On Performance Evaluation Metrics for Lane Estimation

Ravi Kumar Satzoda and Mohan M. Trivedi
Laboratory for Intelligent and Safe Automobiles, University of California San Diego, La Jolla, CA-92093
Email: rsatzoda@eng.ucsd.edu, mtrivedi@ucsd.edu

Abstract—Accurate and efficient lane estimation is a critical component of active safety systems in automobiles such as lane departure warning systems. A substantial number of lane estimation methods have been proposed and evaluated in literature. However, a common set of evaluation metrics that assess different components of lane estimation process has not been addressed. This paper proposes a set of performance evaluation metrics for lane estimation process, that can be deployed to evaluate different kinds of lane estimation algorithms. Evaluation by applying the proposed metrics is demonstrated using a recent lane estimation method.

I. INTRODUCTION

Active safety using Intelligent Driver Safety Systems (IDSS) is increasingly becoming a vital component of modern automobiles [1]. Visual perception of the road and lanes is considered to be the first step in most vision-based IDSS such as lane departure warning, lane change assistance, lane keeping [2]–[4]. Additionally, lanes are also been recently used in detecting obstacles more efficiently. For example [5], [6] have used lanes to detect vehicles more accurately. Considering that lanes play a significant role in most IDSS, there is a substantial volume of literature on lane estimation and related applications such as lane departure detection etc. for structured or urban roads [3], [4], [6]–[10]. A detailed survey of these algorithms can be found in [3], [7].

Lane estimation involves three main steps [7], [11]. In the first step, lane features are extracted from the image scene using operators such as steerable filters [7], Gabor filters [10] etc. Thereafter an outlier removal process is applied using road models and techniques such as RANSAC [8] etc. to eliminate incorrect lane features. Finally, the selected lane features are used to fit a lane model. Additionally, lane tracking is applied that aids in outlier removal process. Fig. 1 shows the typical computational steps in lane estimation techniques.

Existing literature on lane estimation shows that lanes are robustly extracted in varying complex road scenarios using different combinations of the above steps. Although each contribution in this topic evaluates the proposed techniques for different performance metrics, an objective assessment of lane estimation process is still not adequately addressed.

In this paper, we propose and present different kinds of performance metrics that can be used to evaluate the different aspects of a lane estimation technique. The proposed metrics will be shown to quantify the performance of a lane estimation technique in terms of robustness (at multiple levels), applicability and computational complexity. The proposed metrics are also demonstrated using one of the recent lane analysis algorithms. The rest of the paper is organized as follows. We present a brief survey of existing metrics for lane estimation in Section II. In Section III, the different performance metrics are presented in detail. An evaluation of the lane analysis method in [12] using the proposed metrics is presented in Section IV before concluding in Section V.

![Fig. 1. Typical steps in lane estimation](image)

II. RELATED WORK

Unlike [3], [7] which present a detailed survey of lane analysis techniques, we briefly survey the performance evaluation metrics that have been used in recent lane estimation studies.

| TABLE I. SURVEY OF PERFORMANCE EVALUATION IN RECENT LANE ANALYSIS STUDIES |
|-----------------|-----------------|-----------------|
| Metrics & Evaluation | 2006 | 2007 | 2012 |
| McCall [7] | Rate of change of position | Lane detection in varying road scenarios (visual outputs) |  |
| Cheng [10] | Ego-vehicle position with respect to lane/center of the lane | Lane detection in varying road scenarios (visual outputs) | Lane detection in varying road scenarios (visual outputs) |
| Borkar [8] | Lane position error estimated as \( (p, \theta) \) tuple from Hough transform | Lane detection in varying road scenarios (visual outputs) | Lane detection in varying road scenarios (visual outputs) |
| Gopalan [9] | Detection accuracy of detected lane markings with varying neighborhood sizes around ground truth | Ego-vehicle position with respect to lane/center of the lane for omni and rectilinear cases | Lane detection in varying road scenarios (visual outputs) |
| Nedevschi [13] | Qualitative evaluation of lane identification by checking Bayesian network probabilities | Lane detection in varying road scenarios (visual outputs) | Lane detection in varying road scenarios (visual outputs) |
| Srivastava [6] | Right lane marker position with respect to ego-vehicle | Lane detection in varying road scenarios (visual outputs) | Lane detection in varying road scenarios (visual outputs) |
| Satzoda [12] | Lane detection in varying road scenarios (visual outputs) | Lane detection in varying road scenarios (visual outputs) | Lane detection in varying road scenarios (visual outputs) |

Table I lists the different evaluation techniques that are deployed in recent studies. Visual inspection of lane estimation is deployed in all studies. However, this is only a qualitative
metric. In order to quantify the performance of a lane estimation technique, one of the most common metrics is ego-vehicle position localization within the lane [7], [10], [12]. Lower error in ego-vehicle localization is a result of more accurate lane detection process itself. This is because ego-vehicle localization is computed using the estimated lane positions in the close proximity of the ego-vehicle. It can also be seen in Table I that most lane estimation studies have evaluated the robustness/accuracy of the algorithm. Considering that the lane estimation modules will be deployed in real-time systems, computational efficiency of the algorithms is also critical. Such a metric must be evaluated in a more generic way and not be restricted to a specific implementation platform.

In the following sections, we will show that the evaluation of lane estimation needs to be performed at different levels, which cater to different aspects of a lane estimation algorithm. A set of performance metrics will be proposed that can be used in an objective way to benchmark lane estimation algorithms.

III. PERFORMANCE EVALUATION METRICS

The top view of a typical lane is shown in Fig. 2. The actual lanes are given by dashed thick lines denoted by (h). The lane features detected by any algorithm are given by the black bullets denoted by (b) and (c). Fig. 2 will be used to demonstrate the significance of the different metrics that we present in this section.

A. Accuracy of Lane Feature Extraction

As shown in Fig. 1, extraction of lane features is the first step in most lane estimation methods. The accuracy of lane feature extraction directly affects the rest of the lane analysis process. This is shown in Fig. 2, wherein the detected lane (d) is determined by the lane features (b) that are extracted by a lane estimation algorithm. Inaccurate lane features will not fit the road model accurately, and hence the estimated lane will not follow the actual lane. For example if too many false positives (like (c) or (d) in Fig. 2) are detected, the road model cannot be fit accurately on the lane features. Additionally, the performance of the lane tracker is also affected by the incorrect lane features. Therefore analyzing the accuracy of the lane feature extraction step aids in improving the performance of the entire lane estimation process.

The evaluation of the accuracy of lane features itself is less addressed in literature. Conventional accuracy metrics such as true positive rates (TPR), accuracy (ACC) etc. can be used to evaluate this accuracy. Let $P_{GT}(x,y)$ be the ground truth coordinate of the lane marker in an image. If the lane estimation algorithm estimates the position of lane feature at $P_{LE}(x,y)$, then it is a true positive if:

$$||P_{GT} - P_{LE}|| < \delta$$

where $\delta$ is an acceptable tolerance error in the position of the lane marker. The accuracy of lane feature extraction step for an image frame is therefore defined as follows:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

where $TP$, $TN$, $FP$ and $FN$ represent the number of true positives, true negatives, false positives and false negatives respectively, with respect to the detected lane features in an input image frame.

B. Ego-vehicle localization

This metric has been used to evaluate the performance of the lane estimation methods in [6], [7]. This is illustrated by distance $d_e$ denoted by (a) in Fig. 2, which is the distance between the center of the camera on the ego-vehicle with respect to the lane markings detected by the algorithm. This distance $d_e$ can also be computed as the deviation of the vehicle position with respect to the center of the lane.

Although this quantifies the ego-vehicle localization, this does not completely evaluate the lane estimation process itself. The accuracy of $d_e$ is influenced by the accuracy of the lane detection in the near depth of view (given by the curve (f) in Fig. 2) only. It does not completely capture the accuracy of the lane detection in the far depth of view (given by the curve (g) in Fig. 2). Therefore, this metric evaluates the lane-based contextual information about the ego-vehicle, i.e. its motion trajectory or deviation, and applicable in scenarios like lane keeping, wherein the position of the vehicle with respect to the lane center is critical. However, in applications such as lane change assistance, it is necessary to accurately determine lanes that are in the far depth of view of the ego-vehicle. This will be considered by the next metric.

C. Lane Position Deviation (LPD)

We introduce lane position deviation (LPD) metric to determine the accuracy of the estimated lane in the road scene.
especially in the far depth of view from the ego-vehicle. Referring to Fig. 2, the lane features ((b) in Fig. 2) are used to estimate the lane, which is denoted by the dashed line (d) in Fig. 2. The lane position deviation ((e) in Fig. 2) measures the deviation of detected lane (d) from the actual lane that is obtained by joining the actual lane markings ((h) in Fig. 2). It can be seen from Fig. 2 that it captures the accuracy of the lane estimation process in both the near and far depths of view of the ego-vehicle. Moreover, this measure also evaluates the accuracy of the road model that is used to determine the lane curvature. For a given lane estimation algorithm, the different parameters of the algorithm can be used to determine their effect on LPD.

![Image](image.png)

**Fig. 3.** Illustration of lane position deviation (LPD) estimation.

LPD is computed in the following manner. Referring to Fig. 3, which shows a possible input image scene, let us consider the solid line to be the actual lane in the ground truth in the input image. Let us consider the left lane for the formulation below. Let the dashed line indicate the left lane position that is determined by a lane estimation method. The LPD metric determines the average deviation $\delta_{LPD}$ in the $x$-direction between the actual and detected lane positions, between $y_{max}$ and $y_{min}$ positions in the image scene, i.e.,

$$
\delta_{LPD} = \frac{1}{y_{max} - y_{min}} \sum_{y_{min}}^{y_{max}} \delta_y
$$

(3)

where $y_{min} \leq y \leq y_{max}$ is the selected region of the road scene in which the lane deviation is measured. The same metric can be scaled with an appropriate factor (based on camera calibration) to determine the lane position deviation in the world coordinate system.

**D. Computation Efficiency and Accuracy**

In addition to accuracy, computational efficiency is another important metric, that is often less addressed in literature. There is always a tradeoff between accuracy and computational efficiency [14]. Having a high accuracy of detection at the cost of high computational resources is not always a desirable solution, especially in the case of embedded IDSS which are powered by an on-board battery. Similarly, reducing the computational cost by trading accuracy is also not desirable because of the active safety nature of IDSS. Therefore, both these metrics, i.e., computational cost and accuracy, cannot be studied in isolation but must be evaluated together to understand the tradeoffs between the two. In this paper, we consider a simplified formulation for the computational cost of an algorithm, which is defined using the following proportionality:

$$
C \propto N_p
$$

(4)

where $C$ is the computation cost and $N_p$ is the number of the number of pixels on which lane estimation algorithm is applied. As shown in [15], the amount of data for processing and their movement (such as memory access) are two main contributing factors to the computation cost of an algorithm. The value of $N_p$ is dependent on the algorithm design and complexity, and this directly affects the accuracy also. Therefore, the computational efficiency of a lane estimation algorithm must also be evaluated keeping the accuracy in view. In order to this, we will show the tradeoff between accuracy and computational efficiency of a lane estimation algorithm in Section IV.

**E. Cumulative Deviation in Time**

In this metric, we study how the accuracy of the lane estimation process varies in the last $p$ seconds (indicated by (i) in Fig. 2). This is important to study because according to [1], the critical response time in most active safety driver assistance systems is approximately 2-3 seconds. This metric helps to determine the maximum amount of time for which a lane analysis method results a “given” accuracy of lane deviation. This will help to indicate the reliability of the algorithms for active safety systems. In order to find this metric, the following formulation is used:

$$
d_e(t_e) = \frac{1}{F_p} \sum_{i=0}^{i=F_p-1} d_i
$$

(5)

where $F_p$ is the number of frames evaluated within the $t_e$ seconds time window from the current frame, and $d_i$ is the vehicle position deviation within the lane in the $i$-th frame.

**IV. EVALUATION OF LANE ESTIMATION TECHNIQUES**

In this section, we will first demonstrate the use of the proposed performance metrics on a recent lane estimation method described in [12]. Thereafter, equivalent metrics from recent lane estimation studies are tabulated for future benchmarking purposes.

The lane estimation technique in [12] called LASeR (lane analysis using selected region) involves selective scan bands from the inverse perspective mapping (IPM) [16] of the input image to extract lane features. Each scan band is $h_B$ pixels in height and a modified filter based on steerable filters [7] is applied to $N_h$ such scan bands. A shift-and-match operation (SMO) is applied to extract lane features. A clothoid lane model and Kalman tracking is used for outlier removal and lane visualization. More details about LASeR are elaborated in [12].

We evaluate the accuracy of the lane feature extraction process in LASeR using an accuracy map as shown in Fig. 4. In order to generate this map, two sets of images were chosen corresponding to daytime and nighttime scenarios. Scan bands were selected and ground truth information was generated by manually marking the lane positions of the lane marker if it passes through the selected scan bands. Equations (1) and (2) are used to determine the accuracy of LASeR for different
settings of scan band height $h_B$ and shift in SMO of LASeR. It can be seen that an accuracy of over 95% is achieved (for both daytime and nighttime scenarios) in the feature extraction step of LASeR. Also, having a lower scan band height $h_B = 5$ gives better accuracy for daytime scenario, whereas $h_B = 10$ is better for lane feature extraction during nighttime. Additionally, shift of 2 or 3 pixels seems to be an optimal choice. Such analysis can be performed for other lane estimation algorithms also.

We will now illustrate the effect of some of the parameters in LASeR on the lane position deviation (LPD) metric. In order to perform this evaluation, we selected five different datasets, each containing about 400-500 images. Table II lists the five datasets and their labels that will be used in the rest of the paper.

![Fig. 4. Accuracy map of lane feature extraction in LASeR.](image)

For each dataset, we manually marked the lanes using a GUI in MATLAB to generate the ground truth information for each input image in the dataset. $\delta_{LPD}$ (equation (3)) is computed for each image using the lane positions estimated by LASeR and the ground truth positions. For a given dataset, the $\delta_{LPD}$ of all the images are used to compute the following two measures: (a) mean absolute LPD ($\mu_{LPD}$), (b) standard deviation of the absolute LPDs ($\sigma_{LPD}$). The two main parameters governing the accuracy of LASeR in terms of LPD are the number of scan bands $N_B$ and the height of each scan band $h_B$. ($\mu_{LPD}$) and ($\sigma_{LPD}$) are computed for the different datasets listed in Table II with two settings of $N_B = 16$, 8, and $h_B = 10, 5$. These values of LPD are listed in Table IV.

The maximum and minimum values for ($\mu_{LPD}$) and ($\sigma_{LPD}$) for each of the cases of $N_B$ and $h_B$ are indicated in red and green respectively. It can be seen from Table IV that the maximum absolute LPD is less than 9.8cm, with a maximum standard deviation being less than 5.5cm. It can be seen that the maximum deviations occur usually in either LISA $S_3$ (dashed curved roads) or LISA $S_3$ (night dataset).

![Table III. Mean absolute LPD ($\mu_{LPD}$) and standard deviation of the absolute LPDs ($\sigma_{LPD}$) in centimeters for left lanes in different data sets for varying number $N_B$ and heights $h_B$ of scan bands](image)

The mean absolute lane position deviation obtained from different frames of a dataset can also be used to plot histograms. Fig. 5 illustrates such histograms for dataset $S_1$ for $N_B = 16$ and $h_B = 10$ and 5. These distributions can be used to understand the frequency of lane position deviations in a given dataset. For example, the distributions in Fig. 5 show that for a curved road scene, it is best to use $N_B = 16$ and $h_B = 10$ because the error is less than 11cm for more than 88% of the frames in the dataset.

![Fig. 5. Distribution of mean absolute lane position deviation.](image)

Let us now look at the ego-vehicle localization, i.e. the position of the vehicle with respect to the lanes given by $d_c$ in Fig. 2. In order to evaluate $d_c$, a ground truth generation method is described in [7], in which a downward looking camera is installed above either of the side view mirrors. This camera captures the lane marking that passes by the ego-vehicle, without any perspective distortions. This makes the calibration of the lane positions obtained from this camera a simpler process. These lane positions are considered as ground truth and are used to evaluate the ego-vehicle localization obtained from the lane positions from front facing camera.

The distance of the ego-vehicle from the left lane is plotted in Fig. 6. The positions of the lane markings obtained from the nearest scan band in LASeR are considered to find the distance $d_c$ between the camera center on the ego-vehicle and the left lane position. A video sequence comprising 3000 frames (from Frame 14000 to Frame 17000, captured at 10-15 frames per
second) is used to evaluate LASeR against the ground truth that is manually generated using the side downward facing camera described above. These 3000 frames capture the ego-vehicle making three lane changes and passing over different kinds of road surfaces. The three lane changes occur between Frame 14200 to Frame 14300, Frame 14700 to Frame 14800, and Frame 15600 to Frame 15800 as shown by the thumbnails in Fig. 6. There is a left lane boundary split occurring at 16650 shown by the last set of thumbnails in Fig. 6. It can be seen that the distance $d_e$ closely follows the ground truth measurement during the frames when the ego-vehicle is not changing lanes. When the vehicle changes lanes in the three instances listed above, the localization measure follows the ground truth closely till it switches over and starts following again when the lane tracker settles down with the left lane of the new lane. The disturbance in $d_e$ during frames 16650 and above is because LASeR follows the original lane boundary for a few frames before it realizes that the lane has split and the tracker readjusts itself to follow the new lane boundary.

We will now evaluate computational efficiency of LASeR using the formulation in equation (4) and how it is affected by the different parameters of LASeR. In LASeR, we process $N_B$ number of scan bands, each of height $h_B$, that are sampled at specific positions in the $m_W \times n_W$ sized IPM image, where $m_W$ is the height and $n_W$ is the width of the IPM image. Therefore, the cost of processing in LASeR using (4) is given by:

$$C_{LASeR} = (h_B \times n_W)N_B$$

(6)

where the constant of proportionality in (4) is taken as 1. Most existing methods like [7], [8] process the entire IPM giving the cost as $C_{exist} = m_W \times n_W$. As discussed previously in Section III-D, the tradeoff between computational cost and accuracy is more critical in lane estimation process as compared to evaluating computation cost alone. Therefore, we plot the computation cost for the different cases of $h_B$ and $N_B$ versus mean absolute deviation on applying LASeR to dataset $S_1$ in Fig. 7. This graph gives the option to choose the parameters of the algorithm keeping computation cost as one of the criteria for performance evaluation. For example, it can be seen that we get lesser mean absolute deviation for 50% of the computation cost when scan bands of height 5 are applied instead of 10.

The cumulative mean absolute LPD and standard deviation of LPD for the lane positions that are estimated during the last $t = 1sec, 2sec,$ and so on till $t = 5sec$ are plotted in Fig. 8 for LISA $S_1$ dataset. It is observed that the mean and standard deviation is minimal when it is computed using the lane information available in the first second. These errors increase as frames for longer duration, i.e. 2 or 3 seconds, are used.

These trends can help to indicate the reliability of the lane estimation algorithm for active safety systems. For example, if 10cm is considered to be an acceptable lane position deviation for an active safety system, then the minimum response time that is available for the active safety system is between 2 and 3 seconds (from Fig. 8).

Finally, in Table IV, we list the values of the performance metrics that are available for recent lane estimation techniques. Considering that it is unfair to compare the different techniques using the values listed in the respective papers, the values in Table IV should be considered as a consolidation of performance metrics of different techniques. This is especially applicable to the feature extraction accuracy (ACC) and ego-vehicle localization ($d_e$) metrics. It can be seen that LPD metric is defined by the models each method uses. Considering that the LPDs defined in [9] and [8] are applicable to only
quadratic and Hough based models respectively, they cannot be generically compared with other lane estimation techniques using other models. The pixel level based LPD defined in this paper can applied to any estimation technique irrespective of the lane model that is used. The computation cost is derived based on the amount of data that is being processed in $M \times N$ image and the cost of lane detection operations (denoted by $C_F$ for filtering operation and $C_C$ for classification operation).

### TABLE IV. PERFORMANCE METRICS IN EXISTING LANE ESTIMATION METHODS

<table>
<thead>
<tr>
<th>Model</th>
<th>ACC</th>
<th>$d_e$</th>
<th>LPD</th>
<th>$C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>McCall [7]</td>
<td>-NR-</td>
<td>9.65</td>
<td>-NR-</td>
<td>$M \times N \times C_F$</td>
</tr>
<tr>
<td>Gopalan [9]</td>
<td>0.947</td>
<td>-NR-</td>
<td>Quadratic Model</td>
<td>$M \times N \times C_C$</td>
</tr>
<tr>
<td>Borkar [8]</td>
<td>-NR-</td>
<td>-NR-</td>
<td>Hough Model</td>
<td>$M \times N \times C_F$</td>
</tr>
<tr>
<td>Satzoda [12]</td>
<td>0.95</td>
<td>9.59</td>
<td>Pixel level</td>
<td>$N_B \times h_B \times N \times C_F$</td>
</tr>
</tbody>
</table>

ACC: feature extraction accuracy, $d_e$: ego-vehicle localization LPD: lane position deviation, $C$: computation cost -NR-: not reported

### V. CONCLUDING REMARKS

In this paper, we presented five different metrics that can be used to evaluate different aspects of a lane estimation algorithm. This includes evaluating the accuracy of the lane feature extraction process itself, followed by metrics to evaluate the accuracy of lane estimation in near and far fields of view from the ego-vehicle. The proposed lane position deviation metric is shown to be a critical metric that evaluates the efficiency of a lane estimation technique for far fields of view. Additionally, we have also considered the tradeoff between computational efficiency and accuracy, and the cumulative deviation in time as necessary performance metrics in embedded active safety systems such as ADAS. By demonstrating the proposed metrics on LASeR technique, we have shown that the different aspects of LASeR can be thoroughly evaluated.

### ACKNOWLEDGEMENTS

The authors would like to acknowledge the sponsors of our research especially, Toyota Collaborate Safety Research Center (CSRC), Korea Electronics Technology Institute (KETI), NextChip, and UC Discovery Program. We would also like to thank our colleagues from Laboratory for Intelligent and Safe Automobiles, UCSD for their constructive comments and support.

### REFERENCES


