LASER-CAMERA COMPOSITE SENSING FOR ROAD DETECTION AND TRACKING

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

School of Electrical and Electronic Engineering

Nanyang Technological University, Singapore 639798

e-mail: {eswwijesoma, ps6983882j, earjuna}@ntu.edu.sg

Abstract:

An important feature in most urban roads, and similar environments such as in theme parks, campus sites, industrial estates, science parks and the like is the existence of pavements or curbs on either side defining the road boundaries. These curbs, which are mostly parallel to the road, can be harnessed to extract useful features of the road for implementing autonomous navigation or driver assistance systems. However, vision alone methods for extraction of such curbs or road edge features with accurate depth information is a formidable task as the curb is not conspicuous in the vision image and also it requires the use of stereo images. Further, bad lighting, adverse weather conditions, non-linear lens aberrations, lens glare due to sun and other bright light sources can severely impair the road image quality and thus the operation of vision alone methods. In this paper an alternative and a novel approach involving the fusion of 2D laser range and monochrome vision image data is proposed to improve the robustness and reliability. Experimental results are presented to demonstrate the viability, and effectiveness, of the proposed methodology and its robustness to different road configurations and shadows.

Key Words—Intelligent sensors, robot vision systems, laser measurement systems, transportation

1. Introduction

The development of techniques for high performance autonomous navigation and control of vehicles in outdoor road environments has become an important and active research endeavor in the face of emerging markets for advanced Autonomously Guided Vehicles (AGVs), mobile robots, and smart vehicles. An intelligent system for the detection of road boundaries and painted lane markings are indispensable elements in the realization of such autonomous systems as well as in driver assistance systems. Some of the well known road detection methods can be cited as, RALPH (Rapidly Adapting Lateral Position Handler) [1], Dickmanns et al [2], GOLD (Generic Obstacle and Lane Detection system) [3], ALVINN (Autonomous Land Vehicle In a Neural Network) [4], ARCADE (Automated Road Curvature And Direction Estimation) [5], LOIS (Likelihood Of Image Shape) [6] and work by Aufrere et al [7].

Detection of pavement boundaries and range estimation are valuable for number of intelligent applications including autonomous navigation, drowsy driver warning and forward collision warning. Work by Lakshmanan et al [8], Kaliyaperumal et al [9], Ma et al [10] and Nikolova et al [11] are few instances where millimeter wave radar (MMWR) have been applied for road boundary detection. laser measurement systems (LMS) have also been used for road boundary detection using reflection posts [12], guardrails and posts [13]. To the best of our knowledge application of two-dimensional (2D) LMS and monocular camera for road curb detection and the manner in which, it is utilized here have not been reported in the literature.

An assembly involving a 2D LMS and a CCD camera and appropriate image, vision and laser range data processing algorithms are proposed in this paper as an

effective intelligent system for compact and real-time extraction of road features useful for autonomous navigation and driver assistance systems. In *Section 2*, the 2D ladar for road curb detection is described. The extracted curb data mapped on the image plane is used to estimate the midline (or lanes) of the road in *Section 3*. In both cases, i.e., the curb detection using laser and midline extraction using vision, are formulated as an Unscented Kalman filtering problem. In *Section 4*, we develop a Kalman filter framework for temporal tracking of extracted road curbs. Experimental results gathered from the composite sensor mounted on our in-house built test-bed (Figure 10), GenOME (Generic Outdoor Mobile Explorer) is used in *Section 5* to illustrate the effectiveness of the proposed algorithms. *Section 6* concludes the paper.

2. Road Boundary Extraction using a 2D LMS

In the past decade, interests in the use of LMSs in autonomous navigation systems have been on the increase [12-18]. LMSs can provide a low cost alternative to MMWR systems for depth/range measurement under poor lighting, visibility and bad weather conditions [14]. In autonomous navigation, LMSs have been used in obstacle detection [15], navigation [16], localization [17-18], and road boundary detection (using reflection posts [12], using guard rails and posts [13]). In our application, the LMS is used mainly to detect the road boundaries or edges as defined by the curbs. The use of laser measurement device to detect the curbs is more robust and reliable as compared to camera-based vision systems especially in bad weather and poor lighting conditions.

Curb extraction is achieved by a front mounted 2D ladar sensor looking down the road ahead (10-15m) with a small tilt, zero pan and swing angles (Figure 1). The LMS scans a laser spot beam from right to left through 180 degrees on a plane

inclined at an angle α_L (equal to the tilt angle) to the road surface. The idealized road model and sensing scenario is shown in Figure 2. For a given region (e.g. road surface (CD), curb surface (BC or DE) or pavement surface (AB or EF)) the evolution of the range data, (d_i), provided by the LMS can be described approximately by a straight-line over a small window. In the Unscented Kalman Filter (UKF) based technique, a straight-line process model is used to predict the next range data (d_{i+2}) given the past two range measurements (d_i, d_{i+1}) obtained at equal angular separation (see Figure 3). The filtering is effective and valid provided the model of the process adequately describes the evolution of the range data points, i.e. for data lying on the same region, for example, flat pavement surface and curb surface. The prediction error would be significant from the measured data at the boundary separating two contiguous regions (e.g. pavement surface to curb surface). Thus, the magnitude of the prediction error computed at a particular data point can be used to validate the authenticity of the process model describing the evolution of the range data points. If the prediction error exceeds a threshold at a particular data point it indicates start of a new process model, and hence a candidate end-point of a segment. These segmented edge lines are analyzed for possible curb edges using bank of filters.

2.1 Data Filtering and Segmentation

The Unscented Kalman filter (UKF) methodology, which is based on a stochastic framework, filters (random errors and outliers), effectively segments the data and estimate parameters simultaneously. Selection of threshold for segmenting the data is based on the statistical properties of the data. Although, methods based on the Hough Transform can be used for line extraction, it is unsuitable for this application since even a few points (at least three) can correspond to a valid curb. Thus, many

candidate edge lines result from its application requiring the use of extensive processing to filter out the lines corresponding to the curbs. Further, the selection of threshold is quite arbitrary.

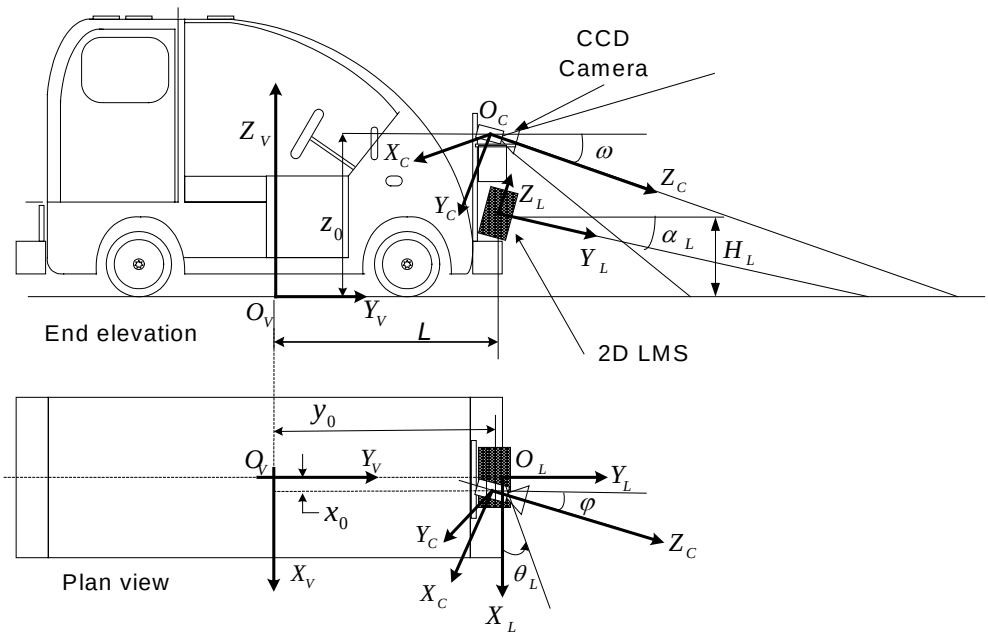


Figure 1: LMS and CCD camera sensor assembly and reference coordinate frames

Process model:

Consider the three points *P1*, *P2* and *P3*, lying on a line segment *L1*, at a range of d_i, d_{i+1}, and d_{i+2} respectively from the ladar measuring system, as shown in Figure 3. The range measurements are obtained at equal angular separation of γ degrees. Using elementary trigonometry it can be shown that [19]:

$$d_{i+2} = \frac{d_i d_{i+1}}{2d_i \cos\gamma - d_{i+1}} \tag{1}$$

Now, a three state process model can be set up by choosing state variables as, $x_1(k+1) = d_{i+2}$ and $x_2(k+1) = d_{i+1}$.

$$
\begin{aligned}
x_1(k+1) &= \frac{x_2(k)x_1(k)}{2x_2(k)\cos\gamma - x_1(k)} + v_1(k) \\
x_2(k+1) &= x_1(k) + v_2(k) \\
x_3(k+1) &= \tan^{-1}\left\{\frac{x_1(k)\sin\beta - x_2(k)\sin(\beta-\gamma)}{x_1(k)\cos\beta - x_2(k)\cos(\beta-\gamma)}\right\} + v_3(k)
\end{aligned} \tag{2}
$$

Represented in vector form,

$$
\mathbf{x}(k+1) = f(\mathbf{x}(k)) + \mathbf{v}(k) \tag{3}
$$

$\mathbf{v}(k)$ is the process noise and is assumed to be $N \sim (0, \mathbf{Q}_L)$.

Observation model:

Linear measurement model is,

$$
z(k) = d_{i+1} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + w(k) = \mathbf{H}\mathbf{x}(k) + w(k) \tag{4}
$$

where, $\mathbf{H}$ is the observation matrix, and $w(k)$ is the measurement noise assumed to be $N \sim (0, \sigma_L^2)$ and uncorrelated.

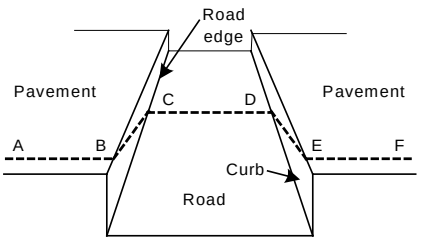


Figure 2: Ideal model of road, curbs and pavements

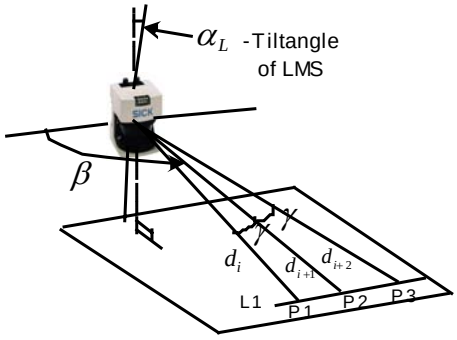


Figure 3: Three consecutive laser data points on a flat road surface

It is to be noted that the process model is nonlinear. A well-known approach is to linearize it and use the Extended Kalman Filter (EKF). Although EKF is a widely used modality, it has the disadvantages of difficulty in implementation, difficulty in

tuning, and only reliable for systems that are almost linear on the time scale of the update intervals [20]. Therefore, Unscented Kalman Filter [20] is adopted for the filtering and segmentation.

Filter Initialization:

A reasonable initial value for the state estimate, $\hat{\mathbf{x}}(k/k)$ and posteriori error covariance, $\mathbf{P}(k|k)$ can approximately be chosen using two range measurements, at time k=1 and k=2. From equation (4), we have $\hat{x}_1(1|1) = z(1)$, $\hat{x}_1(2|2) = z(2)$. From equation (2), we have $\hat{x}_2(2|2) = \hat{x}_1(1|1)$ and hence,

$$\hat{\mathbf{x}}(2|2) = \begin{bmatrix} z(2) \\ z(1) \\ \tan^{-1}\left\{\dfrac{z(2)\sin\beta - z(1)\sin(\beta-\gamma)}{z(2)\cos\beta - z(1)\cos(\beta-\gamma)}\right\} \end{bmatrix} \tag{5}$$

Since the measurement noise is assumed uncorrelated,

$$\mathbf{P}_L(2|2) = \begin{bmatrix} \sigma_L^2 & 0 & 0 \\ 0 & \sigma_L^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{6}$$

Data Segmentation:

The filtered data is segmented using an appropriate threshold, and is based on [21]. Suppose the true values of the predicted measurement, $\mathbf{H}\hat{\mathbf{x}}(k+1/k)$ and observation, $\mathbf{z}(k+1)$, are $\mathbf{H}\hat{\mathbf{x}}'(k+1/k)$ and $\mathbf{z}'(k+1)$ respectively. Then their difference is, $\mathbf{y}'_k \triangleq \mathbf{H}\hat{\mathbf{x}}'(k+1/k) - \mathbf{z}'(k+1)$ and its estimate is, $\mathbf{y}_k = \mathbf{H}\hat{\mathbf{x}}(k+1/k) - \mathbf{z}(k+1)$. Now, it is to be tested for the hypothesis, H_o, such that

the predicted and the observed values represent the same point. i.e., $H_0: \ \mathbf{y}'_k = 0$, against the alternative hypothesis H_1 as, $H_1: \ \mathbf{y}'_k \neq 0$. Assuming that $\mathbf{H}\hat{\mathbf{x}}(k+1/k)$ and $\mathbf{z}(k+1)$ have statistically independent errors and their covariance matrices are $\mathbf{\Lambda}_{\hat{\mathbf{x}}}$ and $\mathbf{\Lambda}_{\mathbf{z}}$ respectively, then under the hypothesis H_0, the covariance matrix of the difference $\mathbf{y}_k$ is given by, $\mathbf{\Lambda}_{\mathbf{y}} = \mathbf{\Lambda}_{\hat{\mathbf{x}}} + \mathbf{\Lambda}_{\mathbf{z}}$

Now, the square of the *Mahalanobis distance*, δ_k^M between the predicted and observed measurement, i.e. $\mathbf{H}\hat{\mathbf{x}}(k+1/k)$ and $\mathbf{z}(k+1)$ is,

$$\begin{aligned} \delta_k^M &= \mathbf{y}_k^T \mathbf{\Lambda}_y^{-1} \mathbf{y}_k \\ &= \left(\mathbf{z}(k+1) - \mathbf{H}\hat{\mathbf{x}}(k+1/k)\right)^T \left(\mathbf{\Lambda}_{\mathbf{z}} + \mathbf{\Lambda}_{\hat{\mathbf{x}}}\right)^{-1} \left(\mathbf{z}(k+1) - \mathbf{H}\hat{\mathbf{x}}(k+1/k)\right) \end{aligned} \tag{7}$$

The hypothesis, H_0, is accepted against H_1, with confidence $\bar{\alpha}$ by choosing a threshold ε, such that $P\{\delta_k^M \leq \varepsilon \mid H_0\} = \bar{\alpha}$. Hence, at a confidence level of $\bar{\alpha}$, the new measurement point $\mathbf{z}(k+1)$ is regarded as belonging to the same process, if $\delta_k^M \leq \varepsilon$. It is to be noted that δ_k^M is having a χ^2 distribution and therefore, an appropriate threshold ε for a given $\bar{\alpha}$ can easily be determined using a χ^2 table. In the event, $\delta_k^M > \varepsilon$ the new measurement point is regarded as an initial observation of a new process. In such an event the previously filtered point is declared as the end of current collinear segment and re-initialize and execute the algorithm using the new measurement point as the starting point of a new segment (process model).

Here, a line is represented by a point and orientation $\{x_L, y_L, \phi_L\}$. $\{x_L, y_L\}$ is estimated as the mean of each data (in Cartesian coordinates) in a particular segmented data set and ϕ_L is directly estimated through UKF.

2.2 Road Boundary Extraction

Similar to painted lane markings, curbs (height is about 14cm) are features that parallel to the road and provide useful information for road detection. This property is used to filter the candidate line segments correspond to the road boundaries among those extracted in section 2.1.

Orientation Filter:

The straight-line segments extracted are first analyzed for their orientation. If it is assumed that the vehicle is initially, approximately parallel to the road edges (curbs), then only those edge segments whose orientations are constrained by equation (8) are chosen for further processing.

$$\phi_L = \frac{\pi}{2} \pm \Delta\phi_{\max} \tag{8}$$

where, $\Delta\phi_{\max} = \sin^{-1}\left(\bar{y} \Big/ \left(R_{\min} - w_r\right) \right)$, $R_{\min}$ and w_r are the minimum radius of the road curb and the road width respectively (see Figure 4). w_r is assumed to be known *a priori* and $R_{\min}$ is usually a road design parameter [22].

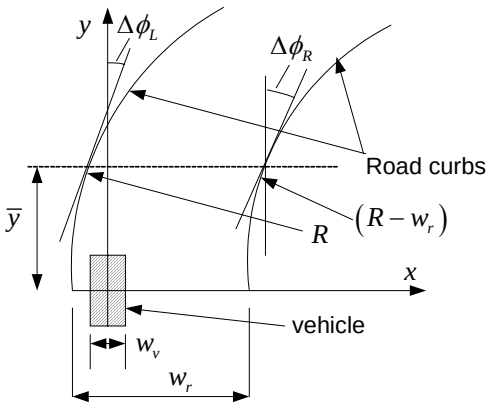


Figure 4: Orientation of curbs with respect to vehicle

Neighborhood Filter:

Subsequent to orientation filtering, edge segments lying in the neighborhood of the road ahead at a distance corresponding to the look-ahead distance of the laser ($\overline{y}$) are extracted. The neighborhood is defined as the lateral interval between the two points $(\overline{x}_L, \overline{y})$ and $(\overline{x}_R, \overline{y})$ shown in Figure 5. Points corresponding to the maximum possible left and right lateral positions of the curbs at a look-ahead distance of $\overline{y}$ can be calculated as,

$$\overline{x}_L = \frac{w_v}{2} - \frac{w_r}{2} - \frac{\overline{y}^2}{2\left(R_{\min} - w_r\right)}$$
$$\overline{x}_R = w_r - \frac{w_v}{2} + \frac{\overline{y}^2}{2\left(R_{\min} - w_r\right)} \tag{9}$$

where, $R_{\min} > w_r$ and w_v is the width of the vehicle. Therefore, the line segments obtained in orientation filtering are further filtered in neighborhood filtering based on where the edge segments intersect the line $y = \overline{y}$. That is, the edge segments whose x-coordinates, $\overline{x}$, that satisfy the following constraint (10) are chosen for further processing.

$$\overline{x}_L \le \overline{x} \le \overline{x}_R \tag{10}$$

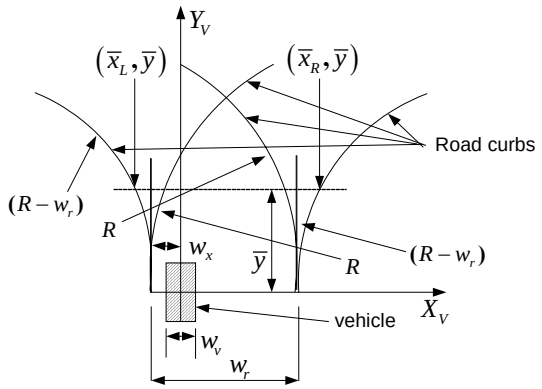


Figure 5: Vehicle in a left/right road curvature

Road Width Filter:

With the proliferation of GPS and digital maps, it is reasonable to assume the availability of approximate road widths. In this stage, the road width constraint is applied to extract a pair of edge segments that correspond to the road curbs. The interline distance at a look-ahead distance ($\bar{y}$), between all pairs of edge segments extracted in neighborhood filter, are determined. Now we choose the single pair of edge segments whose interline distance (l_d), closely matches the *a priori* road width ($\bar{w}_r$) subject to a tolerance ε_w as given in equation (11).

$$\left(\bar{w}_r - \varepsilon_w\right) \le l_d \le \left(\bar{w}_r + \varepsilon_w\right) \tag{11}$$

3. Road Boundary Extraction using Vision

The computational complexity of image processing can drastically be reduced by restricting the computations to a region of interest (ROI). The extracted curbs using LMS data can effectively be used to define an ROI. This needs a calibrated laser-camera system.

3.1 Laser-Camera Calibration

Vehicle, camera and laser reference coordinate frames are chosen as shown in Figure 1. It can be shown that if a point in 3D space is given in homogenous coordinates as $(x, y, z, 1)'$ with respect to the vehicle reference frame, $\{V\}$, its representation in the image, (u, v) is [23]:

$$u = u_0 + g_x f \frac{\left(x - x_0\right)}{\left(y - y_0\right)\sin\left(\omega\right) + \left(z - z_0\right)\cos\left(\omega\right)} \tag{12}$$

$$v = v_0 + g_y f \frac{\left(y - y_0\right)\cos\left(\omega\right) - \left(z - z_0\right)\sin\left(\omega\right)}{\left(y - y_0\right)\sin\left(\omega\right) + \left(z - z_0\right)\cos\left(\omega\right)} \tag{13}$$

where, $(u_0, v_0)'$ is the principal point, g_x and g_y (pixels/m) are the gains along x and y axes, f is the focal distance, $(x_0, y_0, z_0)'$ is the center of perspectivity or the position of the camera lens (origin of camera frame $\{C\}$) relative to the vehicle frame, $\{V\}$) and ω is the tilt angle of the camera. The camera mounting is fixed with zero pan and swing angles.

If the 3D vehicle coordinates of a point on the road surface is $(x, y, z)'$, and its coordinates in the laser measurement domain is (θ_{Li}, d_i), then assuming a flat road surface and a laser tilt angle of α_L we have (see Figure 1):

$$\begin{aligned} x &= \left(d_i \cos\alpha_L \cos\theta_{Li}\right) \\ y &= (d_i \cos\alpha_L \sin\theta_{Li} + L) \\ z &= 0 \end{aligned} \tag{14}$$

3.2 Road Boundary Extraction

Lane markings provide more reliable edges than that of pavement boundaries in vision images. Therefore, it is proposed to use lane markings in road boundary extraction. In Figure 6, suppose the curve, C_i is the image of the middle lane marking of the road ahead. Now, consider the points labeled P_i, P_{i+1} and P_{i+2} lying on the curve C_i at the polar coordinates (d_i, θ_i), (d_{i+1}, θ_{i+1}), and (d_{i+2}, θ_{i+2}) respectively. If the three points are chosen such that $\theta_i - \theta_{i+1} = \theta_{i+1} - \theta_{i+2} = \gamma$, where γ and the curvature of the curve, C_i near the points are sufficiently small, then the points can be assumed approximately collinear and their relationship given by equation (1). Here the range of a point on the image plane is interpreted to be the radial distance to the point from the origin. Thus, given the range measurements d_i, and d_{i+1} of the previous consecutive points P_i, and P_{i+1}, on the curve, C_i, one may predict the next

point P_{i+2}'s range measurement d_{i+2}. A filtered estimate of the predicted point P_{i+2}'s range measurement d_{i+2}, can be computed if a measurement of it is available. Such a measurement can be obtained by searching for an edge pixel along the line *OA* in the neighborhood of the point P_{i+2}. This process of prediction, measurement and filtering can be applied in sequence to all measured points in order to extract the curve, C_i. Similar to the procedure described in *Section 2.1*, we use the non-linear process (equation (2)), the measurement model (equation (4)), however without the state $x_3(k)$ and UKF to extract the middle lane marking of the road.

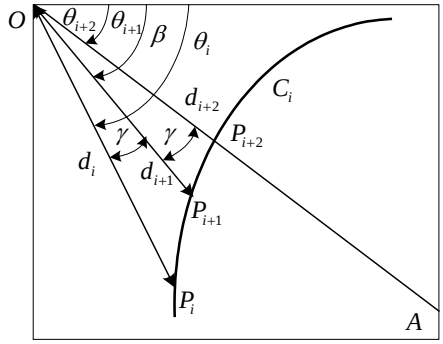


Figure 6: Image points on the middle lane image curve.

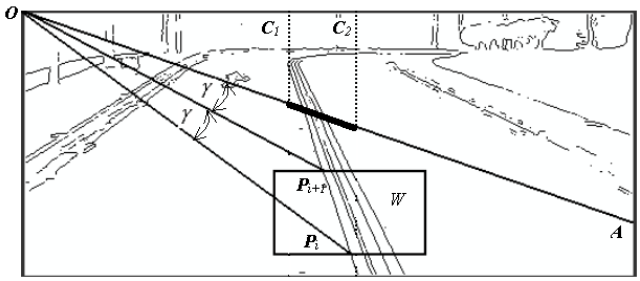


Figure 7: Search area on an edge image

Filter implementation:

To start the UKF it is important that we have the polar coordinates of two initial points P_1, and P_2 on lane curve C_i to be extracted. The two initial points are determined in a robust manner. Hough Transformation (HT) is used for the edges in a rectangular window (see Figure 7), W, centered approximately at a point (obtained using laser data) on the middle lane image curve C_i to extract candidate line segments. The HT procedure is made computationally efficient using the predicted midline segment using linear curb segments extracted by laser. Then candidate line

segments of HT are clustered and line corresponding to white lane marking is extracted. Intersection of that line with top and bottom edges of the window W are chosen as initial two points. The choice of width and height of the rectangular window, W are influenced by a combination of factors including, computational efficiency, curvature of the lane marking, flatness and gradient of road and camera pose.

Knowing the coordinates (row, column) of the two points on the mid line, it is possible to calculate the coordinates, (d_1,θ_1), and (d_2,θ_2) of the points, P_1, and P_2. These range values computed for the two initial coordinate points can be considered as the first two measurements $z(1)$, and $z(2)$ of d_1 and d_2. Hence, a reasonable initial state estimate of the UKF is:

$$\hat{\mathbf{x}}(2\,|\,2)=\begin{bmatrix} z(2) \\ z(1) \end{bmatrix} \tag{15}$$

Since the measurement noise is assumed uncorrelated, normally distributed with zero mean and variance σ_c^2, an initial estimate of the state error covariance matrix can be:

$$\mathbf{P}(2\,|\,2)==\begin{bmatrix} \sigma_c^2 & 0 \\ 0 & \sigma_c^2 \end{bmatrix} \tag{16}$$

For $k\geq 2$, we may use the UKF to obtain the next point's range d_{i+2}'s prediction, $\hat{x}_1(k+1\,|\,k)$, given the previous two points' (P_i and P_{i+1}) filtered range data, $\hat{x}_1(k\,|\,k)$ and $\hat{x}_2(k\,|\,k)$. To obtain the filter update, $\hat{x}_1(k+1\,|\,k+1)$, $k\geq 2$ of the prediction, $\hat{x}_1(k+1\,|\,k)$, using the UKF, measurement $z(k+1)$ is used. Repeating the process for all available points within a window defined by the look ahead distance, $k\geq 2$, the mid lane image curve of the road can be extracted. Next, the general

procedure for determining a range measurement $z(k+1)$ of the predicted range $\hat{x}_1(k+1|k)$ for $k \geq 2$ is described.

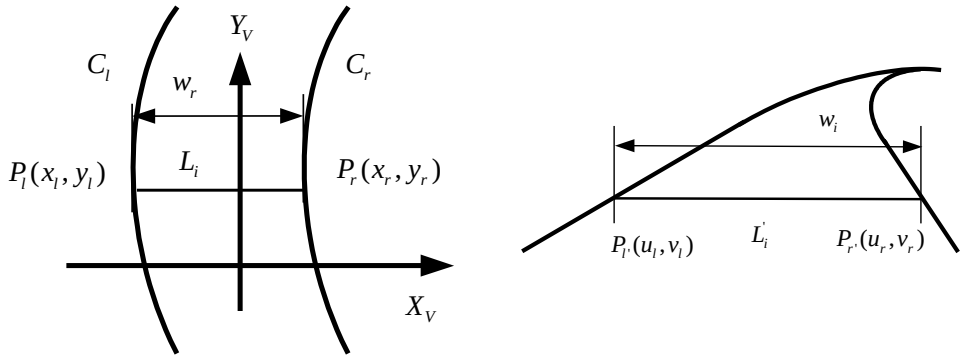
As was described at the outset, by searching along the line segment *OA* (Figure 6, and 7, straight line at an angle of γ with line OP_{i+1}) of the edge-image one can detect edge pixels that could possibly define the range measurement, $z(k+1)$ of point P_{i+2}. In actual practice, it is only necessary to search a segment of the line *OA*, such as between intersection points of lines C_1 and C_2 with *OA* of Figure 7, thus reducing the computation time. C_1, and C_2 are determined based on the 3σ limit of the error covariance. Since, apart from the correct edge pixel data corresponding to the mid lane marking, there can be spurious white pixels that may be detected as possible measurements of P_{i+2}, along the line *OA*, we choose the data point that yields the minimum absolute innovation ($e = | z(k+1) - \mathbf{H}\hat{\mathrm{x}}(k+1/k+1) |$) at each iteration. In other words, the measurement value $z(k+1)$ for update of state $\hat{\mathbf{x}}(k+1/k+1)$, covariance $\mathbf{P}(k+1/k+1)$ and Kalman gain $\bar{K}(k+1)$, is chosen to be the range measurement that is closest to the predicted range measurement based on the process and measurement models. This process will be repeated for all points defined by, $\theta_{i+2} = \theta_2 - i.\gamma$, where $i = 1, ..., n$. The number of points n considered is a function of the look-ahead distance.

Road Boundary Estimation:

Now the boundary of the road is obtained in image coordinates by mapping the trajectories C_l and C_r (Figure 8) of the intersection points or the end-points of the line segment L_i as it is swept along the road in a direction perpendicular to the *x*-axis of the camera frame. The road boundary in the image plane is defined by the

trajectories of the end-points of the line segments L_i' (projections of line L_i) of length w_i centered at the mid line on the image plane. Knowing the width of the line segment w_r from the laser range measurements (*section 2*), equation (17) can be used to generate the length w_i at a given row coordinate *v*, and hence the road boundary in image coordinates [23].

$$w_i = \frac{g_x D \sin(\alpha_C)}{g_y}\left(v_\infty - v\right) \tag{17}$$



(a) Parallel road edge curves on the ground plane

(b) Projection of road edge curves on the image plane

Figure 8: Road in ground plane and its projection in the image plane

4. Road Boundary Tracking

So far, we have dealt with road feature detection and in this section we begin with the formulation of the problem of road feature tracking. In the case of using laser data for road curb tracking, the formulation and its solution are complicated by the fact that there are no obvious moving target(s) to detect/track and the detection/tracking sensors are on board a moving platform (vehicle). Since the objective is to detect and track curbs, a pseudo target is defined as the line segment that results from the

intersection of the planar laser scan with the planar curb surface. Thus when the vehicle is in motion the line segment or target moves along the curb.

4.1 Problem Formulation

Coordinate Frames:

There are three coordinate frames of interest. A fixed world coordinate frame $\{X_W, Y_W\}$ is used mainly for visualization purposes, and two coordinate frames $\{{}^{k}X_V, {}^{k}Y_V\}$ and $\{{}^{k+1}X_V, {}^{k+1}Y_V\}$ denoting the vehicle frame at the time instances k and $k+1$ respectively (see Figure 9) are for derivation of the curb tracking algorithm. The relationship between the vehicle frames at the $(k+1)^{th}$ instant and the k^{th} instant is a simple rotation ($\Delta\phi_v$) followed by a translation (t_x, t_y).

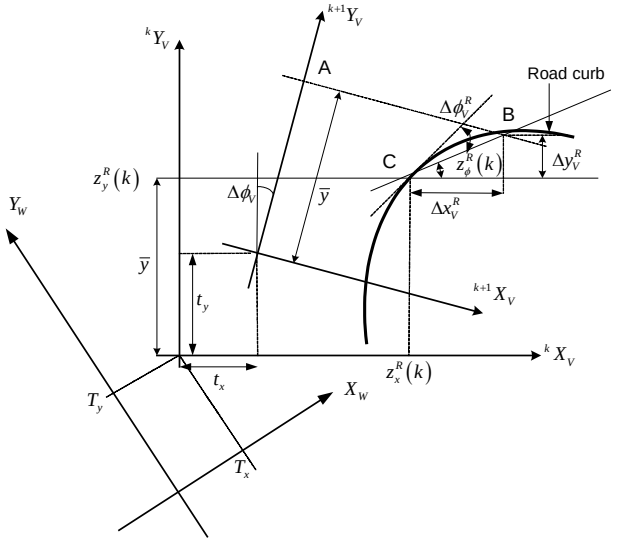


Figure 9: Road curb and coordinate frames

Target Model:

The state vector, $\mathbf{x}$, is defined as $\mathbf{x} \in \{x^R(k), y^R(k), \phi^R(k), x^L(k), y^L(k), \phi^L(k)\}$,

where $\left(x^R(k), y^R(k), \phi^R(k)\right)$ and $\left(x^L(k), y^L(k), \phi^L(k)\right)$ denote the right and left curb segments with respect to the k^{th} vehicle frame. Since the road is locally flat, the target (LMS- scan and curb intersection) would be approximately at a known distance $\bar{y}$ in the direction of the vehicle *y*-axis. From Figure 9, it may be easily deduced that the evolution of the curb segments is described the by equation (18).

$$\mathbf{x}(k+1) = \mathbf{f}\left[\mathbf{x}(k),\ \mathbf{u}(k)\right] + \mathbf{v}(k) \tag{18}$$

where, $\mathbf{x}(k) = \left[\, x^R(k),\ y^R(k),\ \phi^R(k),\ x^L(k),\ y^L(k),\ \phi^L(k) \right]^T$,

$$\mathbf{f}\left[\mathbf{x}(k),\ \mathbf{u}(k)\right] = \mathbf{F}(k)\ \mathbf{x}(k) + \ \mathbf{u}(k),$$

$$\mathbf{F}(k) = \begin{bmatrix} \mathbf{R}_v & \mathbf{O}_{3x3} \\ \mathbf{O}_{3x3} & \mathbf{R}_v \end{bmatrix},\ \mathbf{R}_v = \begin{bmatrix} \cos\Delta\phi_v & -\sin\Delta\phi_v & 0 \\ \sin\Delta\phi_v & \cos\Delta\phi_v & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{u}(k) = \begin{bmatrix} (\Delta x_V^R - t_x)\cos\Delta\phi_v - (\Delta y_V^R - t_y)\sin\Delta\phi_v \\ (\Delta x_V^R - t_x)\sin\Delta\phi_v + (\Delta y_V^R - t_y)\cos\Delta\phi_v \\ \Delta\phi_V^R - \Delta\phi_v \\ (\Delta x_V^L - t_x)\cos\Delta\phi_v - (\Delta y_V^L - t_y)\sin\Delta\phi_v \\ (\Delta x_V^L - t_x)\sin\Delta\phi_v + (\Delta y_V^L - t_y)\cos\Delta\phi_v \\ \Delta\phi_V^L - \Delta\phi_v \end{bmatrix}$$

$\mathbf{v}(k)$ is zero-mean white Gaussian process noise with covariance, $E\left[\mathbf{v}(k)\mathbf{v}(k)^T\right] = \mathbf{Q}_{Vt}(k)$. In Figure 9, point C is the position measurement $\left(\mathbf{z}_x^R(k),\ \mathbf{z}_y^R(k)\right)$ of the right curb segment $\left(\mathbf{z}_x^R(k),\ \mathbf{z}_y^R(k),\ \mathbf{z}_\phi^R(k)\right)$ obtained at k^{th} instant.

Observation Model:

The observations correspond directly to the left and right straight curb segments, each represented in terms of a point (z_x, z_y) and orientation (z_φ). The observation model in the presence of two curbs (left and right) is,

$$\mathbf{z}(k) = \mathbf{h}\left[\ \mathbf{x}(k)\right] + \mathbf{w}(k) \tag{19}$$

where, $\mathbf{z}(k) = \left[\ z_x^R(k),\ z_y^R(k),\ z_\phi^R(k),\ z_x^L(k),\ z_y^L(k),\ z_\phi^L(k)\right]^T$ superscript R and L denote right and left sides respectively. $\mathbf{h}\left[\ \mathbf{x}(k)\right] = \left[\mathbf{H}\right]\mathbf{x}(k)$, where, $\mathbf{H} = \mathbf{I}_{6x6}$. $\mathbf{w}(k)$ is the sequence of zero-mean white Gaussian measurement noise with block diagonal covariance matrix $E\left[\mathbf{w}(k)\mathbf{w}(k)^T\right] = \mathbf{R}_{Vt}$. It is important to note that depending on the number of curbs that exist, the measurement vector dimension (equation (19)) is modified accordingly.

4.2 Filter Implementation

The extracted image features (*Section 3*) with LMS- data (*Section 2*) are used in curb tracking. The target model and the measurement model defined by equations (18) and (19) are used with following parameter values (see Figure 9).

$$\begin{aligned}
\Delta x^i &= \frac{\left(c_{BC}^i - c_{AB}\right)}{\left(m_{AB} - m_{BC}^i\right)} - \mathbf{z}_V^i(1) \\
\Delta y^i &= m_{AB}\frac{\left(c_{BC}^i - c_{AB}\right)}{\left(m_{AB} - m_{BC}^i\right)} + c_{AB} - \mathbf{z}_V^i(2) \\
\Delta\phi^i &= \mathbf{tan}^{-1} m_{BC}^i - \mathbf{z}_V^i(3) \\
m_{AB} &= -\mathbf{tan}\,\Delta\phi_v \\
c_{AB} &= t_y + \bar{y}\,\mathbf{cos}\,\Delta\phi_v + \mathbf{tan}\,\Delta\phi_v(t_x + \bar{y}\,\mathbf{sin}\,\Delta\phi_v)
\end{aligned} \tag{20}$$

where, $i = L, R$; meaning *left* or *right*. m_{BC}^i and c_{BC}^i values are obtained from the fitted lines for the vision data along the curb segment BC. It is to be noted that it can be made more accurate by fitting a polynomial to the vision data of the segment BC. However, due to the limitation of the maximum road curvature that can have, it is reasonable to use a straight line fit rather than a more complex polynomial fit. Having defined the process and sensor models, the Kalman filtering approach [24] is used in

tracking.

4.3 Data Association

One of the challenging problems in curb tracking is that the LMS may provide cluttered measurements with partially observable measurements corresponding to the actual target. Partial observability is mainly due to the non-existing curbs due to road branching, x-junctions and curbs blocked by other vehicles. The image sensor can be effectively used to overcome those problems.

Let us assume the measurements of vision data ((x, y)) in vehicle coordinates are having the same covariance, σ_V^2, where *V* refers to vision data. Now, the variance of best-fitted line parameters can be calculated as follows.

Using the maximum likelihood, in the case of Gaussian distributed uncertainties, the best fit for parameters, $\mathbf{a} = \{a_k\}$ with measurement, $y = f(\mathbf{x}, \mathbf{a})$ is the one that minimizes the χ^2 function,

$$\chi^2 = \sum_i \frac{\left(y_i - f(\mathbf{x}_i, \mathbf{a})\right)^2}{\sigma_i^2} \tag{21}$$

where, σ_i^2 is the variance of data. The information matrix, $\mathbf{H}^I$ can be defined as,

$$H_{jk}^I = \frac{1}{2}\frac{\partial^2 \chi^2}{\partial a_j \partial a_k} \tag{22}$$

Then, variances and covariances of $\mathbf{a} = \{a_k\}$ are given by,

$$V_{a_j, a_k} = \left(H_{jk}^I\right)^{-1} \tag{23}$$

If the line is represented by, $y = \tan\phi_l x + c_l$, the perpendicular distance from a point (x_i, y_i) to the line is given by,

$$d_i = x_i \sin\phi_l - y_i \cos\phi_l + c_l \cos\phi_l \tag{24}$$

Finding the best-fit parameters, which minimizes d_i in equation (24) can be formulated as a minimization of χ^2 as,

$$\chi^2(\phi_l, c_l) = \frac{1}{\sigma_V^2} \sum_i \left(x_i \sin\phi_l - y_i \cos\phi_l + c_l \cos\phi_l \right)^2 \tag{25}$$

and therefore, variance of the angle of the fitted line, $\sigma_{\phi_V}^2$ can be calculated. It is assumed that the fitted line is passing through the mean of the data set and hence the variance of the (x, y) corodinates can be calculated as, σ_V^2 / N, where *N* is the number of data in the data set. Let's assume the line parameters at time *k* is $\psi_k = \left[x_l, y_l, \phi_l \right]^T$ and its covariance matrix is,

$$\Lambda_k^{\psi} = \begin{pmatrix} \sigma_V^2/N & 0 & 0 \\ 0 & \sigma_V^2/N & 0 \\ 0 & 0 & \sigma_{\phi l}^2 \end{pmatrix} \tag{26}$$

Let's consider the transformation of line data from k^{th} frame $\psi_k = \left[x_l, y_l, \phi_l \right]^T$ to $k+1^{\text{th}}$ frame, ψ_{k+1}, (see Figure 9),

$$\hat{\psi}_{k+1} = \mathbf{R}_v \psi_k + \mathbf{u} \tag{27}$$

and propagation of uncertainties,

$$\Lambda_{k+1}^{\hat{\psi}} = \mathbf{R}_v \Lambda_k^{\psi} \mathbf{R}_v{}^T + \left(\frac{\partial \mathbf{R}_v}{\partial \Delta\phi_v} \right) \Lambda_{\mathbf{R}_v} \left(\frac{\partial \mathbf{R}_v}{\partial \Delta\phi_v} \right)^T + \Lambda_{\mathbf{u}} \tag{28}$$

where, $\Lambda_{\mathbf{R}_v}$ and $\Lambda_{\mathbf{u}}$ are uncertainties in rotation and translation respectively. The *Mahalanobis* distances are calculated for all the candidate LMS line segments in

$k+1^{\text{th}}$ time instant with the transformed vision line data from the k^{th} time instant to $k+1^{\text{th}}$ time instant as follows.

$$\delta_{k+1}^{M} = \left(\hat{\psi}_{k+1}^{V} - \psi_{k+1}^{L}\right)^{T} \left(\Lambda_{k+1}^{\hat{\psi}} + \Lambda_{k}^{\psi}\right)^{-1} \left(\hat{\psi}_{k+1}^{V} - \psi_{k+1}^{L}\right) \tag{29}$$

For the values of δ_{k+1}^{M}, which are less than a threshold κ, (determined using χ^2 distribution tables), we choose the LMS line segment corresponding to the minimum value of δ_{k+1}^{M} as the candidate LMS line to be used in the Kalman filter as a measurement. The transformed vision line data also be used in the Kalman filter as the second sensor measurement.

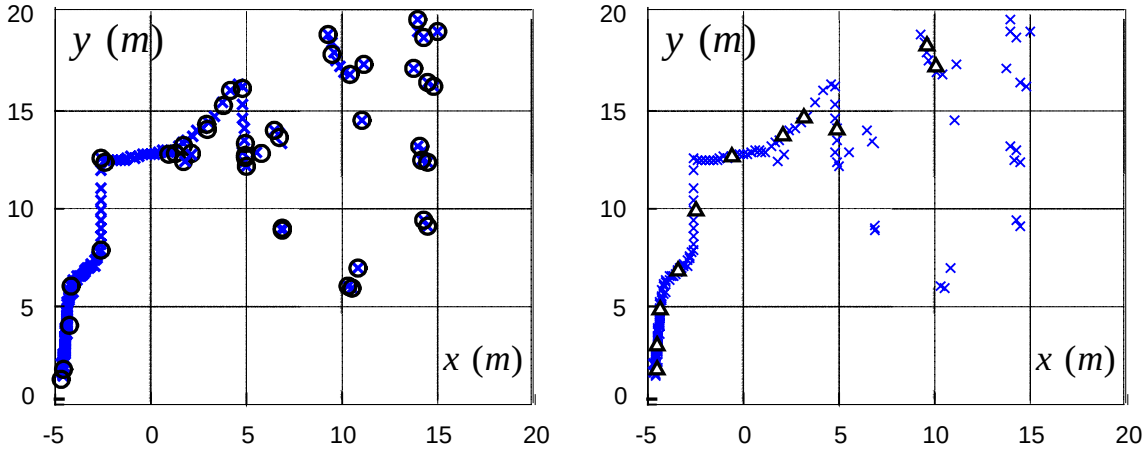
5. Experimental Results

Experiments were carried out on roads in a campus environment with a SICK [14] 2D laser measurement system and a CCD camera mounted on an in-house built test-bed vehicle as shown in Figure 10. The SICK LMS 290 provides range data at 1^0 intervals over 180^0 with accuracy of $\pm$5cm over 80m. The data transfer rate is 500kbaud. A dedicated image processing hardware (IP5000-CD [25]) was used for low-level real-time image processing tasks of the boundary detection algorithm. All the sensor data acquisition and algorithmic processing were implemented over three distributed single board Pentium III, 800MHz computers connected in a 100Mbs LAN [26]. The combination of the distributed processing hardware and the dedicated image processing hardware enabled the system to be operated at 200ms sampling time, which was quite adequate for relatively moderate speeds of the vehicle (4-5m/s).

Figure 10: Laser/vision sensor assembly mounted on the GenOME

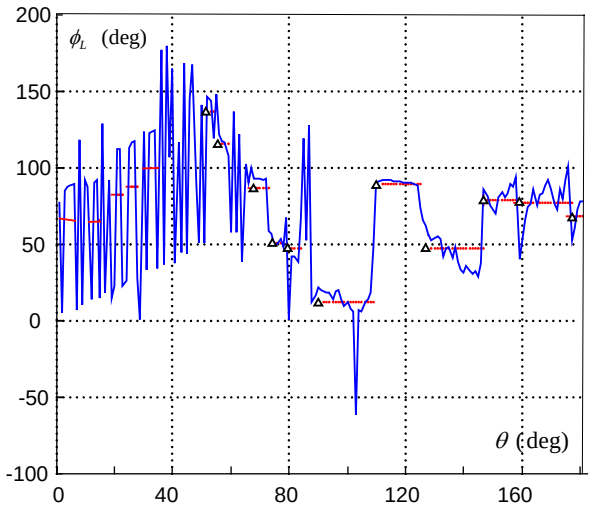
5.1 Road Boundary Extraction using an LMS

Laser data extracted from the looking down laser (tilt angle of 2.6^{o}) in a road environment (in vehicle coordinates) is shown in Figure 11(a). The *crosses* refer to the actual laser data and *circles* denote the detected discontinuity points using the UKF described in *Section 2*. The approximately vertical two segments around 7 *m* apart (in the middle part of the plot) are the data corresponding to the road curbs. Due to the convex nature of the road surface, the data corresponding to the road surface form a "V" shape rather than a straight line. Left most data correspond to a bank on the left side of the road. Scatter of data on the right hand side are due to the trees, lampposts and other man-made structures. In Figure 11(b) the mean value ($\left\{x_L^c, y_L^c\right\}$) of each data segment is shown using a '*traingle*'. The orientation (ϕ_L^c) of each line segment passing through the mean ($\left\{x_L^c, y_L^c\right\}$) are shown in Figure 11(c) using a '*triangle'*.



(a) Laser data (*crosses*) and detected discontinuity points (*circles*)

(b) *Triangle* represent the mean of each data segment

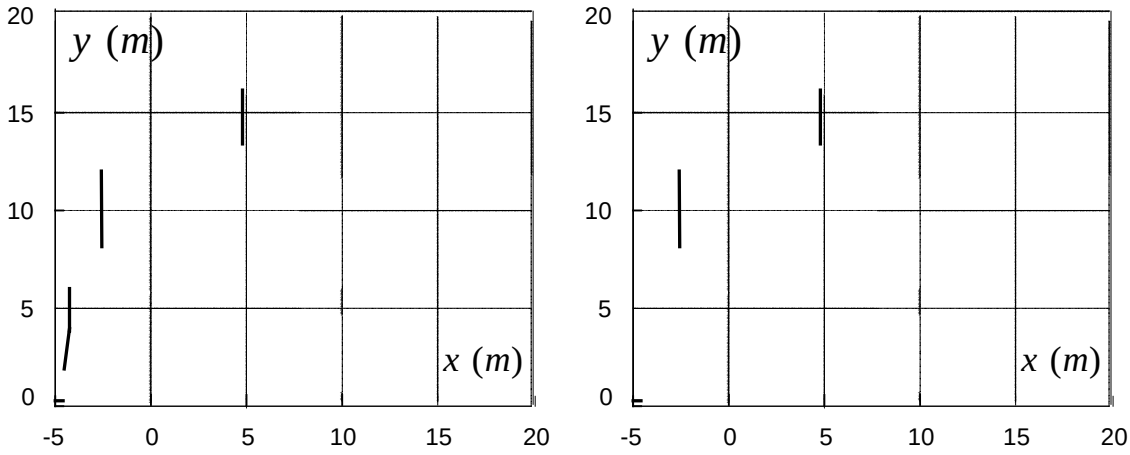


(c) *x-axis*: laser sweeping angel, *y-axis*: Estimated orientation angle of each data point and *Triangle* represents the orientations corresponding to *Triangle* in (b)

Figure 11: Line feature extraction using 2D LMS

The result of filtering the lines using Orientation and neighborhood filtering is depicted in Figure 12(a). It may be observed that there are four line segments satisfying the orientation and neighborhood constraints, although their relative locations are not the same. Results after road width filtering is shown in Figure 12(b).

For qualitative assessment of results, laser data corresponding to the detected edge-lines are mapped to the visual image of the road scene (see Figure 12(c)) captured by the camera mounted on the vehicle. It could be noted that the curbs are correctly detected as shown in Figure 12(c).



(a) Line segments after orientation and neighborhood filter (b) Line segments after road width filter

(c) Detected curbs on the image plane

Figure 12: Curb extraction results

5.2 Road Boundary Estimation using Vision

Images were captured in various road scenarios and assessed the robustness of the feature extraction algorithm as shown in Figure 13. Although, a straight-line model is assumed in the synthesis of the UKF, it is capable of detecting curved lane markings with shadows, letters and speed regulating strips effectively.

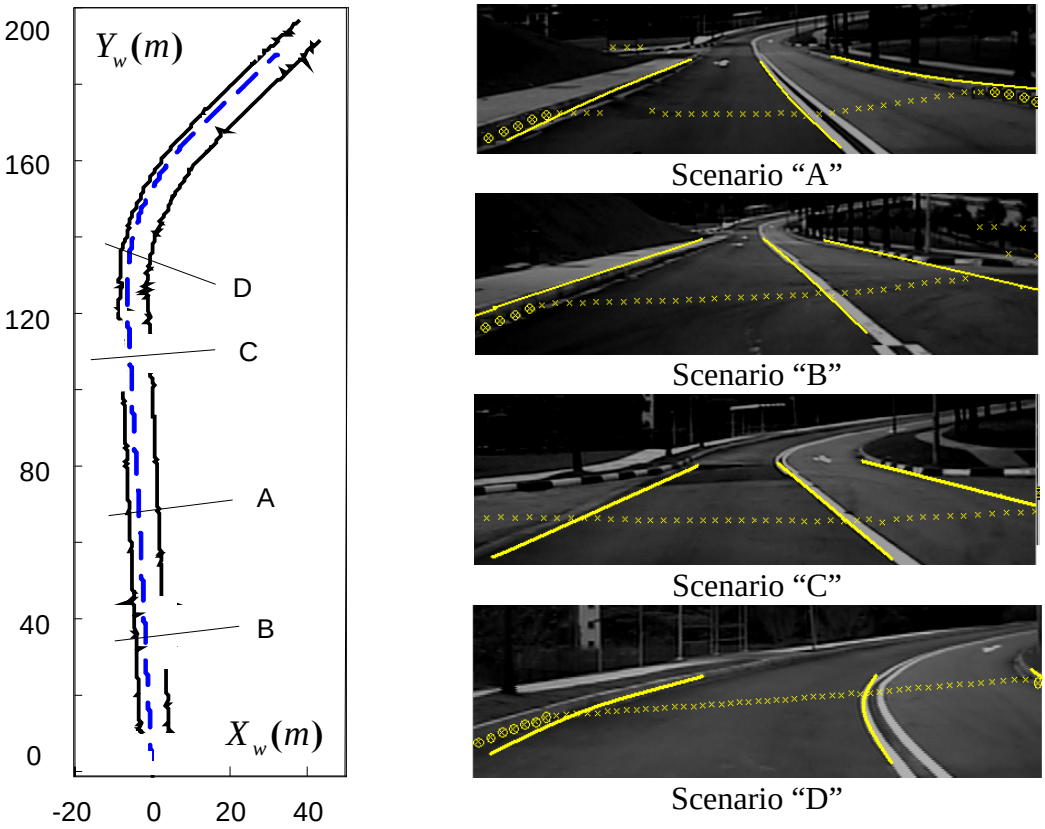
Figure 13: Feature extraction results in extreme images

5.3 Road Boundary Tracking

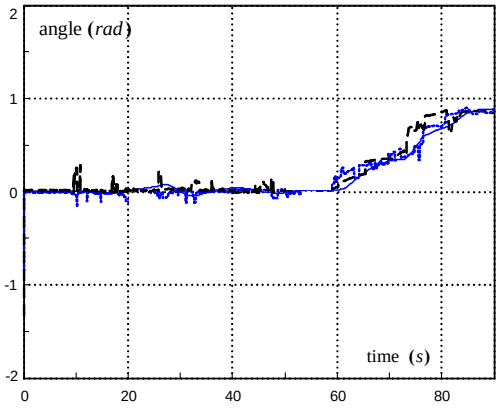
In this experiment, the vehicle was driven approximately at a constant speed of 4 ms^{-1} along different road segments. The vehicle speed and orientation (yaw angle) measurements were obtained using wheel encoders and a fiber optic gyroscope respectively. The detection and tracking results obtained in different road scenarios, referenced to the global coordinate frame, are shown in Figure 14. For qualitative assessment of the results, the detected and tracked road boundaries are shown in Figure 14(a). In Figure 14(a) and (b), scenario 'A' corresponds to a situation where both left and right curbs exists, 'B' only the left curb (a new road branches to the right from the main road; y-junction), 'C' no curbs on either side (vehicle entering or in the midst of an x-junction), and 'D' a sharp bend to the right. In Figure 14(a), *solid lines*

denote the curbs while *dashed line* denotes the vehicle path. In Figure 14(b), *crosses* denote the raw ladar data, *circled and crosses* denote the detected curbs using ladar and *solid lines* correspond to estimated midline and road boundaries using vision sensor. The vehicle's orientation (*solid line*), estimated left (*dashed line*) and right (*dotted line*) curb angles as the vehicle moves, referenced to the global frame, are shown in Figure 14(c).

Figure 15 shows different but typical scenarios in a road environment. In situations as shown in Figures 15(a), 15(b), and 15(c) the separation of curbs from complex lane markings and other signs imprinted on the road using image alone methods even using complex image processing supplemented by heuristics would be unreliable and time consuming. The vision alone methods also fail in extreme images due to low lighting (rainy weather in Figure 15(d) and nighttime in Figure 15(e)) and saturation (reflection of sunlight in Figure 15(f)). Being an active sensor, ladar is capable of handling such complex road scenarios due to lighting and weather quite effectively improving the robustness of the road boundary tracking algorithm.



(a) Position tracking (*solid line* – curbs, *dashed line* – vehicle path)
(b) Various road scenarios (*crosses* - laser data, *circled crosses* - laser data corresponding to the curbs, *solid lines* – midline and road boundaries using vision)



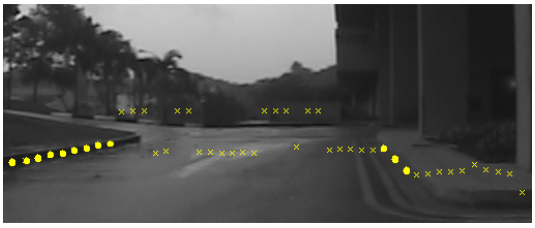
(c) Curb angle tracking (*dashed*- left curb, *dotted* – right curb, *solid* – vehicle's orientation)

Figure 14: Road boundary tracking results

(a) Road with zigzag markings

(b) Road with pedestrian crossings

(c) Road view blocked by an overtaking vehicle



(d) Road view in rainy weather

(e) Road view at nights

(f) Saturated images due to reflections

Figure 15: Extreme situations in roadway environments (*crosses* - laser data, *circled crosses* - laser data corresponding to the curbs)

6. Conclusion

Sensing and detection of lanes and road boundary in real time are important tasks in the synthesis of autonomous vehicles and driver assistance systems. In this paper, we have proposed a laser/camera composite sensing methodology for robust road extraction of pavement curbs with accurate depth information. On one hand vision alone methods, being passive, suffer from degradation of performance due to bad lighting, adverse whether conditions, reflections from bright source, etc. On the other hand, using active sensors such as a laser 2D scanner can over come such limitations,

although it can only provide limited information. However, by fusing both sensor data in a complementary fashion the limitations of one can be compensated by the other to achieve better performance. This has been demonstrated through experimental results for the case of detection and tracking road curbs under different road scenarios such as road branching, sharp curvatures and other vehicles obstructing road curbs.

References

[1] D. Pomerleau and T. Jochem, Rapidly adapting machine vision for autonomated vehicle steering, *IEEE Expert*, *11*(2) , 1996, 19 –27.

[2] E. D. Dickmanns and B. D. Mysliwetz, Recursive 3D road and relative ego-state recognition, *IEEE Trans. of Pattern Analysis and Machine Intelligence, 14*(2), 1992, 199-213.

[3] M. Bertozzi and A. Broggi, GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection, *IEEE Trans. on Image Processing, 7*(1), 1998, 62-81.

[4] D. A. Pomerleau, Progress in neural network-based vision for autonomous robot driving, *Proc. of the Intelligent Vehicles '92 Symposium*, 1992, 391-396.

[5] K. Kluge, Extracting road curvature and orientation from image edge points without perceptual grouping into features, *Proc. of the Intelligent Vehicles'94 Symposium*, 1994, 109-114.

[6] K. Kluge and S. Lakshmanan, A deformable-template approach to lane detection, *Proc. of the Intelligent Vehicles'95 Symposium*, 1995, 54-59.

[7] R. Aufrère, R. Chapuis, and F. Chausse, A model driven approach for real time road recognition, *Int. J. of Machine Vision and Applications, 13*, 2001, 93-107.

[8] S. Lakshmanan and D. Grimmer, A deformable template approach to detecting straight edges in radar images, *IEEE Trans. on Pattern Analysis and Machine Intelligence, 18*(4), 1996, 438-443.

[9] K. Kaliyaperumal, S. Lakshmanan and K. Kluge, An algorithm for detecting roads and obstacles in radar images, *IEEE Trans. on Vehicular Technology, 50*(1), 2001, 170-182.

[10] B. Ma, S. Lakshmanan and A. O. Hero, Simultaneous detection of lane and pavement boundaries using model-based multi-sensor fusion, *The IEEE Trans. on Intelligent Transportation Systems, 1*(3), 2000, 135-147.

[11] M. Nikolova and A. Hero, Segmentation of a road from a vehicle-mounted radar and accuracy of the estimation, *Proc. of the IEEE Intelligent Vehicles Symposium*, 2000, 284-289.

[12] J. Sparbert, K. Dietmayer and D. Streller, Lane detection and street type classification using laser range images, *Proc. of IEEE Intelligent Transportation Systems Conference*, 2001, 454-459.

[13] Kirchner A. and Heinrich T, Model based detection of road boundaries with a laser scanner, *Proc. of the Int. Conf. on Intelligent Vehicles*, 1998, 93-98.

[14] *Laser measurement systems: Production overview,* SICK AG, Auto Ident, Germany, 2002.

[15] T. Dunlay, Obstacle avoidance perception processing for the autonomous land vehicle, *Proc. of the Conf. of IEEE Robotics and Automation", 2,* 1998, 912-917.

[16] D. Dedieu, V. Cadenat and P. Soueres, Mixed camera-laser based control for mobile robot navigation, *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, *2*, 2000, 1081-1086.

[17] J. A. Castellanos, J. Neira and D. Tardos, Multisensor fusion for simultaneous localization and map building, *IEEE Trans. on Robotics and Automation, 17*(6), 2001, 908-914.

[18] J. Neira, J. D. Tardos, J. Horn and G. Schmidt, Fusing range and intensity images for mobile robot localization, *IEEE Trans. on Robotics and Automation, 15*(1), 1999, 76-84.

[19] M. D. Adams, *Sensor modelling, design and data processing for autonomous navigation* (World Scientific Publishing Co. Pte. Ltd, 1999).

[20] S. Julier, J. Uhlmann, H. F. Durrant-Whyte, A new method for the nonlinear transformation of means and covariances in filters and estimators, *IEEE Trans. on Automatic Control*, 2000, 477-482.

[21] Z. Zhang and O. Faugeras, *3D dynamic scene analysis* (Springer Velberg, New York, 1992).

[22] Connecticut Department of Transportation Highway Design Manual, January 1999.

[23] W.S. Wijesoma, K.R.S. Kodagoda and A.P. Balasuriya, A Laser-vision sensing for road detection and reconstruction, *Proc. of the IEEE 5th Intl. Conf. on Intelligent Transportation Systems (ITSC 2002),* Singapore, pp.248-253.

[24] Y. Bar-Shalom and X. R. Li, Estimation and tracking, principles, techniques, and software, (Artech House, Boston, 1993).

[25] http://www.hitachi.co.jp

[26] E. K. Teoh, W.S. Wijesoma, D. W. Wang, A. P. Balasuriya, K.R.S. Kodagoda, Z. K. Lu, C. Ye and C. S. Tan, Development of a Test-Bed for Outdoor AGV Research, *Electrical and Electronic Engineering Research*, Nanyang Technological University, Singapore, January 2001, pp70-71.