A Novel Lane Detection System With Efficient Ground Truth Generation

Amol Borkar, Monson Hayes, *Fellow, IEEE*, and Mark T. Smith

***Abstract*—A new night-time lane detection system and its accompanying framework are presented in this paper. The accompanying framework consists of an automated ground truth process and systematic storage of captured videos that will be used for training and testing. The proposed *Advanced Lane Detector 2.0* (ALD 2.0) is an improvement over the ALD 1.0 or *Layered Approach* with integration of pixel remapping, outlier removal, and prediction with tracking. Additionally, a novel procedure to generate the ground truth data for lane marker locations is also proposed. The procedure consists of an original process called *time slicing*, which provides the user with unique visualization of the captured video and enables quick generation of ground truth information. Finally, the setup and implementation of a database hosting lane detection videos and standardized data sets for testing are also described. The ALD 2.0 is evaluated by means of the user-created annotations accompanying the videos. Finally, the planned improvements and remaining work are addressed.**

***Index Terms*—Database, ground truth, inverse perspective mapping (IPM), Kalman filter, lane detection, lane tracking.**

I. Introduction

AROUND the world, high fatality rates from traffic accidents are evident. As vehicles commute at high speeds on freeways, the consequences of a distraction can be tragic. In 2002, the Secretary for the Transport Ministry of Malaysia cited 4.9% of traffic related accidents as fatal [3]. The Ministry of Public Safety of China reported 667 507 of traffic-related accidents in 2003 as fatal [4]. In the United States, the Fatal Analysis Reporting system of the National Highway Traffic Safety Administration (NHTSA) reported 41 059 casualties and more than 2.5 million injuries in 2007 [5]. Multiple studies performed by the NHTSA have shown that 20% of traffic accidents were the result of driver distractions. Traffic accidents account for a vast majority of fatalities worldwide; consequently, improving public safety on roads has become an important area of research.

Since the early 1990s, small footprint microprocessors and microcontrollers have been making their way into specialized consumer applications. One of these applications is driver assistance (DA) systems for automobiles. DA systems, as the name suggests, are systems aimed at aiding the driver of the vehicle. One example of a relatively simple DA system can be one that automatically turns on or off windshield wipers when water is detected. On the other hand, pedestrian detection is an example of an advanced DA system that is implemented in a few high-end automobiles. A variety of sensors such as microphones, Global Positioning System (GPS) devices, and radar allow the DA system to interface with the world; however, the research in this paper purely focuses on a specific camera-based application.

Camera-based lane change assistance (LCA) is an important area of automotive research and development. A vital element of LCA is the lane detection module. By continually monitoring the position of a car within a lane, collisions due to unintentional lane departure caused by driver distractions, fatigue, or driving under the influence of a controlled substance can be avoided.

Lane detection is a problem of locating painted lane markings or boundaries on the road surface with little knowledge of the road geometry [6]–[8]. Vision-based lane detection systems involve the use of one or more calibrated cameras looking out of the front windshield. Data in the form of video or an image sequence are acquired, after which they are analyzed to extract features that closely correspond to the desired lane markers. Lane detection is of much interest to many applications such as lane bobbing detection, lane departure warning (LDW), and blind spot monitoring as they strongly rely on a good-quality lane detector as a building block to make their assessment of the situation. In addition, it is one of the fundamental modules required by autonomous guided vehicles that enables them to independently drive on roads in urban environments [9]–[12].

Following the introduction, the methods commonly used to detect lanes markers and create ground truth, and details regarding database setup are described. Next, the new lane detection system is introduced, and its core elements are explained. This is followed by a detailed explanation of a new automated ground-truth generation process, and its benefits over the traditional approach are discussed. Then, the need for a publicly accessible database motivated by standardized data sets for testing is described. The proposed lane detector is evaluated using the annotations accompanying the videos used in testing. Quantitative analysis is then used to determine

Manuscript received June 9, 2010; revised March 3, 2011, June 23, 2011, and September 2, 2011; accepted September 17, 2011. Date of publication December 8, 2011; date of current version March 5, 2012. The Associate Editor for this paper was R. I. Hammoud.

A. Borkar is with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250 USA (e-mail: amol@gatech.edu).

M. Hayes is with the Graduate School of Advanced Imaging Science, Multimedia, and Film, Chung-Ang University, Seoul 156-756, Korea (e-mail: mhh3@gatech.edu).

M. T. Smith is with the Department of Communication Systems, School of Information and Communication Technology, Swedish Royal Institute of Technology, 114 28 Stockholm, Sweden (e-mail: msmith@kth.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

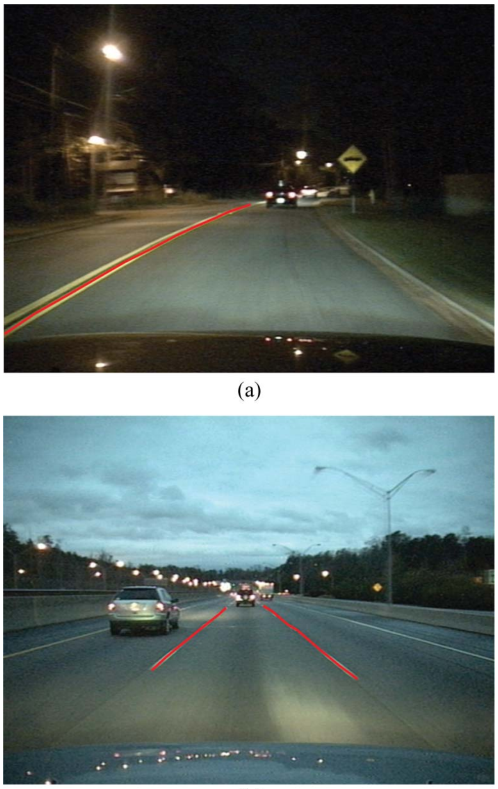
Digital Object Identifier 10.1109/TITS.2011.2173196

the accuracy and error in the performance of the lane detector. Finally, the conclusion and work remaining are addressed.

II. Prior Research

Lane detection is still a fertile area of machine vision research; consequently, many approaches have been proposed to accomplish this task. However, variations of the Hough transform are still among the most popular and commonly used methods. In these approaches, the input images are first preprocessed to find edges using a Canny edge detector [13] or steerable filters [14], followed by a threshold. The classical Hough transform is then used to find straight lines in the binary image, which often correspond to lane boundaries. The randomized Hough transform [15], which is a quicker and a more memory efficient counterpart of the classical Hough transform, has also been also been used for lane detection [16], [17]. The classical Hough transform for line finding works well when the roads are mostly straight; however, for curved roads, splines [18] and hyperbola fitting [19] are often used to provide support. A piecewise line fit shows some improvement by performing the Hough transform on sections of the road image to produce a curve and handles many problems regarding shadows and road patterns [20], [21]. Additionally, the incorporation of edge directions has also been used to remove some false signaling [22], [23]. Unfortunately, with the Hough transform, it is often difficult to determine if a line corresponds to an artifact or a lane boundary. In color segmentation approaches, RGB images are frequently converted to YCbCr, HSI, or custom color spaces. In these alternate color spaces, the luminance and chrominance components of a pixel are separately modeled. As a result, the effects of shadows and dynamic illumination in the color components can be greatly reduced. Consequently, the detection of colored objects such as yellow lane markers is often enhanced from this transformation [6], [24]. However, since these approaches operate at pixel level, they are often sensitive to change in ambient light color from street lights or similar illumination sources. The use of histograms to segment lane markers has been shown in [25]. However, some portion of a lane marker is required to be present in the horizontal bands for histogram calculations. Stereovision and 3-D have also been used in lane detection. In [26] and [27], stereo cameras are used to provide two vantage points of the road surface in hopes of improving the results over the single-camera approach. Specifically, lane markers are detected in each view, and then, the results are combined using epipolar geometry and camera calibration information. In [26] and [27], fixed search regions are used to find lane markers, assuming that the vehicle is always at the center of the lane. Learning methods such as artificial neural networks [10] and support vector machines [28] have also been used in lane detection; however, they may not perform well when road conditions that were not encountered in training are encountered [29]. Finally, a comprehensive literature review found in [30] summarizes most of the prominent lane detection techniques used today.

In addition to lane detection, generating ground truth information for testing the quality of any system is generally difficult. In lane detection applications, the most common approach is to individually mark the lane locations in every frame. This is the simplest procedure to produce the ground truth and works well for full or continuous lines, as shown in Fig. 1(a). However, it is an extremely slow process and can takes close to 1 min to annotate the lane locations in a single frame. Consequently, it could take a couple of months to complete the ground truth for 1 h of video. In addition, marking the ground truth locations in the gaps of broken or dashed lane markers becomes very difficult, as shown in Fig. 1(b). As generating this type of data is very time consuming and extremely tedious, it is commonly avoided.


(a)

(b)

Fig. 1. Marked ground truth locations. (a) Continuous lane. (b) Dashed or broken lanes.

A profound problem associated with testing lane detection systems is the access to common resources. Even though lane detection is a prime area of research in the automotive industry, most of the work appears to be largely proprietary. As a result, access to data sets used in development and testing can be difficult. Since the quality of data used in testing could be questionable, it becomes extremely difficult to validate the performance of a system. Consequently, in the absence of identical test data or standardized test sets, performing an unbiased assessment between the different systems becomes practically impossible.

Lane detection is a crucial component of many DA systems because it serves as the foundation for applications such as LDW, LCA, and autonomous driving. Existing lane detectors are generally restricted to operation under certain illumination and road conditions. In addition, without ground truth information, the lane detection results are qualitative and based purely on visual inspection by either one or multiple users. With visual inspection, problems of subjectivity arise where one user might grade the performance of a particular system better than another user. Another issue is the lack of standardized data sets that are used for testing. This can further lead to discrepancy in results when the lane detection systems are evaluated using nonidentical test data. Hence, it is necessary to develop a systematic framework that will coexist with existing lane detection systems and will be useful in performing unbiased assessments with the help of ground truth information and standardized test sets.

III. Methodology

A. Lane Detection

The layout of the new night-time lane detection system dubbed *Advanced Lane Detector 2.0* (ALD 2.0) is shown in Fig. 2. The original work of lane detection using a *Layered Approach* [1] or ALD 1.0 is improved with the incorporation of certain new features. These features include the following:

1) image transformations using inverse perspective mapping (IPM);
2) outlier removal with random sample consensus (RANSAC);
3) prediction and smoothing using the Kalman filter.

The following sections will detail each component of this improved system.

1) Preprocessing: First, temporal blurring is used to generate an *average image* to help connect dashed lane markers, giving them the appearance of a near continuous line [1]. The average image is formed as follows:

$$Average\ Image = \sum_{i=0}^{N} \frac{I(n-i)}{N} \qquad (1)$$

where $I(n)$ is the current frame, and i is the frame index offset. Since the commuting speed, steering angle, and GPS coordinate of the vehicle were not recorded at the time of data acquisition, the temporal blurring does not incorporate ego dynamics.

Since detecting dashed lane markers in a single image can be difficult due to their sparseness, the enhancement provided by the temporal blurring shown in Fig. 3(b) allows these dashed lane markers to appear as connected lines, making them easier to detect. Temporal blurring does not significantly alter the perceived width of the lane markers as only a few frames from the past are used in the averaging. This is also supported by the fact that the lateral moving speed of the vehicle e.g., during a normal lane change, is usually much slower than its commuting speed. As a result, the lane markers appear to less laterally stretch. A standard pixelwise transformation [31] is then used to transform the average image into a grayscale image.

2) IPM: One of the improvements of ALD 2.0 over the previous lane detector was the incorporation of the IPM [32]. IPM is used to transform the captured images from a camera perspective to a bird's-eye view. A forward-looking camera will capture images in which lane markers are not parallel and have line widths that vary as a function of the distance from the camera. In Fig. 4(a), a side view or a view along the Y-axis shows the camera configuration along with its related calibration parameters, and $\widehat{o}$ is a unit vector representing the camera's optical axis. Similarly, in Fig. 4(b), the camera configuration is shown along the Z-axis or top view. With IPM, lane detection now becomes a problem of locating a pair of near-constant-width parallel lines that are generally separated by a predetermined fixed distance. In addition, IPM also enables a mapping between pixels in the image plane to world coordinates (feet) [33]. This intrinsic property of the IPM image is particularly useful as a fixed-size template can be used in template matching. The road surface is also assumed to be flat in the vicinity of the vehicle.

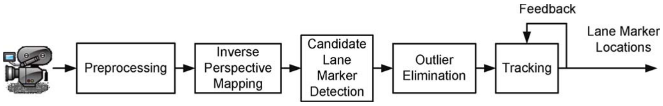
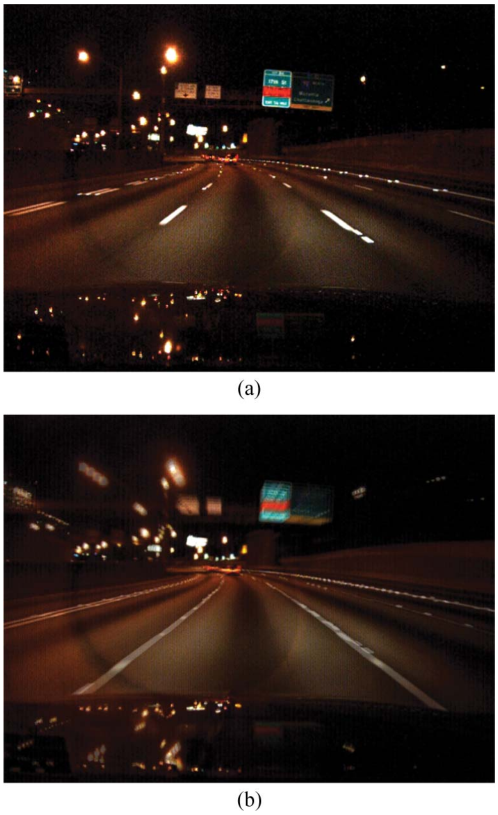


Fig. 2. Design model of the *Advanced Lane Detector 2.0*.

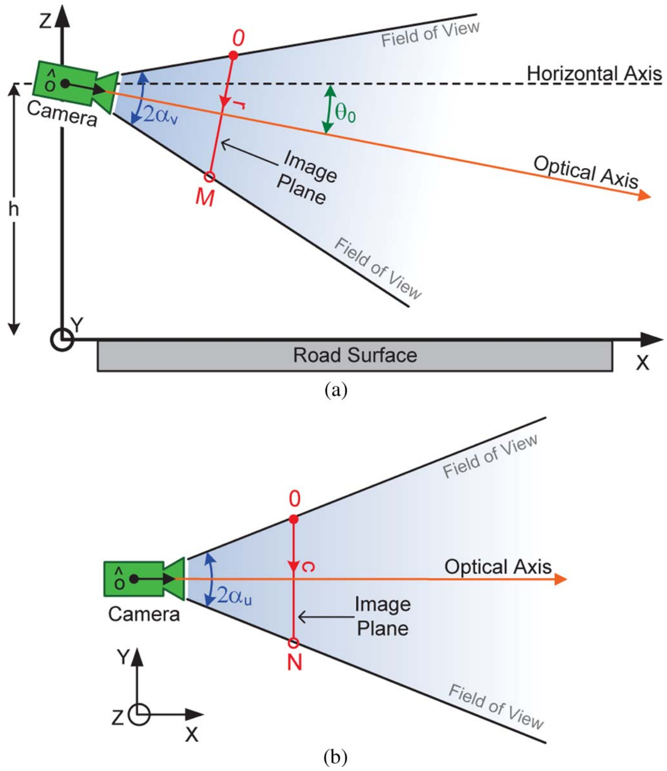

(a)

(b)

Fig. 3. Intermediate results of preprocessing. (a) Original image. (b) Average image.

Algorithm 1 Adaptive Threshold

```
$minrange = 255/10$
 /*Minimum likely difference between the lane markers and the road*/
**for** (Each Pixel) **do**
  $P0$ = Actual pixel value
  $Q0$ = New pixel value after threshold
  min = Minimum value of neighborhood around P0
  max = Maximum value of neighborhood around P0
  $range = \max - \min$;
  **if** ($range > minrange$) **then**
```



(a)

(b)

Fig. 4. Camera configuration. (a) Along the Y-axis. (b) Along the Z-axis.

```
        $T = (\min + \max)/2$;
    else
        $T = \max - minrange/2$;
    end if
    if $P0 > T$ then
        $Q0 = 255$;
    else
        $Q0 = 0$;
    end if
end for
```

3) Candidate Lane Marker Detection: The bird's-eye view image is converted into a binary image using the adaptive threshold shown in Algorithm 1 [1], [34]. The bottom half of the binary image is shown in Fig. 5(a). Using the low-resolution Hough transform, the ten highest scoring lines are found in the half-image [1], as shown in Fig. 5(b). Each of these lines is sampled at the *Sampling Columns*, as shown in green in Fig. 5(c). For illustration, the sampled points of only one line are shown in Fig. 5(d). The corresponding sample points are then found in the grayscale image, as shown in Fig. 5(e).

The transition from dark to light to dark intensity values along the yellow window in Fig. 5(f) would ideally produce a rectangular pulse; however, temporal smoothing in the average image leads to a Gaussian-like shape in one dimension [1], [33]. Consequently, template matching with normalized cross correlation is performed using a collection of predefined templates inside the search windows centered at each sample point, as shown in Fig. 5(g). In this particular example, only narrow lane markers are present in Fig. 5(g). However, the Federal Highway Administration (FHA) has a large catalog of lane marker specifications, where each has a unique purpose. However, the narrow, wide, and double lane markers are the most common, occurring on the interstate and city roads of Atlanta, GA. As a result, templates are created only for these three markers. The FHA states that narrow lane markers are 6 in wide, wide markers are close to 10 in wide, and double lane markers consist of two narrow markers with a discernable gap in between [35]. The templates for narrow and wide markers are created using Gaussian kernels, each with a variance such that two standard deviations equal 6 and 10 in, respectively, in the IPM image. For double lane markers, two narrow lane-marker-type Gaussians with means that are separated by 5 in are used to create the template. Finally, the pixel location inside the search window corresponding to the largest correlation coefficient among the three templates is selected as the best estimate of the center of the lane marker.

The matching process using the collection of templates is repeated on the remaining nine lines. The pixel with the largest correlation coefficient in each *Sampling Column* that also exceeds a minimum threshold is selected as the best estimate of the center of the lane marker, as shown in Fig. 5(h). After acquiring the correlation coefficients of several positive and negative training examples, the threshold was chosen as the point of maximum accuracy on the ROC curve.

4) Outlier Elimination and Data Modeling: After finding a set of lane marker candidates, the RANSAC algorithm is performed to eliminate outliers. The RANSAC algorithm uses an iterative method to estimate parameters for fitting a mathematical model to a set of observed data that may contain outliers [36], [37]. After rejecting the outliers, linear least-squares estimation (LSE) is used to find the best fit of a line to the inliers. The fitted line is then parameterized in terms of ρ and θ, where ρ is the distance from the origin (top-left corner pixel) to the line and θ is the angle as shown in Fig. 6 (generally is close to 90°). Hyperbola and quadratic models are often used to produce smooth curves to portray lane boundaries; however, minor perturbations significantly affect the computation of the LSE. Although a straight line is more rigid, its orientation is less affected by these minor perturbations.

5) Tracking: Finally, a Kalman filter is used to track and smooth the estimates of ρ and θ based on the measurements. While tracking, if lane markers are intermittently not detected, then the Kalman filter relies on its prediction to produce estimates. However, if the lane markers go undetected for more than a few seconds, then tracking is disabled until the next detection. This is to avoid producing incorrect estimates when markers may not present be on the road.

B. Ground Truth Generation

1) Procedure: Marking the individual lane marker locations in a camera-captured image is a common approach to generate ground truth for a particular frame. However, it is a very slow process in the absence of any sort of automation or assistance. Consequently, generating the ground truth for a few hours of video footage could take several weeks to complete. Therefore, a technique to efficiently generate ground truth is essential. With *time-sliced (TS)* images, this process can be sped up [2].

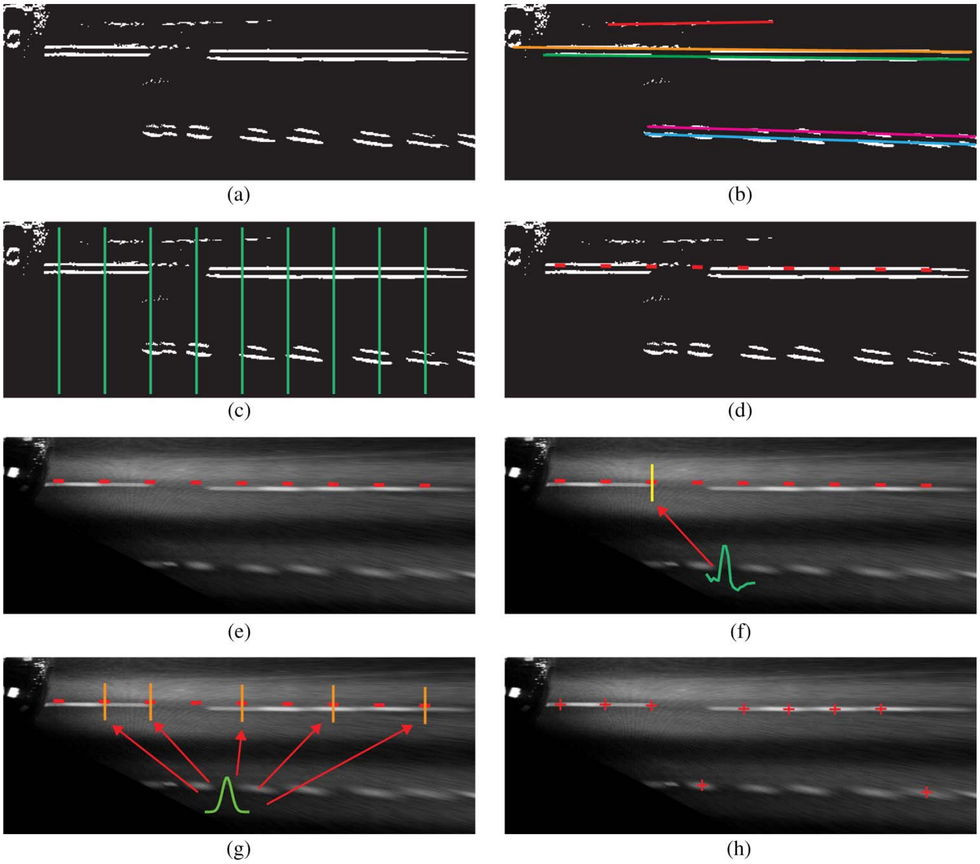


Fig. 5. Intermediate stages for estimating ideal lane marker locations. (a) Bottom half binary image. (b) Best fitting lines (shown in different colors). (c) Specific columns where the sampling occurs. (d) Line sampled along its length. (e) Correspondence in the grayscale image. (f) Gaussian-like shape for the 1-D intensity profile of a lane marker within the yellow window. (g) Search windows centered at sample points for template matching. (h) Best lane marker estimates within each column.

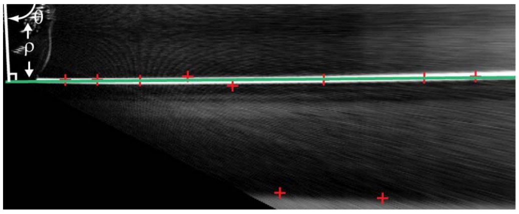


Fig. 6. Model fitting and line parametrization.

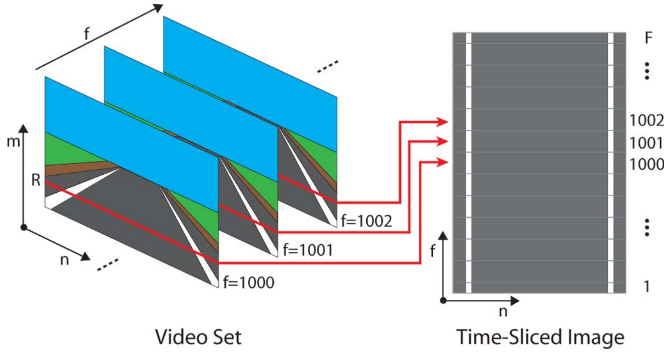


Fig. 7. Rows of pixels from a video set are stacked to create a TS image.

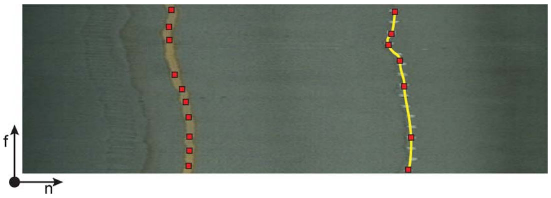


Fig. 8. Portion of an annotated TS image. The red boxes represent points marked by the user. The yellow line represents the interpolated values between the control points.

A TS image is created by stacking a specific row of pixels from each frame in a video sequence. For example, consider a video sequence consisting of F frames, where each frame $f = 1, 2, \ldots, F$ is an $M \times N$ image. Slicing through the stack of F frames in time at a specific row, e.g., row R, generates an $F \times N$ TS image, as shown in Fig. 7. The rows in the TS image are indexed by f, which is the frame number, since moving from one row to the next is equivalent to traversing through the video sequence in time.

Creating the ground truth then begins by generating two or more TS images from different rows in the video sequence. For example, Fig. 9 shows three TS images that are generated from rows 300, 340, and 370 in the video sequence. A user then establishes ground truth points by marking the centers of the lane markers at a number of points in each TS image, as shown in Fig. 8. In this figure, the red boxes represent ground truth points marked by the user, and the yellow line represents the points that are generated by interpolation between the marked points. Cubic spline is the preferred interpolation method since it is C^2 continuous and produces a smooth curve between each

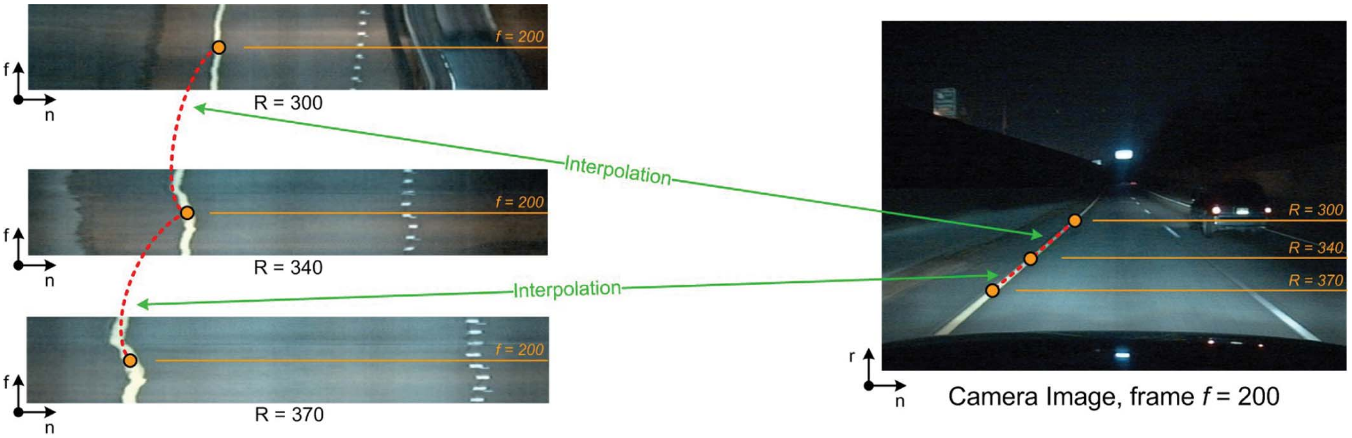


Fig. 9. Overview of the automated ground truth generation process.

control point. These interpolated values represent interpolated lane marker locations for frames that are between two marked points, and if the interpolation is not precise enough, more ground truth points may be added.

After each TS image has been sufficiently annotated, the ground truth for each frame is then created as follows: First, a row in the TS images is selected, which is equivalent to selecting a specific frame in the video sequence. For example, row 200 in TS image $R = 300$ corresponds to row 300 in frame $f = 200$ in the video sequence and similarly for TS images $R = 340$ and $R = 370$. The orange circles marked in row 200 of the TS images in Fig. 9 are either user marker points or spline interpolated points, and these points are directly mapped to points in frame 200 shown in the camera image in Fig. 9 (right). A second cubic spline interpolation is then performed between the points in the camera image to *fill in* the lane marker locations between the rows in the camera image, as indicated by the dotted red line in Fig. 9. This process is repeated for all frames of the video sequence.

The automation provided by interpolation greatly helps in improving efficiency in ground truth creation and, at the same time, reduces the requirement of detailed input from the user.

2) Error Calculation: The FHA states the official width of the lane markers as 6 in [35]. In an IPM image, this width is translated to a distance in pixel units and used as an interval around the ground truth locations. Consequently, lane marker estimates that fall within this interval are categorized as having no error. As a result, the error $E(f)$ in each frame is computed as the average distance between the ground truth locations and the estimated lane markers for all defined rows in the IPM image, i.e.,

$$\lambda_{(i,f)} = \mathbf{max}\left(\left|Gt_{(i,f)} - X_{(i,f)}\right| - \frac{W}{2}, 0\right) \qquad (2)$$

$$E(f) = \sum_{i=1}^{N} \frac{\lambda_{(i,f)}}{N} \text{ ft.} \qquad (3)$$

In (2), $Gt_{(i,f)}$ is the ground truth location of the lane marker, and $X_{(i,f)}$ is its estimate on row i of frame f. W is the width of the interval around the ground truth location, and λ is the measured distance. Based on the FHA specifications, W is set to the number of pixels that represent 6 inches in the IPM image. The λ values are measured in the IPM image instead of in the camera-captured images as interpixel distances remain uniform in the perspective-free image [2], [33].

TABLE I
EVALUATION OF THE AUTOMATED GROUND TRUTH

Type of Test	Avg. E(f) ft.
Straight Road	0.167
Curving Road	0.182
Lane Change	0.189

3) Automated Ground Truth Evaluation: Since the generation of the ground truth data is largely automated with the help of *time slicing*, it is essential to compare its accuracy to reference ground truth data that are manually created in its entirety. A few minutes of video consisting of driving on straight roads, curving roads, and lane changes were used in the evaluation. Tests were performed by marking the ground truth locations of lane markers in every frame of the video clips and computing $E(f)$. In these tests, X in (2) is the lane marker estimate calculated using the automated ground truth process, and W is set to 0 to determine the absolute error generated as a result of the automation in ground truth generation. After extensive testing, it was determined that an $E(f) < 0.25$ implies that the estimates are within a 3 in of the ground truth, portraying good accuracy. In addition, the automated process is not restricted to specific driving scenarios.

Table I shows the errors produced by the automated process when used to generate the ground truth on different types of roads.

C. Database

Although lane detection is a research area of interest to the automotive industry, the research tends to be proprietary, and the results and resources may not be publicly accessible. These resources mainly consist of hardware specifications, data sets used in development and testing, and source code for implementation details. Unlike research areas such as face recognition or character recognition where standardized data sets are readily available for training and testing, lane detection

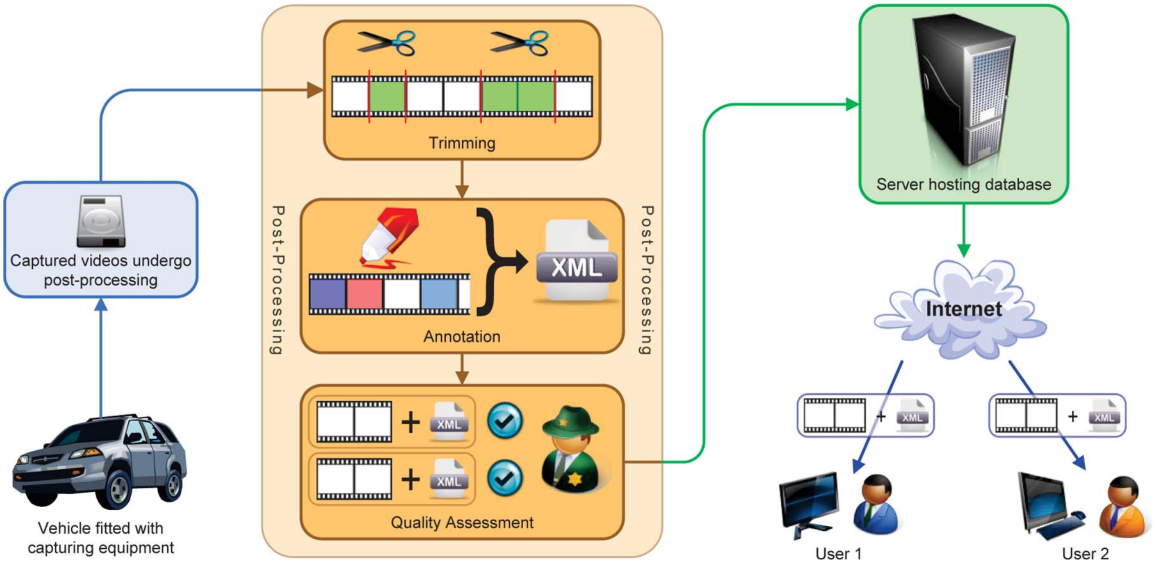


Fig. 10. Overview of the database operations.

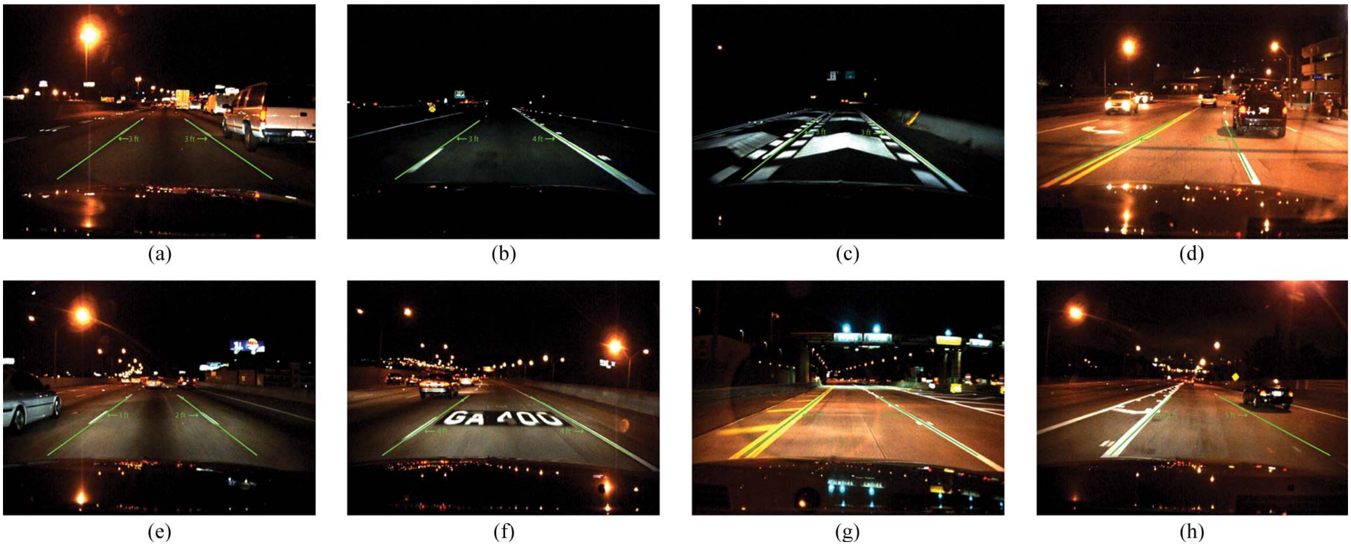


Fig. 11. Examples of correct lane detection. (a) Normal highway. (b) Dark highway. (c) Road patterns. (d) Urban street. (e) Other vehicles present. (f) Navigational information. (g) Toll plaza. (h) On-ramp.

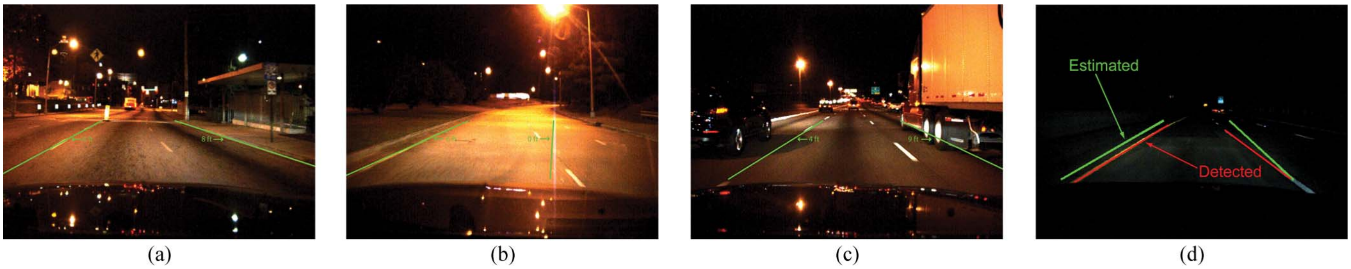


Fig. 12. Examples of missed and incorrect lane detection. (a) Worn markers and cracks. (b) Lens flare. (c) Step ramp on truck. (d) Effects of bumps.

appears to have no such resources. Consequently, it is nearly impossible to perform a fair comparison between the different algorithms. Therefore, it becomes extremely difficult to validate the performance of an algorithm as the data used in testing could be questionable in terms of its neutrality. Therefore, to remove some of the issues regarding bias in performance evaluation, a lane detection video database is being made available.

A test vehicle equipped with a camera installed under the rear-view mirror is used to record videos of the roadways. Videos are recorded in 24-bit RGB in VGA (640 × 480)

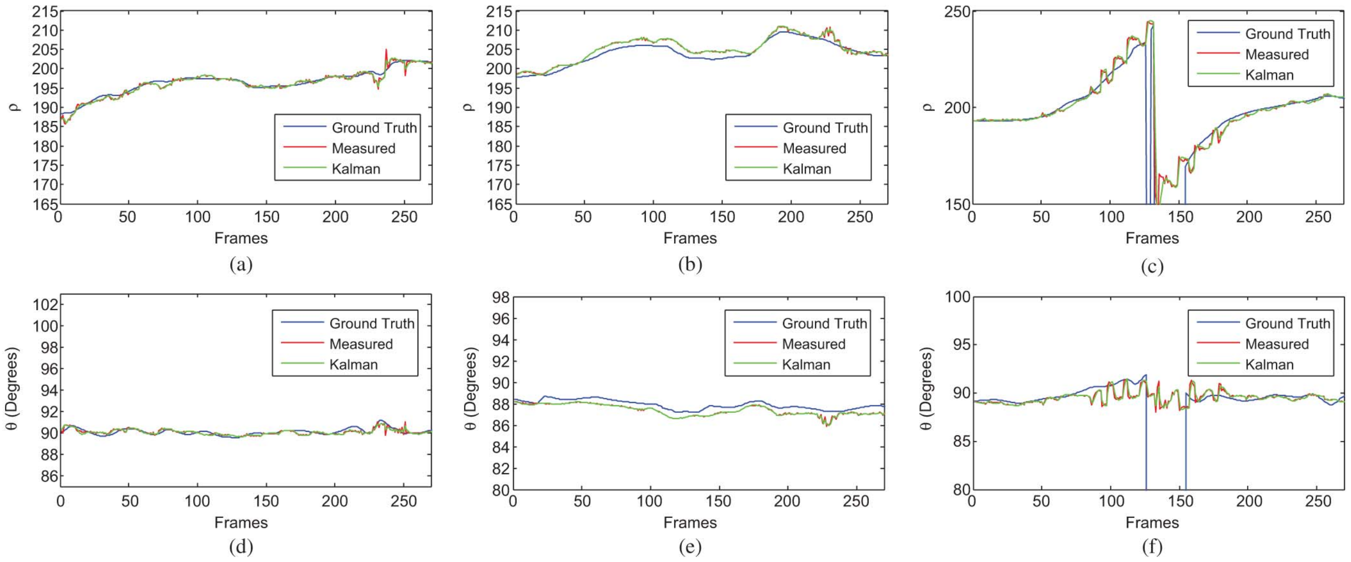


Fig. 13. Comparison between the ρ and θ values on straight roads, on curving roads, and during a lane change. The first row shows the changes in ρ values, and the second row shows the changes in θ values. The blue line represents the ground truth, the red line represents the measured values, and the green line represents estimates determined by the Kalman filter. (a) Straight road. (b) Curving road. (c) Lane change (d) Straight road. (e) Curving road. (f) Lane change.

resolution at 30 frame/s. The data are captured as a sequence of images and loselessly stored in Portable-Network-Graphics format. Additionally, each file is tagged with a millisecond time stamp and commuting speed of the vehicle.

The videos stored in the database are captured with the intention of generating a realistic set of conditions that one would encounter while driving in the real world. To make this possible, recording sessions were conducted at various hours during the day and night to portray diversity under illumination and traffic conditions. Additionally, videos are recorded on a number of interstate highways and city streets to generate variety in the road surface textures.

After each recording session, the captured video undergoes a set of tasks before it is made available for training and testing.

1) Recorded videos undergo postprocessing in the form of trimming and annotation.

   a) Trimming: keeping only relevant segments from the recorded video clip;
   b) Annotation: marking ground truth in the video segments.

   Annotation data are stored as Extensible Markup Language (XML) files.

2) Video clips and annotations are evaluated for quality and accuracy.
3) After quality assessment, the files are checked into the database and made available for download.

This idea is clearly shown in Fig. 10. The primary purpose of creating such a database is to attempt to standardize the data sets that are used for testing and evaluating different systems. The database is located at `ftp://ftpx.ece.gatech.edu/Db_Mhh_Aab/LaneData`, and login information can be acquired by contacting the authors.

IV. Results

Three rules were used to quantify the results into different categories.

1) Correct detection occurs when at least $N/2$ lane marker estimates are within the ground truth interval (W).
2) Missed detection occurs when more than $N/2$ lane marker estimates are outside the ground truth interval (W).
3) Incorrect or false detection occurs when no lane marker is present, yet one is detected.

Here, N is the total number of λ values used in error calculation in (3). Despite the presence of navigational text, cracks, tar patches, and skid marks on the road, the proposed system is still able to accurately locate lane markers, as shown in Fig. 11. In addition, the distance to each lane marker relative to the host vehicle is also calculated.

A few instances of missed and incorrect lane detections are shown in Fig. 12. The most common cause of errors was the absence of lane markers due to age and wear on local city roads. This occasionally leads to the detection and tracking of false signals such as cracks, tar patches, and lens flares. We observed that the camera pitch was not affected by normal vibrations encountered while driving. The few times that the camera pitch was modified were when small bumps were encountered or at points where there was a significant change in surface inclination. In these cases, the modifications were temporary, and only in a few cases did it affect the detection of lane markers, e.g., in Fig. 12(d). However, the Kalman filter does help in smoothing out these bumps. Fig. 13 shows a comparison between ground truth, measurement and Kalman filter estimates for ρ and θ on different road conditions. The measurement and Kalman estimates showed some error during certain portions of a lane change; however, the Kalman estimated ρ and θ were very close to the ground truth at most times.

TABLE II
QUANTITATIVE ANALYSIS OF THE ADVANCED LANE DETECTOR 2.0

Road Type	Traffic	Avg. Rate Per Minute			
		Correct	E(f) ft.	Missed	Incorrect
Isolated Highway	Light	98.35%	0.032	1.64%	0.31%
	Moderate	97.94%	0.061	2.05%	0%
Metro Highway	Light	98.47%	0.14	1.53%	0.41%
	Moderate	97.77%	0.21	2.22%	0.55%
City	Variable	87.21%	1.49	12.79%	6.61%

Table II shows the performance of the ALD 2.0. The results are quantified in terms of accuracy per minute, which allows normalization of the results when data are captured using cameras with different frame rates. The data set used for testing consisted of close to 30 min of videos of driving scenarios on local city roads and highways. The detection rates of left and right markers were averaged to produce the numbers in Table II.

As expected, the ALD 2.0 correctly displayed overall good performance in detecting and classifying the lane markers. Some of the elements that played a key role in its performance were the following:

1) adaptive local threshold;
2) correlation threshold.

The adaptive local threshold used in both ALD 1.0 and 2.0 modifies its threshold value locally based on neighborhood pixel intensities, which, in turn, allows detection of weak edges [1], [34]. This approach tends to work better than the popular Canny edge detector, which uses an adaptive global threshold to convert the gradient edge image to binary [13]. The global threshold may prohibit the detection of weak edges like worn-out lane markers in the presence of dominant edges in the image.

In addition, the predetermined threshold enables the template-matching stage to further verify the candidacy of a pixel as part of a lane marker and helps in rejecting noise [1], [33]. Methods that rely on voting or finding highest score are unable to reliably tell if a binary pixel belongs to an actual lane marker or it is noise. Finally, the error values of $E(f) < 0.25$ implies that the lane marker estimates are, in fact, very close to the ground truth locations.

V. Conclusion

The *Advanced Lane Detector 2.0*, which is an extension of the original ALD 1.0 [1], has been presented in this paper. The integration of IPM for image transformation, RANSAC for outlier removal, and Kalman filtering for prediction and tracking into the ALD 1.0 has shown innovation in technology and considerable performance gain. The current Matlab implementation operates at approximately 0.8 s/frame on current Intel-based personal computer hardware. In addition, a new and efficient process to generate ground truth has been presented. With the introduction of time slicing, the process benefits from automation via spline interpolation, which greatly reduces the time spent to produce ground truth [2]. Finally, the proposed implementation of the database will greatly help in performing unbiased tests as data sets will be available for download; consequently, different algorithms can evaluated on identical data.

The videos from the database that were used for testing have been recorded on interstate highways and city streets in and around Atlanta, GA. The results are quantitatively deduced using set categories and error calculations rather than visual inspection as a result of the available ground truth information. Even in the presence of a diversity of road surface artifacts such as cracks, navigational text, skid marks, and tar patches, the ALD 2.0 was able to show high accuracy, as portrayed in Table II.

VI. Future Work

Speed, GPS location, and steering angle data acquired from the vehicle during a recording session will be made available in the database, a well as used in estimating ego motion. The temporal blurring algorithm will be made adaptive by incorporating ego motion. In addition, ego motion could be used to provide redundant information regarding lane marker locations and improve the existing lane detector. Lane borders will be modeled using quadratic or parametric curves to provide a more accurate representation. Changes in camera pitch while maneuvering and its effects on lane detection will be investigated in a controlled environment. Image-processing techniques will also be used to improve the ground truth annotations by accurately determining the exact widths of the lane markers rather than using a static interval. Additional investigations will be conducted to implement daytime lane detection, along with operability during inclement weather conditions.

References

[1] A. Borkar, M. Hayes, M. Smith, and S. Pankanti, “A layered approach to robust lane detection at night,” in *Proc. IEEE Workshop Comput. Intell. Vehicles Veh. Syst.*, 2009, pp. 51–57.

[2] A. Borkar, M. Hayes, and M. T. Smith, “An efficient method to generate ground truth for evaluating lane detection systems,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2010, pp. 1090–1093.

[3] A. M. Muad, A. Hussain, S. A. Samad, M. M. Mustaffa, and B. Y. Majlis, “Implementation of inverse perspective mapping algorithm for the development of an automatic lane tracking system,” in *Proc. IEEE TENCON Region 10 Conf.*, 2004, pp. 207–210.

[4] M. Amemiya, K. Ishikawa, K. Kobayashi, and K. Watanabe, “Lane detection for intelligent vehicle employing omni-directional camera,” in *Proc. SICE Annu. Conf.*, 2004, vol. 3, pp. 2166–2170.

[5] Nat. Highway Traffic Safety Admin., [Online]. Available: Fatality Analysis Reporting System (FARs) Encyclopedia, 2008. [Online]. Available: http://www-fars.nhtsa.dot.gov/Main/index.aspx

[6] T. Y. Sun, S. J. Tsai, and V. Chan, “HSI color model based lane-marking detection,” in *Proc. IEEE Conf. Intell. Transp. Syst.*, 2006, pp. 1168–1172.

[7] K. Y. Chin and S. F. Lin, “Lane detection using color-based segmentation,” in *Proc. IEEE Intell. Vehicles Symp.*, 2005, pp. 706–711.

[8] Q. Li, N. Zheng, and H. Cheng, “An adaptive approach to lane markings detection,” in *Proc. IEEE Conf. Intell. Transp. Syst.*, 2003, pp. 510–514.

[9] S. Kammel and B. Pitzer, “Lidar-based lane marker detection and mapping,” in *Proc. IEEE Intell. Vehicles Symp.*, 2008, pp. 1137–1142.

[10] D. Pomerleau, “ALVINN: An autonomous land vehicle in a neural network,” in *Advances in Neural Information Processing Systems*. San Mateo, CA: Morgan Kaufmann, 1989.

[11] S. G. Jeong, C. S. Kim, D. Y. Lee, S. K. Ha, D. H. Lee, M. H. Lee, and H. Hashimoto, “Real-time lane detection for autonomous vehicle,” in *Proc. IEEE Int. Symp. Ind. Electron.*, 2001, pp. 1466–1471.

[12] Y. Shu and Z. Tan, "Vision based lane detection in autonomous vehicle," in *Proc. World Congr. Int. Control Autom.*, 2004, vol. 6, pp. 5258–5260.

[13] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[14] J. McCall, D. Wipf, M. Trivedi, and B. Rao, "Lane change intent analysis using robust operators and sparse Bayesian learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 431–440, Sep. 2007.

[15] L. Xu, E. Oja, and P. Kultanen, "A new curve detection method: Randomized hough transform (RHT)," *Pattern Recognit. Lett.*, vol. 11, no. 5, pp. 331–338, 1990.

[16] A. T. Saudi, J. Hijazi, and J. Sulaiman, "Fast lane detection with randomized hough transform," in *Proc. Symp. Inf. Technol.*, 2008, vol. 4, pp. 1–5.

[17] J. Wang, Y. Wu, Z. Liang, and Y. Xi, "Lane detection based on random hough transform on region of interesting," in *Proc. IEEE Conf. Inform. Autom.*, 2010, pp. 1735–1740.

[18] M. Aly, "Real time detection of lane markers in urban streets," in *Proc. IEEE Intell. Vehicles Symp.*, 2008, pp. 7–12.

[19] O. Khalifa, A. Assidiq, and A. Hashim, "Vision-based lane detection for autonomous artificial intelligent vehicles," in *Proc. IEEE Int. Conf. Semantic Comput.*, 2009, pp. 636–641.

[20] T. Taoka, M. Manabe, and M. Fukui, "An efficient curvature lane recognition algorithm by piecewise linear approach," in *Proc. IEEE Veh. Technol. Conf.*, 2007, pp. 2530–2534.

[21] Q. Truong, B. Lee, N. Heo, Y. Yum, and J. Kim, "Lane boundaries detection algorithm using vector lane concept," in *Proc. Conf. Control, Autom., Robot. Vis.*, 2008, pp. 2319–2325.

[22] J. W. Lee and J. Cho, "Effective lane detection and tracking method using statistical modeling of color and lane edge-orientation," in *Proc. Conf. Comput. Sci. Convergence Inf. Technol.*, Nov. 2009, pp. 1586–1591.

[23] J. Gong, A. Wang, Y. Zhai, G. Xiong, P. Zhou, and H. Chen, "High speed lane recognition under complex road conditions," in *Proc. IEEE Int. Vehicles Symp.*, 2008, pp. 566–570.

[24] H. Cheng, B. Jeng, P. Tseng, and K. Fan, "Lane detection with moving vehicles in the traffic scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 571–582, Dec. 2006.

[25] J. P. Gonzalez and U. Ozguner, "Lane detection using histogram-based segmentation and decision trees," in *Proc. IEEE Conf. Intell. Transp. Syst.*, 2000, pp. 346–351.

[26] N. Benmansour, R. Labayrade, D. Aubert, and S. Glaser, "Stereovision-based 3D lane detection system: A model driven approach," in *Proc. IEEE Conf. Intell. Transp. Syst.*, 2008, pp. 182–188.

[27] S. Nedevschi, R. Schmidt, T. Graf, R. Danescu, D. Frentiu, T. Marita, F. Oniga, and C. Pocol, "3D lane detection system based on stereovision," in *Proc. IEEE Conf. Intell. Transp. Syst.*, 2004, pp. 161–166.

[28] Z. Kim, "Robust lane detection and tracking in challenging scenarios," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 1, pp. 16–26, Mar. 2008.

[29] J. C. McCall and M. M. Trivedi, "An integrated, robust approach to lane marking detection and lane tracking," in *Proc. IEEE Int. Vehicles Symp.*, 2004, pp. 533–537.

[30] J. McCall and M. Trivedi, "Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 20–37, Mar. 2006.

[31] G. Hoffmann, Luminance models for the grayscale conversion, Mar. 2002. [Online]. Available: http://www.fho-emden.de/_hoffmann/gray10012001.pdf

[32] M. Bertozzi and A. Broggi, "GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 62–81, Jan. 1998.

[33] A. Borkar, M. Hayes, and M. Smith, "Robust lane detection and tracking with ransac and kalman filter," in *Proc. IEEE Int. Conf. Image Process.*, 2009, pp. 3261–3264.

[34] E. R. Davies, *Machine Vision : Theory, Algorithms, Practicalities*, 3rd ed. San Mateo, CA: Morgan Kaufmann, Dec. 2004.

[35] Fed. Highway Admin., Manual Uniform Traffic Control Devices, Nov. 2009. [Online]. Available: http://mutcd.fhwa.dot.gov/

[36] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, Apr. 2004.

[37] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

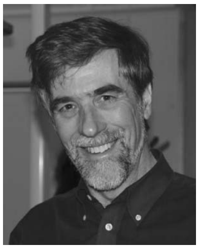
Amol Borkar received the B.S. degree in computer engineering from North Carolina State University, Raleigh, in 2004 and the M.S. degree in electrical engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, where he is currently working toward the Ph.D. degree with the School of Electrical and Computer Engineering. His Ph.D. dissertation focuses on developing novel driver assistance systems for vehicles using a single- or multicamera approach.

He is currently a Research Assistant with the Center for Signal and Image Processing, Department of Electrical and Computer Engineering, Georgia Tech. He has multidisciplinary research experience in the fields of automotive vision, face detection and recognition, and medical imaging and has spent eight months at IBM's Watson Research Center, NY, working on various computer vision projects. His research interests are applied computer vision, image processing, pattern recognition, and machine learning.

Monson Hayes (F'92) received the B.S. degree from the University of California, Berkeley, in 1971 and the D.Sc. degree in electrical engineering and computer science from Massachusetts Institute of Technology, Cambridge, in 1981.

He was a Systems Engineer with Aerojet Electrosystems until 1974. He then joined the faculty with the Georgia Institute of Technology (Georgia Tech), Atlanta, where he is currently a Professor of electrical and computer engineering. He is currently an Associate Chair with the School of Electrical and Computer Engineering, Georgia Tech, and an Associate Director for Georgia Tech Savannah. Since joining the faculty at Georgia Tech, he has become internationally recognized for his contributions to the field of digital signal processing, image and video processing, and engineering education. He is currently a Distinguished Foreign Professor with the Graduate School of Advanced Imaging Science, Multimedia, and Film, Chung-Ang University, Seoul, Korea. He has published more than 150 papers. He is the author of two textbooks. His research interests are face recognition, image and video processing, adaptive signal processing, and engineering education.

Dr. Hayes has served the IEEE Signal Processing Society in numerous positions, including as Chairman of the Digital Signal Processing Technical Committee (1995–1997), Associate Editor for the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ASSP) (1984–1988) and Associate Editor for the IEEE TRANSACTIONS ON EDUCATION (2000–2010), Secretary–Treasurer of the ASSP Publications Board (1986–1988), Chairman of the ASSP Publications Board (1992–1994), General Chairman of the 1996 IEEE International Conference on ASSP, and General Chairman of the 2006 International Conference on Image Processing. He has received numerous awards and distinctions from professional societies, as well as from Georgia Tech, e.g., the Presidential Young Investigator Award.



Mark T. Smith received the Ph.D. degree in bioengineering from the University of Utah, Salt Lake City, in 1983.

He is currently a Professor of information technology product development with the Department of Communication Systems, School of Information and Communication Technology, Swedish Royal Institute of Technology (KTH), Stockholm, Sweden. He joined the faculty of the KTH in October 2005 and is the Coordinator of the KTH Masters Program in Design and Implementation of IT Products and Systems. This program focuses on how an ever-increasingly broad array of electronics, computer architecture, embedded systems, and communication technologies are integrated together into products, systems, and solutions. He spent more than 20 years at Hewlett Packard Laboratories, Palo Alto, CA. There, he performed applied research in a large number of IT areas, including fixed and mobile information devices, streaming media systems, context-aware technologies, and sensors. His current research interests are architectures for future Information Communications Technology devices, context information measurement, and data fusion technologies, particularly for use in mobile systems and enterprise applications.