

© 2006 IEEE. Reprinted, with permission, from Sarath Kodagoda, CuTE: curb tracking and estimation . Control Systems Technology, IEEE Transactions on (Volume:14 , Issue: 5 ), September 2006. This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Technology, Sydney's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org). By choosing to view this document, you agree to all provisions of the copyright laws protecting it

# CuTE: Curb Tracking and Estimation

K. R. S. Kodagoda, W. S. Wijesoma, *Member, IEEE*, and A. P. Balasuriya, *Member, IEEE*

**Abstract**— The number of road accident related fatalities and damages are reduced substantially by improving road infrastructure and enacting and imposing laws. Further reduction is possible through embedding intelligence on to the vehicles for safe decision making. Road boundary information plays a major role in developing such intelligent vehicles. A prominent feature of roads in urban, semi-urban and similar environments is curbs on either side defining the road's boundary. In this paper a novel methodology of tracking curbs is proposed. The problem of tracking a curb from a moving vehicle is formulated as tracking of a maneuvering target in clutter from a mobile platform using on board sensors. A curb segment is presumed to be the maneuvering target, and is modeled as a *nonlinear Markov switching process*. The target's (curb's) orientation and location measurements are simultaneously obtained using a 2D scanning LADAR and a CCD monocular camera and are modeled as traditional base state observations. Camera images are also used to estimate the target's mode which is modeled as a *discrete-time point process*. An effective curb tracking algorithm, known as CuTE (Curb Tracking and Estimation) using multiple modal sensor information is thus synthesized in an image enhanced interactive multiple model filtering framework. The use and fusion of camera vision and LADAR within this frame provide for efficient, effective and robust tracking of curbs. Extensive experiments conducted in a campus road network demonstrate the viability, effectiveness and robustness of the proposed method.

**Index Terms**— laser radar, multi sensor systems, road transportation, robot sensing systems, robot vision system

## I. INTRODUCTION

IN the literature, several methodologies based on single exteroceptive sensors including camera [1]-[2], millimeter wave radar (MMWR) [3]-[4], and laser-radar (LADAR) [5]-[6] have been applied to the problem of road boundary tracking. However, using a single mode of sensing for road boundary tracking has its limitations [16]. In the recent years we have witnessed a growing interest in the use of multiple sensors applied to this problem. An important contribution is due to [7]. The methodology is based on active structured lighting techniques used to construct 3D object geometry,

which is very popular in computer vision. The sensors used are a laser line stripper and a camera. The main drawbacks of the methodology are the limited range of operation (up to 3m from the vehicle), and its susceptibility to bright sunlight.

In this paper an effective, reliable and robust methodology is proposed for curb tracking using LADAR and vision sensing. The approach is based on tracking a maneuvering target in clutter using sensors on a moving platform. A multiple model approach, as implemented by the Interacting Multiple Model (IMM) method, provides one of the most effective frameworks for tracking maneuvering targets [8]. However, with the complexity of the maneuver and presence of clutter, IMM can fail mainly due to the significant time lag between the time of the target's actual change of maneuver and the time that maneuver is detected unambiguously, which causes inadequate timely adjustments to the tracking filter. To minimize the detection time of target mode (maneuver) an image based approach is proposed in [9]. The target's mode is inferred by an image processing system based on orientation information provided by a 2D imaging sensor. The output of the image processor is modeled as a sequence of classifications of the target into one of the known modes of operation. The discrete-time counterpart of [9] is developed in [10]. In [10] history of modal observations is used to optimally estimate the target's modes using point process filtering theory [11]. Further enhancements are made in [12] and [13] by fusing those estimated target modes (obtained using the image observations) with traditional point mass state observations obtained using radar measurements. The main idea is to use conditional mode probabilities calculated in [9] and [10] to select the most appropriate model to be used in the extended Kalman filter (EKF) for target state estimation. These methods do not utilize all the information available, and rely on *ad hoc* methods of incorporating uncertainty about the mode estimates. In [14] a filter framework known as Image Enhanced Interactive Multiple Model Filter (IEIMM) is proposed which incorporates all the information from point mass sensors and image sensors in a unified manner to more accurately estimate target state and its mode simultaneously. This method has provided the inspiration and basis of the proposed CuTE algorithm for curb tracking using multiple sensors. In this paper, the curb tracking problem is first formulated as tracking a maneuvering target and is solved by adapting IEIMM [14] framework. The key differences with [14] are the utilization of a moving observer, nonlinear measurement and process models and extensive experimental evaluations. Target or curb measurements are made using a monocular CCD camera and a 2D scanning LADAR. The use and fusion of CCD vision and LADAR in an IEIMM

Manuscript received January 5, 2005. K. R. S. Kodagoda is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798. e-mail: sarath@pmail.ntu.edu.sg.

W. S. Wijesoma is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798. phone: (065) 6790 4518; fax: (065) 6790 .....; e-mail: eswwijesoma@ntu.edu.sg.

A. P. Balasuriya is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798. e-mail: arjunab@mit.edu.

framework provides for efficient, effective and robust tracking of curbs over a longer operating region ahead of the vehicle.

In Section II curb tracking problem is formulated and solved within an IEIMM framework. Section III describes the use of LADAR and vision sensing for extracting road features. Results of extensive experiments carried out in a campus site are presented in Section IV followed by the concluding remarks given in Section V.

## II. CURB TRACKING AND ESTIMATION (CUTE)

### A. Problem Formulation

The targets, in our case are the line segments corresponding to vertical curb surfaces as observed by the 2D scanning LADAR and the CCD camera. When the vehicle is in motion, the line segments or targets move along the curbs (left/right). Thus on a straight segment of a road the targets (curbs) can be considered in a non-maneuvering state, whilst at a bend in a maneuvering state.

### B. Hybrid Vehicle and Target Model

Suppose  $\{m_i(k+1)\}; k \geq 0$  represents the target (curb) mode at time  $k+1$  and assume that  $m_i$  evolves as a homogeneous, discrete-time Markov process in the state space  $\{1, \dots, d\}$  with transition probability matrix  $\mathbf{T}_{ij} = P(m_i(k+1) = i | m_i(k) = j)$  with initial conditions,  $P(m_i(0) = i) = \pi_0(i)$ . The composite nonlinear vehicle (with state  $\mathbf{s}^v(k)$ ) and target dynamics (with state  $\mathbf{s}^t(k)$ ) can be represented as,

$$\begin{aligned} \mathbf{s}(k+1) &= [\mathbf{s}^v(k+1)^T \quad \mathbf{s}^t(k+1)^T]^T \\ &= \mathbf{f}[\mathbf{s}(k), m_i(k+1)] + \mathbf{B}(m_i(k+1))\mathbf{v}(k+1) \end{aligned} \quad (1)$$

where,  $\mathbf{s}(k), \mathbf{v}(k)$  are  $n$  dimensional vectors.  $\mathbf{s}(k)$  is assumed to be a Gaussian random vector with covariance matrix  $\mathbf{C}(0)$ .  $\mathbf{f}[\cdot]$  is the nonlinear process model.  $\mathbf{B}(m_i(k+1))$  is a mode dependent matrix. Noise term,  $\mathbf{v}(k)$  is a sequence of independent zero mean Gaussian random vectors with covariance matrix,  $\mathbf{Q}$ . The noise,  $\mathbf{v}(k)$  and  $\mathbf{s}(0)$  are assumed to be uncorrelated.

Vehicle process model  $\mathbf{s}^v = [x_w^v \quad y_w^v \quad \phi_w^v]^T$  can be derived using Fig. 1 as,

$$\mathbf{s}^v(k+1) = \begin{bmatrix} x_w^v(k) + \Delta TV(k) \cos \phi_w^v(k) \\ y_w^v(k) + \Delta TV(k) \sin \phi_w^v(k) \\ \phi_w^v(k) + \frac{\Delta TV(k) \tan \gamma(k)}{L} \end{bmatrix} + \mathbf{v}^v(k+1) \quad (2)$$

where,  $\{x_w^v, y_w^v\}$  are the coordinates of the center of the rear axel of the vehicle and  $\phi_w^v$  is the orientation of the vehicle axis with respect to a world coordinate system,  $(X_w, Y_w)$  in Fig. 1.  $V, \gamma, \Delta T$  and  $L$  are speed, steer angle, sampling time and wheel base length respectively.  $\mathbf{v}^v(k)$  is the noise related to the vehicle model.

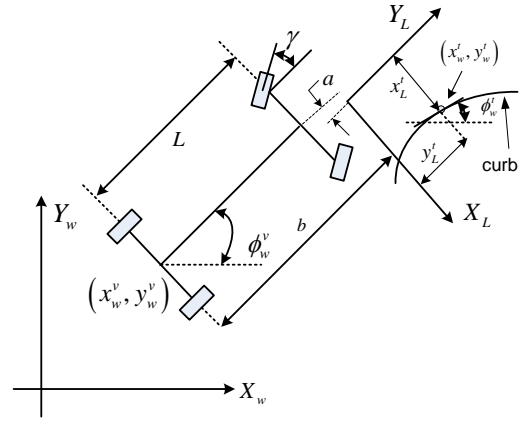


Fig. 1. Vehicle kinematics

The road boundary or curb is assumed locally straight and hence approximated by a straight line segment and represented by its mid point  $(x_w^t, y_w^t)$  and orientation  $(\phi_w^t)$ . Therefore, the target (curb) model corresponding to a straight curb scenario, i.e the non-maneuvering mode of target, can be assumed to be the usual constant speed model [15] with  $\mathbf{s}^t = [x_w^t \quad \dot{x}_w^t \quad y_w^t \quad \dot{y}_w^t \quad \phi_w^t]^T$ .

$$\mathbf{s}^t(k+1) = \begin{bmatrix} x_w^t(k) + \Delta T \dot{x}_w^t(k) \\ \dot{x}_w^t(k) \\ y_w^t(k) + \Delta T \dot{y}_w^t(k) \\ \dot{y}_w^t(k) \\ \phi_w^t(k) \end{bmatrix} + \begin{bmatrix} 0.5\Delta T^2 & 0 \\ \Delta T & 0 \\ 0 & 0.5\Delta T^2 \\ 0 & \Delta T \\ 0 & 0 \end{bmatrix} \mathbf{v}^t(k+1) \quad (3)$$

where,  $(x_w^t, y_w^t)$  are the position coordinates of the center of the target (mid point of the line segment) and  $\phi_w^t$  is the orientation of the target (line segment) with respect to the world coordinate system.  $\mathbf{v}^t(k)$  is the noise related to the target model.

The usual turn rate model [15] is used for left-bend and right-bend curb scenarios, which are equivalent to the distinct modes or states of the maneuvering target.

$$\mathbf{s}^t(k+1) = \begin{bmatrix} x_w^t(k) + \frac{\sin \omega \Delta T}{\omega} \dot{x}_w^t(k) - \frac{(1 - \cos \omega \Delta T)}{\omega} \dot{y}_w^t(k) \\ \cos(\omega \Delta T) \dot{x}_w^t(k) - \sin(\omega \Delta T) \dot{y}_w^t(k) \\ y_w^t(k) + \frac{\sin \omega \Delta T}{\omega} \dot{y}_w^t(k) + \frac{(1 - \cos \omega \Delta T)}{\omega} \dot{x}_w^t(k) \\ \cos(\omega \Delta T) \dot{y}_w^t(k) + \sin(\omega \Delta T) \dot{x}_w^t(k) \\ \omega \Delta T + \phi_w^t(k) \end{bmatrix} + \begin{bmatrix} 0.5\Delta T^2 & 0 \\ \Delta T & 0 \\ 0 & 0.5\Delta T^2 \\ 0 & \Delta T \\ 0 & 0 \end{bmatrix} \mathbf{v}^t(k+1) \quad (4)$$

where,  $\omega$  is the turn rate of the target.

### C. Hybrid Vehicle and Target Sensor Model

The composite moving observation model for the vehicle and sensor hybrid is,

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{s}(k)] + \mathbf{w}(k) \quad (5)$$

where,  $\mathbf{z}(k), \mathbf{w}(k)$  are  $m$  dimensional vectors.  $\mathbf{w}(k)$  is a sequence of zero mean Gaussian random vectors with covariance matrix  $\mathbf{R}$ . The process  $\mathbf{v}(k)$  is uncorrelated with  $\mathbf{w}(k)$  and  $\mathbf{s}(0)$ . The observation model can be derived from Fig. 1 as,

$$\begin{aligned} & \begin{bmatrix} x_w(k) & y_w(k) & \phi_w(k) & x_L'(k) & y_L'(k) & \phi_L'(k) & x_C'(k) & y_C'(k) & \phi_C'(k) \end{bmatrix}^T \\ & = \begin{bmatrix} x_w^v(k) \\ y_w^v(k) \\ \phi_w^v(k) \\ (x_w'(k) - x_w^v(k))\cos\phi_w^v(k) + (y_w'(k) - y_w^v(k))\sin\phi_w^v(k) - a \\ -(x_w'(k) - x_w^v(k))\sin\phi_w^v(k) + (y_w'(k) - y_w^v(k))\cos\phi_w^v(k) - b \\ \pi/2 + \phi_w^v(k) - \phi_w^v(k) \\ (x_w'(k) - x_w^v(k))\cos\phi_w^v(k) + (y_w'(k) - y_w^v(k))\sin\phi_w^v(k) \\ -(x_w'(k) - x_w^v(k))\sin\phi_w^v(k) + (y_w'(k) - y_w^v(k))\cos\phi_w^v(k) \\ \pi/2 + \phi_w^v(k) - \phi_w^v(k) \end{bmatrix} + \mathbf{w}(k) \end{aligned} \quad (6)$$

$(x_w, y_w, \phi_w)$  are the measurements of vehicle position using Differential Global Positioning System (DGPS) along  $x$ -axis,  $y$ -axis and vehicle orientation measured using gyroscope with respect to the world coordinate system respectively.  $(x_L', y_L', \phi_L')$  are the curb data extracted by the laser scanner [5] in laser coordinate system.  $(x_C', y_C', \phi_C')$  denotes the line segments describing the road boundaries in vehicle coordinate frame as determined by the vision system [16]. The error covariance matrix  $\mathbf{R}$  can be determined by the covariance matrices of DGPS, gyroscope data, pose parameters through LADAR and vision sensing.

#### D. Image Sensor Model for Mode Estimation

The mode of target (curb) as detected by the CCD camera (section III) is denoted by  $n(k) \in \{0, 1, 2, \dots, d\}$ : where  $d$  is the number of modes the target could possibly exist in and  $n(k) = 0$ , represents no useful information. Lets introduce an indicator vector  $\rho(k+1)$  describing the binary mode observations of image sensor output,  $n(k+1)$ , with the  $i^{\text{th}}$  element defined as,

$$\rho_i(k+1) = \begin{cases} 1 & n_i(k+1) = i \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Similarly  $\varphi_i(k+1)$ , the  $i^{\text{th}}$  element of the indicator vector of the target mode,  $m_i(k+1) \in \{1, \dots, d\}$  is defined. Due to fast sampling, slow processing or complicated scene interpretations, image sensor may not yield an output at all sampling times. This is handled using a parameter  $\delta$  as follows;

$$\delta(k+1) = \begin{cases} 1 & \text{an image output is generated} \\ 0 & \text{no image output is generated} \end{cases} \quad (8)$$

Probability that an image output is generated given the target is in mode  $i$  is given by

$\lambda_i = P\{\delta(k+1) = 1 | \varphi_i(k+1) = 1\} = E\{\delta(k+1) | \varphi_i(k+1) = 1\}$ . The discernibility matrix  $\mathbf{D}^v = [d_{ij}]$  is defined as the probability of mode  $j$  is reported given the actual mode is  $i$  at the reporting time, where  $d_{ij} = P\{\rho_j(k+1) = 1 | \delta(k+1) = 1, \varphi_i(k+1) = 1\}$  with  $\sum_{j=1}^d d_{ij} = 1$ . This discernibility matrix provides a measure of the quality of the sensor, for example, large diagonal elements represent a sensor, which can distinguish different modes unambiguously. A rate matrix,  $\Lambda^v$ , is defined as  $\Lambda^v = (\mathbf{D}^v)^T \text{diag}(\lambda)$  [10]. The notation  $\text{diag}(\lambda)$  is used to denote the matrix whose diagonal elements are the components of the vector  $\lambda$ . It is assumed that  $n(k+1)$  is conditionally independent of all other random variables. Given  $\delta(k+1) = 1$  and  $\varphi_i(k+1) = 1$ ,  $\Lambda^v$  fully describes the image-based observation process.

#### E. Image Enhanced Curb Tracking

Tracking of line segments corresponding to the curbs is non trivial due the significant orientation changes of the curbs and their detection in the presence of clutter. This requires mode adaptive estimation techniques. One of the ways is to use the Interacting Multiple Model (IMM) [15] technique. If mode observations can be obtained, an image-enhanced IMM algorithm (IEIMM) can provide better tracking performance [14]. The IEIMM in [14] utilizes a linear process model and a static observer. As shown in section II, the process model (1) and the observation model (5) are nonlinear and the observations are made from a mobile vehicle. Thus IEIMM in [14] is suitably adapted by utilizing an extended Kalman filter (EKF) with a modified observation model (6) to accommodate for the nonlinearity and effects of moving observer. Therefore, the optimal vehicle and target hybrid state estimate,  $\mathbf{s}_{k+1}$ , given the image based mode observations up to and including time  $k+1$ , i.e.  $\mathbf{N}_{k+1} = \{n(1), \dots, n(k+1)\}$  and the base state laser/vision measurements up to and including time  $k+1$ , i.e.  $\mathbf{Z}_{k+1} = \{z(1), \dots, z(k+1)\}$  can be derived as [14],

$$\begin{aligned} & p(\mathbf{s}_{k+1}, m_i(k+1) = i | \mathbf{N}_{k+1}, \mathbf{Z}_{k+1}) \\ & = \beta_{k+1}^{-1} \left[ (1 - \lambda_i) \mathbf{p}(n_{k+1} = 0) + \lambda_i \sum_{l=1}^d d_{il} \mathbf{p}(n_{k+1} = l) \right] \phi^m(\mathbf{z}_{k+1} - \mathbf{h}[\mathbf{s}_{k+1}], \mathbf{D}_i \mathbf{R} \mathbf{D}_i^T) \\ & \quad \times p(\mathbf{s}_{k+1}, m_i(k+1) = i | \mathbf{N}_k, \mathbf{Z}_k) \end{aligned} \quad (9)$$

where,  $\beta_{k+1} = p(n_{k+1}, z_{k+1} | \mathbf{N}_k, \mathbf{Z}_k)$  is a normalization constant.

The optimal filter (9) is computationally prohibitive to implement. Therefore, we utilize an EKF based implementation. Fig. 2 shows the detailed functional block diagram of the Curb Tracking and Estimation (CuTE) algorithm. Firstly the laser data is processed to extract the road boundaries (Section III). Secondly, the image sequences are analyzed to extract the road boundaries using an Unscented Kalman Filter (UKF). Then the images are further processed to determine the type of road or target mode by analyzing the road curvature (Section III). To minimize the effects of noisy images on mode estimation, a Discrete Time Point Process

Filter (DTPPF) [10], which uses mode observation history, is utilized. Now the extracted image and laser based observations together with encoder, DGPS and gyroscope data are used in a bank of Kalman filters for target and mode estimation as shown in Fig. 2.

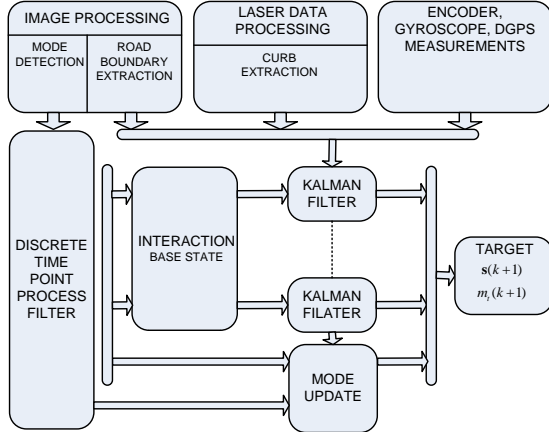


Fig. 2. Block diagram of the CuTE algorithm

### III. FEATURE EXTRACTION AND MODE ESTIMATION

The 2D scanning LADAR is mounted on the front of the vehicle with a tilt angle of  $\alpha_L$  to the flat horizontal road surface [5]. When the laser sweeps and intercepts a surface it will generate data point sets which are co-linear. A UKF is used for filtering, segmentation and line parameter estimation of these collinear point sets corresponding to the road surface, curb surface and pavements [5]. The line parameters estimated are used to represent the curbs described in section II. Painted lane markings are common on most roads and are highly correlated with curbs as they are mostly parallel to the curbs. The lane markings imaged from a CCD camera are extracted and transformed onto the vehicle coordinates [16]. These estimates together with a priori information of road widths are used for road boundary estimation.

Target maneuver, and hence its mode of operation, is slowly reflected in position measurements. In the case of curbs this corresponds to the observations from the LADAR (possibly complemented by vision) as described above. However, target (curb) maneuver can be quickly detected by observing the sudden variations of overall curb orientation in a single CCD camera image. Many modes of operation of the target may be observed depending on the complexity of the road. Here, only three frequent modes of operation of target or curb are considered; straight (*mode 1*), left-bend (*mode 2*) and right-bend (*mode 3*). These modes are determined by analyzing the curvature of road. The predicted curb positions in Cartesian coordinates of the vehicle frame are transformed on to the image plane of camera to aid in defining Regions of Interest (ROIs) for the fast detection and extraction of possible lane markings. These lane markings extracted are analyzed for curvature. Obviously these curvature based mode detections are noisy. Hence a “quality measure” defined as the number of pixels with good curvature information to total number of pixels in a given ROI is utilized to quantify quality of image observations. Curvature estimates resulting from poor quality

measures are discarded and yields no output from the image processing system. However, the outputs of the image processing system may still produce inconsistent outputs. In order to improve on the quality of the mode output estimates of the image processing system, they are further filtered using a discrete time point process filter (DTPPF) [10]. The DTPPF uses all the mode outputs of image processing system up to time  $k+1$  to estimate the current mode.

### IV. EXPERIMENTAL RESULTS

Using a test bed vehicle known as GenOME [5], extensive experiments were carried out on roads in a campus environment over a stretch of 4km. A front mounted SICK© 2D scanning LADAR and a monocular CCD camera were used to obtain curb and lane marking measurements. A high accuracy Hitachi© Fiber Optic Gyroscope with a bias stability of 0.005deg/s was used for vehicle orientation measurements. However for greater accuracy, its bias was estimated off-line, and used to compensate the online measurements. Further, the orientation was estimated by fusing the Gyro measurements with the steering angle measurements obtained through a steering encoder and vehicle kinematics using a Kalman filter. The speed of the vehicle was estimated using odometers mounted on the wheels of the vehicle. All the sensor data acquisition including camera, LADAR, encoders and gyroscope were synchronized to 100 ms sampling time.

Target maneuvers or in this case curb mode (also the road type) were classified into three modes empirically based on the curvature,  $\kappa$ :  $-0.01 < \kappa < 0.01$ : straight (*mode 1*),  $\kappa > 0.01$ : left-bend (*mode 2*),  $\kappa < -0.01$ : right-bend (*mode 3*). In each case,  $\lambda_i$  (probability that a valid output is generated given the mode “ $i$ ”), associated with  $i^{\text{th}}$  mode is chosen to be equal to the quality measure of  $\kappa$  at that instant. For the multi-dimensional DTPPF, the transition probability matrix  $\mathbf{T}$  and the discernibility matrix  $\mathbf{D}^v$  were empirically chosen as follows;

$$\mathbf{T} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.85 & 0.05 \\ 0.1 & 0.05 & 0.85 \end{pmatrix} \text{ and } \mathbf{D}^v = \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.2 & 0.79 & 0.01 \\ 0.2 & 0.01 & 0.79 \end{pmatrix}$$

Since the sampling rate is high, the likelihood of the target being in the same state in the next sampling instant is high and hence larger values of the diagonal terms of  $\mathbf{T}$ . Transition probabilities from *mode 1* (straight curb ahead) to 2 (left-bend ahead) and 3 (right-bend ahead) and vice versa have been chosen to be equal and smaller than the diagonal elements. Further, transition from *mode 2* (left-bend ahead) to 3 (right-bend ahead) and 3 to 2 have been chosen to be even smaller to reflect the fact that these scenarios are less likely. The values in  $\mathbf{D}^v$  suggest that the *modes 2* and *3* are relatively easier to distinguish from each other. Extensive experiments were conducted over a 4km run with the test-bed vehicle driven at an approximate speed of 4m/s. The track covers all four possible observation scenarios: (i) simultaneous availability of image and LADAR observations, (ii) unavailability of LADAR observations, (iii) unavailability of vision based observations, and (iv) intermittent unavailability of both image and LADAR observations. Fig. 3 shows the tracking performance for the

entire 4km run. It may be noted that the tracked curbs corresponds very closely to the road boundaries as shown on the map. Following discusses in detail the performance of the algorithm in different sections of the road network under different observation scenarios.

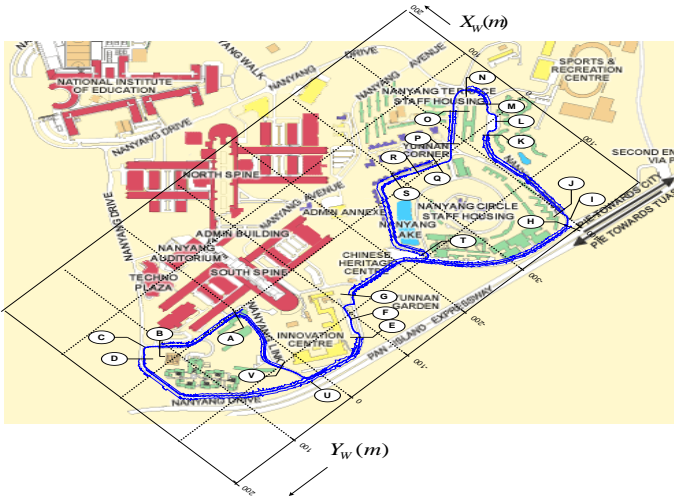


Fig. 3 Curb tracking results in a campus environment: a 4 km drive (dotted line – vehicle path, solid lines – curbs

Both image and LADAR based observations are available in most of the road scenarios. For example, the curbs extracted on a straight section (F-G in Fig. 4) and a left bend (H-I) based on LADAR [5] and vision [16] are shown in Fig. 6(a) and 7(a) respectively. It is seen that road feature extraction is quite robust to the interfering white strips, letters and complex shadows. The mode output of the image processing system and the output of DTPPF for 100 consecutive frames of measurements on the straight road and left turn are shown in Fig 6(b) and 7(b) respectively. Although the mode outputs (vertical lines) of the image processor are noisy, the DTPPF correctly classifies the mode of curb in each case.

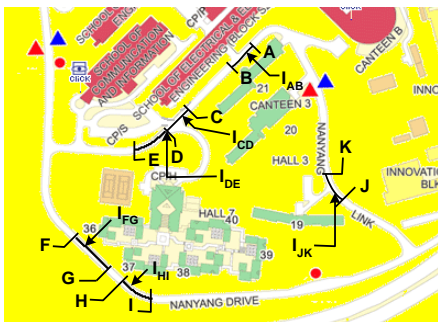


Fig. 4 Part of the map of the experimental test site

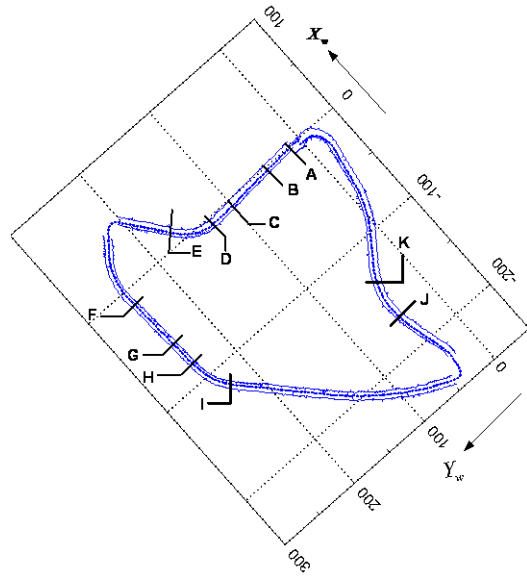
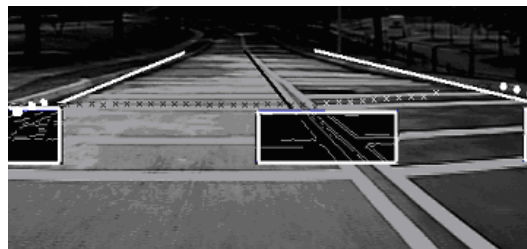
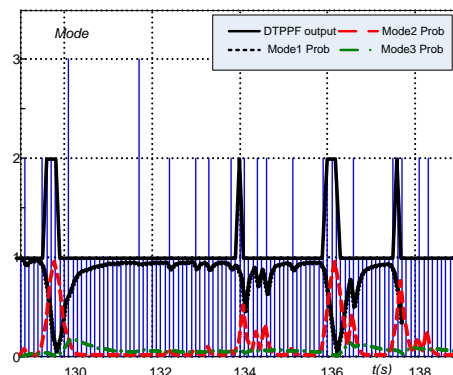


Fig. 5. Curb tracking results (axes are in meters): dashed line – vehicle path, solid line – curbs

At or near y-intersections (A-B in Fig. 4) and x-intersections (C-D in Fig. 4), it is noted that only vision information is available. For example, Fig. 8 (a) shows an x-intersection (C-D in Fig. 4). Although there is visual information, there are no returns from the LADAR due to the absence of curbs on either side. However, the painted lane markings extracted from camera images still provides information of the road boundaries. The output modes as determined by the image processing system and its filtered output via the DTPPF when driving in the section C-D in Fig. 4 are as shown in Fig. 8 (b). It can be seen that the DTPPF correctly estimates the mode of curb despite the fluctuations in the output of the image processor output.



(a) lines – vision based method, dots – laser based method, crosses – raw laser data

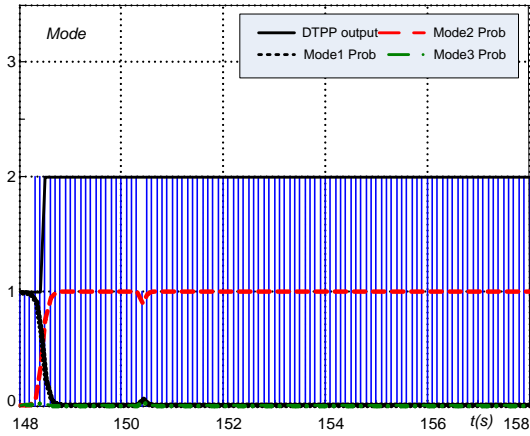


(b) Mode and point-process filter outputs

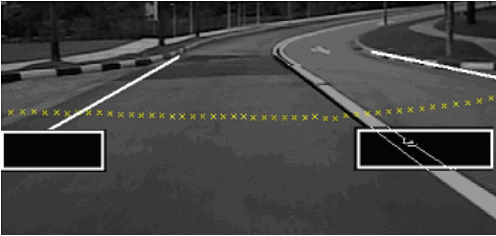
Fig. 6. Road boundary detection in a straight road



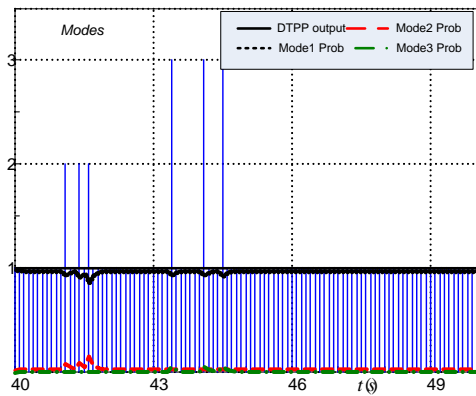
(a) *lines* – vision based method, *dots* – laser based method, *crosses* – raw laser data



(b) Mode and point-process filter outputs  
Fig. 7. Road boundary detection in a left turn



(a) *lines* – vision based method, *crosses* – raw laser data



(b) Mode and point-process filter outputs  
Fig. 8. Road boundary detection in an X-intersection

Although, CuTE is effective in most situations, its performance can be compromised by the failure of the image processing system. This is mainly due to noisy image observations resulting from bright or poor lighting, occlusions or complex road markings. In these circumstances the quality measure for estimation of road curvature is below the threshold and thus image observations are discarded in mode

estimation. In these cases the IEIMM filter collapses to an IMM as there is no additional information from vision sensing. The occurrences of these image failure modes at certain road sections (eg. J-K in Fig. 4) are shown in Fig. 9. Despite the absence of vision in these circumstances, LADAR can still be effective in providing information for curb extraction and tracking in an IMM framework.



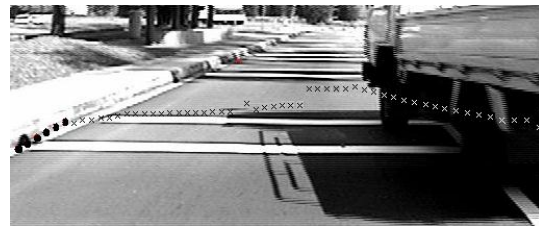
(a) Over illumination (sun's glare)



(b) Ill illumination (rainy)



(c) Road with zebra crossings



(d) Overtaking vehicles

Fig. 9 Failure modes of image based road boundary detection: *dots* – laser based detected curbs, *crosses* – raw laser data

Another challenging scenario is the intermittent unavailability of both image and LADAR measurements. Two such examples are shown in Fig. 10. In Fig. 10 (a), views of the laser and camera are partially blocked by another vehicle, whereas Fig. 10 (b) corresponds to a situation where the front wheels of the vehicle are traveling on top of a road hump causing the LADAR and camera to tilt up. In latter case, the laser beams do not intersect with curbs and hence no laser based curb observations. Image observations obtained are noisy and cannot be utilized. Since these instances are intermittent and the fact that image based road boundaries carry much more information ahead of the vehicle, the CuTE algorithm, however, is still capable of tracking curbs. In the case of no observations for a long time, the CuTE algorithm keeps

predicting curbs using the most probable target model without updating until measurements are available. The quality of estimates depends on correctness of the target model and how quickly valid observations are available.



(a) Road view blocked by front vehicle



(b) Traveling on top of a road hump

Fig. 10. Failure modes of image/ laser road boundary detection : crosses – raw laser data

## V. CONCLUSIONS

Tracking of road curbs using on board 2D LADAR and camera can be formulated as tracking of maneuvering targets in clutter. The curb is perceived and modeled as a maneuvering target. The curb's maneuvers correspond to different road scenarios or modes such as straight, left-bend, right-bend, etc. The maneuvering target or curb and the vehicle composite are thus modeled as a nonlinear Markov switching system. The principal difficulty of tracking a maneuvering target is due to the significant time lag between the instant a target's actual change of maneuver occurs and the instant that change of maneuver is detected unambiguously. In this particular curb tracking application, curvature of curbs is used as an indicator of type of maneuver (mode) and is determined indirectly through analysis of camera images of lane markings using an image processing system. The mode output of the image processing system is susceptible to poor/bright ambient illumination, complex shadows, and occlusions. Those effects are minimized using a discrete-time point process filter. Extensive experiments carried out on an actual 4 km run on a campus road environment demonstrated the effectiveness and robustness of the CuTE algorithm for curb tracking. The performance was evaluated for different road scenarios under different conditions including, presence of both image and LADAR based observations, image only observations, LADAR only observations and intermittent unavailability of both image and LADAR observations. Overall, it was shown that the CuTE algorithm performs well in all cases except when the LADAR and image observations

are unavailable over a long period. Although, only three commonly occurring types of curb and road scenarios were considered, viz. straight curb ahead, left-bend ahead, right-bend ahead, it is straightforward to incorporate other cases within CuTE. This would, however, mean complex image and laser data processing for detecting the different scenarios.

## REFERENCES

- [1] D. Pomerleau, and T. Jochem, "Rapidly adapting machine vision for automated vehicle steering", *IEEE Expert*, Vol.11, Issue: 2, pp.19–27, April 1996
- [2] E. D. Dickmanns, and B. D. Mysliwetz, "Recursive 3D road and relative ego-state recognition", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol.14, No.2, pp.199-213, February 1992.
- [3] K. Kaliyaperumal, S. Lakshmanan, and K. Kluge, "An algorithm for detecting roads and obstacles in radar images", *IEEE Trans. Veh. Technol.*, Vol.50, No.1, pp.170-182, January 2001.
- [4] S. Lakshmanan, and D. Grimmer, "A deformable template approach to detecting straight edges in radar images", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol.18, No.4, pp.438-443, April 1996.
- [5] W.S. Wijesoma, K.R.S. Kodagoda, and A.P. Balasuriya, "Road-boundary detection and tracking using ladar sensing", *IEEE Trans. Robot. Automat.*, Vol.20, No.3, pp.456-464, June 2004.
- [6] A. Kirchner, and T. Heinrich, "Model based detection of road boundaries with a laser scanner", in *Proc. Int. Conf. Intelligent Vehicles*, Stuttgart, 1998, pp.93-98.
- [7] R. Aufrere, C. Mertz, and C. Thorpe, "Multiple sensor fusion for detecting location of curbs, walls, and barriers", in *Proc. IEEE Intelligent. Vehicles Symposium*, Columbus, 2003, pp. 126 – 131.
- [8] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Boston, Artech House, 1999, ch.4.
- [9] D. D. Sworder, and R. G. Hutchins, "Image-enhanced tracking", *IEEE Trans. Aerosp. Electron. Syst.*, Vol 25, No 2, pp701-709, September 1989.
- [10] C. Yang, Y. Bar-Shalom, and C. F. Liu, "Discrete-time point process filter for mode estimation", *IEEE Trans. Automat. Contr.*, Vol.37, No. 11, pp.1812-1816, November 1992.
- [11] A. Segall, "Recursive estimation from discrete-time point processes", *IEEE Trans. Inform. Theory*, Vol.22, No.4, pp.422-431, July 1976.
- [12] R. G. Hutchins, and D. D. Sworder, "Image fusion algorithms for tracking maneuvering targets", *Journal of Guidance, Control and Dynamics*, Vol.15, No.1, pp.175-184, 1992.
- [13] C. Yang, Y. Bar-Shalom, and C. F. Liu, "Maneuvering target tracking with image enhanced measurements", in *Proc. American Control Conference*, Boston, 1991, pp.2286-2291.
- [14] J. S. Evans, and R. J. Evans, "Image-enhanced multiple model tracking", *Automatica*, Vol.35, No.11, pp.1769-1786, November 1999.
- [15] Y. Bar-Shalom and X. R. Li, *Estimation and Tracking: Principles, Techniques, and Software*, Boston, Artech House, 1993, ch.11.
- [16] W. S. Wijesoma, K. R. S. Kodagoda, and A. P. Balasuriya, "A laser and a camera for mobile robot navigation", in *Proc. 7th Int. Conf. Control, Automation, Robotics and Vision*, Singapore, 2002. Vol 2, pp. 740 – 745.