COMPUTER VISION SYSTEMS FOR 'CONTEXT-AWARE' ACTIVE VEHICLE SAFETY AND DRIVER ASSISTANCE

Pinar Boyraz, Xuebo Yang, Amardeep Sathyanarayana, John H.L. Hansen
University of Texas at Dallas, Department of Electrical Engineering
Richardson, TX, USA
Paper No: 09-0266

ABSTRACT

The last decade has witnessed the introduction of several driver assistance and active safety systems in modern vehicles. Considering only systems that depend on computer vision, several independent applications have emerged such as lane tracking, road/traffic sign recognition, and pedestrian/vehicle detection. Although these methods can assist the driver for lane keeping, navigation, and collision avoidance with vehicles/pedestrians, conflict warnings of individual systems may expose the driver to greater risk due to information overload, especially in cluttered city driving conditions. As a solution to this problem, these individual systems can be combined to form an overall higher level of knowledge on traffic scenarios in real time. The integration of these computer vision modules for a 'context-aware' vehicle is desired to resolve conflicts between sub-systems as well as simplifying in-vehicle computer vision system design with a low cost approach. In this study, the video database is a subset of the UTDrive Corpus, which contains driver monitoring video, road scene video, driver audio capture and CAN-Bus modalities for vehicle dynamics. The corpus includes at present 77 drivers' realistic driving data under neutral and distracted conditions. In this study, a monocular color camera output is used to serve as a single sensor for lane tracking and road sign recognition. Finally, higher level traffic scene analysis will be demonstrated, reporting on the integrated system in terms of reliability and accuracy.

INTRODUCTION

It has been reported that the fatality rate is 40,000-45,000 per year with annual cost of 6M [1] in USA. Although each year the number and the scope of vehicle safety systems (ABS, brake assist, traction control, EPS) are increased, these figures have reached a plateau, and further reduction in the number of accidents and associated costs require a pro-active approach. It is known that 90% of the accidents have drivers as a contributing factor while 57% are solely caused by human errors [2]. These figures reveal another explanation for the constant level of accident rates despite the increased efforts for safety systems: However well equipped the vehicles, and however well-designed the infrastructure, accidents generally happen due to human error. This does not necessarily mean that most accidents are inevitable. On the contrary, if all available road scene analysis, lane tracking, collision avoidance and driver monitoring technologies are applied, it is estimated that at least 30% of the fatal accidents (including several accident types) can be averted [3]. The obstacle preventing these systems from being applied is primarily cost-related. Therefore, resolving this problem lies in system integration to reduce the cost in order to obtain more compact and reliable active safety systems. In this study, a proposed integration is studied in the scope of in-vehicle computer vision systems. Video streams, whether processed on-line or off-line contains rich information content regarding road scene. It is possible to detect and track vehicle, lane markings, and pedestrians and recognize the road signs using a frontal camera and some additional sensors such as radar. It is of crucial importance to be able to detect, recognize and track road objects for effective collision avoidance. However, we believe that higher level context information can be extracted as well from the video stream for better integration of these individual computer vision systems. By interpreting the traffic scene for the driver, these systems can be employed for driver assistance systems as well as providing input for collision systems. In the scope of the work presented here, a lane tracking and road sign recognition algorithm is combined for future context-aware active vehicle safety applications. This paper first presents the proposed system, and then provides details on each individual computer vision algorithm. Next, the individual algorithms are combined to (1) extract higher level context

information on traffic scene, (2) lower the computational cost of road sign recognition using lane position cues from lane tracking algorithm. Experimental results and performance evaluation of the system is proposed in the following section. Finally, prospective applications for active safety and driver assistance are discussed together with conclusions.

Proposed system

The proposed system offers the ability to extract high-level context information using a video stream of the road scene. The videos considered form a sub-database of multi-media naturalistic driving corpus UTDrive. In fact, CAN-Bus information from OBD-II port is used to improve the performance of the designed lane tracker. The system distinguishes itself from existing in-vehicle computer vision approaches in two ways: (1) the lane tracking algorithm is used to feed the road sign recognition algorithm with spatial cues providing a top-down search strategy which improves process speed, (2) detected lane and road signs are combined using a rule-base interpretation to represent the traffic scene ahead at a higher level. The signal flow of the final system is presented in Fig.1. Three main blocks in this diagram (i.e. lane tracking, road sign recognition and rule base) are addressed in the next sections.

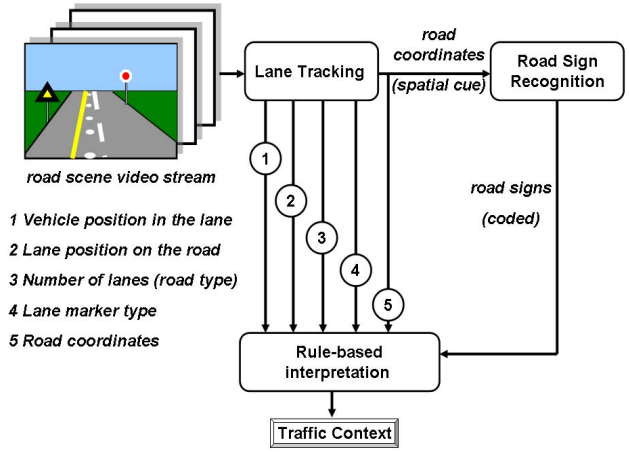


Figure 1. Signal flow in computer vision system

<u>Lane Tracking Algorithm</u>

There has been extensive work in developing lane tracking systems in the area of computer vision. These systems can be potentially utilized in driver assistance systems related to lane keeping and lane change. In [4], a comprehensive comparison of various lane-position detection and tracking techniques is presented. From that comparison, it is clearly seen that most lane tracking algorithms do not perform adequately to be employed in actual safety-related systems, however, there are encouraging advancements towards obtaining a robust lane tracker. A generic lane tracking algorithm has the following modules: a road model, feature extraction, post-processing (verification), and tracking. The road model can be implicitly incorporated as in [5] using features such as starting position, direction and gray level intensity. Model based approaches are found to be more robust compared to feature-based methods, for example in [6] B-snake is used to represent the road. Tracking lanes in real traffic environment is an extremely difficult problem due to moving vehicles, unclear/degraded lane markings, and variation of lane marks, illumination changes, and weather conditions. In [7] a probabilistic framework and particle filtering was suggested to track the lane candidates selected from a group of lane hypotheses. A color based scheme is used in [8]; shape and motion cues are employed to deal with moving vehicles in the traffic scene. Based on this previous research, we propose a lane tracking algorithm as presented in Fig. 2 to operate robustly in real traffic environment.

The lane tracking algorithm presented here uses conventional methods seen in the literature; however, the advancement here is the combination of a number of them in a unique way to obtain robust tracking performance. It utilizes a geometric road model represented by two lines. These lines represent the road edges and are updated by a combination of road color probability and road spatial coordinate probability distributions which are built using road pixels from 9 videos each at least containing 200 frames. The detection module of the algorithm employs three different operators, namely, steerable filters [9] to detect line-type lane markers, yellow color filter, and circle detectors. These three operators are chosen to capture three different lane-marker types existing in Dallas, TX where data collection took place. After application of the operators, the resultant images are combined since it is bow believed to have all the lane-mark features that could be extracted. This combined image is shown as the 'lane mark image' block in Fig.2. Next, Hough transform is applied to extract the lines together with their angles and sizes in the image. The extracted lines go through a post-processing step name as 'lane

verification' in Fig.2. Verification involves elimination of lane candidates which do not comply with the angle restrictions imposed by the geometric road model. From the lane tracking algorithm, 5 individual outputs are obtained: (i) lane mark type, (ii)lane positions and (iii)number of lanes in the image plane, (iv) relative vehicle position within its lane, and (v) road coordinates.

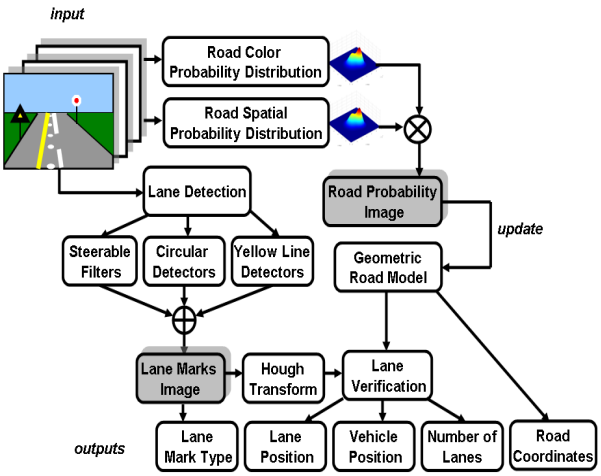


Figure2. Lane Tracking Algorithm

Two different types of updates are needed in the lane tracking algorithm. First, the road probability image updates the geometric road model for each frame and runs simultaneously with lane detection. Some of the false lane candidates are eliminated by the imposed angle limits of the geometrical road model; however, it is not adequate for robustness. The prominent lane candidates which exist in each *n* consecutive frame are held in a track-list, supplying the system with a time coherency or memory. A lateral vehicle dynamics model is applied together with Kalman filter to smooth the lane position measurement given by the lane candidates in the tracking list.

A sample output of the lane detection algorithm is given in Fig. 3(a). To observe the performance of the lane detection part the raw lane position measurement is compared with the lane position calculated by the vehicle model (see Fig.3(b)). In Fig.3 (c), the road area and possible lane candidates are marked by the algorithm.

The full evaluation will be presented in section 'Results and Performance Analysis'.

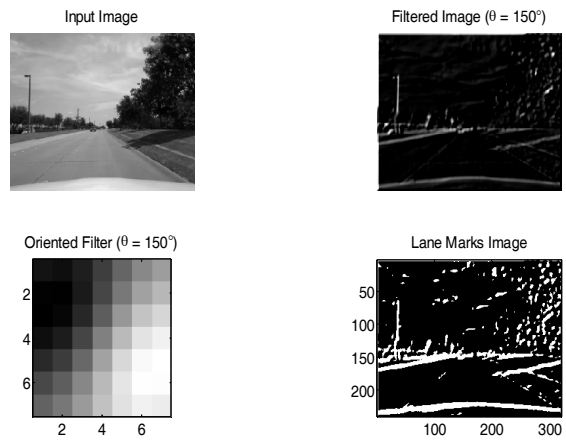


Figure3(a) .Sample output from lane detection algorithm

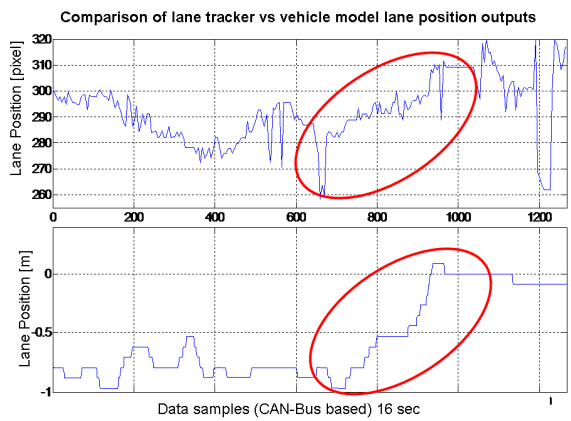


Figure3(b). Comparison of lane position measurement from raw lane detection (top) and vehicle model calculation (bottom) before outlier removal and tracking

Figure3(c). Road area and prominent lane candidates are projected on image (top), detected lanes after outlier & false candidate removal

The performance of the lane tracker algorithm depends heavily on feature (i.e. lane mark) detection; and feature detection based vision systems are vulnerable since they are affected by illumination change, occlusion, imperfect representations (i.e. degraded lane marks) plane distortion and vibration in our case since the road surface is not flat. Some of these disadvantages of feature based lane detection are overcome by employing a vehicle model for correcting the measurements in time domain and road model helps removing the outliers in spatial domain in the image. Since vehicle and road model are crucial for reliability and accuracy of the resultant system, they are briefly presented here.

Vehicle Model

The vehicle model used is known as 'bicycle model' [10] for calculating the lateral dynamics of the vehicle and linearized at a constant longitudinal velocity. In order to obtain more realistic behaviour from this model, it is updated by changing longitudinal velocity at discrete time steps. Therefore the resulting vehicle model is a hybrid system containing a set of continuous time LTI vehicle models switched by a discrete update driven by longitudinal speed changes. As a consequence the resultant model is non-linear and closer to realistic vehicle response. The model inputs are steering wheel angle, longitudinal vehicle velocity and outputs are lateral acceleration and side slip angle. The lateral acceleration output of this model is used to calculate the lateral speed and finally lateral position of the vehicle employing numerical integration by trapezoids. The variables of model are given in Table 1.

Table1. Variables of vehicle model

Symbol	Meaning
c_f or c_r	cornering stiffness coefficients for front and back tire
J	Yaw moment of inertia about z-axis passing at CG
m	Mass of the vehicle
r	Yaw rate of vehicle at CG
U	Vehicle speed at CG
y_c	Lateral offset or deviation at CG
τ	Wheel steering angle of the front tyre
α_f or α_r	Slip angle of front or rear tyre
β	Vehicle side slip angle at CG
ρ_{ref}	Reference road curvature
ψ	Yaw/ heading angle
ψ_d	Desired yaw angle
ω	Angular frequency of the vehicle

The equations of motion using the variables given in Table 1 are presented in equations (1-2).

$$\begin{bmatrix} \dot{\beta} \\ r \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \beta \\ r \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tau \qquad (1)$$

Where, in a format $\dot{x} = Ax + Bu$,

$$\dot{x} = \begin{bmatrix} \dot{\beta} \\ r \end{bmatrix}, \; A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \; B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \; u = \tau$$

And in the output format, $y = Cx + Du$

$$\begin{bmatrix} a_s \\ \beta \\ r \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta \\ r \end{bmatrix} + \begin{bmatrix} d_1 \\ 0 \\ 0 \end{bmatrix} \tau$$

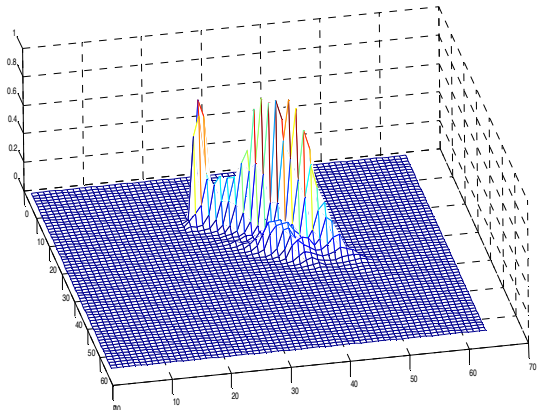
Where, $y = \begin{bmatrix} a_s \\ \beta \\ r \end{bmatrix}, C = \begin{bmatrix} c_{11} & c_{12} \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} d_1 \\ 0 \\ 0 \end{bmatrix}$.

The elements of the matrices are given explicitly here:

$$a_{11} = -(C_f + C_r)/mU \quad ,$$
$$a_{12} = -1 + [(h_r C_r - h_f C_f)/mU^2]$$
$$a_{21} = (h_r C_r - h_f C_f)/J \; , \; a_{22} = -(h_f^2 C_f + h_r^2)/JU$$
$$b_1 = C_f/mU \; , \; b_2 = h_f C_f / J$$
$$c_{11} = -[(C_f + C_r)/m] + [l_s(h_r C_r - h_f C_f)/J]$$
$$c_{12} = [(h_r C_r - h_f C_f)/mU] - l_s[(h_f^2 C_f + h_r^2 C_r)/JU]$$
$$d_1 = [(C_f/m) + l_s(h_f C_f / J)] \qquad (2)$$

Road Model

The road area is the main region of interest (ROI) in the image since it contains important cues on where the lane marks and road signs might be located. In other words, it allows performing a top-down search for the road objects that needs to be segmented. The color of the road is an important cue for detecting the road area, however it is affected by illumination and the surface does not always have the same reflection/color properties on all roads. Therefore, road and non-road color histograms using R, G and B channels are formed using 9 videos. The color histogram is normalized to obtain road color probability which is a joint measure in 3-D color space. The probability surface is shown in 2D R-B space in Figure 4. Also, probability image and the extracted road area is shown together with updated road model.



(a) 2-D color probability surface in R-B space

(b) Road area extraction and update process: updated road model(top), probability image (bottom left), extracted road area (bottom right)

Figure4. Road color probability model and its implementation

In addition to color probability, spatial coordinates of the road area is also used to obtain a location probability surface.

Road Sign Recognition Algorithm

The methods used for automatic road sign recognition can be classified into three groups: color based, shape based and others. The challenges in recognition of road signs from real traffic scenes using a camera in a moving vehicle has been listed as lighting condition, blurring effect, sign distortion, occlusion by other objects, sensor limitations. In [11] a nonlinear correlation scheme using filter banks is proposed to tolerate in/out of plane distortion, illumination variance, background noise and partial occlusions. However, the method has not been tested on different signs in a moving vehicle. [12] has addressed real time road signs recognition in three steps: color segmentation, shape detection and classification via neural networks, however, vehicle motion problem is not explicitly addressed. [13] used FFT signatures of the road sign shapes and SVM based classifier. The algorithm is claimed to be robust in adverse conditions such as scaling, rotations, and projection deformations and occlusions.

The road sign recognition algorithm employed in this paper (see Fig.5) uses a spatial coordinate cue from the lane tracking algorithm to achieve a top-down processing of the image; therefore, it does not need the whole frame but particular regions of interest, which are sides of the road area. This provides faster processing of the image and reduces the computational cost as well as eliminating the number of false candidates that might exist after color filtering.

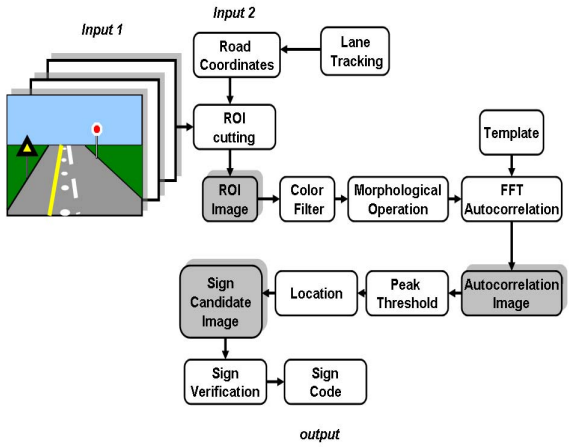


Figure 5. Road Sign Recognition Algorithm

After ROI cutting, color filters (designed for red, yellow and white) are used to perform the initial segmentation. Since they are critical for safety applications, we include only stop, pedestrian crossing and speed limit signs in our analysis for this stage. These signs have distinctive features related to color and shapes that are not shared, therefore classification becomes an easier problem since determining color and the shape is enough for recognizing the particular sign. Most of the time the speed limit sign segmentation ends with multiple sign candidates from the scene since the road scene might

contain road objects in that color range. Furthermore, the challenges exist due to moving vehicle vibrations, illumination change, scaling and distortion. An example of color segmentation for speed sign is presented in Fig.6.

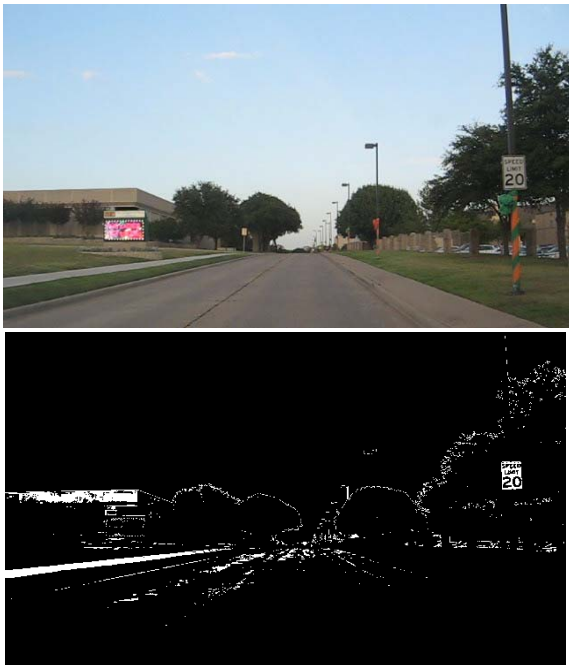


Figure 6. Sample output for white color filter segmenting the speed limit sign

Since the road sign shape edges are not clear, morphological operators (dilation and erosion) are used to correct this with a structural element of line or disk; an example of resultant image is given in Fig.7 for a road scene containing stop signs.

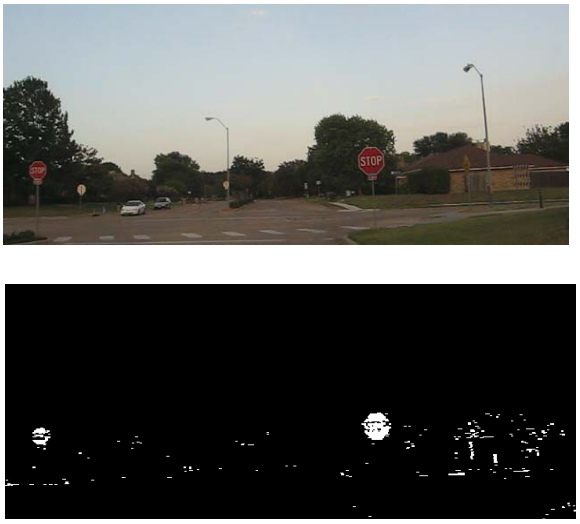


Figure7. Color segmented stop signs after dilation

The next step involves FFT correlation of the predetermined road sign shape templates (octagon, rectangle and triangle) with the ROIs in the image. The resultant image can be interpreted as location map of prospective sign. After application of a threshold the peak value of correlation image yields the location of the searched template. An example FFT correlation between the original frame seen in Figure 5 and rectangle template is given in Fig. 8. The speed limit sign gives a peak in this image and the location of the sign can be extracted easily by a threshold.

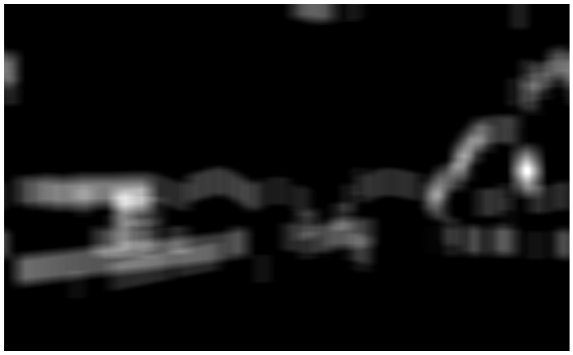


Figure8. FFT Correlation result with a rectangular template showing a peak in the location of speed limit sign

Although rare, false sign candidates may exist in the image after the threshold operation or multiple versions of the same sign may require distinguishing (i.e., speed limit signs are the same although the limits written on them are different) therefore an subsequent verification step/classification is needed. This is achieved via Artificial Neural Networks. For each sign a separate ANN is trained for negative and positive examples using at least 50 real world images of the aforementioned road signs and 50 negative candidates. The candidate sign images are obtained by cutting the image area around the row and column suggested by threshold image attained after FFT correlation. In order to obtain false candidates the threshold is lowered, therefore obtaining more candidates from FFT correlation image. The output of the initial sign classification and verification by ANN is a sign code (1, 2, 3) corresponding to speed limit, stop and pedestrian crossing signs in order. If there is no sign in the scene or if ANN verifies the sign is a false candidate the output is 0.

Integration of Algorithms for Context-aware Computer Vision

Previous studies such as [14] has combined different computer vision algorithms for a vehicle guidance system using road detection, obstacle detection and sign recognition at the same time. However, to the best knowledge of the authors very few studies if any focused on the traffic scene interpretation for extracting driving-related context information in real-time. This study proposes a framework for the fusion of outputs from individual computer vision algorithms to obtain higher level information on the context. The awareness of the context can pave the way in design of truly adaptive and intelligent vehicles. The fusion of information can be realized using a rule based expert system as a first step. We present a set of rules combining the outputs of vision algorithms with output options of warning, information message, and activation of safety features.

Case 1: *If road sign is 0, standard deviation of lane position< 10 pixels, standard deviation of vehicle speed <10 km/h, context: normal cruise.*

Case 2: *If road sign is 0, standard deviation of lane position<10 pixels, standard deviation of vehicle speed >10 km/h, context: stop-go traffic, likely congestion, and output: send information to traffic control center.*

Case 2: *If road sign is 1, vehicle speed >20 km/h, context: speed limit is approaching, output: warning.*

Case 3: *If road sign is 2, vehicle speed >20 km/h, context: stop sign is approaching and the driver did not reduce the vehicle speed yet, output: warning and activation of speed control and brake assist.*

Case 4: *If road sign is 3, vehicle speed >20 km/h, context: pedestrian sign is approaching, output: warning and activation of brake assist.*

The cases represented here are only a subset of the rule-base and it can be possible to use more advanced rule-base construction methods such as fuzzy logic. An inference engine acting as a co-pilot monitoring the context and driver status continuously can be designed with more inputs from other vision systems and in-vehicle sensors such as pedestrian and vehicle tracking, CAN-Bus analyzer, audio signals.

Results and Performance Analysis

Lane tracker and road sign recognition algorithms are tested on 9 videos each having at least 200 frames. The lane tracker with the help of probabilistic road model can overcome difficult situations such as passing cars as demonstrated in Fig.8. A similar recovery is observed when there is shadow on the road surface.

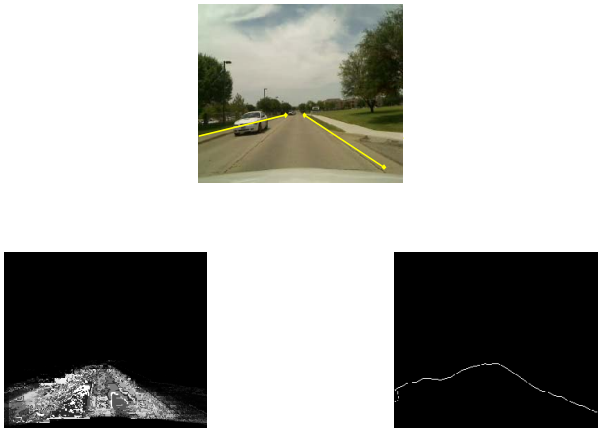


Figure8. Road model overcoming the occlusion

In order to assess the accuracy of the lane tracking algorithm the error between the lane position measured by the algorithm and the ground truth lane position marked manually on the videos is calculated. As an error measurement the angle of the lanes is taken. Table 2 shows the mean square error and standard deviation of the error in lane position measurement.

Table 2. Mean square error and standard deviation of the error in lane measurement

	MSE	STD
Left Lanes	15.24	3.86
Right Lanes	28.02	4.67

The algorithm runs close to real-time and the speed is increased by using lane position measurement as a cue for top-down search for road signs. The improvement in speed can go up to 3-fold based on the processed image area. The road sign recognition algorithm is capable of detecting the signs in 5-30 m range without failure. If the road sign occupies less than 10x20 pixel area, the recognition is not possible.

CONCLUSIONS

A context-aware computer vision system for active safety applications is proposed. The system is able to detect and track the lane marks and road area with acceptable accuracy. The system is observed to be highly robust to shadows, occlusion and illumination change thanks to road and vehicle model based feedback and correction. Road sign algorithm uses multiple cues such as color, shape and location cues. Although it can recognize the signs after the correlation result, a verification step is added to eliminate false candidates that might exist due to existence of similar road objects. Lastly, two algorithms are combined under a rule-based system to interpret the current traffic/driving context. This system has at least three possible end uses: driver assistance, adaptive active safety applications and lastly it can be used as mobile probes reporting back to the centre responsible for traffic management.

In our future work, more modules are planned to be added to detect and track more road objects and fuse additional in-vehicle signals (audio, CAN-Bus) to obtain a fully developed cyber co-pilot system to monitor the context and driver status.

REFERENCES

[1] http://www-fars.nhtsa.dot.gov/Main/index.aspx

[2] Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R. E., Stanisfer, R. L. and Castellan, N. J. (1977) Tri-level study of the causes of traffic accidents. *Report No. DOT-HS-034-3-535-77 (TAC)*.

[3] Eidehall, A., Pohl, J., Gustafsson, F., Ekmark, J., 'Towards Autonomous Collision Avoidance by Steering', *IEEE Transactions on Intelligent Transportation Systems,* vol. 8, no.1, pp. 84-94, March 2007.

[4] McCall, J.C., Trivedi, M.M., 'Video-Based Lane Estimation and Tracking for Driver Assistance: Survey, System and Evaluation', *IEEE Transactions on Intelligent Transportation Systems,* vol. 7, no.1, pp. 20-37, March 2006.

[5] Kim, Y. U., Oh, S-Y., 'Three-Feature Based Automatic Lane Detection Algorithm (TFALDA) for Autonomous Driving', *IEEE Transactions on Intelligent Transportation Systems,* vol. 4, no.4, pp. 219-225, Dec 2003.

[6] Wang, Y., Teoh, E.K., Shen, D., 'Lane detection and tracking using B-snake', *Image and Vision Computing,* vol 22, pp.269 -280, 2004.

[7] Kim, Z., 'Robust Lane Detection and Tracking in Challanging Scenarios', , *IEEE Transactions on Intelligent Transportation Systems,* vol. 9, no.1, pp. 16-26, March 2008.

[8] Cheng, H-Y., Jeng, B-S., Tseng, P-T., Fan, K-C., 'Lane Detection with Moving Vehicles in Traffic Scenes', *IEEE Transactions on Intelligent Transportation Systems,* vol. 7, no.4, pp. 16-26, Dec 2006.

[9] Jacob, M., Unser, M., 'Design of Steerable Filters for Feature Detection Using Canny-Like Edge Criteria', *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 26, no.8, pp. 1007-1019, Aug 2004.

[10] You, S-S, Kim, H-S, Lateral dynamics and robust control synthesis for automated car steering, Proc. of IMechE, Part D, Vol. 215, pp. 31-43, 2001.

[11] Perez, E., Javidi, B., 'Non-linear Distortion-Tolerant Filters for Detection of Road Signs in Background Noise', *IEEE Transactions on Vehicular Technology,* vol. 51, no.3, pp. 567-576, May 2002.

[12] Broggi, A., Cerri, P., Medici, P., Porta, P.P., Ghisio, G., 'Real Time Road Signs Recognition', *Proceeedings of 2007 IEEE Intelligent Vehicles Symposium,* Istanbul, Turkey, June 13-15, 2007.

[13] Jilmenez, P.G., Moreno, G., Siegmann, P., Arroyo, S.L., Bascon, S.M., 'Traffic sign shape classification based on Support Vector Machines and the FFT of the signature blobs', *Proceeedings of 2007 IEEE Intelligent Vehicles Symposium,* Istanbul, Turkey, June 13-15, 2007.

[14] De La Escalera, A., Moreno, L.E., Salichs, M.A., Armingol, J.M., 'Road Traffic Sign Detection and Classification', *IEEE Transactions on Industrial Electronics,* vol. 44, no.6, Dec 1997.