

Interactive Road Situation Analysis for Driver Assistance and Safety Warning Systems: Framework and Algorithms

Hong Cheng, *Member, IEEE*, Nanning Zheng, *Fellow, IEEE*, Xuetao Zhang, Junjie Qin, and Huub van de Wetering

Abstract—Road situation analysis in Interactive Intelligent Driver-Assistance and Safety Warning (I²DASW) systems involves estimation and prediction of the position and size of various on-road obstacles. Real-time processing, given incomplete and uncertain information, is a challenge for current object detection and tracking technologies. This paper proposed a development framework and novel algorithms for road situation analysis based on driving action behavior, where the safety situation is analyzed by simulating real driving action behaviors. First, we review recent development and trends in road situation analysis to provide perspective for the related research. Second, we introduce a road situation analysis framework, where onboard sensors provide information about drivers, traffic environment, and vehicles. Finally, on the basis of the previous frameworks, we proposed multiple-obstacle detection and tracking algorithms using multiple sensors including radar, lidar, and a camera, where a decentralized track-to-track fusion approach is introduced to fuse these sensors. In order to reduce the effect of obstacle shape and appearance, we cluster lidar data and then classify obstacles into two categories: static and moving objects. Future collisions are assessed by computation of local tracks of moving obstacles using extended Kalman filter, maximum likelihood estimation to fuse distributed local tracks into global tracks, and finally, computation of future collision distribution from the global tracks. Our experimental results show that our approach is efficient for road situation evaluation and prediction.

Index Terms—Driver assistance systems, environment modeling, object detection and tracking, sensor fusion, situation assessment.

I. INTRODUCTION

THE ASIAN Development Bank states that “In the five years 2000–2004, more than 500 000 people were killed and around 2.6 million injured in road accidents in the People’s Republic of China (PRC), equivalent to one death every 5 minutes—the highest rate in the world.” and estimates a

Manuscript received February 1, 2006; revised November 30, 2006, December 5, 2006, and December 11, 2006. This work was supported by the National Natural Science Foundation of China under Grant 60021302. The Associate Editor for this paper was F.-Y. Wang.

H. Cheng, N. Zheng, and X. Zhang are with the Institute of Artificial Intelligence and Robotics, Xi’an Jiaotong University, Xi’an 710049, China (e-mail: hcheng@mail.xjtu.edu.cn; nzheng@aiar.xjtu.edu.cn; xtzhang@aiar.xjtu.edu.cn).

J. Qin is with the Realsil Microelectronics Incorporation, Suzhou 215021, China.

H. van de Wetering is with the Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, 5600 MB Eindhoven, The Netherlands (e-mail: wstahw@win.tue.nl; h.v.d.wetering@tue.nl).

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.2006.890073

yearly economic loss of \$12.5 billion. Driver assistance and safety warning systems promise to provide partial solutions to these problems, and consequently, many research efforts aim at developing algorithms and building frameworks for them [1]–[6].

Road situation analysis requires not only obstacle information at the current time but also predicted obstacle information at the future time. Indeed, an experienced driver looks several seconds along the road and bases his actions on information so obtained. This previewing of the road is necessary to avoid accidents since vehicle dynamics limits the car in making speed or direction changes.

Interactive Intelligent Driver-Assistance and Safety Warning (I²DASW) uses more than one kind of sensor: image sensors, lidar, and radar. No single sensor can provide input as complete, robust, and accurate as required by I²DASW. Image sensors have some problems, such as low ability of sensing depth and higher ability of discrimination than lidar and radar. Radar shows limited lateral spatial information because it is not available at all, the field of view is narrow, or the resolution is reduced at large distances. Although lidar has a wide view field that solves part of the previous problems, there are other problems such as low ability of discrimination, clustering error, and recognition latency. These restrictions of the different sensor types explain the attention given to sensor fusion in research on object detection and tracking [2], [7], [8], resulting in a wide spectrum of promising applications in assistance driving, including multisensor Adaptive Cruise Control (ACC), fusion of advanced ACC and lane keeping systems [9], and smart airbag systems.

On the basis of [1], we proposed a road safety situation and threat analysis framework and algorithms based on driver behavior and vehicle dynamics. In current environment modeling, obstacles are detected and tracked in future situation assessment; we use the position and size of obstacles at the current time, together with vehicle dynamics equations to predict the future road situation. For lidar data, we distinguish the object types, namely, static or moving objects, by estimating object speed. Here, a decentralized track-to-track fusion algorithm is proposed to detect and track obstacles.

The remainder of this paper is given as follows: Section II introduces the state of the art related to road safety frameworks and multiple-sensor multiple-object tracking algorithms. In Section III, we provide a detailed description of our interactive safety analysis framework. We describe how to judge

obstacle types by estimating its speed in Section IV, where a speed estimation algorithm for obstacles is proposed. In Section V, we implement obstacle tracking in global tracks and then evaluate the safety situation by combining the current obstacle information into vehicle dynamics equation. In Section VI, we state the conclusions of this paper and provide perspective for future work.

II. PREVIOUS WORK

Road situation analysis for driver assistance and safety warning is an interdisciplinary endeavor involving many research fields, for instance, computer science, automobile engineering, cognitive science, and psychology. We classify existing frameworks that analyze obstacles in a traffic scene into two categories: The first one is a current situation analysis framework, which attempts to provide the vehicle and the driver with the obstacles' information in the current time by fusing different sensors [2], [3], [10]. The other one is obstacle situation prediction in the future [11]. To assess the future situation, many prediction approaches have been used, such as the extended Kalman filter (EKF), Monte Carlo method [11], and Bayesian network [7].

Real-time safety analysis in a traffic environment involving driver, vehicle, the traffic environment itself, and their interaction is a challenge for perception, modeling, decision, and control. Several safety analysis systems have been proposed to address different aspects in road situation analysis [10]–[13]. In [11], a Monte Carlo reasoning framework evaluates the probability of a future collision and use a Monte Carlo important sampling for the approximation of a collision integral. The looking-in and looking-out framework proposed by McCall *et al.* is a system-oriented safer driving framework [12], which consists of driving ecology sensing, hierarchical context processing, and modeling of drivers, vehicles, and environment. They build the Human-Centered Intelligent Driving Support System to emphasize the role of driver. In context of an earth-work vehicle, a distributed sensor network aims at processing data acquired by different sensors, integrating them, and producing an interpretation of the environment observed [10]; its main objectives of low-level and high-level data fusion are to obtain a rough and an accurate estimate of the number of objects present in the observed scene and their three-dimensional (3-D) positions, respectively. In addition, intersection and intersection-related crashes consistently make up a high proportion of total crashes, and the main causes include inappropriate speed, absence of traffic signs, wrong turning, and pedestrian mistake, etc. The INTERSAFE project shows the need of driver assistance systems for intersection safety [13], where two approaches in parallel, namely 1) top-down approach and 2) bottom-up approach, are realized. Here, a dynamic risk is assessed and analyzed to support the driver at the intersection based on object tracking and classification, and the intent of the driver. Consequently, potential conflicts with other road users can be reported only a few seconds before.

Multisensor multiobject detection and tracking systems have received considerable attention over the last five years [2], [3], [7], [8], [14]–[18]. In [14], a strategy that distinguishes between

a static object and a moving object by estimating object speed has been proposed, where both the speed and the direction of the objects and the host vehicle are used to estimate the speed. In [15], three different geometric-object models are designed for small objects, objects described by rectangular cars and the like, and free-form objects, respectively. In terms of obstacle classification and tracking, the most generally used combination manner consists of a camera and a range sensor [14]–[17]. An approach that simplifies the fusion between range and vision sensors using corresponding sets of hypothesis was proposed [16]. In this system, a radar device and a monocular camera are fused by sharing sets of hypotheses for detection of vehicles. In [2], a decentralized multiple-sensor multiple-target tracking approach for the Autotaxi system is considered for avoiding collisions, where the tracking involves three stages, namely 1) data alignment; 2) track-to-track association; and 3) track fusion. A sensor fusion strategy that introduces depth cue into the segmentation algorithm improves the target segmentation performance due to the complement of radar and vision [18]. As a precrash system, the SAVE-U project aims at protecting pedestrians and bicyclists and avoiding collisions between pedestrians and vehicles, where the sensor platform consists of radar sensors, normal cameras, and infrared cameras. Alternatively, other combination manners, such as infrared camera and radar [3], have good performance in driver assistance systems.

This paper proposes an integrated current and future safety situation analysis framework as general as possible, where we model not only the sensing phase but also the control phase. In this framework, a speed estimation algorithm based on lidar data is used to distinguish two types of obstacles, namely 1) static objects and 2) moving objects. On the basis of the speed and type of obstacles, we form obstacle tracks using only a single sensor and following a track fusion approach that is used to yield accurate and robust global tracks. We use a camera to detect lanes and obstacles, such as vehicles and pedestrians, in the regions of interest (ROIs) generated by range sensors. Combining the lane structure with obstacle tracks, we can model the traffic environment and assess road situation at both the current and near-future time.

III. INTERACTIVE SAFETY ANALYSIS FRAMEWORK

Many existing robotics technologies apply to intelligent assistance driving [20]; however, much research work neglects the preview of a driver and driver response delay. Moreover, the behavior of high-speed vehicles differs greatly from other robots. To obtain safe driving, a driver should be in the center of the safety analysis [12]; driver response delay, together with other factors, restricts the driving path of a vehicle. On the basis of these factors, we proposed an integrated interactive road safety analysis framework, where the system consists of the safety analysis framework consisting of the following modules: onboard sensor network, environment modeling and sensor fusion, vehicle ego-state and vehicle dynamics module, future situation assessment, decision-making agents, human–computer interface (HCI), and a preview-following model-based control module (see Fig. 1). In this framework, we consider a driver assistance system as a vehicle–driver–environment interactive

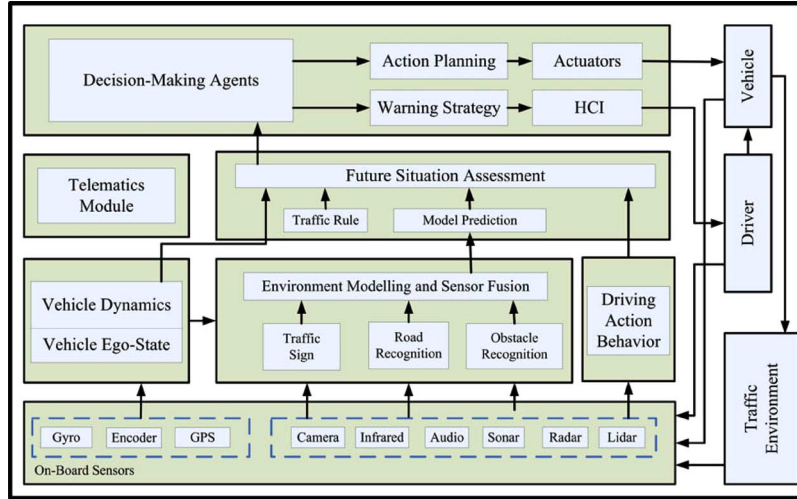


Fig. 1. Interactive road situation analysis framework.

closed-loop system; moreover, we focus on not only the current situation but also the future situation by predicting the potential collision probability distribution.

In our framework, onboard sensors provide real-time information about drivers, traffic environment, and vehicles. How to configure these sensors is closely related to the application domain. For example, for multisensor ACC systems, radar and camera often suffice, but for pedestrian protection systems, an infrared sensor is essential to robust detection under various weather conditions. In general, the external sensors capture object appearance, range, and sound outside a vehicle, and the interior sensors collect vehicle state, such as speed, acceleration, and steering angle.

The main functions of environment modeling and sensor fusion are to sense and recognize obstacles, lanes and traffic signs, etc., where sensor information of different kinds is fused to model the environment. Lane detection is the problem of locating lane boundaries, and robust detection for a variety of road types and under a variety of illumination conditions is necessary. Fig. 2 shows various lane scenarios including general conditions and hard conditions. The general conditions include different weather conditions, lane types, lane mark types, etc. The hard conditions include “against the sun,” “exit at highway,” lane-number change, etc. For moving objects, robust and reliable detection and tracking is a challenging task under the conditions of variable size and distance, various weather and illumination, and cluttered background [21], [22]. For instance, in visible ACC, detecting and tracking cut-in and overtaking vehicles in a dynamic background is an important task for avoiding a crash. Fig. 3 shows various cut-in and overtaking-vehicle scenes.

In the future situation assessment module, future road safety situations are assessed by combining traffic rules, vehicle dynamics, and environment prediction. Since the safety distance varies with the speed of a host vehicle, we adopt preview time rather than safety distance as the measurement of safety response. The safety response time is given by

$$T_o = \frac{d_r + d_v + d_s}{v} \quad (1)$$

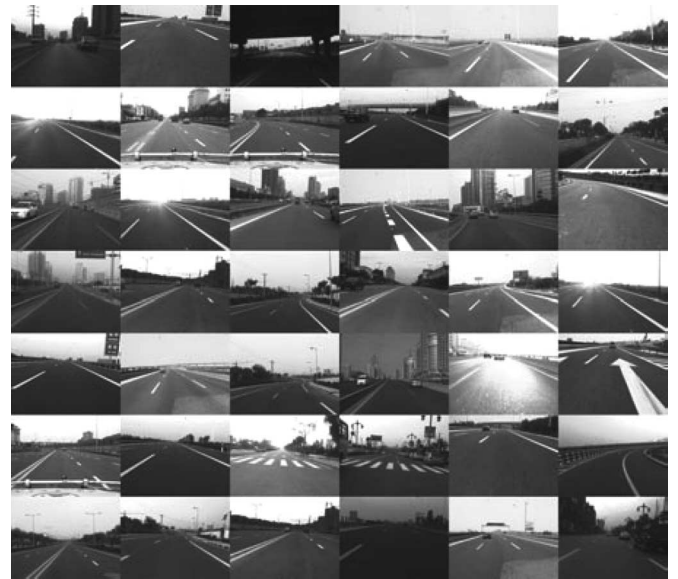


Fig. 2. Various lane scenario in Xi'an, China.



Fig. 3. Close cut-in and overtaking vehicles with different sizes, colors, and styles.

where d_r is the distance required to respond to the nearest object due to driver response delay, d_v is the distance to slow down, d_s is the safety distance between the host vehicle and obstacles, and v is the speed of the host vehicle.

Decision-making agents have two functions: One is to generate warning strategies for warning systems, such as route guide systems and warning display devices, and the other one is to plan the actions of the actuators to control the path of the host vehicle. As we know, the parameters of a vehicle model are affected by many factors and vary with the outside environment. Therefore, real-time learning algorithms using fuzzy logic have been applied to solving lateral and longitudinal control problems for a long time and have shown good performance; we use the fuzzy self-adjusted and real-time learning algorithm to estimate the parameters in this framework [23]. Moreover, a preview optimal curvature model based on the preview follow model and driver behavior characteristic is utilized to control a vehicle's velocity and direction, where the key problem is to establish fuzzy evaluation indexes and their membership functions that represent the front road geometry shape, traffic rules, and driver behavior. For details, see [24]. In our framework, decision-making agents use rigid kinematics and vehicle dynamics stable-state response properties to yield the expected path, and then action planning updates the ideal path by using vehicle dynamics dynamic-state response properties. Their main interaction activities are that a driver operates a vehicle and that a vehicle produces lateral motion and longitudinal motion, resulting in change in the vehicle.

Furthermore, this framework involves in-vehicle telematics modules and HCI modules. As a consequence of increase in traffic accident, the driver-centered HCI is becoming critically important. An in-vehicle HCI module presents warning information from the decision module and the path planning module to the driver. Meanwhile, the visual and audio messages of driver warning and assistance should not distract the attention of drivers. Telematics modules implement information sharing between vehicles and between a vehicle and a base station. Future vehicles will be able to share the information about the environment to provide cooperative, convenient, and safer driving.

For reasoning about the future motion of obstacles in a traffic environment, some road safety reasoning rules are concluded in [11]. Similarly, here, we sum up the factors of the future road situation.

- 1) *Traffic rule.* While a driver implements a driving activity, traffic rules have a potential effect on the expected path. Traffic rules in highway and urban road ensure safe comfortable collision-free driving. To achieve these objectives, the driver and the onboard sensors must recognize the traffic sign.
- 2) *Vehicle dynamics.* The motion of a vehicle is restricted by vehicle dynamics including two factors: One is the internal factor involving the tires, steering systems, and acceleration and deceleration systems, and the other is the external factor involving driver instruction. In our framework, we consider it as a whole to affect safe driving rather than look at each influencing subfactor and focus on vehicle stable-state and dynamic-state response properties.
- 3) *Driver behavior.* The aim of the I²DASW systems is to develop an automatic system that can replace fully

or partly a professional and experienced driver. Clearly, driver behavior characteristics are inevitably involved in safe driving. Except road conditions and vehicle mechanical failure, most of traffic accidents are caused by the actions of drivers, e.g., driving in an inappropriate speed, ignoring right of way, overtaking, and following too closely. In real-life traffic scenes, driver behavior, such as driver response delay, deeply affects the driving operation.

- 4) *Sensor uncertainty.* Sensor noise causes uncertainty in data. The incomplete information is used to assess the scene, and modeling of the background and dynamic object is a challenge, given the incomplete and uncertainty information.
- 5) *Vehicle state.* Vehicle state includes position, velocity, acceleration, direction angle, yaw angle, etc. Yaw angle affects the dynamic properties of a vehicle greatly. Given the basic state parameters, we can generate a predicted path.

As general as possible, the I²DASW systems have three major functions:

- 1) *Provide appropriate just-in-time information regarding the vehicle, driver, and traffic environment for safer and better driving.* For example, real-time traffic and traveler information aims at facilitating the access to public data and providing drivers with information about the traffic environment and other vehicles.
- 2) *Safety warning and assistance systems.* The system warns the driver proactively about possible hazardous situations on the basis of the vehicle's current position, orientation, speed, and the road situation; moreover, steps can be taken to control the driver when his or her vehicle is in a hazardous situation. The safety warning systems monitor the driving situation and provide the traffic situation, e.g., potential collision information, including route guide systems, Lane-Change Decision Aid Systems, Traffic Impediment Warning Systems, Forward Vehicle Collision Warning Systems, etc. Safety assistance systems use the warning information to generate the expected path and control the vehicle directly. Typical systems are Forward Collision-Avoidance Assistance Systems, ACC systems, Low-Speed Following Systems, Stop & Go systems, etc.
- 3) *In-vehicle safety protection device for drivers and passengers.* Such a system can protect drivers and passengers from the impact between humans and vehicle bodies, e.g., smart airbag systems.

IV. LIDAR CLASSIFICATION INTO STATIC OR MOVING OBSTACLES

A lidar sensor is often used as an onboard sensor for driver assistance systems. Much effort about clustering the original data and classifying the resulting objects have been made [14], [25]. A strategy was proposed for distinguishing all objects by lidar and dividing them into three categories, namely 1) moving objects; 2) roadside reflectors; and 3) overhead signs [14], where the motion of detected objects are judged by the

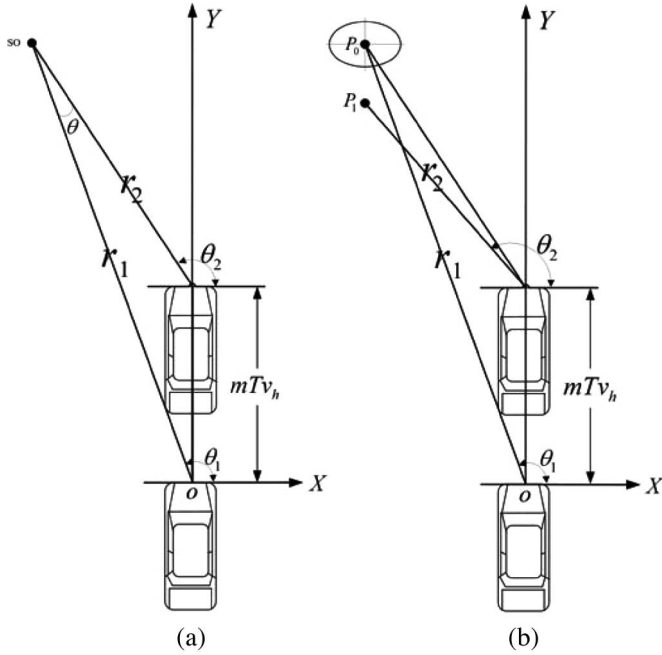


Fig. 4. Our speed estimation approach.

relationship between the path of the host vehicle and the changes in the positions of the objects. A method consisting of three modules, namely 1) scan segmentation; 2) object classification; and 3) object tracking by lidar, is used to detect and track multiple objects [25]. The position and size of obstacles are not sufficient to assess safety in driving assistive systems, and the various behaviors of all the obstacles on the road, such as speed and acceleration, should also be considered. We have developed an algorithm to estimate the speed of all obstacles.

First, we segment the lidar data into several clusters, and each cluster represents one target. We say that the lidar measurements r_k and r_{k+1} of distances to the laser device belong to points of the same target [25] iff

$$|r_k - r_{k+1}| \leq r_{\min} \frac{2 \tan \beta \cdot \sin\left(\frac{\phi}{2}\right)}{\cos\left(\frac{\phi}{2}\right) - \sin\left(\frac{\phi}{2}\right) \tan \beta} \quad (2)$$

where $r_{\min} = \min\{r_k, r_{k+1}\}$, ϕ is the angular resolution, and β represents the maximum absolute angle between the object face and the incidence line at time k . In our experiments, $\phi = 0.25^\circ$, and $\beta = 85^\circ$.

In many similar systems, the vehicle speed is measured by the encoder [26]. In contrast, we proposed a vehicle speed estimation algorithm by using a static object given the two observation values (r_1, θ_1) and (r_2, θ_2) as follows:

$$v_h = \frac{r_1^2 + r_2^2 - 2r_1r_2 \cos(\theta_1 - \theta_2)}{mT} \quad (3)$$

where T is the sampling interval, and m is the number of the consecutive frames and is generally larger than 1 for improving the accuracy of velocity. Fig. 4(a) shows the velocity computation approach for a host vehicle.

Here, we assume that over a small interval of time mT , the driving direction of the host vehicle is consistent with the

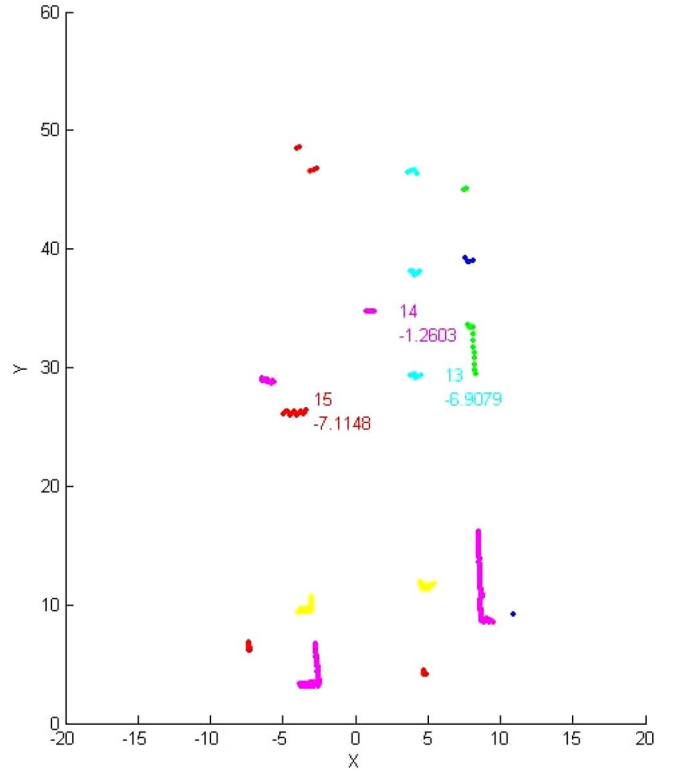


Fig. 5. Lidar clustering and speed estimation.

Y -axis in the Cartesian coordinates system XOY , as shown in Fig. 4(b). Therefore, after finishing the speed estimation of a host vehicle, we can obtain the coordinates of two segments $P_0 = (x_0, y_0)$ and $P_1 = (x_1, y_1)$, i.e.,

$$\begin{cases} x_0 = -r_1 \cos \theta_1 \\ y_0 = r_1 \sin \theta_1 \end{cases} \quad (4)$$

$$\begin{cases} x_1 = -r_2 \cos \theta_2 \\ y_1 = r_2 mT v_h + r_2 \sin \theta_2 \end{cases} \quad (5)$$

Then, we can estimate the absolute speed of stable objects in the scene by the following equation:

$$v_o = \frac{\sqrt{(y_1 - y_0)^2 + (x_1 - x_0)^2}}{mT} \quad (6)$$

Considering the noise and vibration of the lidar, we can judge whether an object is moving or stationary by Mahalanobis distance, given two segments P_0 and P_1 , i.e.,

$$d = (P_0 - P_1)^T \Sigma^{-1} (P_0 - P_1)$$

where Σ is a covariance matrix reflecting the uncertainty characteristics of lidar data. If $d < d_0$, the object is static, or else the object is moving. Here, the decision rule can be interpreted geometrically as saying that the distance between the two points is less than d_0 , taking into account the variance. Fig. 5 shows the results of velocity estimation using our algorithm, where the 13th to 15th objects are stable for several consecutive frames.

V. MULTIPLE-OBSTACLE TRACKING AND SITUATION ASSESSMENT

A. Multiple-Obstacles Tracking Based on EKF Using a Single Sensor

1) *Probability Framework of Tracking*: From the viewpoint of probability, tracking is a statistical inference; in other words, given the observation values at time 1 and extending up to and including time k : $Z_{1:k} = \{z_1, z_2, \dots, z_k\}$, we may construct the posterior probability $P(X_k|Z_{1:k})$ and then obtain estimation \hat{X}_k and covariance matrix \mathbf{C} of the state vector X_k at time k .

For the sake of simplicity, we make two assumptions.

1) The state at current time k only depends on the state at previous time $k - 1$, which is called a first-order Markov process. Consequently, it yields the following equation:

$$P(X_k|X_{k-1}, X_{k-2}, \dots) = P(X_k|X_{k-1}).$$

2) The observation at current time k depends only on the current state $P(Z_{1:k}|X_k) = P(Z_k|X_k)P(Z_{1:k-1}|X_k)$.

By hypothesis, we can deduce the Bayesian posterior probability as

$$P(X_k|Z_{1:k}) = \frac{P(Z_k|X_k)P(X_k|Z_{1:k-1})}{P(Z_{1:k}|Z_{1:k-1})} \quad (7)$$

where $P(X_k|Z_{1:k})$ is the posterior probability, $P(Z_k|X_k)$ is the likelihood, $P(X_k|Z_{1:k-1})$ is the prior probability, and $P(Z_{1:k}|Z_{1:k-1}) = \int P(Z_k|X_k)P(X_k|Z_{1:k-1})dX_k$ is the belief.

For the probability framework of a tracking problem, we may proceed in the manner described next.

- *Prediction Step*. Given $P(X_{k-1}|Z_{1:k-1})$, we can obtain $P(X_k|Z_{1:k-1})$ and $\hat{X}_{k|k-1}$.
- *Update Step*. Given $P(X_{k-1}|Z_{1:k-1})$ and Z_k , we can obtain $P(X_k|Z_{1:k})$ and $\hat{X}_{k|k}$.

2) *System Model*: In this system, we adopt the constant acceleration model to build the system equation

$$\begin{cases} X_k = \mathbf{F}X_{k-1} + \mathbf{G} \cdot v \\ Z_k = h(X_k) + w. \end{cases} \quad (8)$$

The state vector at time k is defined as

$$X_k = [x_k, \dot{x}_k, \ddot{x}_k, y_k, \dot{y}_k, \ddot{y}_k]^T$$

where x_k, \dot{x}_k , and \ddot{x}_k are the position, velocity, and acceleration in the X direction at time k , respectively, and y_k, \dot{y}_k , and \ddot{y}_k are the position, velocity, and acceleration in Y direction at time k , respectively.

The state transition matrix can be written as

$$\mathbf{F} = \begin{bmatrix} 1 & T & T^2/2 & 0 & 0 & 0 \\ 0 & 1 & T & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & T & T^2/2 \\ 0 & 0 & 1 & 0 & 1 & T \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

where T is the sampling interval of a sensor. In (8), $v = [v_{a_x}, v_{a_y}]^T$ is the process noise, which is modeled as zero-mean white noise whose correlation matrix is defined by

$$\mathbf{E} \{v_k v_j^T\} = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \delta_{kj}$$

where the dimension of v is $n = 2$. The process noise distribution matrix corresponding to the preceding expression is

$$\mathbf{G} = \begin{bmatrix} T^2/2 & 0 \\ T & 0 \\ 1 & 0 \\ 0 & T^2/2 \\ 0 & T \\ 0 & 1 \end{bmatrix}.$$

We define the observation value at time k as

$$Z_k = \begin{bmatrix} r_k \\ \theta_k \end{bmatrix}$$

and its observation function is

$$h(x_k) = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ \arctan[y_k/x_k] \end{bmatrix}.$$

Therefore, we can get the following form:

$$\begin{cases} x_k = r_k \cos \theta_k \\ y_k = r_k \sin \theta_k. \end{cases} \quad (9)$$

In this system, the state of objects is expressed in Euclidean coordinates systems, while the observation values are expressed in polar coordinates systems. Consequently, the observation equation is nonlinear. We may now linearize $h(X)$ around $X = \hat{X}_{k|k-1}$ and obtain the observation matrix

$$\begin{aligned} H_k &= \left. \frac{\partial h}{\partial X} \right|_{X=\hat{X}_{k|k-1}} \\ &= \begin{bmatrix} \frac{\hat{x}_{k|k-1}}{\sqrt{\hat{x}_{k|k-1}^2 + \hat{y}_{k|k-1}^2}} & 0 & 0 & \frac{\hat{y}_{k|k-1}}{\sqrt{\hat{x}_{k|k-1}^2 + \hat{y}_{k|k-1}^2}} & 0 & 0 \\ \frac{-\hat{y}_{k|k-1}}{\hat{x}_{k|k-1}^2 + \hat{y}_{k|k-1}^2} & 0 & 0 & \frac{\hat{y}_{k|k-1}}{\hat{x}_{k|k-1}^2 + \hat{y}_{k|k-1}^2} & 0 & 0 \end{bmatrix}. \end{aligned}$$

In (8), $w = [w_r, w_\theta]$ is the observation noise, which is modeled as zero-mean white noise whose correlation matrix is defined by

$$\mathbf{E} \{w_k w_j^T\} = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\theta^2 \end{bmatrix} \delta_{kj} = \mathbf{R} \delta_{kj}.$$

Here, the dimension of w is 2.

Combining the probability framework of tracking with the minimum mean square error, we can yield the EKF's prediction equation

$$\begin{cases} \hat{X}_{k|k-1} = \mathbf{F}X_{k-1|k-1} \\ P_{k|k-1} = \mathbf{F}P_{k-1|k-1}\mathbf{F}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T \end{cases} \quad (10)$$

where $\hat{X}_{k|k-1}$ is the state prediction at time k given the state at time $k - 1$, and $P(k|k - 1)$ is the prediction covariance.

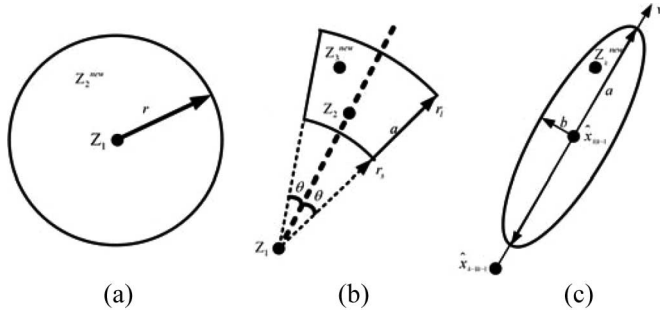


Fig. 6. Various association gates. (a) Circle. (b) Sector. (c) Ellipse.

The update equation given z_k and $\hat{X}_{k|k-1}$ at time k can be written in the form

$$\begin{cases} \hat{X}_{k|k} = \hat{X}_{k|k-1} + W_k [Z_k - H_k \hat{X}_{k|k-1}] \\ P_{k|k} = P_{k|k-1} - W_k S W_k^T \end{cases} \quad (11)$$

$$\begin{cases} S = \mathbf{H} P_{k|k-1} \mathbf{H}^T + \mathbf{R} \\ W_k = P_{k|k-1} \mathbf{H}^T S^{-1} \end{cases} \quad (12)$$

where S is the observation prediction covariance, W_k is the Kalman gain, $\hat{X}_{k|k}$ is the output of the state update, and $P_{k|k}$ is the update state covariance.

3) *Initial Conditions*: Concerning the initialization of the EKF, we determine the local tracks by using the acceleration of three points where it is assumed that the motion of an object is modeled as constant acceleration, finishing the initialization operation. For details, see [27].

4) *Data Association for a Single Sensor*: For lidar and radar data, data association is the first step at judging the correspondence relation between the current observation and the previous track. Our data association includes two categories, namely 1) observation-to-observation and 2) observation-to-track. The main objective of the association between observations is to initialize tracks correctly, while the association between observation and track aims at holding and updating the existing tracks.

a) *Observation-to-observation association*: When a new object appears, we can hold the observation directly. For a single point, we do not know the moving direction of the object. In that case, when the object has an observation z_1 , we use the circle association gate to judge the correlation between z_1 and z_2^{new} without moving direction [see Fig. 6(a)].

To associate a new observation, the next problem is to compute the radius r of the association gate. The radius of associate gate is defined as

$$\Delta r_{\max} = (v_h - v_o) \cdot T \quad (13)$$

where v_h is the velocity of the host vehicle and v_o is the velocity of an obstacle. If $|z_2^{\text{new}} - z_1| \leq \Delta r_{\max}$, it means that z_2^{new} is correlated with z_1 ; otherwise, they do not correlate.

If there are two existing observation points, i.e., z_1 and z_2 , of a certain object, we can use a sector association gate to judge the correlation between z_3^{new} and z_2 [see Fig. 6(b)]. If the following inequality is satisfied:

$$\begin{cases} r_s \leq |z_1 z_3^{\text{new}}| \leq r_l \\ \left| \arg(z_1 z_3^{\text{new}}) - \arg(z_1 z_2) \right| \leq \theta \end{cases} \quad (14)$$

Then, z_3^{new} is located inside the association gate, which represents the correlation between z_3^{new} and z_2 . Here, θ is a threshold value.

b) *Observation-to-track association*: During the period of tracking, we obtain the state update value of a track $\hat{X}_{k-1|k-1}$ at time $k-1$ and the state prediction value of a track $\hat{X}_{k|k-1}$ at time k . Combining observation value z_k^{new} at time k with the two previous state values judges whether z_k^{new} is associated with $\hat{X}_{k|k-1}$ or not.

After the initialization of tracks, it yields the state estimation of objects by using a prediction and update model. In general, a longer exiting period results in less estimation covariance. In this paper, we set an ellipse association gate with its center $\hat{X}_{k|k-1}$ [see Fig. 6(c)] and choose the motion direction of an object as a major axis.

Define the new observation as

$$z_k^{\text{new}} = [r_k^{\text{new}}, \theta_k^{\text{new}}].$$

Consequently, we obtain the Euclidean coordinates of the observation given the observation value z_k^{new} in the form

$$\begin{cases} x_k^{\text{new}} = r_k^{\text{new}} \cos(\theta_k^{\text{new}}) \\ y_k^{\text{new}} = r_k^{\text{new}} \sin(\theta_k^{\text{new}}) \end{cases} \quad (15)$$

Then, we can obtain the state prediction value

$$\hat{X}_{k|k-1} = [x_{k|k-1}, \dot{x}_{k|k-1}, \ddot{x}_{k|k-1}, y_{k|k-1}, \dot{y}_{k|k-1}, \ddot{y}_{k|k-1}]^T$$

and the motion direction of the object

$$\theta_o = \arctan\left(\frac{\dot{y}_{k|k-1}}{\dot{x}_{k|k-1}}\right).$$

Here, θ_o is the rotated angle of the ellipse association gate.

On the basis of the previous results, we get the ellipse equation of the association gate

$$\frac{x_e^2}{a^2} + \frac{y_e^2}{b^2} = 1 \quad (16)$$

where

$$\begin{cases} x_e^2 = (x - x_{k|k-1}) \cos \theta_o + (y - y_{k|k-1}) \sin \theta_o \\ y_e^2 = (x - x_{k|k-1}) (-\sin \theta_o) + (y - y_{k|k-1}) \cos \theta_o \end{cases}$$

We may now define a distance function

$$d_{X,Z} = \frac{x_e^2}{a^2} + \frac{y_e^2}{b^2} \quad (17)$$

where a and b are the half length of two axes of an ellipse.

Substituting the observation value $z_k^{\text{new}} = (x_k^{\text{new}}, y_k^{\text{new}})$ into (17), $d_{X,Z} \leq 1$ indicates that z_k^{new} and $\hat{X}_{k|k-1}$ are correlated, while $d_{X,Z} > 1$ indicates that z_k^{new} is uncorrelated with $\hat{X}_{k|k-1}$.

On the basis of the distance function of the preceding observation-to-observation and observation-to-track, we build a distance matrix for all the passing points and use the Global

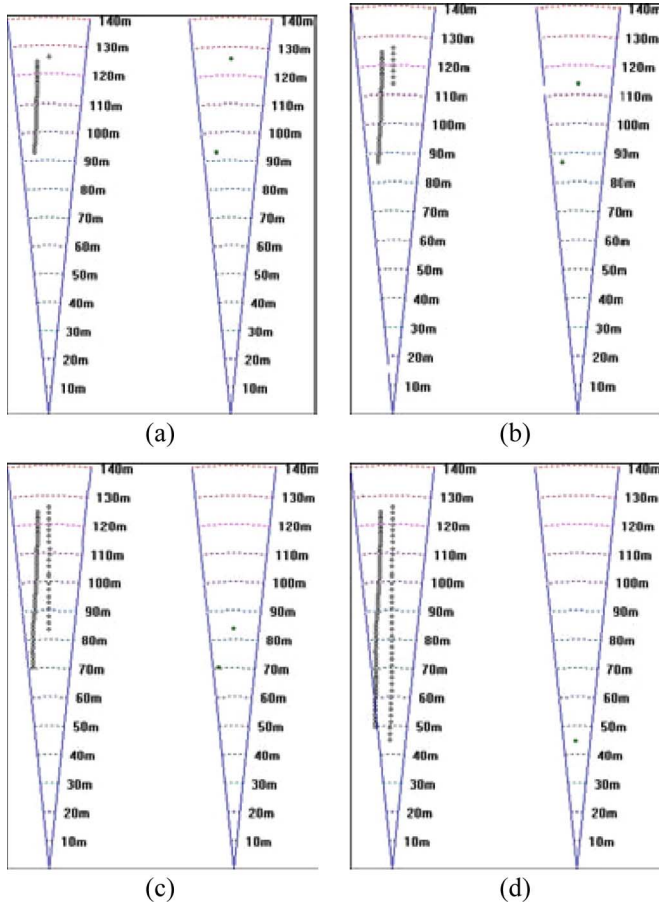


Fig. 7. Multiple-obstacle tracking using a radar sensor. (a) $t = 2200.7823$ s. (b) $t = 2201.1762$ s. (c) $t = 2202.0933$ s. (d) $t = 2203.2786$ s.

Nearest Neighbor (GNN) algorithm to associate an observation with another observation or a track.

5) *Single-Track Management*: Single-track management is an important step of object tracking. In this paper, for every observation point, if there are three correlated observation values among five consecutive values, we can initialize the EKF finishing the start of a track.

Tracks holding is done to keep the tracks of objects continuously by the beforehand stated rules after the start of the track. We use a sliding window detector to hold the tracks, where an N/M rule is used to judge whether these tracks exist. In other words, N correlated observation values out of M observation values are considered to show that the track can exist. With the increase on the holding time of tracks, the belief of this track is getting bigger and bigger. Consequently, in terms of actual implementation, M and N/M , during the start period of a track, can be set to a smaller value than that of the later period of tracking. In this paper, $M = 8$, and $N = 5$.

To process a vanishing object, cancellation of tracks is necessary. There are three categories in which tracks are required to be canceled. The first one is the point without initialization; if there are no three correlated consecutive observation values, the track is canceled. The second one is a start track; if the N/M rule is violated, the track is canceled. The third one is a track made by a reverse-direction object; when the object moves behind the host car, the track can be canceled immediately.

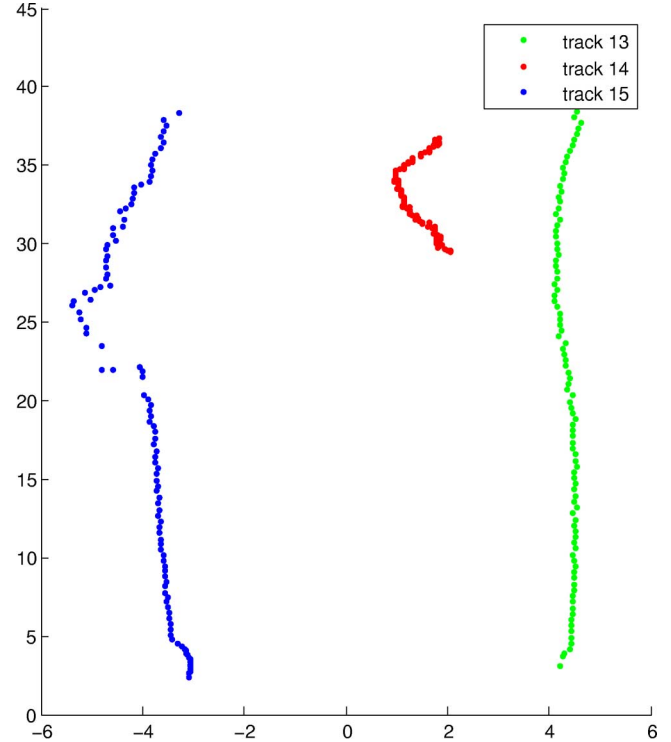


Fig. 8. Obstacles tracks using our approach.

Fig. 7 shows the tracks of multiple obstacles using a single sensor.

B. Lidar and Radar Track Fusion

1) *Data Alignment*: Lidar and radar work independently and unsynchronized; consequently, data alignment is needed. For multisensor fusion, we must first transform the different coordinates into a same-coordinates system and then fuse the local tracks. Here, we map the lidar coordinates and radar coordinates into the vehicle coordinates to solve the position alignment. Moreover, we synchronize the time between lidar and radar by using the prediction equation of EKF.

2) *Track Association*: On the basis of the two local tracks of lidar and radar \hat{X}_l and \hat{X}_r , we can yield the corresponding relation between the two local tracks. The distance function is defined as [2], [28]

$$d_{lr} = (\hat{X}_l - \hat{X}_r)^T (P_l + P_r - P_{lr} - P_{rl})^{-1} (\hat{X}_l - \hat{X}_r). \quad (18)$$

In actual implementation, we neglect the cross-covariance matrices between lidar and radar: P_{lr} and P_{rl} . That is, $P_{lr} = P_{rl} = 0$.

Let $x = d_{lr}$, and let it have a χ^2 distribution with M degrees of freedom in the form [28]

$$f(x) = \frac{1}{2^{\frac{M}{2}} \Gamma(\frac{M}{2})} x^{\frac{M-2}{2}} e^{-\frac{x}{2}} \quad (19)$$

where Γ is the gamma function with the following properties:

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \quad \Gamma(1) = 1 \quad \Gamma(m+1) = m\Gamma(m).$$

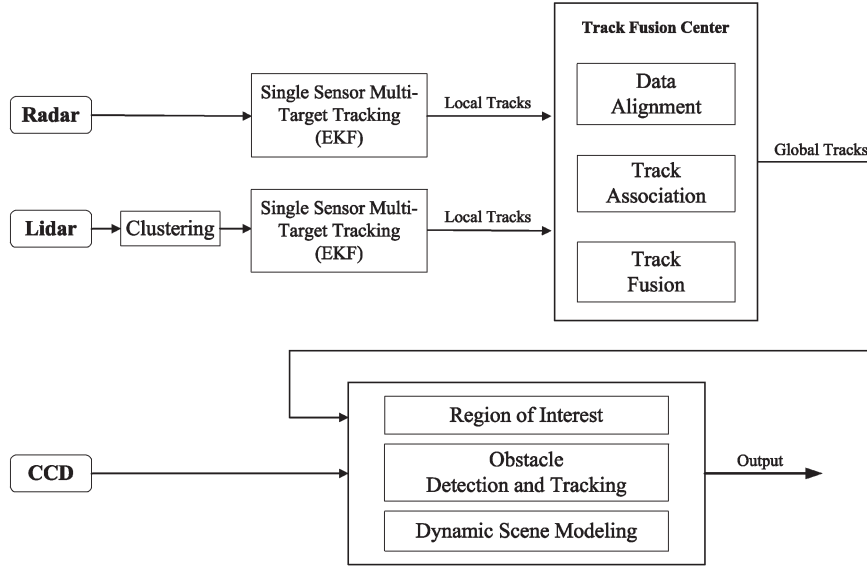


Fig. 9. Multiple-sensor multiple-obstacle detection and tracking system.

The probability of $x \in (0, \sigma)$ may be written as

$$\alpha = \int_0^{\sigma} f(x) dx.$$

Based on a σ corresponding to a given α , we can set the ellipse association gate as

$$\begin{cases} H_0 : d_{lr} \leq \sigma \\ H_1 : d_{lr} > \sigma \end{cases} \quad (20)$$

where H_0 indicates that \hat{X}_l and \hat{X}_r come from the same object; H_1 indicates that \hat{X}_l and \hat{X}_r come from different objects.

Assuming that N track pairs pass the association gate, we rank the track pairs by the corresponding distance value d_{lr} . Since one object has only one track pair, we take the track pair with minimum distance value d_{lr} .

3) *Track Fusion Algorithm*: There now remains the problem of track fusion given the local tracks of lidar and radar \hat{X}_l and \hat{X}_r , and its covariance matrices P_r and P_l . To solve the track fusion, we use maximum likelihood estimation to fuse the tracks [2]. First, we assume that the state estimation error has a Gaussian distribution and then obtain the state estimation value and its covariance of a local track in the form [2]

$$\begin{cases} X_{ml} = P_{ml} (P_l^{-1} \hat{X}_l + P_r^{-1} \hat{X}_r) \\ P_{ml} = (P_l^{-1} + P_r^{-1})^{-1} \end{cases} \quad (21)$$

Fig. 8 shows the fusion results of the lidar and radar sensors.

Through the preceding process, we can yield the ROIs using global tracks. Moreover, we can extract a more accurate environment structure using visual information. Fig. 9 shows the structure of multiple-sensor and multiple-obstacle detection and tracking system. In this paper, two image sensors are used for lane recognition and vehicle detection. Our lane recognition approach is an Adaptive Randomized Hough Transform [19], with robust and accurate detection of lane markings without manual

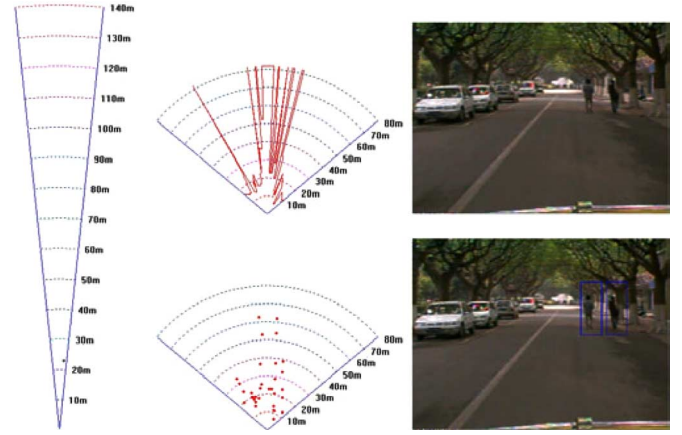


Fig. 10. Software interface of fusion using the three sensors.

initialization or *a priori* information under road environment. The results of lane recognition provide the road structure and limit the region of obstacles. In terms of vehicle detection, we use Gabor features to represent and detect vehicles in ROIs [21]. The software interface of multiple-sensor fusion is shown in Fig. 10 on our Springrobot platform (see Fig. 11).

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed an interactive road situation analysis framework and its algorithms, especially for multiple-sensor multiple-obstacle detection and tracking. We emphasize the future-situation evaluation rather than the current-obstacle situation. Vehicle dynamics and driving action behavior are considered to be two influencing factors for the I²DASW system. In addition, compared to other similar systems, our framework is a more integrated one, where a control module based on preview following is involved, which causes a concise and efficient framework.

There are still several questions in need of further investigation. For a special driving assistance application, how to select and setup the sensor network is important. Currently, we



Fig. 11. I²DASW platform-Springrobot.

manually calibrate the sensors in our system, respectively. Needless to say, a joint calibration of multiple sensors, including the camera, lidar, and radar, is desired to be automatically made in driver assistance and safety warning systems.

ACKNOWLEDGMENT

The authors would like to thank a large group of staff and students who participated in Springrobot intelligent vehicle development.

REFERENCES

- [1] N. Zheng, S. Tang, H. Cheng, Q. Li, G. Lai, and F. Y. Wang, "Toward intelligent driver-assistance and safety warning system," *IEEE Intell. Syst.*, vol. 19, no. 2, pp. 8–11, Mar./Apr. 2004.
- [2] P. J. Escamilla-Ambrosio and N. Lieven, "A multiple-sensor multiple-target tracking approach for the autotaxi system," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2004, pp. 601–606.
- [3] R. Mobus, "Multi-target multi-object tracking, sensor fusion of radar and infrared," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2004, pp. 732–737.
- [4] S. Park and M. M. Trivedi, "Driver activity analysis for intelligent vehicles: Issues and development framework," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2005, pp. 644–649.
- [5] L. Li, J. Y. Song, F. Y. Wang, W. Niehsen, and N. N. Zheng, "IVS05: New developments and research trends for intelligent vehicles," *IEEE Intell. Syst.*, vol. 20, no. 4, pp. 10–14, Jul./Aug. 2005.
- [6] M. Bertozzi, A. Broggi, A. Fascioli, T. Graf, and M. M. Meinecke, "Pedestrian detection for driver assistance using multi-resolution infrared vision," *IEEE Trans. Veh. Technol.*, vol. 53, no. 6, pp. 1666–1678, Nov. 2004.
- [7] N. Kawasaki and U. Kiencke, "Standard platform for sensor fusion on advanced driver assistance system using Bayesian network," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2004, pp. 250–255.
- [8] N. Srinivasa, Y. Chen, and C. Daniell, "A fusion system for real-time forward collision warning in automobiles," in *Proc. IEEE Intell. Transp. Syst. Conf.*, 2003, vol. 1, pp. 457–462.
- [9] H. Fritz, A. Gem, and H. Schiemenz, "CHAUFFEUR assistant: A driver assistance system for commercial vehicles based on fusion of advanced ACC and lane keeping," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 14–17, 2004, pp. 495–500.
- [10] G. L. Foresti and C. S. Regazzoni, "Multisensor data fusion for autonomous vehicle navigation in risky environments," *IEEE Trans. Veh. Technol.*, vol. 51, no. 5, pp. 812–817, Sep. 2002.
- [11] A. Broadhurst, S. Baker, and T. Kanade, "Monte Carlo road safety reasoning," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2005, pp. 319–324.
- [12] J. C. McCall, O. Achler, M. M. Trivedi, J. B. Haue, P. Fastrez, D. Forster, J. D. Hollan, and E. Boer, "A collaborative approach for human-centered driver assistance systems," in *Proc. IEEE Intell. Transp. Syst. Conf.*, Oct. 3–6, 2004, pp. 663–667.
- [13] K. C. Fuestenberg, "A new European approach for intersection safety—the EC-Project INTERSAFE," in *Proc. IEEE Intell. Transp. Syst. Conf.*, Sep. 2005, pp. 343–347.
- [14] N. Shimonura, K. Fujimoto, T. Oki, and H. Muro, "An algorithm for distinguishing the types of objects on the road using laser radar and vision," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 3, pp. 189–195, Sep. 2002.
- [15] N. Kaempchen, M. Buehler, and K. Dietmayer, "Feature-level fusion for free-form object tracking using laserscanner and video," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2005, pp. 453–458.
- [16] B. Alefs, D. Schreiber, and M. Clabian, "Hypothesis based vehicle detection for increased simplicity in multi sensor ACC," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2005, pp. 261–266.
- [17] J. Kolodko and L. Vlacic, "Fusion of range and vision for real-time motion estimation," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2004, pp. 256–261.
- [18] Y. J. Fang, I. Masaki, and B. Horn, "Depth-based target segmentation for intelligent vehicles: Fusion of radar and binocular stereo," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 3, pp. 196–202, Sep. 2002.
- [19] Q. Li, N. N. Zheng, and H. Cheng, "Springrobot: A prototype autonomous vehicle and its algorithm for lane detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 4, pp. 300–308, Dec. 2004.
- [20] C. Urmson, "Navigation regimes for off-road autonomy," Robotics Institute, Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-RI-TR-05-23, May 2005.
- [21] H. Cheng, N. N. Zheng, C. Sun, and H. van de Wetering, "Vanishing point and Gabor feature based multi-resolution on-road vehicle detection," in *Proc. Int. Symp. Neural Netw.*, 2006, pp. 46–51.
- [22] —, "Boosted Gabor features applied to vehicle detection," in *Proc. Int. Conf. Pattern Recog.*, 2006, pp. 662–666.
- [23] J. Yang, N. Zheng, X. Lv, and H. Cheng, "Fuzzy self-adjusted and real time learning algorithm for vehicle lateral control," in *Proc. IEEE Conf. Ind. Electron. and Appl.*, 2006, pp. 164–168.
- [24] Z. H. Gao, N. N. Zheng, H. Guan, and K. H. Guo, "Application of driver direction control model in intelligent vehicle's decision and control algorithm," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 17–21, 2002, vol. 2, pp. 413–418.
- [25] A. Mendes, L. C. Bento, and U. Nunes, "Multi-target detection and tracking with a laserscanner," in *Proc. IEEE Int. Intell. Veh. Symp.*, Parma, Italy, Jun. 14–17, 2004, pp. 796–801.
- [26] D. Baehring, S. Simon, W. Niehsen, and C. Stiller, "Detection of close cut-in and overtaking vehicles for driver assistance based on planar parallax," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2005, pp. 261–266.
- [27] J. J. Qin, "Study on obstacle detection and tracking based on multi-sensor fusion," M.S. thesis, Inst. Artificial Intell. Robotics, Xi'an Jiaotong Univ., Xi'an, China, 2005.
- [28] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: Wiley, 2001.



Hong Cheng (M'06) received the Ph.D. degree from Xi'an Jiaotong University, Xi'an, China, in 2003.

Since 2000, he has been a Teaching Assistant with the Department of Automation, Xi'an Jiaotong University. He is currently an Associate Professor with the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University. His research interests include intelligent vehicle, machine vision, pattern recognition, Intelligent Driver Assistance and Safe Warning systems, and human-computer interaction.

Dr. Cheng is a member of the IEEE Intelligent Transportation Systems Society and the IEEE Computational Intelligence Society. He is a Reviewer for the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS; the *IEE Proceedings of Vision, Image and Signal Processing*; the *Journal of Electronics* (in Chinese); *Intelligent Automation and Soft Computing*; the 2005 IEEE International Intelligent Symposium; and the 8th IEEE International Conference on Intelligent Transportation Systems. He served as Registration Chair of the 2005 IEEE International Conference on Vehicular Electronics and Safety and as Finance Secretary of the 2002 International Symposium on Nonlinear Theory and Its Applications.



Nanning Zheng (SM'93–F'06) received the B.S. and M.E. degrees in information and control engineering from the Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, in 1975 and 1981, respectively, and the Ph.D. degree in electrical engineering from Keio University, Tokyo, Japan, in 1985.

He is currently a Professor and the Director of the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University. His research interests include computer vision, pattern recognition, com-

putational intelligence, image processing, and hardware implementation of intelligent systems.

Prof. Zheng became a member of the Chinese Academy Engineering in 1999. He served as the General Chair for the International Symposium on Information Theory and Its Applications in 2002 and the General Cochair for the International Symposium on Nonlinear Theory and Its Applications in 2002. Since 2000, he has been the Chinese representative on the Governing Board of the International Association for Pattern Recognition. He currently serves as Executive Editor of the *Chinese Science Bulletin*.



Junjie Qin received the B.E. degree in information and communication engineering from Xi'an Jiaotong University, Xi'an, China, in 2002 and the M.E. degree from the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University.

He has been with Realsil Microelectronics Incorporation, Suzhou, China, since 2005. His research interests include intelligent vehicles, multisensor fusion, and image processing.



Xuetao Zhang received the B.E. degree in information and communication engineering and the M.E. degree in pattern recognition and intelligent systems from Xi'an Jiaotong University, Xi'an, China, in 2003 and 2006, respectively. He is currently working toward the Ph.D. degree at the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University.

His research interests include computer vision, pattern recognition, and intelligent vehicles.



Huub van de Wetering studied mathematics at the Technische Universiteit Eindhoven, Eindhoven, The Netherlands.

He is currently an Assistant Professor with the Department of Mathematics and Computer Science, Technische Universiteit Eindhoven. His research interests include computer graphics, implicit surfaces, and information visualization.