49].

Note that there is no guarantee that the closest point (8) is the right corresponding point. In many cases, two maps are partially overlapped, or scan points may not be exactly matched due to the physical limitation of a sensor, e.g., different sensing range or resolution. From an implementation perspective, the closest point should be limited by some threshold distance. Now, (7) can be rewritten as follows:

$$q_{i,j}^* = \arg \min_{q_{i,j}} \sum_{m_j \in \mathcal{M}_j} d_{\mathcal{M}_i}(m_j) \|C_{\mathcal{M}_i}(m_j) - m_j \otimes q_{i,j}\| \quad (9)$$

where

$$d_{\mathcal{M}_i}(m_j) = \begin{cases} 1 & \text{if } \min_{m_i \in \mathcal{M}_i} \|m_i - m_j\| < d_{\text{th}} \\ 0 & \text{otherwise,} \end{cases} \quad (10)$$

where d_{th} is the threshold distance. d_{th} is one of the key design parameters to decide the performance, which has a tradeoff between accuracy and convergence speed. Using the concept of closest points, the similarity metric S can be reformulated as follows:

$$S(\mathcal{M}_i, \mathcal{M}_j, q_{i,j}) = \sum_p d_{\mathcal{M}_i}(m_j) \|\mathcal{M}_i[p] - C_{\mathcal{M}_j \otimes q_{i,j}}(p)\|. \quad (11)$$

2) *Probabilistic Scan Matching*: Probabilistic scan matching methods approach the problem differently from the closed-form solutions such as ICP, in particular, where several uncertainties are considered, such as sensor measurement errors. The key principle is to find a pose at which the current observation is maximally likely or most probable based on the previous observation, such as a map, which can be formulated as $x^* = \arg \max_x \mathbb{P}(z|x, \mathcal{M})$, where $\mathbb{P}(z|x, \mathcal{M})$ is an observation model, x is the pose of an observer, z is the sensor reading, and \mathcal{M} is a map [50]. The formulation can be used for cooperative perception as follows:

$$q_{i,j}^* = \arg \max_{q_{i,j}} \mathbb{P}(\mathcal{M}_j | q_{i,j}, \mathcal{M}_i) \quad (12)$$

where \mathcal{M}_j corresponds to observation from a sensor located away from the observer. To efficiently compute the probability of (12), a lookup table is usually built in advance [51].

There are many solutions to obtain the maximum-likelihood pose $q_{i,j}^*$ of (12). In this paper, the CSM method is considered,

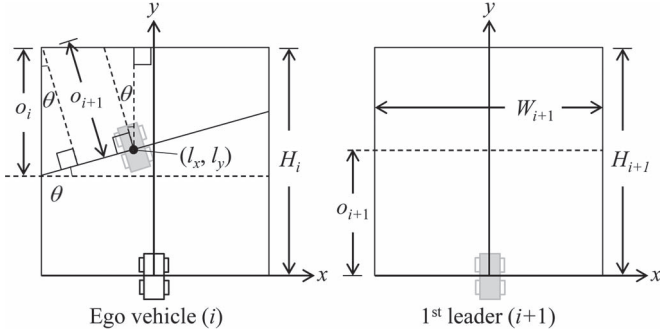


Fig. 4. Coordinates and overlapping areas of the ego vehicle and its leader. H_i and W_i are the height and width of the local sensing area of a vehicle i . $o_{i,i+1}$ is the height of overlapping sensing area between a vehicle i and $i+1$ from the perspective of a vehicle i . Likewise, $o_{i+1,i}$ is the height of overlapping area between a vehicle i and $i+1$ from the perspective of a vehicle $i+1$.

because of several advantages for our purposes. One key aspect of CSM is the ability to find the global maximum instead of trusting a local search algorithm. By evaluating the entire search space, the scan matching algorithm is able to provide a robust solution even with a large initialization error.

Because of this, CSM approaches the problem using multilevel resolution. Specifically, two lookup tables are used for avoiding local maxima and achieving lower computational complexity: a low- and a high-resolution table. First, the likelihood is quickly evaluated over the entire 3-D search window by using the low-resolution table. When the maximum-likelihood voxel is found in the process, it starts to find more accurate values by evaluating inside the voxel using the high-resolution table. One can find a specific description in detail in [20].

C. On-Road Map Merging Method

In this paper, we consider a 2-D map for map merging, i.e., $\mathcal{M}[m] \rightarrow \mathbb{R}$, $m \in \mathbb{R}^2$. Fig. 4 shows the spatial coordinates of the ego vehicle, and its predecessor vehicle, where a *leader* is used for the predecessor vehicle. We assume that range sensors are installed in a y -axis direction. (l_x, l_y) represents the position of the leader. Algorithm 1 shows the overall procedure of our map merging method. Two maps from different vehicles are taken as input parameters, and a merged map is generated from the algorithm.

Algorithm 1: On-Road Map Merging Method

```

input :  $\mathcal{M}_i, \mathcal{M}_j$ 
output: Merged map

begin
   $q_0 \leftarrow \text{LeaderVehicleDetection}(\mathcal{M}_i)$ 
   $\mathcal{M}_i \leftarrow \{m | m > l_y - (l_x + W_i/2) \tan \theta, m \in \mathcal{M}_i\}$ 
   $\mathcal{M}_j \leftarrow \{m | m < (l_y - (l_x + W_i/2) \tan \theta) \cos \theta, m \in \mathcal{M}_j \otimes q_0\}$ 
   $\mathcal{M}_j \leftarrow \mathcal{M}_j \cup \text{VirtualLeaderScan}()$ 
   $q_{i,j}^* = \text{ScanMatching}(\mathcal{M}_i, \mathcal{M}_j, q_0)$ 
  return  $\mathcal{M}_i \cup (\mathcal{M}_j \otimes q_{i,j}^*)$ 
end

```

1) *Leader Vehicle Detection*: One of the distinct characteristics of scan data on the road is that features are not obvious to match due to trees, bushes, or pedestrians. However, scan

data of a leader has relatively clear features such as the side or backside of vehicles. The initial pose of a leader can be guessed from the features. More specifically, the proposed system keeps checking whether there are consecutive points lying in the desired path ahead.

For this task, let \mathcal{D}_i be the set of positions that a vehicle i intends to move to, typically the set of the center positions of the current lane, which can be generated by combination of vehicle localization, lane detection and path planning methods. We use curb-intersection features based Monte Carlo localization [52] for lane-keeping, and a path planner for lane-changing as described in [42].

In Fig. 24(a), the yellow horizontal line passing through the vehicles depicts a simple \mathcal{D}_i to keep within a lane. In Fig. 24(c), the red curve depicts \mathcal{D}_i as a result of path replanning for lane changing. To find the leader position, the proposed system keeps watching straight or slightly curved lines in the following point set:

$$\{p | r_{th} > \|p - p_l\|, p_l \in \mathcal{D}_i\} \quad (13)$$

where r_{th} is the maximum boundary that a vehicle can deviate from the desired path, typically half of the lane width. This method considers the current lane position to extrapolate the desired path of the leader vehicle, which can significantly reduce the search space and occurrence of false positives in leader detection compared with relying on scan matching only.

If 1) the detected points compose a straight or slightly curved line, and 2) the line is almost orthogonal to the desired path, the center position of the line and the orthogonal angle to the line are used as t_0 and θ_0 , respectively, which is used as the initial transformation $q_0 = (t_0, \theta_0)$.

Note that reliable leader vehicle detection is important to avoid circular reasoning [7]. Our method is to merge information that comes from the leader to information gathered by the ego vehicle. To identify the leader, vehicle identification is necessary, which is dealt with in [3] and [48].

In addition, with the support of vehicle identification and rear sensors, the following vehicle can be identified by the ego vehicle. Accordingly, the map merging process can be started from the following vehicle with information delivered from the following vehicle. This map merging from the following vehicle will be evaluated in Section VIII-B.2.

2) *Overlapping Area Extraction*: The size of overlapping area between two maps is relatively small during driving. Computational complexity can be reduced by restricting matching operation to the overlapping area instead of whole maps. Fig. 4 shows the overlapping area between two vehicles. In case of the ego vehicle, scan points in overlapping area are collected into

$$\mu_i = \{m | m > l_y - (l_x + W_i/2) \tan \theta_0, m \in \mathcal{M}_i\} \quad (14)$$

where the leader position (l_x, l_y) is obtained from t_0 . Likewise, scan points of the leader in overlapping area, μ_j , are

$$\{m | m < (l_y - (l_x + W_i/2) \tan \theta_0) \cos \theta_0, m \in \mathcal{M}_j \otimes q_0\}. \quad (15)$$

3) *Virtual Leader Scan Points Recovery*: We assume that a vehicle cannot scan itself. The map of the ego vehicle has

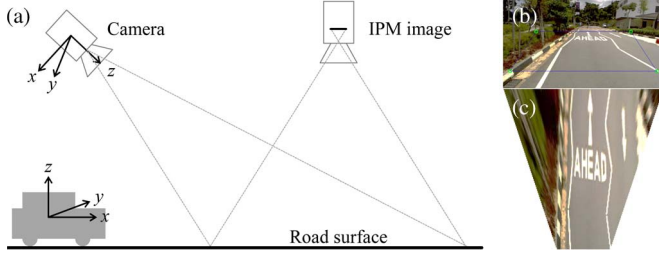


Fig. 5. (a) Concept of IPM; (b) an example of projection on-road image captured by a forward-looking camera; and (c) its corresponding IPM result.

scan points detecting its leader vehicle, but the map delivered from the leader includes no scan point of the leader itself. Since a sensor is typically installed at a fixed position, the vehicle knows how it would be shown on its own map. Therefore, the vehicle can recover virtual scan points reflecting its shape as if the vehicle detects itself. The virtual laser scan points are one of the most obvious features to match, so that they can play a important key role to provide good matching points, especially when a few scan points are collected or scan points are not obvious to match.

4) *Scan Matching*: Now, we are ready to perform scan matching operation with M_i , M_j , and an initial transformation q_0 . In this paper, we use ICP and CSM for our implementation. As a result, the optimal relative pose $q_{i,j}^*$ is obtained. In Section VIII-B, we will evaluate the performance in terms of computation time and accuracy with the comparison of two scan matching approaches.

5) *Final Merged Map*: The final map is the union of M_i and transformed M_j with the relative pose $q_{i,j}^*$, which can be represented as

$$M_i \cup (M_j \otimes q_{i,j}^*). \quad (16)$$

D. Multimodal Cooperative Perception

Our proposed system supports multiple sensory modalities for cooperative perception. In case of range sensors, such a laser scanner, it is relatively easy to perform map merging operations, because their readings are a set of physical quantities recorded with their spatial coordinates. However, in the case of vision sensors whose reading is an image, the map merging is not a straightforward task, because the vision image is the result of perspective projection. Therefore, the images should be mapped into the spatial coordinate of the ego vehicle.

We use the inverse perspective mapping (IPM) [53] method to deal with this problem. IPM is an image transformation method to get a satellite view of the road surface. When a forward-looking camera is capturing an image, the shape of the road surface is usually distorted due to perspective projection. An IPM process can be applied to remove this distortion and recover original road shapes for the map merging purpose.

Fig. 5(a) illustrates the basic idea of IPM. For a fixed-mount on-board camera, its pose in the vehicle coordinate frame is usually known, and hence the pose relative to the road surface can be also obtained. Together with its intrinsic parameters from a calibration process, the perspective transformation matrix from the road surface to the camera image can be

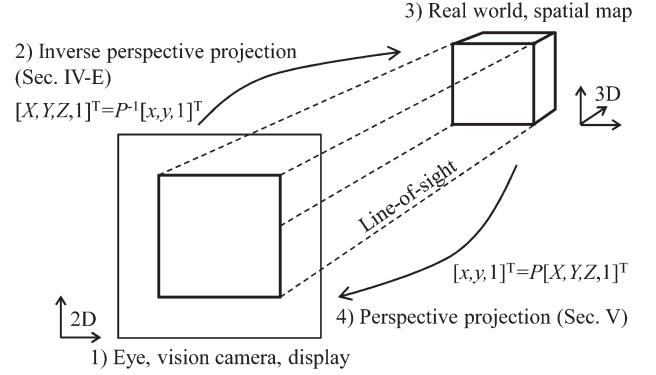


Fig. 6. Concept of perspective projection and inverse perspective projection.

calculated. An inverse operation of this perspective matrix will restore the original road surface, and store it in an IPM image. Fig. 5(b) and (c) shows an example of projection on-road image captured by a forward-looking camera, and its corresponding IPM result, respectively. One can find more details on IPM in ITS-related applications in [54].

From a mathematical point of view, IPM is an inverse operation to camera projection transformation. In the projection process, a 3-D point $[X, Y, Z]$ is projected onto an image with its projection pixel as $[x, y]$, where $[x, y, 1]^T = P[X, Y, Z, 1]^T$, and P is the 3×4 camera projection matrix. In the IPM process, we assume that the height of points on the road surface is 0, where their 3-D coordinates $[X, Y, 0]$ can be easily recovered with the P matrix, and their image pixels $[x, y]$. Fig. 6 describes these operations. Due to the assumption of IPM where points lie on the road surface, it is only suitable to recover the object shapes with $Z = 0$ height. In our experiment, we use it to facilitate our map merging for the road surface. Other objects such as pedestrians and vehicles, which violate the IPM assumption, will be taken care with other sensory modalities. One can find a general framework for road marking detection and analysis in [55].

Although visual odometry has been proposed and used [15], we primarily use laser scan points to obtain the relative pose $q_{i,j}^*$ in the implementation of this paper, because our system is targeting on-road scenarios, and it is much faster to deal with scan data rather than images with the consideration of this purpose. Using (16) and the obtained relative pose $q_{i,j}^*$, the recovered road surfaces can be merged into one.

Finally, three types of information can be represented in the merged map: 1) vehicle poses, including poses of the ego vehicle and other vehicles; 2) laser scan points; and 3) color points of the road surface from vision. Fig. 14(a) shows an example of merging multimodal sensing data from a vision sensor and laser scanner. By supporting multiple sensory modalities, the cooperative perception system helps the ego vehicle to perceive not only occluded vehicles and obstacles but also road surfaces that may be out of its sensing ability.

V. PERSPECTIVE VISUALIZATION

The cooperative perception process generates a merged map with various types of information. To provide the information as intuitive driving guidance, our system represents the information in a perspective projection.

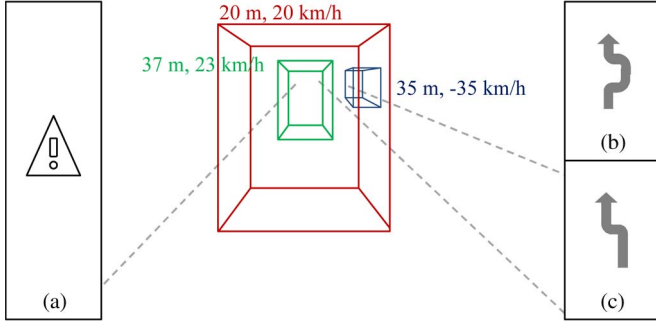


Fig. 7. Example of a perspective see-through visualization on the road. The two numbers next to the cuboid are the distance from the ego vehicle and its speed, respectively. The gray dotted lines represent road lanes. The cuboid represents a simplified vehicle model. (a), (b), and (c) are examples of visual warning signs for driving assistance, where (a) is a forward collision warning, (b) is a sign for overtaking assistance and (c) is a sign for lane-changing assistance. These visual warnings can be activated along with audio or tactile warning.

A. First-Person See-Through Visualization

Through the map merging steps presented in Section IV, a 3-D spatial map is obtained from different sensor information, where x, y along with z as an occupancy probability or height of (x, y) . Contrary to the way of the IPM method, the perspective projection method can map 3-D points to a 2-D plane, as shown in 4 of Fig. 6. Based on the perspective projection method, we can provide a first-person see-through visualization consisting of the perspective visualization of the road surface and vehicle skeletons. The concept of the first-person see-through system using cooperative perception has been proposed in [32]. Here, we present our implementation in detail.

Fig. 7 shows an example of the first-person see-through visualization on the road. The leftmost and rightmost boxes are used for providing visual warning to a driver, which is explained in detail in Section VI.

1) *Road Surface Projection*: As a first step, the road surface is restored from vision images of the ego vehicle or leader vehicles by the IPM method and represented as a cloud of color points, as addressed in Section IV-D. To visualize the road surface information for a human driver, we project the surface points onto the camera image and label the projection pixels to the color of their corresponding points.

2) *Vehicle Skeleton Visualization*: Vehicle detection and tracking is realized with laser scan data readings, with vehicle poses stored in the merged map. To visualize the detected vehicles more clearly, we use a cuboid as the simplified vehicle model, and then plot the skeleton of this cuboid. In a bad situation, e.g., at night, under shadow, in heavy rain, such as an explicit cuboid definitely helps to identify vehicles. The cuboid has 8 vertices and 12 edges. Its vertices are projected onto the camera image, and its edges are drawn by connecting the projection pixels of neighboring vertices.

B. Adjustable Third-Person Visualization

While the first-person see-through visualization provides an intuitive representation of the merged data, various kinds of information is squeezed and overlapped at a small region of its

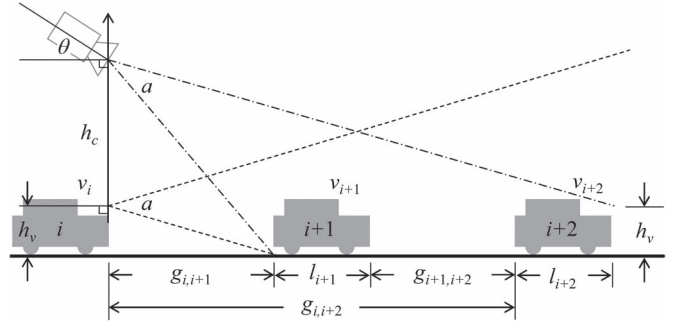


Fig. 8. Schematic of a third-person visualization. A virtual camera is mounted above the roof of ego-vehicle. θ is its tilted-down angle, and D is its mounting height. The tilted-down angle and mounting height of this virtual camera can be adjusted to find an optimal view to visualize the data from cooperative perception.

image, due to the low mounting position of the vehicle camera, which makes them difficult to recognize ahead traffic situations, as we can see in Fig. 7. For this reason, we propose to use an adjustable third-person view to visualize the data. Fig. 8 illustrates the basic idea of third-person visualization. In Fig. 8, h_v , $g_{i,j}$, and h_c are the camera installation height, the distance from a vehicle i to a preceding vehicle j , and the uplifted vision height, respectively. While the mounting height $h_v + h_c$ and tilted-down angle θ of this virtual camera are both adjustable, the best view point can always be selected to visualize the merged data in different scenarios. Note that the camera is typically mounted with some tilted-down angle, e.g., 12° in our experiments.

For determining a proper θ , we use the following boundary for deciding a tilt-down angle w.r.t. to camera height h_c and distance $g_{i,j}$

$$\theta \geq \arctan\left(\frac{h_v + h_c}{g_{i,i+1}}\right) - \frac{a}{2} \quad (17)$$

which can be derived from Fig. 8, where a is a horizontal angle of view. Likewise, if we want to restrict the maximum number of tracking leaders as s , the upper boundary of θ is

$$\theta < \arctan\left(\frac{h_c}{g_{i,i+1} + \sum_{j=i+1}^s (g_{j,j+1} + l_j)}\right) + \frac{a}{2} \quad (18)$$

where $g_{i,i+1}$ is a distance between a vehicle i and $i+1$, and l_i is the width of a vehicle i . For example, the minimum tilted-down angle θ is $0.8 \approx 45.8^\circ$ to obtain a bird's eye view at $20 \text{ m} (= h_c + h_v)$ if a vision camera is installed at $h_v = 1 \text{ m}$, a target vehicle is preceding at $L = 10 \text{ m}$ ahead, and the vertical field-of-view of the camera is $a = 35^\circ$. One may want to watch the display only when necessary, instead of watching it continuously. For this purpose, the next section addresses a see-through forward collision warning.

VI. COOPERATIVE DRIVING ASSISTANCE

Here, we present how cooperative perception results presented in Section IV and V can be applied to specific driving assistances, such as forward collision warning, overtaking/lane-changing assistance with an all-around view.

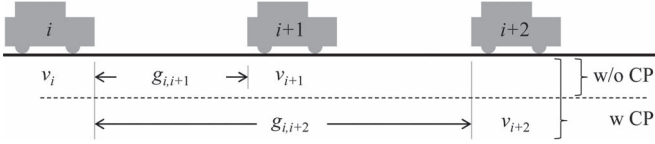


Fig. 9. Collision warning with and without cooperative perception.

A. See-Through Forward Collision Warning

A driver can focus on driving itself without having to keep watching a driving assistance display, if the system can let a driver know the moment when the driver has to drive carefully through proper visual, audio, or tactile feedback.

In this context, one of the most desired functions is forward collision warning. Numerous forward collision warning algorithms have been proposed along with risk assessment methods, which include time-based and distance-based approaches. First, the time-based methods usually use time-to-collision (TTC), which can be formulated as

$$TTC_{i,j} = \frac{g_{i,j}}{v_i - v_j}, \quad j \in \mathcal{V}_i \quad (19)$$

where $TTC_{i,j}$ is corresponding to TTC, and v_i is the speed of a vehicle i . A forward collision warning is activated if $TTC_{i,j}$ is less than a certain threshold time, typically 2–3 s according to safety requirements [56].

In distance-based methods, a forward collision warning is activated if, given two vehicles i and j , the recommended safety gap $r_{i,j}$ is larger than the distance between two vehicles $g_{i,j}$. Formally, a forward collision warning is activated if

$$g_{i,j} < r_{i,j}, \quad j \in \mathcal{V}_i \quad (20)$$

where $r_{i,j}$ is a minimum distance to avoid collision to a preceding vehicle j , which can be formulated as follows [57]:

$$r_{i,j} = (v_i - v_j)/2\gamma^2 + (v_i - v_j) * T_{rs}, \quad j \in \mathcal{V}_i \quad (21)$$

where γ is the deceleration of the ego vehicle, e.g., -0.2 g ($\approx -2 \text{ m/s}^2$). T_{rs} is the response time of a driver, e.g., $0.5\text{--}1.5 \text{ s}$.

Note that one of the great advantages of our system is that j is not limited to only $i + 1$. In our system, due to the capability of a see-through view, j can be beyond the first leader according to the connectivity via wireless communications, as shown in Fig. 9. This see-through characteristic enables the driver to avoid hidden obstacles earlier than without cooperative perception systems. We will investigate this early hidden obstacle detection in Section VIII-C.

B. All-Around View Using Cooperative Perception

Another benefit of our system is that it enables an all-around view without having to install sensors covering all sides of a vehicle. In particular, a following vehicle can see the ego vehicle in a third-person view, including its blind spots. Fig. 16(e) shows one snapshot of the all-around view of the ego vehicle. This all-around view aspect of our system is investigated using real experiments in Section VIII-B.2.

Furthermore, our system can detect a vehicle approaching toward the ego vehicle using a spatio-temporal moving obstacle detection and tracking method. From the following, we investigate the benefit of an all-around view in terms of overtaking assistance and lane-changing assistance.

C. Overtaking/Lane-Changing Assistance

When a driver should drive slowly or stop due to a slow-moving truck or obstacle ahead, the driver should decide whether to wait longer in the current lane or change lanes. In particular, this is an important issue for overtaking on a single-lane road.

Various overtaking assistance methods are possible according to the sensor configurations, requirements, and sensing capability. The overtaking decision should be determined with the consideration of: 1) the number of lanes; 2) the speed of a preceding vehicle; 3) cut-in space availability; 4) distance from an on-coming vehicle; and 5) the existence of another overtaking vehicle from behind as in Fig. 16(h). Our system can inform 3), 4), and 5) that are difficult to be provided without cooperative perception due to the limitation of line-of-sight.

Equations (19) and (20) can be applied to read-end collision warning at lane changing, which are corresponding to $TTC_{j,i}$ and $r_{j,i}$, respectively, where the vehicle j is approaching toward the ego vehicle behind in an adjacent lane, and $j < i$.

D. Feedback to a Driver

Once a forward collision is expected, or overtaking is not possible, or lane changing is not possible, it should be notified to a driver through visual, audio, or tactile feedback. Fig. 7(a)–(c) are examples of visual warnings for a forward collision warning, overtaking possibility, and lane-changing possibility, respectively. In Section VIII, we evaluate how our system can contribute to the see-through forward collision warning, overtaking assistance, and lane-changing assistance are, through real experiments on the road using vehicles equipped with the proposed system.

VII. SCALABILITY ANALYSIS

Information delivered via wireless communication inherently has uncertainty. In this paper, we quantify communication uncertainty by using communication delay, because the communication delay highly affects the overall system performance of moving vehicles.

Let $d_{i,j}^c$ be communication delay for message delivery from a vehicle i to j . For simplicity, we define $d_i^c \triangleq d_{i+1,i}^c$. Let d_i^p be the processing delay of a vehicle i . In general, d_i^c and d_i^p are not fully controllable, so they should be measured and estimated. The leader $i + 1$ is moving, whereas the information is being processed and delivered from $i + 1$ to i . From the perspective of a vehicle i , the estimated position \tilde{p}_{i+1} is

$$\tilde{p}_{i+1} = p_{i+1} + v_{i+1} (d_i^c + d_i^p) \quad (22)$$

where p_{i+1} and $v_{i+1} = \dot{p}_{i+1}$ can be obtained by vehicle i , as addressed in Section IV. By using the definition of

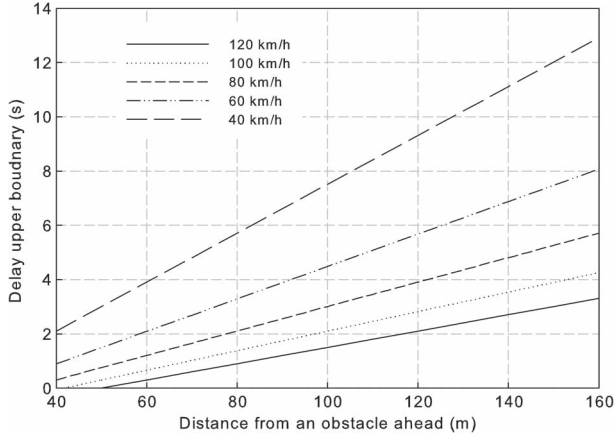


Fig. 10. Delay upper boundary including processing time w.r.t. distance from an obstacle ahead, where a driver's reaction time is set to 1.5 s.

communication and processing delay and (22), we derive the worst case deadline for information delivery as follows.

Assuming that no vehicle moves backward on a given path, information delivered from a vehicle $k \in \mathcal{V}_i$ has an impact to path planning and vehicle control of an ego vehicle i whose maximum speed is \bar{v}_i , if the following inequality is satisfied:

$$d_s \triangleq \sum_{j=i}^{k-1} d_j^c + d_j^p < \frac{\|p_k - p_i\|}{\min(\bar{v}_i, \|v_i + d_s \bar{u}_i\|)} - T_{rs} \quad (23)$$

where d_s is a total delay for delivering k 's information to i , and \bar{u}_i is the maximum acceleration of a vehicle i . A sketch of proof is as follows.

Suppose that k is an obstacle staying at a position p_k as a worst case, because we assume that no vehicle move backward. For intuitive understanding, suppose that $i + 2$ is a stopped vehicle or a static obstacle in Fig. 3. In this case, k is $i + 2$. Note that the position p_k is recognized by a vehicle $k - 1$. In (23), $\min(\bar{v}_i, \|v_i + d_s \bar{u}_i\|)$ represents the maximum speed that the vehicle i can perform during information delivery of the vehicle k . Therefore, the second rightmost term of (23) represents the minimum time that the vehicle i can reach the position p_k . The rightmost term of (23), i.e., T_{rs} , is a reaction time to steer or slow down for collision avoidance. As a result, the rightmost two terms represent the minimum time that the driver can react to a dangerous situation ahead. The ahead traffic information should be delivered within the time.

Depending on the applications, an average delay boundary can be more useful. Suppose that a group of vehicles achieves velocity consensus at v_i on highway. In this case, the delay boundary can be relaxed as follows:

$$d_s < \frac{\|p_k - p_i\|}{\|v_i\|} - T_{rs}. \quad (24)$$

As long as k satisfies (24), k can be increased. In this way, the maximum size of leader string can be obtained by using (23) or (24). Note that k is highly affected by message profile, radio characteristics, and vehicle speed. Fig. 10 shows the upper boundary of total communication and processing delay under the assumption that the speed of the ego vehicle is constant

and the driver's reaction time is 1.5 s. For example, total delay should be less than 2 s to detect and avoid the obstacle at 100 m ahead, if the vehicle runs at 100 km/h, and driver's reaction time is 1.5 s. The delay boundary, in this case 2 s, is a key parameter to determine message profile and radio interface. We evaluate this using measurement data from experiments on the road in Section VIII-B.3.

VIII. EXPERIMENTAL RESULTS

Here, we provide experimental results to evaluate our proposed system with four automotive vehicles.

A. Experimental Setup

In the experiments, we used several mass-produced automotive vehicles including one Mitsubishi iMiEV and three Yamaha golf carts. The vehicles are equipped with 2-D LIDARs (Light Detection And Ranging), a vision camera, and wireless interface IEEE 802.11g, IEEE 802.11n, 3G HSDPA (High-Speed Downlink Packet Access), and 4G LTE (Long-Term Evolution). The devices of 3G HSDPA and 4G LTE are Huawei E153 and Huawei E3276, respectively, which are both USB-type. The vision cameras were mounted with a tilted-down angle of 12° from horizontal. The software architecture of this system was established on robot operating system (ROS) suite [58] using only open source libraries. Detail specifications are available in [26].

The setup of three vehicles is represented in Fig. 9. i , $i + 1$, and $i + 2$ are corresponding to the ego vehicle, the first leader and the second leader, respectively. In Fig. 9, the second leader transmits its sensing information to the first leader via wireless communications, while moving forward. With the support of our system, the first leader merges the remote information with its local sensing information, and then transmits the combined information to the ego vehicle via wireless communications, while moving forward as well. In this experiment, we assumed vehicle identification is supported before map merging.

B. General Evaluation

The experiments was performed on a campus road at National University of Singapore (NUS), Singapore, where the rule of the road is to drive on the left and the speed limit is 40 km/h. Fig. 11 shows the test road. The fleet of test vehicles start to move at *Start* toward *End* via *P-turn*. The road from *Start* to *P-turn* is an urban road with buildings. The road from *P-turn* to *End* is a road often lined with trees and bushes on either side, and includes a hill road. Experiments were conducted on sunny or cloudy days between 12 P.M. and 5 P.M.

1) *Map Merging Evaluation*: Fig. 12 shows the comparison of the different scan matching methods used to perform map merging during driving on the test road. The initial pose is obtained from LIDAR-based vehicle detection using the method presented in Section IV-C.1. The obtained pose is used as the initial condition for both the ICP and CSM scan matching algorithm. In this paper, the ICP algorithm from [59] is used. In the ICP method, $d_{th} = 2.5$ m is used. For CSM, two levels of

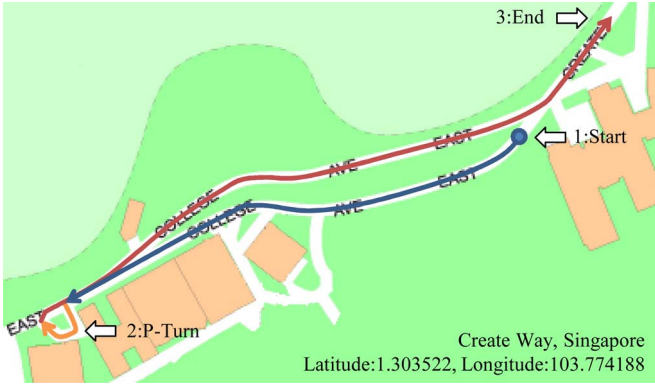


Fig. 11. Test road at NUS. The fleet of test vehicles start to move at *Start* toward *End* via *P-turn*.

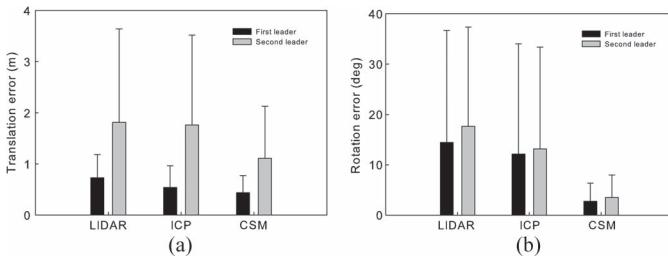


Fig. 12. (a) Translation errors and (b) rotation errors of map merging, according to the method. Average computation time of LIDAR, ICP, and CSM is 6, 6, and 108 ms, respectively.

0.5 and 0.1 m search grids are used. It is also set to search within a window of $10 \text{ m} \times 10 \text{ m} \times 60^\circ$. In order to obtain the ground truth, a particle filtered-based localization is used. We built a map using a 2-D laser scanner for localization on the test road before all experiments. Each of the three vehicles is localized by using the map and the relative position is extracted. We evaluated the whole trajectory of the vehicles on the test road. Details on the mapping method and its evaluation are available in [60].

In Fig. 12, the average error increases from the first leader to second leader across different map merging methods. The CSM, although slower, is able to ensure a high-quality scan matching and therefore able to achieve slightly over 1-m errors. This is also true when the rotation error is considered. There are several factors that contribute to the rather large standard deviation. Among others, the noisy observation obtained from LIDAR in an urban environment, LIDAR-based vehicle detection algorithm uncertainty and communication delays are the main contributing factors. Through Fig. 12, it was found that there are significant improvements in the merged map accuracy from the second leader by ICP and CSM. While CSM performed slower than ICP, it is able to robustly estimate the relative pose since CSM is robust to initialization error. We conclude that the results using the both scan matching approaches perform sufficient for our requirements.

Fig. 13 shows the leader vehicle detection addressed in Section IV-C.1. The method works well in a situation where there are a number of vehicles in the adjacent lane.

2) *Visualization Evaluation*: Fig. 14 shows a variety of visualizations made possible through our system including a satel-

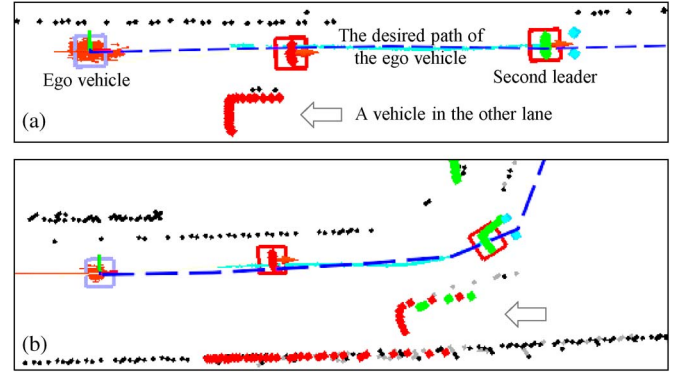


Fig. 13. Leader vehicle detection (a) in multilane roads and (b) at a curb. The horizontal blue dotted line is the desired path of the ego vehicle. Only vehicles detected near the desired path can be classified as a leader vehicle.

lite view, see-through views and third-person views. Fig. 14(a) is one snapshot of a merged map of the ego vehicle resulted from the proposed system. The red, green, and blue dots represent laser scan points of the ego vehicle, the first and second leader, respectively. Note that the ego vehicle cannot see the second leader due to the limitation of line-of-sight. Fig. 14(b) is the raw vision image from the second leader via wireless communications, where four small green circle indicate the region of interest for IPM. Fig. 14(c) and (d) are the raw vision image from the first leader via wireless communications and of the ego vehicle via a local vision sensor, respectively.

Fig. 14(e) is the see-through view from the first leader where the cuboid is a simplified vehicle model. Comparing to Fig. 14(c), the otherwise obstructed road marking “AHEAD”, as shown in Fig. 14(e) in a see-through manner. Fig. 14(f) is the see-through view from the ego vehicle, where the blue and green cuboid represent the second and first leader, respectively. Fig. 14(g) and (h) are the third-person views of the first leader and the ego vehicle, respectively, where viewpoint is lifted at 4 m from the ground in Fig. 14(g) and 30 m from the ground in Fig. 14(h). In brief, we can see that our system provides ahead traffic information even beyond line-of-sight, which can be applied to a see-through forward collision warning or an overtaking assistance.

At a sharp curve, such as T-junction or intersection, on-coming cars at blind spots can be potentially dangerous. The proposed system can help see an hidden on-coming vehicle at blind spots. Fig. 1(b) is a good example of this. As like the sharp curve, a downhill also includes a hidden obstacle problem. In Fig. 15, the vehicle i cannot see the vehicle $i + 2$ due to the limitation of line-of-sight. However, the vehicle $i + 1$ can let the vehicle i know the situation of the vehicle $i + 2$, as long as the vehicles are connected via wireless connectivity. Instead of the vehicle $i + 1$, an infrastructure equipped with cooperative perception system can let the vehicle i know the traffic situation at a downhill ahead.

Fig. 16(a)–(d) shows how our system aides in driving over a hill. Fig. 16(a) is the second leader’s camera view, where the preceding vehicle is descending from the top of hill. Fig. 16(b) and (c) are the camera views of the first leader and the ego vehicle, respectively. The ego vehicle is ascending a hill. Fig. 16(d) is a third-person view lifted at 30 m from the ground.

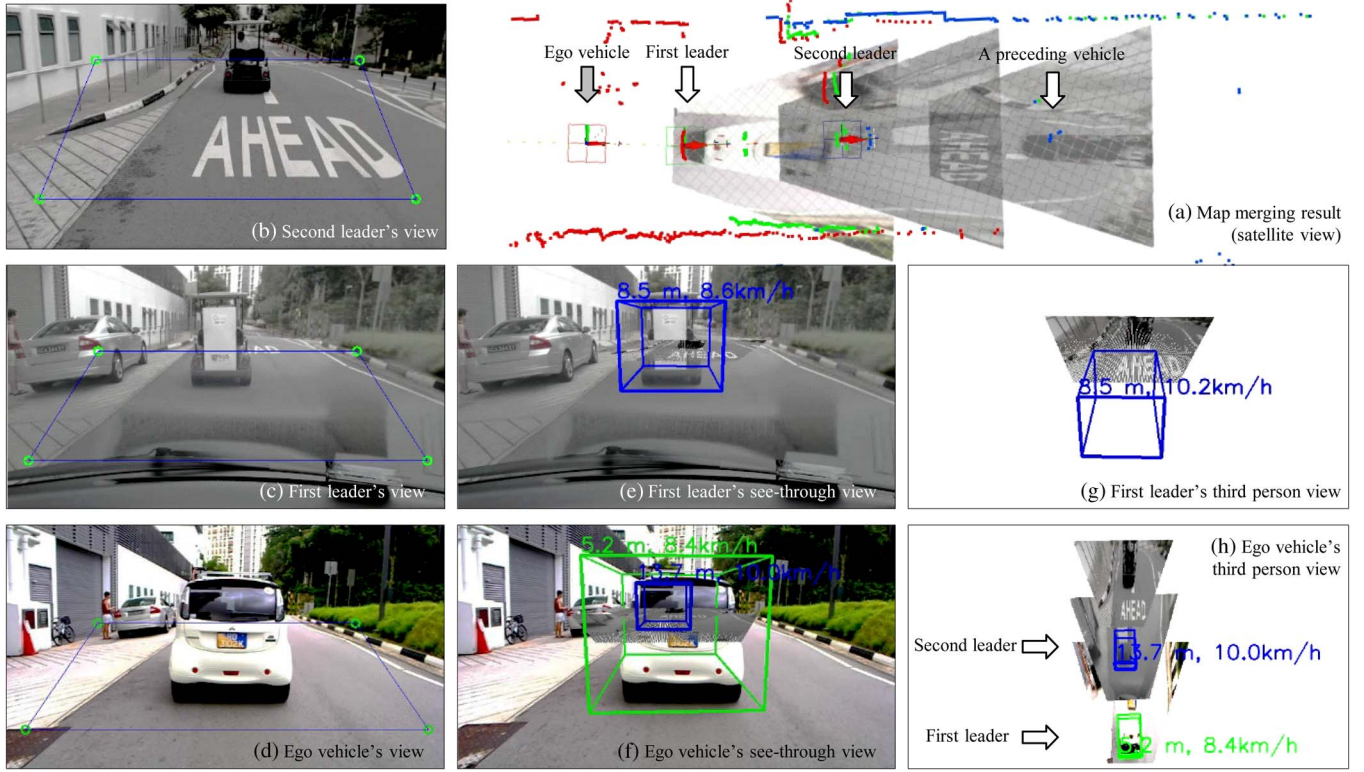


Fig. 14. (a) Merged map of the ego vehicle, where the red, green, and blue dots represent laser scan points of the ego vehicle, the first and the second leader, respectively. (b) Raw vision data from the second leader, where four small green circles represent the region of interest for IPM. (c) Raw vision image from the first leader. (d) Raw vision image of the ego vehicle. (e) See-through view of the first leader where the cuboid is a simplified vehicle model. Comparing to (c), a traffic sign on the road surface can be seen in (e) in a see-through manner. (f) See-through view of the ego vehicle, where the blue and green cuboids represent the second and first leaders, respectively. (g) Third-person view of the first leader, where driver's seat is lifted at 4 m from the ground. (h) Third-person view of the ego vehicle, where driver's seat is lifted at 30 m from the ground.

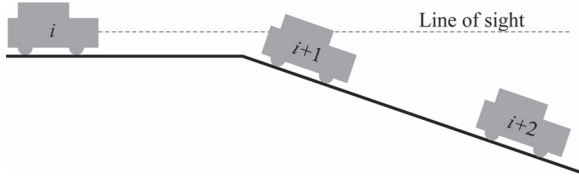


Fig. 15. Hidden obstacle problem at downhill.

Fig. 16(e)–(h) shows how all-around view is supported by a following vehicle connected via wireless communications. For this reason, the vehicle order is different between Fig. 16(a)–(d) and (e)–(h). Fig. 16(e) is the satellite view of the ego vehicle, where the gray arrow indicates the ego vehicle. The ego vehicle is moving between the first leader and the following vehicle. In Fig. 16(e), the ego vehicle can detect a vehicle approaching behind, because the following vehicle connected via wireless communications can tell the ego vehicle about what is going on at blind spot of the ego vehicle, which can greatly improve situational awareness.

Fig. 16(f) and (g) are a normal camera view and see-through view of the ego vehicle, respectively. Fig. 16(h) is a view of ego vehicle from behind given by the following vehicle. All-around view characteristics of Fig. 16(e) and (h) can be applied to the lane-changing and overtaking assistance. In addition, a normal vision camera can sense the adjacent left and right lane, as we can see Fig. 16(e). A fisheye lens can be considered for multiple lane detection.

Fig. 17 shows the perspective projection results according to mounting height and tilted-down angle, as presented in Section V-B. In the figure, the left-bottom $(h_c + h_v, \theta)$ indicates a camera mounting height (unit: meter), and a tilted-down angle (unit: radian), respectively. Fig. 17(a)–(i) shows a third-person view according to the different $h_c + h_v$ and different θ . In case of $h_c = 0$, our system performs similarly to see-through systems. According to the growth of h_c , the projection result generated by our system becomes close to a satellite view via a bird's eye view. Consequently, our proposed system supports drivers to be able to choose optimal parameters according to their preferences.

3) *Communication Evaluation*: Wireless communications are one of the key factors that enable cooperative perception, and at the same time, significantly affect overall system performances. In this paper, we suggest or propose no new communication method. Instead, we use IEEE 802.11g, IEEE 802.11n, 3G HSDPA and 4G LTE, which are off-the-shelf commercially ready solutions. Therefore, this evaluation focuses on how communication profile and uncertainty affect system performance in a given radio interface.

Table I provides a message profile we used in the experiment. The message is transmitted to other vehicles at 20 Hz. More specifically, we use standard message types in ROS, to transmit our sensing data. First, laser scan data is transmitted in the form of *LaserScan*, where 180 laser beams are contained in each scan. The total amount of information

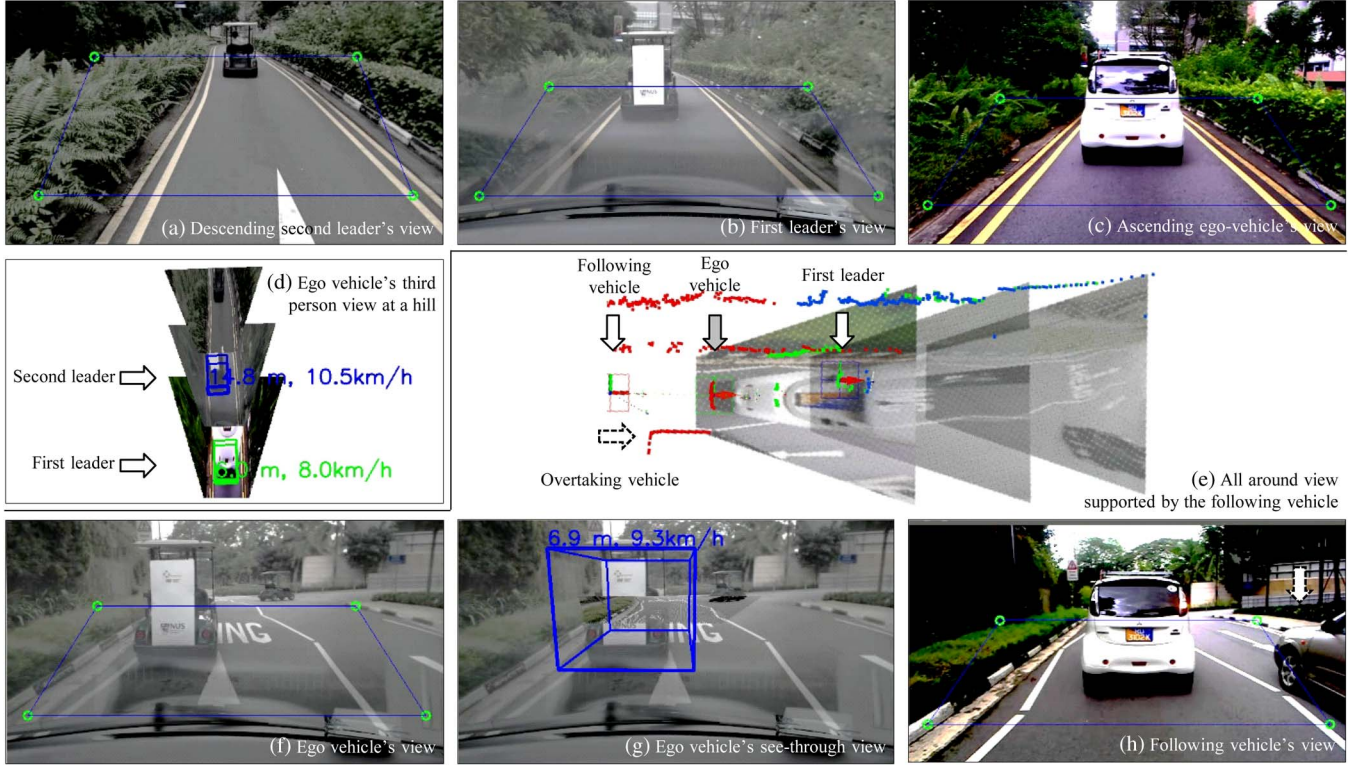


Fig. 16. [Driving over a hill] (a) Second leader's camera view, where the preceding vehicle is descending from the top of a hill. (b) Camera views of the first leader. (c) Camera views of the ego vehicle that is ascending the hill. (d) Third-person view lifted at 30 m from the ground. [All-around view] (e) Satellite view of the ego vehicle. Note that the ego vehicle is moving between the first leader and the following vehicle. In (e), the ego vehicle can detect a vehicle approaching from behind. (f) and (g) are a normal vision image and see-through image of the ego vehicle, respectively. (h) is view of the ego vehicle from behind.

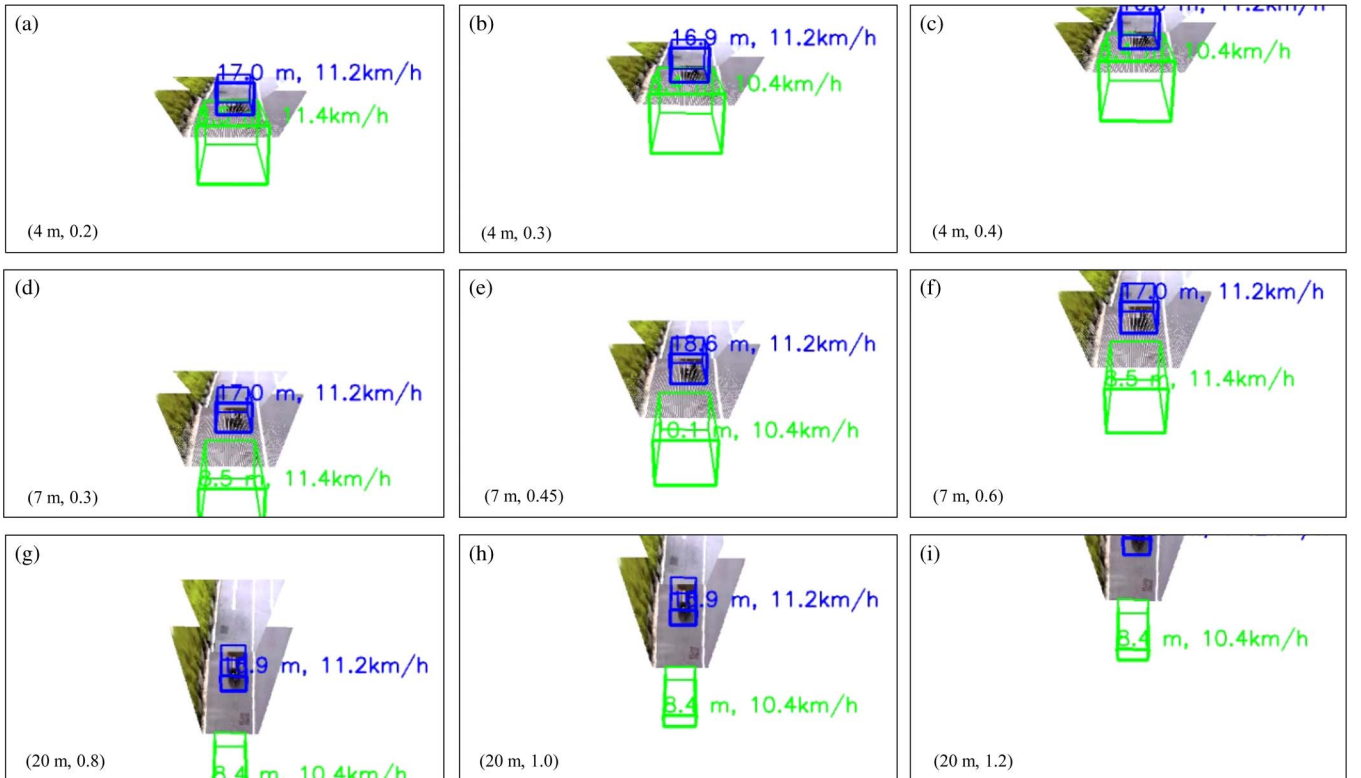


Fig. 17. Left-bottom indicates (camera height $h_c + h_v$, tilted-down angle θ). (a)–(i) show third-person view according to the different $h_c + h_v$ and different θ . In case of $D = 0$, our system performs similarly to the conventional see-through systems. As D increases, the projection result becomes closer to a satellite view. Drivers can choose optimal parameters according to their preferences.

TABLE I
MESSAGE PROFILE AND DATA SIZE

Message type	Size (Byte)	Desc
Laser scan	752	180 Beams
Raw image	691K	640×360×3
Compressed image	≈ 50K - 110K	640×360×3, PNG
Processed image	≈ 5,750	Point cloud
Processed image + Laser scan	≈ 6,500	-

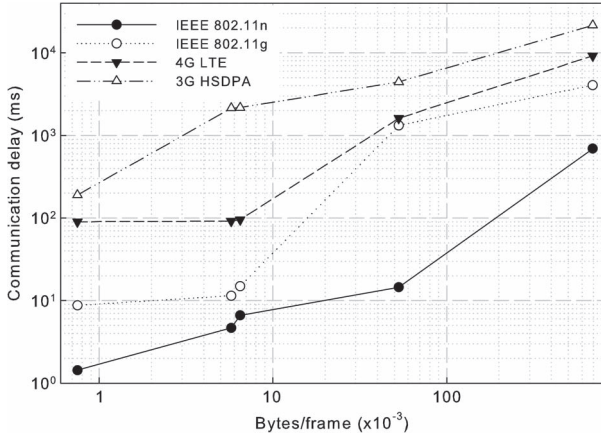


Fig. 18. Average delay for one single message delivery between two nodes away from 10 m each other via IEEE 802.11g, 4G LTE, and 3G HSDPA.

per frame is 752 bytes/frame = 25 bytes message header) + 7 bytes message description) + 180 beams \times float32 (4 bytes). Second, the raw data is 691,200 ($640 \times 360 \times 3$ (24-bit color)) bytes. The size of compressed images is roughly range from 50 to 110 kB in these experiments. The processed vision contains metadata, such as lane information, which is represented by a point cloud and transmitted as *PointCloud* in ROS. The message size varies depending on the extracted information, which is usually less than 6 kB, much smaller than raw images. Finally, we evaluate sensor multimodality with laser and processed vision data.

To monitor communication delay according to the message profile and radio interface, we set up two nodes, 10 m away from each other, and measured round-trip time (RTT). More specifically, the sender attached its CPU clock to the message. Once the message arrived at the receiver, the receiver resends the message to the sender without any message modification, such as an echo server. The sender can compare current local CPU clock with the CPU clock included in the message retransmitted from the receiver. In this evaluation, we use a half of RTT as a communication delay. The measurements were conducted for around one hour according to the message profile and radio interface. We measured communication delay in open public area, which means the delay can be affected by uncontrollable environmental interferences.

Fig. 18 shows the delay measurement via IEEE 802.11g, IEEE 802.11n, 3G HSDPA and 4G LTE. In terms of average delay, IEEE 802.11g and IEEE 802.11n outperforms 3G HSDPA and 4G LTE. However, the communication performance decrease as the distance $g_{i,j}$ increases. In this experiment, it was observed that 3G HSDPA and 4G LTE were not affected by the distance between communication correspondences.

Since the target of the proposed system is a moving vehicle on the road, the communication should be investigated from the perspective of impact on vehicle control, specifically position errors in this evaluation. Fig. 19(a) shows the impact of IEEE 802.11g communication delay on position error according to vehicle speed. From the perspective of average delay, the position error of one hop is 24 cm at 100 km/h in case of laser scan data only. The worst case position error becomes 5 m at 100 km/h. This may be good enough to be used as control-purpose information, depending on the goal or task. However, communication delay becomes significantly uncertain, as the size of data increases. Depending on the application and purpose, the message profile must be carefully determined. For example, in case of fusion with laser scan and processed vision of this measurement data, the worst case position error is 8 m. Fig. 19(b) and (c) provides average/worst delay in case 4G LTE and 3G HSDPA, respectively.

We can say that information from other vehicles are less reliable than locally sensed information in terms of delay or loss. From the perspective of vehicle control, time-critical control, such as emergency braking, should be determined based on local sensing information. Remote sensing information can be useful for long-term perspective path planning, such as decision problem between early lane changing and lane keeping for a human driver, as well as an autonomous driver.

C. Usability Test

We conducted usability tests with real drivers using real vehicles equipped with the proposed cooperative driving system. This paper proposes no new H/W to display our visualization results to drivers. Instead, one can use any display H/Ws that have been already used or considered for driver assistance systems such as dash board, center console, head-up display, or smartphone. In this test, we used a 20-in LCD monitor for visualization and installed it at a front passenger seat.

Fig. 20 shows usability test scenarios. In this test, we aimed at evaluating a see-through forward collision warning and overtaking assistance, including lane-changing assistance. Initially, three vehicles move toward the leftmost flag on a single-lane road, where a test driver drives the vehicle i . Then, the vehicle $i + 2$ suddenly stops at a certain position. To avoid collision, the vehicle $i + 1$ and i stop accordingly. After a while, the vehicle $i + 2$ speeds up, but $i + 1$ moves slowly. The test driver should make a decision whether to overtake the slow-moving vehicle $i + 1$ or follow the vehicle in the same lane for safety. In the opposite lane, oncoming vehicles are approaching toward the vehicle i randomly, which makes the test driver hesitate the decision.

In this experiment, the vehicle $i + 2$ was moving forward at 2–5 m/s before sudden braking. Accordingly, TTC was set to 15 s. These are somewhat conservative parameters mainly for safety concerns. Note that the brake lights of the vehicle $i + 2$ was turned off, which makes the vehicle $i + 1$ hard to notice whether the vehicle $i + 2$ decelerates or not. The warning sound includes two loud beeps that last for 2 s per one warning activation. We used IEEE 802.11n as a radio interface of vehicle-to-vehicle communications. LIDAR and compressed

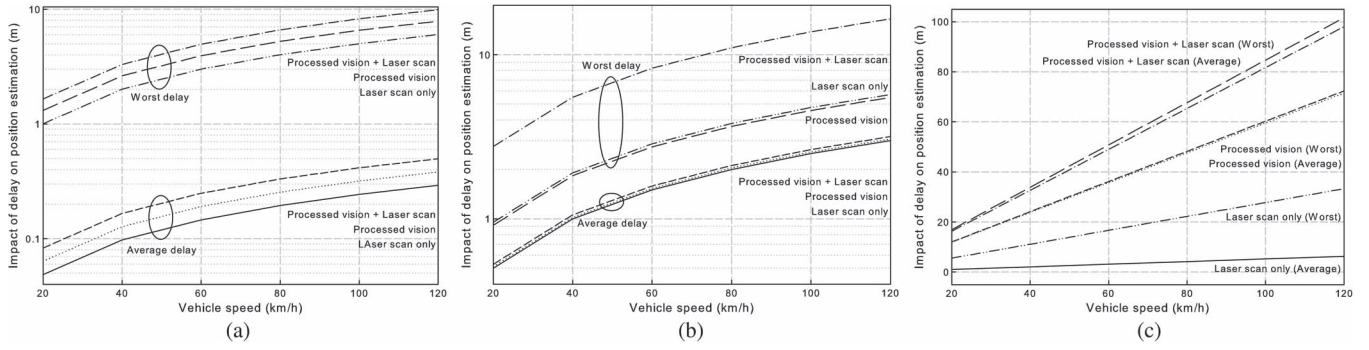


Fig. 19. Impact of communication delay on position error in case of (a) IEEE 802.11, (b) 4G LTE, and (c) 3G HSDPA.

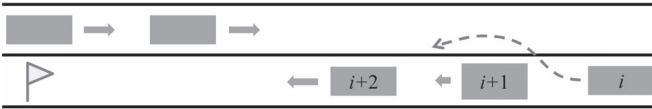


Fig. 20. Usability evaluation scenarios.

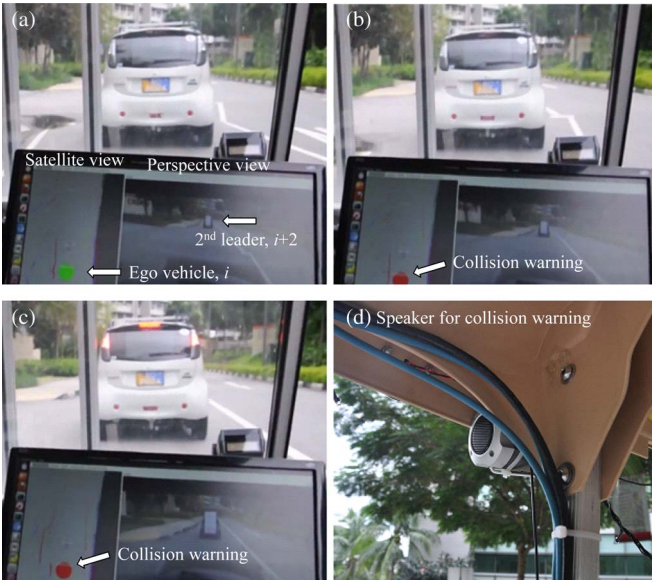


Fig. 21. Usability test with our proposed system on the road. (a) The bottom circle represents the vehicle i on the satellite view, where the green means no possible collision. (b) The green circle turns to red, because the vehicle $i + 2$ decelerates. Note that the vehicle $i + 1$ does not decelerate yet. (c) The vehicle $i + 1$ decelerates due to collision avoidance with the vehicle $i + 2$.

vision were used as a message profile. As shown in Fig. 18, the average communication delay between the vehicle $i + 1$ and i is 16 ms. IEEE 802.11n performed well at more than 30-m distance on the road.

Fig. 21 shows two views from the front passenger seat as seen at different time instances. The left-bottom screen of Fig. 21(a) shows the satellite view, where the bottom arrow indicates the vehicle i . The green circle indicates no collision possibility is detected at the moment. The right bottom screen of Fig. 21(a) is the perspective view. In principle, the test driver cannot see the vehicle $i + 2$ due to the limitation of line-of-sight. However, the test driver can see the vehicle $i + 2$ through our proposed system. Fig. 21(b) shows the snapshot of the activation of a see-through forward collision warning. In Fig. 21(b), the vehicle $i + 2$ suddenly stopped. Accordingly, the forward collision

TABLE II
DECELERATION STARTING TIME AND WARNING ACTIVATING TIME

	Average (s)
Deceleration starting time of $i + 2$	0
Deceleration starting time of $i + 1$	3.3
Forward collision warning activating time	2.15
Deceleration starting time of i	2.97

warning sign turned to red along with a loud sound alarm. Note that the vehicle $i + 1$ does not decelerate yet. In Fig. 21(c), (c) the vehicle $i + 1$ decelerates due to collision avoidance with the vehicle $i + 2$.

We conducted real tests with three human drivers on the road eight times. Table II summarizes the average time of deceleration starting time and forwarding collision warning activating time. The vehicle $i + 1$ decelerated 3.3 s later than the deceleration of the vehicle $i + 2$. 3.3 s can be seen as somewhat a slow reaction time. The time is obtained when the braking lights of the vehicle $i + 1$ is turned on. The driver of the vehicle $i + 1$ released the acceleration pedal before pushing the brake pedal. The brake lights of the vehicle $i + 2$ were turned off, and the vehicles were moving not very fast. All these factors affected the response time of the vehicle $i + 1$.

Average forward collision warning activating time is 2.1 s according to the preset TTC and the distance from the vehicle $i + 2$. The test drivers pushed the brake pedal averagely 0.82 s later after a forward collision warning sound is activating. In this experiment, the test drivers detected the sudden braking of the vehicle $i + 2$ earlier than the vehicle $i + 1$ and accordingly stopped earlier as much as 0.33 s.

Fig. 22 shows the timing diagram of one of usability tests. The speed of the vehicle i , $i + 1$ and $i + 2$ were obtained from odometry, local LIDAR of the vehicle i , and cooperative perception, respectively. In this test, we can see that after the deceleration of the vehicle $i + 2$, a collision warning is activated around 2 s later. Then, both the vehicle $i + 1$ and i decelerate at the almost same time.

Fig. 23 shows overtaking/lane-changing assistance test on the road. In Fig. 23(a), the preceding vehicle moves very slowly. The test driver could know there is enough space in front of the vehicle $i + 1$ and no oncoming vehicle in the opposite lane through the added see-through view. In Fig. 23(b), the test driver is overtaking the vehicle $i + 1$ with the support of our system.

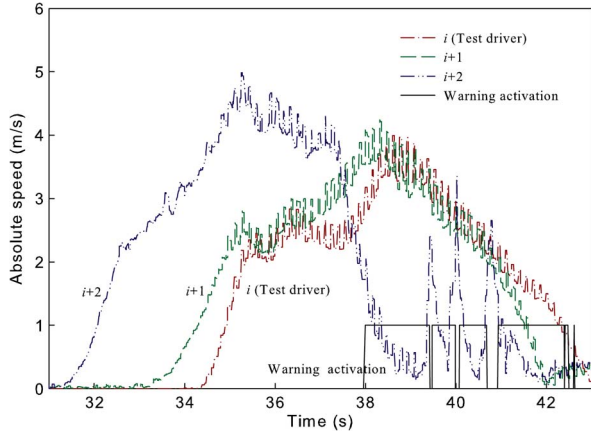


Fig. 22. Timing diagram of speed of three vehicles and collision warning activation. After the deceleration of the vehicle $i + 2$, a collision warning is activated. Then, the vehicle $i + 1$ and i decelerate at the almost same time.

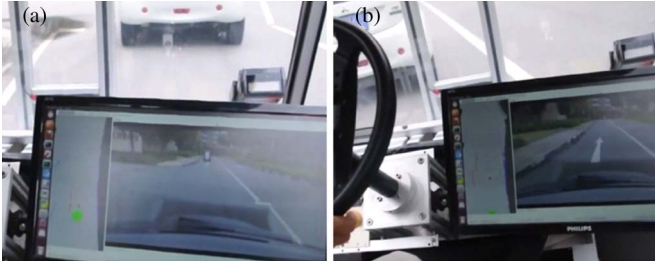


Fig. 23. Overtaking/lane-changing assistance test. (a) The test driver checks enough space in front of the preceding vehicle and no oncoming vehicle in the opposite lane. (b) The test driver overtakes the preceding vehicle.

In summary, it was observed that all test drivers had more confidence of driving by our proposed system in terms of hidden obstacle detection, overtaking, and lane changing, because they could see ahead the traffic situation at their driving seat whenever they wanted, and our system notified them along with visual and sound feedback when they should watch the display. Not surprisingly, our system, particularly a see-through forward collision warning, works well even at night.

D. Automated Lane Changing for Early Obstacle Avoidance

Suppose that in an emergency situation, an autonomous driver could take over vehicle control, so that the vehicle could avoid obstacles earlier by early lane changing with the support of our system. Automated early lane changing could be activated if a forward collision was imminent and lane-changing possibility was positive. Fig. 24 shows simulations of the scenario. These simulations utilized a RRT* path planner [61].

In Fig. 24(a), a sudden obstacle appears in front of the leader vehicle, and only the leader vehicle detects the obstacle due to the limitation of line-of-sight of the ego vehicle. In Fig. 24(b), the both of the leader and the ego vehicle suddenly stop, then the ego vehicle performs path replanning to overtake the stopped leader. However, it is difficult to find the feasible path to overtake the leader in the case of Fig. 24(b), because the ego vehicle and the leader are too close due to the sudden stop. In Fig. 24(c), with the support of cooperative perception, early lane changing is triggered according to (19) or (20). Since

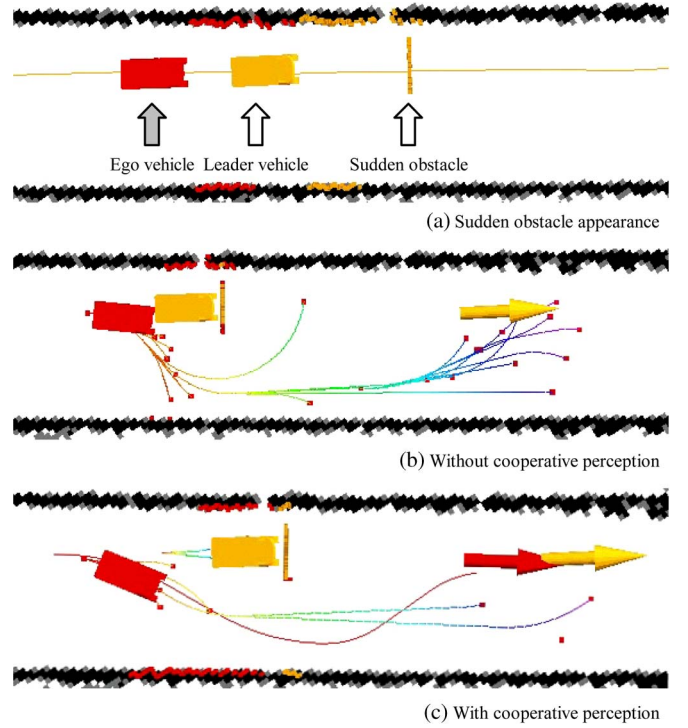


Fig. 24. (a) A sudden obstacle appears in front of leader vehicle, and only the leader vehicle detects the obstacle due to the limitation of line-of-sight of the ego vehicle. (b) The leader and the ego vehicle stop, then the ego vehicle performs path replanning. (c) With the support of cooperative perception, early lane changing is triggered.



Fig. 25. Our self-driving vehicle performs an automated lane-change by a see-through collision warning trigger.

sufficient space is occupied to overtake the leader, the path planner can find the overtaking path quickly.

Fig. 25 shows the moment when our self-driving vehicle performs an automated lane-change by a see-through collision warning trigger during an on-road experiment.

One of the possible benefits of this automated lane changer is to reduce the reaction time T_{rs} mentioned in (21), (23), and (24), as long as path planner and vehicle controller are faster than a human driver.

The proposed method allows only one vehicle to carry out the overtaking at a time and it is not intended for a distributed overtaking. Future work includes further research on multivehicle overtaking in the context of cooperative driving.

E. Impact of Weather Condition

Weather conditions should be considered for potential impact on the performance of wireless communications and sensor measurement.

According to the ITU-R P.676-8 [62], signal attenuation is not significantly worsened by dry weather or water vapor at less than 10 GHz frequency, thus communications over IEEE 802.11, 3G, and 4G networks should be unaffected by rain. In contrary, sensor measurement can be highly affected by weather conditions. First of all, vision can provide abundant information under sunny or cloudy conditions during the day, but much less information during the night. Active sensors, such as LIDAR and radar, are significantly less affected by light conditions. However, LIDAR can be blocked by fog, smoke, or rain. LIDAR may be able to partially overcome this by regarding range readings less than a certain threshold as false positive during a fog/rain situation. Radar is robust against such adverse weather conditions, but provides less resolution in clean air [63]. Conclusively, the sensors should be carefully selected and located such that various sensor types can play complementary roles and help protect against any shortcomings due to adverse weather.

IX. CONCLUSION

In this paper, we have provided our design and experimental validation of cooperative driving using cooperative perception. The extended perception range by the support of cooperative perception enables the driver to know the traffic situation even beyond line-of-sight or beyond field-of-view. The proposed system provides a see-through forward collision warning, overtaking/lane-changing assistance and automated lane-change capability using cooperative perception. We implemented the system and performed real experiments using vehicles on the road. The experimental results quantitatively show that the proposed system can contribute to the improvement of driving safety and non-myopic driving decision-making.

Compared with cooperative driving without sharing of perception data, cooperative perception can better assist to make a driving decision in complex traffic situations by abundant information that cooperative perception provides. Furthermore, the information enables self-driving vehicles to build an online see-through map for navigation, which can be used for non-myopic and safer decisions for cooperative self-driving.

ACKNOWLEDGMENT

The authors would like to thank Z. Cheng for mechanical design of the first vehicles used at our preliminary experiments, Dr. B. Luders and S. Pendleton for reading the manuscript and useful discussions, and anonymous reviewers for insightful and constructive comments, which help to improve this paper.

REFERENCES

- [1] L. Merino, F. Caballero, J. M. de Dios, J. Ferruz, and A. Ollero, "A cooperative perception system for multiple UAVs: Application to automatic detection of forest fires," *J. Field Robot.*, vol. 23, no. 3/4, pp. 165–184, Apr. 2006.
- [2] A. Rauch, F. Klanner, R. Rasshofer, and K. Dietmayer, "Car2X-based perception in a high-level fusion architecture for cooperative perception systems," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2012, pp. 270–275.
- [3] S.-W. Kim *et al.*, "Cooperative perception for autonomous vehicle control on the road: Motivation and experimental results," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013.
- [4] S.-W. Kim *et al.*, "Multiple vehicle driving control for traffic flow efficiency," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2012, pp. 462–468.
- [5] B. van Arem, C. J. G. van Driel, and R. Visser, "The impact of cooperative adaptive cruise control on traffic-flow characteristics," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 429–436, Dec. 2006.
- [6] K. Konolige, D. Fox, B. Limketkai, J. Ko, and B. Stewart, "Map merging for distributed robot navigation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2003, vol. 1, pp. 212–217.
- [7] A. Howard, M. J. Mataric, and G. S. Sukhatme, "Putting the 'I' in 'Team': An ego-centric approach to cooperative localization," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2003, vol. 1, pp. 868–874.
- [8] I. Rekleitis, G. Dudek, and E. Milios, "Experiments in free-space triangulation using cooperative localization," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2003, vol. 2, pp. 1777–1782.
- [9] N. Trawny, X. S. Zhou, K. X. Zhou, and S. I. Roumeliotis, "3D relative pose estimation from distance-only measurements," in *Proc. IEEE/RSJ Int. Conf. IROS*, 2007, pp. 1071–1078.
- [10] G. Dedeoglu and G. Sukhatme, "Landmark-based matching algorithm for cooperative mapping by autonomous robots," in *Distributed Autonomous Robotic Systems 4*. Tokyo, Japan: Springer-Verlag, 2000, pp. 251–260.
- [11] W. Huang and K. Beevers, "Topological map merging," *Int. J. Robot. Res.*, vol. 24, no. 8, pp. 601–613, Aug. 2005.
- [12] A. Birk and S. Carpin, "Merging occupancy grid maps from multiple robots," *Proc. IEEE*, vol. 94, no. 7, pp. 1384–1397, Jul. 2006.
- [13] F. Amigoni, S. Gasparini, and M. Gini, "Merging partial maps without using odometry," in *Multi-Robot Systems. From Swarms to Intelligent Automata Volume III*. Amsterdam, The Netherlands: Springer-Verlag, 2005, pp. 133–144.
- [14] F. Lu and E. Milios, "Robot pose estimation in unknown environments by matching 2D range scans," *J. Intell. Robot. Syst.*, vol. 18, no. 3, pp. 249–275, Mar. 1997.
- [15] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry," in *Proc. IEEE CVPR*, 2004, vol. 1, pp. 652–659.
- [16] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," in *Proc. Int. Soc. Opt. Photon., Robotics-DL Tentative*, 1992, pp. 586–606.
- [17] C. Yang and G. Medioni, "Object modelling by registration of multiple range images," *Image Vis. Comput.*, vol. 10, no. 3, pp. 145–155, Apr. 1992.
- [18] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proc. 3rd Int. Conf. 3-D Digit. Imag. Model.*, 2001, pp. 145–152.
- [19] A. Censi, L. Iocchi, and G. Grisetti, "Scan matching in the hough domain," in *Proc. IEEE ICRA*, 2005, pp. 2739–2744.
- [20] E. B. Olson, "Real-time correlative scan matching," in *Proc. IEEE ICRA*, 2009, pp. 4387–4393.
- [21] T. Rofer, "Using histogram correlation to create consistent laser scan maps," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2002, vol. 1, pp. 625–630.
- [22] M. Bosse and J. Roberts, "Histogram matching and global initialization for laser-only SLAM in large unstructured environments," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2007, pp. 4820–4826.
- [23] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, "Comparing ICP variants on real-world data sets," *Autonom. Robots*, vol. 34, no. 3, pp. 133–148, Apr. 2013.
- [24] H.-S. Tan and J. Huang, "DGPS-based vehicle-to-vehicle cooperative collision warning: Engineering feasibility viewpoints," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 415–428, Dec. 2006.
- [25] S. Wender and K. Dietmayer, "Extending onboard sensor information by wireless communication," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2007, pp. 535–540.
- [26] Z. J. Chong *et al.*, "Autonomy for mobility on demand," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 4235–4236.
- [27] B. Steux, C. Laugier, L. Salesse, and D. Wautier, "Fade: A vehicle detection and tracking system featuring monocular color vision and radar data fusion," in *Proc. IEEE Intell. Veh. Symp.*, 2002, vol. 2, pp. 632–639.
- [28] R. Labayrade, C. Royere, D. Gruyer, and D. Aubert, "Cooperative fusion for multi-obstacles detection with use of stereovision and laser scanner," *Autonom. Robots*, vol. 19, no. 2, pp. 117–140, Sep. 2005.
- [29] D. Gruyer, A. Cord, and R. Belaroussi, "Vehicle detection and tracking by collaborative fusion between laser scanner and camera," in *Proc. IEEE/RSJ Int. Conf. IROS*, 2013, pp. 5207–5214.
- [30] O. Ludwig, C. Premevida, U. Nunes, and R. Araújo, "Evaluation of boosting-SVM and SRM-SVM cascade classifiers in laser and vision-based pedestrian detection," in *Proc. 14th Int. IEEE ITSC*, 2011, pp. 1574–1579.
- [31] Q. Baig, O. Aycard, T. D. Vu, and T. Fraichard, "Fusion between laser and stereo vision data for moving objects tracking in intersection like scenario," in *Proc. IEEE IV*, 2011, pp. 362–367.

- [32] H. Li and F. Nashashibi, "Multi-vehicle cooperative perception and augmented reality for driver assistance: A possibility to see through front vehicle," in *Proc. 14th Int. IEEE ITSC*, 2011, pp. 242–247.
- [33] C. Olaverri-Monreal, P. Gomes, R. Fernandes, F. Vieira, and M. Ferreira, "The see-through system: A VANET-enabled assistant for overtaking maneuvers," in *Proc. IEEE Intell. Veh. Symp.*, 2010, pp. 123–128.
- [34] P. Gomes, C. Olaverri-Monreal, and M. Ferreira, "Making vehicles transparent through V2V video streaming," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 930–938, Jun. 2012.
- [35] A. Vinel, E. Belyaev, K. Egiazarian, and Y. Koucheryavy, "An overtaking assistance system based on joint beaconing and real-time video transmission," *IEEE Trans. Veh. Technol.*, vol. 61, no. 5, pp. 2319–2329, Jun. 2012.
- [36] E. Belyaev, A. Vinel, K. Egiazarian, and Y. Koucheryavy, "Power control in see-through overtaking assistance system," *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 612–615, Mar. 2013.
- [37] A. Girard, J. de Sousa, J. Misener, and J. Hedrick, "A control architecture for integrated cooperative cruise control and collision warning systems," in *Proc. IEEE Conf. Decis. Control*, Dec. 2001, vol. 2, pp. 1491–1496.
- [38] S. Kato, S. Tsugawa, K. Tokuda, T. Matsui, and H. Fujii, "Vehicle control algorithms for cooperative driving with automated vehicles and intervehicle communications," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 3, pp. 155–161, Sep. 2002.
- [39] S. Tsugawa, S. Kato, T. Matsui, H. Naganawa, and H. Fujii, "An architecture for cooperative driving of automated vehicles," in *Proc. IEEE Intell. Transp. Syst.*, 2000, pp. 422–427.
- [40] J. Kolodko and L. Vlacic, "Cooperative autonomous driving at the intelligent control systems laboratory," *IEEE Intell. Syst.*, vol. 18, no. 4, pp. 8–11, Jul./Aug. 2003.
- [41] J. Baber, J. Kolodko, T. Noel, M. Parent, and L. Vlacic, "Cooperative autonomous driving: Intelligent vehicles sharing city roads," *IEEE Robot. Autom. Mag.*, vol. 12, no. 1, pp. 44–49, Mar. 2005.
- [42] W. Liu, S.-W. Kim, Z. J. Chong, X. Shen, and M. H. Ang, Jr., "Motion planning using cooperative perception on urban road," in *Proc. IEEE Int. Conf. Cybern. Intell. Syst., Robot., Autom. Mechatron.*, Nov. 2013, pp. 130–137.
- [43] B. Rebsamen *et al.*, "Utilizing the infrastructure to assist autonomous vehicles in a mobility on demand context," in *Proc. IEEE TENCON*, Nov. 2012, pp. 1–5.
- [44] D. Jiang and L. Delgrossi, "IEEE 802.11p: Towards an international standard for wireless access in vehicular environments," in *Proc. IEEE VTC Spring*, 2008, pp. 2036–2040.
- [45] S. Biswas, R. Tatchikou, and F. Dion, "Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety," *IEEE Commun. Mag.*, vol. 44, no. 1, pp. 74–82, Jan. 2006.
- [46] A. Rauch, F. Klanner, and K. Dietmayer, "Analysis of V2X communication parameters for the development of a fusion architecture for cooperative perception systems," in *Proc. IEEE Intell. Veh. Symp.*, 2011, pp. 685–690.
- [47] A. Von Arnim, M. Perrollaz, A. Bertrand, and J. Ehrlich, "Vehicle identification using near infrared vision and applications to cooperative perception," in *Proc. IEEE Intell. Veh. Symp.*, 2007, pp. 290–295.
- [48] M. X. Punithan and S.-W. Seo, "King's graph-based neighbor-vehicle mapping framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1313–1330, Sep. 2013.
- [49] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 5, pp. 698–700, Sep. 1987.
- [50] K. Konolige and K. Chou, "Markov localization using correlation," in *Proc. Int. Joint Conf. Artif. Intell.*, 1999, pp. 1154–1159.
- [51] S. Thrun, W. Burgard, and D. Fox, "A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping," in *Proc. IEEE ICRA*, 2000, vol. 1, pp. 321–328.
- [52] B. Qin *et al.*, "Curb-intersection feature based Monte Carlo localization on urban roads," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2012, pp. 2640–2646.
- [53] H. A. Mallot, H. H. Blthoff, J. Little, and S. Bohrer, "Inverse perspective mapping simplifies optical flow computation and obstacle detection," *Biol. Cybern.*, vol. 64, no. 3, pp. 177–185, Jan. 1991.
- [54] M. Bertozzi and A. Broggi, "Gold: A parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 62–81, Jan. 1998.
- [55] B. Qin *et al.*, "A general framework for road marking detection and analysis," in *Proc. 16th Int. IEEE ITSC*, 2013, pp. 619–625.
- [56] E. Dagan, O. Mano, G. P. Stein, and A. Shashua, "Forward collision warning with a single camera," in *Proc. IEEE Intell. Veh. Symp.*, 2004, pp. 37–42.
- [57] R. Parasuraman, P. Hancock, and O. Olofinboba, "Alarm effectiveness in driver-centered collision-warning systems," *Ergonomics*, vol. 40, no. 3, pp. 390–399, 1997.
- [58] M. Quigley *et al.*, "ROS: An open-source robot operating system," in *Proc. ICRA Workshop Open Source Softw.*, 2009, vol. 3, no. 2, pp. 1–6.
- [59] J. L. B. Claraco, *Development of Scientific Applications with the Mobile Robot Programming Toolkit. The MRPT Reference book*. Málaga, Spain: Mach. Perception Intell. Robot. Lab., Univ. Málaga, 2008.
- [60] Z. Chong *et al.*, "Synthetic 2D LIDAR for precise vehicle localization in 3D urban environment," in *Proc. IEEE ICRA*, 2013, pp. 1554–1559.
- [61] S. Karaman, M. R. Walter, A. Perez, E. Frazzoli, and S. Teller, "Anytime motion planning using the RRT*," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 1478–1483.
- [62] "Attenuation by Atmospheric Gases," Geneva, Switzerland, P.676-8, Oct. 2009.
- [63] B. Yamauchi, "All-weather perception for man-portable robots using ultra-wideband radar," in *Proc. IEEE ICRA*, 2010, pp. 3610–3615.



Seong-Woo Kim (M'11) received the B.S. and M.S. degrees in electronics engineering from Korea University, Seoul, Korea, in 2005 and 2007, respectively, and the Ph.D. degree in electrical engineering and computer science from Seoul National University in 2011.

He is currently a Postdoctoral Associate with the Singapore-MIT Alliance for Research and Technology. His research interests include online and offline optimization targeted for perception and control of intelligent and autonomous vehicles.

Dr. Kim was a recipient of the Best Student Paper Award at the 11th IEEE International Symposium on Consumer Electronics and the Outstanding Student Paper Award at the First IEEE International Conference on Wireless Communication, Vehicular Technology, Information Theory, and Aerospace and Electronic Systems Technology.



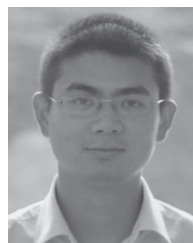
Baoxing Qin received the B.S. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2010. He is currently working toward the Ph.D. degree in the Department of Mechanical Engineering, National University of Singapore, Singapore.

His research interests include autonomous vehicle localization, object recognition, mapping and environment understanding.



Zhuang Jie Chong received the B.S. degree in mechanical engineering (with first-class honors) from Nanyang Technological University, Singapore, in June 2010. He is currently working toward the Ph.D. degree in the Department of Mechanical Engineering, National University of Singapore.

His research interests include autonomous vehicle, localization, and mapping.



Xiaotong Shen received the B.S. degree in mechanical engineering from Harbin Institute of Technology, Harbin, China, in 2012. He is currently working toward the Ph.D. degree in the Department of Mechanical Engineering, National University of Singapore, Singapore.

His research interests include cooperative perception.



Wei Liu received the B.S. degree in mechanical engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2011, and the M.S. degree in mechatronics from National University of Singapore, Singapore, in 2012, respectively. He is currently working toward the Ph.D. degree in the Department of Mechanical Engineering, National University of Singapore, Singapore.

His research interests include motion planning under uncertainties.



Marcelo H. Ang, Jr. received the B.S. degrees (*cum laude*) in mechanical engineering and industrial management engineering from the De La Salle University, Manila, Philippines, in 1981, the M.Sc. degree in mechanical engineering from the University of Hawaii at Manoa, Honolulu, Hawaii, in 1985, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Rochester, Rochester, NY, USA, in 1986 and 1988, respectively.

His work experience includes heading the Technical Training Division of Intel's Assembly and Test Facility in the Philippines, research positions at the East West Center in Hawaii and at the Massachusetts Institute of Technology, and a faculty position as an Assistant Professor of electrical engineering with the University of Rochester, NY. In 1989, he joined the Department of Mechanical Engineering, National University of Singapore, where he is currently an Associate Professor. He also holds a joint appointment with the Division of Engineering and Technology Management as Deputy Head. In addition to academic and research activities, he is actively involved in the Singapore Robotic Games as its founding Chairman. He also chairs the Steering Committee for the World Robot Olympiad (2008–2009) and the World Skills Singapore Competition (2005, 2007, and 2010). His research interests span the areas of robotics, mechatronics, and applications of intelligent systems methodologies. He teaches both at the graduate and undergraduate levels in the following areas: robotics, creativity and innovation, applied electronics and instrumentation, advanced computing, product design and realization, and special topics in mechatronics. He is also active in consulting work in these areas.



Emilio Frazzoli (SM'07) He received the Laurea degree in aerospace engineering from the University of Rome, "Sapienza", Italy, in 1994, and the Ph.D. degree from the Department of Aeronautics and Astronautics of the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2001.

He is a Professor of aeronautics and astronautics with the Laboratory for Information and Decision Systems, and the Operations Research Center with the MIT. Before returning to MIT in 2006, he held faculty positions with the University of Illinois, Urbana-Champaign, and with the University of California, Los Angeles. He is currently the Director of the Transportation@MIT initiative, and the Lead Principal Investigator of the Future Urban Mobility IRG of the Singapore-MIT Alliance for Research and Technology (SMART). He was the recipient of a NSF CAREER award in 2002. He is an Associate Fellow of the American Institute of Aeronautics and Astronautics. His current research interests focus primarily on autonomous vehicles, mobile robotics, and transportation systems, and in general lie in the area of planning and control for mobile cyberphysical systems.



Daniela Rus (F'10) received the Ph.D. degree in computer science from Cornell University, Ithaca, NY, USA, in 1992.

She is currently the Andrew (1956) and Erna Viterbi Professor with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, where she serves as a Director of the Computer Science, and Artificial Intelligence Laboratory. Her research interests include distributed robotics and mobile computing, and her application focus includes transportation, security, environmental modeling and monitoring, underwater exploration, and agriculture.

transportation, security, environmental modeling and monitoring, underwater exploration, and agriculture.

Dr. Rus was the recipient of the National Science Foundation Career Award. She is a Class of 2002 MacArthur Fellow and a fellow of the Association for the Advancement of Artificial Intelligence (AAAI).