Drive Analysis using Vehicle Dynamics and Vision-based Lane Semantics

Ravi Kumar Satzoda, Member, IEEE and Mohan M. Trivedi, Fellow, IEEE

Abstract—Naturalistic driving studies (NDS) capture large volumes of drive data from multiple sensor modalities, which are analyzed for critical information about driver behavior and driving characteristics that lead to crashes and near-crashes. One of the key steps in such studies is data reduction, which is defined as a process by which “trained employees” review segments of driving video and record a taxonomy of variables that provides information regarding the sequence of events leading to crashes. Given the volume of sensor data in NDS, such manual analysis of the drive data can be time consuming. In this paper, we introduce “drive analysis” as one of the first steps towards automating the process of extracting mid-level semantic information from raw sensor data. Techniques are proposed to analyze sensor data from multiple modalities, and extract a set of 23 semantics about lane positions, vehicle localization within lanes, vehicle speed, traffic density and road curvature. The proposed techniques are demonstrated using real-world test drives comprising over 150,000 frames of visual data, that is also accompanied by vehicle dynamics which are captured from the in-vehicle CAN bus, inertial motion unit (IMU) and global positioning system (GPS).

Index Terms—naturalistic driving studies, automatic drive analysis, lane characteristics, lane change detection, speed violation detection, traffic scenario detection

I. INTRODUCTION

Naturalistic Driving Studies (NDS) capture typical day-to-day driving sessions without artificial features that are introduced by controlled experimental studies [1]. NDS such as the 100-car study [2] and the more recent Strategic Highway Research Program (SHRP2) [1], [3] investigate the influence of different contributing factors that lead to specific kinds of driver behaviors and driving styles, which could result in crashes or near-crashes. Such studies not only help to determine the reasons for crashes and accidents but also aid in determining the effectiveness of specific inputs from advanced driver assistance systems (ADAS) to the driver while driving so as to prevent any crashes [4].

In order to conduct such detailed studies, large amounts of naturalistic driving data is collected using hundreds or thousands of instrumented vehicles that are fitted with sensor suites comprising multiple cameras, radars, sensors that capture vehicle dynamics etc. [1], [2], [5]. Data collection and integration using multiple sensors itself is a major effort in many such studies involving accident analysis [5], [6]. In the 100-car study that was completed recently, 2,000,000 vehicle miles, and nearly 43,000 hours of data with over 240 drivers was collected during a 12-13 month duration. In the current SHRP2 study, data from nearly 3100 drivers is being collected over a 3-year duration in six different places across USA [7]. After collecting this data, the next step is to reduce this data and generate the epoch files [7], [8], which will be analyzed by different study groups for further behavioral studies [8]. The data involved in NDS can be organized in the hierarchy levels as shown in Fig. 1. The first level is the raw sensor data that is collected from different sensors in the vehicle. The second level is the reduced information (as it is called in [5], [7], [9], [10]) that is extracted from the raw sensor data. We call this information drive analysis data in this paper because it is obtained by analyzing the data that is captured by the sensors during the drive. This includes semantic information about the lane positions, vehicle position, lane change events, speed analysis etc. The third and the topmost level of data hierarchy in Fig. 1 comprises the higher level inferences such as the behavior of the drivers based on the speeding and braking of the vehicle [9], aggressive driving versus safe driving characteristics [11] etc.

As seen in the pyramid in Fig. 1, the amount of data reduces as we go from the bottom to the top of the pyramid. The sensor data at the bottom is dependent on the sensor suite, and this raw data is not directly useful or understandable by the users of the top of the pyramid such as accidental analysis, cognitive sciences or human factors researchers. Therefore, this necessitates the data reduction step, which not only reduces the amount of data but also generates information at a mid-level of abstraction. NDS such as [4], [5], [7], [10], [12] discuss about data reduction as a process by which trained employees review segments of driving video, and record a taxonomy of variables that provides information regarding the sequence of events leading to crashes or near crashes.
Given the scale and volume of data that is collected in such studies, manual data reduction can be a time-consuming task, and is also subjective to interpretation of the manual subjects performing the data reduction. More recent NDS such as [8], [13] suggest automatic tools that can perform this data reduction step resulting in improved efficiency and accuracy of the NDS.

In this paper, we introduce techniques to analyze the naturalistic driving data, particularly from highway drives, and generate a report of mid-level information of specific events and/or observations that can be further analyzed for behavioral studies. We term this process as Drive Analysis, which was first introduced in [14]. The proposed techniques in this paper automatically generate a drive analysis report with a list of semantic information about the drive by fusing multiple mid-level semantics, which are extracted from multiple sensors in the vehicle.

In the next section (Section II), we identify some of the critical parameters or mid-level information that are being considered in related NDS. The scope of the paper and the data collection process is described in Section III. In Section IV, we first describe the lane feature extraction process which is used in Section V to extract lane-based semantics from the visual and other sensor information. Section VI describes ways to extract semantics from vehicle dynamics obtained from vehicle CAN bus, IMU and GPS. In Section VII, the different kinds of information obtained from the proposed techniques are fused together to generate a list of semantics related to highway drive data before finally motivating future directions in Section VIII.

II. RELATED WORK & MOTIVATION

In this section, we survey recent naturalistic studies to motivate the objectives and scope of the proposed drive analysis presented in this paper. The final objective of most NDS is to identify driver and driving behaviors, and characterizations that can be correlated to crashes or near-crashes [2], [5], [9], [15]. In addition to NDS, works such as [11], [16]–[18] also determine behaviors such as aggressive driving, tactical driver behavior etc. A survey of the NDS [2], [4], [5], [7], [9], [12] and related studies [11], [16]–[18] aids in finding the mid-level semantics that are used to determine such behaviors, some of which will motivate the work in this paper.

Lanes in the road scene are used to extract critical information about the drive in most naturalistic studies [2], [4], [5], [9], [10]. In [9], lane change events were used to determine the driving characteristics leading to crashes or near-crashes. In addition, the behavior of vehicles in adjacent lanes was also considered as a contributing factor for driver attentiveness in [9]. Lane type information can be useful in determining the behavior of such vehicles in adjacent lanes. Lane change maneuvers were studied in significant detail in the Phase-II report of 100-car NDS [5] to identify contributing factors during a lane change maneuver that lead to crashes and incidents. Identifying lane changes of the subject vehicle and the road conditions under which such lane changes have occurred (such as traffic density etc.) was one of the key steps discussed in [5]. In [19], the potential of collision avoidance systems in avoiding rear-end crashes was investigated, and vehicle localization in the lane was studied as one of the main semantic information to study rear-end collisions. Lanes played a pivotal role in the evaluation of road departure crash warning systems (RDCWS) on naturalistic driving behavior in [4]. Similar lane keeping studies were also presented in a recent work in [10]. In [12], [20], lane variability was one of the key metrics that was used to identify levels of driver fatigue and drowsiness.

In addition to information related to lanes, events and information derived from vehicle dynamics and traffic conditions are commonly used in extracting behavioral information [5], [9]–[11], [21]. In [4], the road curvature and vehicle speed are analyzed together to understand the risk posed by incorrect curvature estimates and vehicle speed alerts during curved sections of the road. In [22], the effects of visual and cognitive demand on driving performance and driver state were investigated using simulated and field drives, where the information from vehicle CAN bus and additional mechanical sensors on the steering wheel were used. In [21], [23], lateral and longitudinal accelerations, velocity and yaw rate are used to screen naturalistic driving data based on metrics such as time to collision (TTC) etc.

Traffic density is another semantic that is considered as a major contributing factor in studies such as [5], [7], [9], [17], [18]. Crashes during lane change maneuvers and high traffic density have been correlated in [5], [24]. In [19], the effectiveness of collision avoidance systems is studied under varying traffic conditions. Similarly, traffic density is considered as one of the factors influencing lane keeping capability in [10]. The effect of traffic density on the cognitive capabilities and the attentiveness of the driver is investigated in [25], [26].

Given that NDS are detailed studies covering a number of aspects of driving characteristics and driver behaviors, listing all the details of these works in this paper is not feasible. From the above survey, lane related semantics, such as lane positions, vehicle localization in lanes, lane types and lane change events, have been used extensively in most NDS and other related work. Additionally, information about the traffic and critical vehicle dynamics such as speed characteristics also play significant role in most NDS. It was seen in most current studies that the data reduction step to obtain these semantics is usually a manual step, with little or no automated means to reduce the large volume of sensor data to manageable and useful semantics. In the forthcoming sections, we will demonstrate automated techniques that aid in data reduction and drive analysis using the visual lane information and vehicle dynamics.

III. DRIVE ANALYSIS SEMANTICS FOR DATA REDUCTION & SCOPE OF STUDY

We have instrumented a testbed called LISA-Q2 with a similar set of sensors that are being deployed in on-going NDS called SHRP2. LISA-Q2 is equipped with a forward looking camera that continuously captures the forward-view of the ego-vehicle. The visual information from this camera is the
primary source of information to understand the environment in front of the vehicle. In addition to the camera, the vehicle dynamics such as velocity, braking pressure, acceleration etc. are captured by the in-vehicle sensors and recorded using the vehicle Controller Area Network (CAN) bus. The testbed is also equipped with an inertial motion unit (IMU) comprising accelerometer and magnetometer, that captures the state of the vehicle with respect to the world coordinate system (WCS) using accelerations in all three directions, yaw, pitch, roll, altitude etc. Additionally, we have a global position system (GPS) which continuously records the geo-physical position of the ego-vehicle.

**TABLE I**

<table>
<thead>
<tr>
<th>Drive Analysis Semantics</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lane changes - number and type</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance of the vehicle from the lane center</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position of veh. in left/middle/right lane</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lane change violations</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed related violations/info.</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean speed during different lanes</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deviation in different lanes</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right and left turns</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average speed during left and right curves</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage high traffic time</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average speed during high traffic</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

1-Lane Change Events, 2-Vehicle localization, 3-Lane marker types, 4-Speed, 5-Curvature

Before presenting techniques to reduce NDS data, we first introduce the list of semantics that will be analyzed in this paper. Data reductionists in NDS such as SHRP2 and 100-car study use a reference dictionary manual (such as [27]) to identify the events in raw data. In this paper, we limit the scope of the study to specific events that are related to semantics based on lanes and some vehicle dynamics as listed in the first column of Table I. Considering that each NDS variable or event can be associated with a number of mid-level information, both visual and non-visual, in this paper we will focus on extracting the following: (1) Lane change events, (2) Vehicle localization, (3) Types of lanes (solid or dashed), (4) speed of the vehicle, and (5) vehicle trajectory curvature. Items (1) to (3) are determined using lane information that is extracted from the visual information captured during the drive, whereas items (4) and (5) are related to vehicle dynamics. It can be seen from Table I that combining different information helps to get critical information about the drive as described by the semantics. It is to be noted that this list is not exhaustive and can be expanded with more items by introducing data from other sensors. This report in Table I is one of the first steps in generating automated reports for NDS. The list of semantics in Table I also defines the scope of the study in this paper. Additionally, it is to be noted that the proposed study is catered towards highway driving scenarios, and further extensions need to be presented to address urban and city-limit roads in future studies.

**IV. LANE FEATURE EXTRACTION**

It can be seen from Table I that more than half of the drive analysis semantics are related to lane related information. Therefore, it is necessary to introduce the lane feature extraction method that will be employed to extract the mid-level information related to lanes in the forthcoming sections. We employ an extension of the recently published selective region based lane analysis method (LASER) described in [28], [29] for the drive analysis process in this paper. However, it is to be noted that deploying the positions of the lane features for drive analysis is the focus of the paper, and the not lane feature extraction process itself. We describe LASER in this section for completeness but more details about it can be obtained from [28], [29].

In LASER, selected bands, and not the entire region of interest (ROI), are used to detect lane features. An image I is first converted into its inverse perspective mapping (IPM) [30] image I

\[ I_{IPM} \]

resulting in the top view of the road scene. In this IPM image, \( N_{i} \) scan bands (denoted by \( B_{0} \) and \( B_{1} \) in Fig. 2), each of height \( h_{B} \) pixels, are selected along the vertical \( y \)-axis, i.e. along the road from the ego-vehicle. Each band \( B_{i} \) is then convolved with a vertical filter that is derived from steerable filters [31]. The vertical filter is represented by the following equation:

\[
G^{y}(x,y) = \frac{\partial}{\partial x} G(x,y) = -\frac{2\pi}{\sigma^2} e^{-\frac{x^2+y^2}{\sigma^2}}
\]

where \( G(x,y) = e^{-\frac{x^2+y^2}{\sigma^2}} \) is a 2D Gaussian function. Therefore, the filter response \( B_{i}^{S} \) is given by

\[
B_{i}^{S} = B_{i} * G^{y}(x,y)
\]

where \( * \) represents convolution operation. The significance of the 2D Gaussian function, the implications of the standard deviation \( \sigma \) and the convolution operations are explained in [31]. \( B_{i}^{S} \) from each band is then analyzed for lane features by thresholding using two thresholds to generate two binary maps \( E_{+} \) and \( E_{-} \) for each band. Vertical projections \( p_{+} \) and \( p_{-} \) are computed using \( E_{+} \) and \( E_{-} \), which are then used to detect lane features using shift-and-match operation (SMO) in the following way:

\[
K = (p_{-} << \delta) \circ p_{+}
\]

where \( \odot \) represents point-wise multiplication, \( \delta \) is the amount of left shift (denoted by \( << \) above) that the vector \( p_{-} \) undergoes. The above equation for SMO can be explained using Fig. 2 in the following way. Peaks in \( p_{+} \) and \( p_{-} \) represent dark to light and light to dark transitions in the band respectively. Therefore, shifting \( p_{-} \) by \( \delta \) pixels to the left (denoted by \( << \) in (3)) and then multiplying the resulting vector with \( p_{+} \) will result in detecting adjacent light to dark and dark to light transitions, which characterize lane features. This is shown by \( K_{B_{i}} \) in Fig. 2, wherein the peaks correspond to lane features only. On applying a suitable threshold on \( K \) and road model, we get the positions of left and right lanes in the \( B_{i} \)-th scan band denoted by

\[
P_{B_{i}} = \{P_{L}(x_{L},y_{L}),P_{R}(x_{R},y_{R})\}\]
where \( y \) denotes the \( y \)-coordinate of the \( B_i \)-th scan band, \( x_L \) and \( x_R \) denote the \( x \)-coordinates of the left and right lane markings of the ego-lane in the \( B_i \)-th scan band. These lane marker positions will be used in the forthcoming sections to extract lane related mid-level information for drive analysis. Fig. 2 shows the algorithmic steps of LASeR. More details about LASeR are explained in [28], [29].

![Block diagram illustrating lane analysis using selected regions (LASeR).](image)

**V. Drive Analysis using Lane Semantics**

In this section, we will employ the lane feature extraction algorithm described above to extract lane related mid-level information for drive analysis semantics listed in Table I. As discussed in Section III, lane change events, vehicle localization and lane type detection are three lane related information that will now be discussed in detail.

### A. Lane Change Event Detection

In order to detect lane changes during the drive, the scan bands in the near view of the ego-vehicle are processed. This is because a lane change is characterized by the position of the ego-vehicle with respect to the lane positions in the near-view of the vehicle. We consider the lane feature positions, obtained in the \( N_L \)-th band in the near-view of the vehicle to detect this event. The lane feature positions of the left and right lanes in the \( N_L \)-th band are given by \( P_{L_i}(x_{L_i},y_{N_L}) = \{ P_L(x_L,y_{N_L}), P_R(x_R,y_{N_L}) \} \). Considering that the lanes are usually dashed, this band will not capture the lane marking in every single frame. Therefore, tracking is necessary to predict the lane position in this band. This will also help to eliminate outliers, if any, during lane feature extraction step.

We will now introduce the information from other sensors such as CAN and IMU in addition to the camera to track the lane markings using extended Kalman filter [31]. In order to do this, we initialize two trackers to track the left and right lane markings. The general state space models used for this tracking are given by the following equations:

\[
X^{t_{i+1}|t_i} = AX^{t_i|t_i} \quad \text{and} \quad Y^{t_i} = MX^{t_i}
\]

where \( X \), \( A \) and \( M \) are defined as follows:

\[
X = \begin{bmatrix} \phi \\ \tan \theta \end{bmatrix} \quad A = \begin{bmatrix} 1 & v \Delta t \\ 0 & 1 \end{bmatrix} \quad M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

where \( \phi \) is the lateral drift of the lane marking position that is being tracked. \( \tan \theta \) in \( X \) is used to track the yaw rate change of the vehicle, which is used to determine the lateral drift in the vehicle due to yaw rate change. Considering that we have two different trackers for the left and right lane markings, we will have \( \phi_L \) and \( \phi_R \) for the lane markings respectively. In the above equations \( v \) is the speed of the vehicle obtained from the CAN bus and \( \theta \) is the yaw angle obtained from the IMU. A constant velocity is assumed in the above model between two time instances, i.e. \( \Delta t \). The predicted deviations \( \phi_L \) and \( \phi_R \) are used to track the lane positions in the \( N_L \)-th scan band in the current frame at \( t_i \) using straight lane model. A straight lane model is sufficient because we can assume that the road in the near-view is straight especially in the case of freeways and highways where the lane change events are being detected. Therefore, the positions of the left and right lane markings at \( t_i \) can be tracked using the following equations:

\[
x_L^{t_i} = x_L^{t_{i-1}} + \phi_L + \theta y_{N_L} \\
x_R^{t_i} = x_R^{t_{i-1}} + \phi_R + \theta y_{N_L}
\]

These left and right lane positions \( P_{L_i}(x_{L_i},y_{N_L}) \) and \( P_{R_i}(x_{R_i},y_{N_L}) \) at \( t_i \) are then used to determine the left and right lane changes.

The position of the left lane marking appears to move towards the right as the vehicle makes the left lane change. Similarly, during a right lane change, the right lane marker moves to the left with respect to the vehicle. These observations are used to detect the lane change events in the following way:

\[
LC_L^{t_i} = \begin{cases} 1 & \text{if } x_L^{t_i} > T_{L_{\text{max}}} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad LC_R^{t_i} = \begin{cases} 1 & \text{if } x_R^{t_i} < T_{L_{\text{min}}} \\ 0 & \text{otherwise} \end{cases}
\]

where \( LC_L^{t_i} = 1 \) and \( LC_R^{t_i} = 1 \) indicate a left and right lane change event respectively. The time instances when the lane changes occur are also recorded in:

\[
T_L^E = \{ t_i^E \} \quad \text{and} \quad T_R^E = \{ t_i^R \}
\]

Fig. 3 (a) illustrates the above lane change event detection on a video sequence of a drive comprising 10,000 frames. The \( x \)-axis denotes the frame number or the time instance, and the \( y \)-axis gives the lane position in pixel coordinates along the \( x \)-axis of the IPM image. The lane change events that are detected using the above formulations are also indicated in Fig. 3 (a).

![ROC for left lane change events, (b) ROC for right lane change events.](image)
1) Evaluation of Lane Change Event Detection: LASeR algorithm is evaluated in detail in [36]. In this paper, we evaluate the lane change event detection using the following method. A dataset of 5 input videos with a frame rate of nearly 25 frames per second were considered resulting in a total of 66,540 frames. The videos were manually evaluated to determine the frame numbers where a lane change event occurs. A left lane change event is said to have occurred when the left lane moves to the right and crosses the mid point of the frame. Similarly, a right lane change event is detected using the right lane. This is considered as the ground truth, which is denoted by:

\[ T^L_{GT} = \{ t_i^{GTL} \} \quad \text{and} \quad T^R_{GT} = \{ t_i^{GTR} \} \]  

(11)

where \( t_i^{GTL} \) and \( t_i^{GTR} \) are the time instances where a left or a right lane change event occurs in ground truth respectively. We then consider a window of size \( 2w_L + 1 \) frames around the ground truth to determine if the lane change events detected by the proposed method (that are recorded in \( T^L_E \) and \( T^R_E \)) occur within this window. Therefore, if we consider the lane change event detected by the above technique at time instances \( t_i^L \in T^L_E \), they will be considered as true positive (TP) for the lane change event occurring at \( t_i^{GTL} \) if they occur within the \( 2w_L + 1 \) window around \( t_i^{GTL} \), i.e. \( t_i^{GTL} - w_L \leq t_i^L \leq t_i^{GTL} + w_L \). Otherwise, a false positive (FP) is considered to have been detected. Similarly true negatives and false negatives are defined. Fig. 3 (b)-(c) show the receiver operating curves (ROCs) for the left and right lane change events. It is to be noted that the x-axis shows false positive rate (FPR) for every 1000 frames. The curves are plotted for different window sizes, which are translated into seconds using 25 frames per second. It is evident that increasing the window size increases the accuracy as shown in Fig. 3 (b)-(c). With a window size of 5 seconds around the ground truth event, it can be seen that a true positive rate of nearly 82% is achieved with a false positive rate of one in every thousand frames. This validation setup helps to evaluate the lane change event detection process for NDS.

\[ \chi^L_V = \frac{x^L_i + x^L_{i+k}}{2} \]  

(12)

Fig. 4 shows the deviation of the center of the vehicle with respect to the center of the lane. This information can be used to determine the average deviation of the vehicle from the center of the lane during the drive. Given the time instances \( T^L_E \) and \( T^R_E \) when lane change events have occurred (from the previous section), the average absolute deviation of the vehicle \( D_C \) from the center of the lane can be determined using the following equation for the entire drive:

\[ D_C = \frac{\sum_{i=1}^{N_f} |d^i| - \sum_{k=1}^{N_{LCE}} \sum_{i_k} |d^i|}{N_f - \sum_{k=1}^{N_{LCE}} (j_k - i_k + 1)} \]  

(13)

where \( N_f \) is the total number of frames, \( d^i \) is the scaled deviation of the vehicle from center of the lane (obtained from (12)), and \( i_k \) and \( j_k \) are the indices of the time instances between which the \( k \)-th lane change event occurs. For the vehicle position deviation shown in Fig. 4, \( D_C = 13 \) cm if the above equation is used along with the lane change information as indicated in Fig. 4. This information will be fused with other drive analysis information later in this paper to derive specific observations about the drive.

B. Vehicle Localization in Lanes

The lane positions \( P^L(x^L_i, y^L_i, t) \) and \( P^R(x^R_i, y^R_i, t) \) are used to determine the localization of the vehicle in the lane. This is especially significant during the time periods in the drive when the vehicle is not changing lanes. The x coordinate of the center of the vehicle with respect to the center of the lane is determined by the following equation at every time instant \( t_i \) (given that the camera is placed at the center of the vehicle):

\[ \chi^L_V = \frac{x^L_i + x^R_i}{2} \]  

Fig. 4 shows the deviation of the center of the vehicle with respect to the center of the lane. This information can be used to determine the average deviation of the vehicle from the center of the lane during the drive. Given the time instances \( T^L_E \) and \( T^R_E \) when lane change events have occurred (from the previous section), the average absolute deviation of the vehicle \( D_C \) from the center of the lane can be determined using the following equation for the entire drive:

It is to be noted that the evaluation for lane change event detection presented in the paper is catered for analysis of recorded naturalistic driving data, and for the purpose of reducing large amounts of driving data to specific time instances. This is unlike the lane change event detection in active safety systems where high accuracy is required. Therefore, based on the ROCs in Fig. 3 (b)-(c), by increasing the accuracy rates to near 100% for a 5 second window, we see that the false positive rate is about 3 in 1000 frames for every 1 true positive. Therefore, having such accuracy percentages also would reduce the total amount of time to analyze long driving videos significantly as compared to manual data reduction process which is being followed currently. Additionally, to the best knowledge of the authors, the evaluation presented in this paper is the first of its kind and there are no previously published benchmarks or requirements for analyzing NDS data. Therefore, the proposed evaluations can serve as benchmarks for future studies.

C. Lane Type Detection: Solid or Dashed

We now analyze the type of lane markings to the left and right of ego-vehicle during the drive as solid or dashed lane markings. The knowledge of the type of lane markings can aid in finding drive characteristics such as a slow driving in the innermost lane of the vehicle or vice versa. A straightforward way to determine if the lane marking is solid or dashed is to check the frequency of occurrence of lane marking in a given region [32]. We propose a technique using LASeR to detect the lane type as solid or dashed by monitoring the lane markings appearing in the \( N_k \)-th band. During the lane feature
Given N detection step, we generate a vector $K^i$ at every time instance $t_i$ such that,

$$K^i = \{ k_j \} \text{ where } k_j = \begin{cases} 1 & \text{if lane marking is found in } N_j\text{-th band} \\ 0 & \text{otherwise} \end{cases}$$

and $i - N_k \leq j \leq i$. $K$ is a set of $N_k$ values, each corresponding to the visual presence or absence of a lane marking in the scan band in the last $N_k$ frames from the current $t_i$-th frame. The number of frames that have a lane marking in the previous $N_k$ frames is given by $N_p = \sum_{j=i-N_k}^i k_j$, which is used to define the ratio $\psi = N_p/N_k$. This ratio is used to define if the type of lane marking is solid or dashed in the following way:

$$L_{t_i}^{\psi} = \begin{cases} \text{solid} & \text{if } \psi > T_{\text{type}} \\ \text{dashed} & \text{otherwise} \end{cases}$$

However, the values of $N_k$ and $\psi$ are a function of the vehicle speed $v^i$, length of the gap between the dashed lane markings $d_g$, length of the lane markings $d_m$ and the frame rate of the video capture $f$. The value of $N_k$ will now be derived using these parameters. We first assume the largest values for $d_m = 3.66m$ and $d_g = 10.98m$ for California Highways [33]. If $N_k$ and $\psi$ satisfy these largest values, they will satisfy the lane markings with lesser gaps also. Given frame rate $f$ of a camera configuration in the vehicle, and assuming that the velocity $v$ is constant over $d_m + d_g$, then the number of frames $N_k$ which should be considered at time instant $t_i$ to determine the ratio $\psi^i$ is given by:

$$N_k^i = \left\lfloor \frac{d_m + d_g}{v^i f} \right\rfloor$$

Given $N_k$, we determine $K$, $N_p$ and $\psi$ at every time instant $t_i$ using the visual data in the $N_i$-th band. Fig. 5 shows $\psi_L$ and $\psi_R$ plotted for the left and right lanes during a typical drive captured in about 1000 frames. The thumbnails from the drive corresponding to three different segments of the drive having different types of lane markings are also shown in Fig. 5. In the first segment, the left lane marking is a solid lane which is shown by $\psi_L = 1$, and the right lane is a dashed lane given $\psi_R < 0.6$. In the second segment after a right lane change, the left lane is dashed resulting in $\psi_L < 0.6$ for the rest of the drive in the third segment also. The right lane has two different kinds of dashed lanes as shown in the segments 2 and 3. The dashed lanes are more spaced out in segment 2 as compared segment 3. Therefore $\psi_R < 0.6$ in segment 2, and it increases marginally to 0.7 in segment 3. If we are interested in two classes only, i.e. dashed or solid, a threshold of 0.75 for $\psi$ was found to be a good classification threshold that will satisfy all kinds of dashed lane markings found on California highways.

The information about the lane types can be used to find the periods during the drive when and how many times the vehicle was in inner lanes and outer lanes. This information can be critical information in studying the behavior of the driver. Also, any lane violations during the drive can also be arrested using the lane change events and the lane type information.

1) Note on Use of Non-Vision Sensors for Lane-related Mid-level Information: Given that cameras are becoming pervasive in modern automobiles, and ongoing NDS such as SHRP2 involve 5 different cameras, we chose to analyze the visual data for extracting the lane-related mid-level information. However, it is to be noted that other sensors such as differential GPS and advanced road maps can be used to determine drive analysis semantics. For example, vehicle localization can be determined by using high accuracy differential GPS. However, this may not be cost effective solution for deploying in thousands of test vehicles for NDS such as SHRP2. It is to be noted that differential GPS are being used for generating ground truths for vehicle localization but not for deployment in NDS testbeds. Similarly, advanced digital maps have been used in [34] to determine vehicle localization at lane-level. Details such as types of lane markers can also be determined, given accurate GPS locations of the vehicles. Considering that NDS involve analysis of data that is captured from multiple locations, the availability of such enhanced digital map data always is questionable. Additionally, the GPS locations of the drive need to be accurately recorded to extract accurate information from the digital maps. This is an additional constraint if digital maps alone are relied upon for lane type detection.

High resolution cameras, on the other hand, are cheap, pervasive and readily available. Therefore, analyzing the visual data for extracting such mid-level information is a more suitable option without any uncertainties of capturing the data itself. However, sensors such as GPS and digital maps can be used as complementary sensing and information modalities to cameras so that the visual data can be analyzed more accurately and efficiently.

VI. Drive Analysis using Vehicle Dynamics

In the previous section, visual information from the cameras was the primary sensor data that was being reduced to extract lane related semantics. In this section, we will focus on vehicle

Fig. 5. The ratio $\psi$ plotted for left and right lanes of the ego-vehicle for the drive previously shown in Fig. 4.
dynamics that are obtained from the CAN bus, IMU and GPS modules in the testbed described in Section III.

A. Using Vehicle Speed

As shown in Section II, vehicle speed has been used in many ways in behavioral studies. In this subsection, we will extract information about the drive using vehicle speed from the CAN bus and complemented by other sensors in the testbed.

1) Speed Violations and Average Speed: Firstly, the vehicle speed information along with the road map information can be used to check for any vehicle speed violations. The GPS coordinates of the vehicle can be used to determine the type of road the vehicle is driving on, which can then be used to check if the vehicle is violating any speed limits. Second, given the different sections of the drive, the average speed of the vehicle during these sections can be found. This is useful to understand driver behavior such as aggressive driving.

2) Traffic Conditions using Speed Information: Given the speed information and the GPS coordinates of the drive, the state of the traffic is determined, especially in freeway scenario. This is illustrated using the example shown in Fig. 6, which shows the speed profiles of two different drives that were conducted on a freeway. Assuming that a vehicle is supposed to maintain the speed of the traffic on the freeway, at every time instant, if the speed is less than a minimum speed $v_{min}$, it can be deduced that the traffic is high.

![Fig. 6. Speed profiles for two different drives with thumbnails showing the visual information of the scene during the drive. Distribution of vehicle speed showing two different probability density functions $P_1$ and $P_2$ corresponding to high and low traffic respectively.](image)

Data from the video showing the instances when a road curvature can be correlated with the vehicle trajectory curvature.

![Fig. 7. Curvature profile obtained from vehicle dynamics data with thumbnails from the video showing the instances when a road curvature can be correlated with the vehicle trajectory curvature.](image)

at different points in time can be found using the vehicle dynamics information. Curvature $C$ at any time instance $t_i$ is defined by the following:

$$C(t_i) = \frac{d\theta_i}{ds_i}$$

where $d\theta$ is the change in angle of the tangential motion vector of the vehicle and $ds$ is the distance traversed by the vehicle along the curve. We determine $d\theta$ using the yaw angle from the IMU. The distance traveled in $\Delta t$ is given by $ds$, which is computed using the speed and time information from the CAN bus (assuming constant velocity during $\Delta t$).

The significance of vehicle trajectory curvature is shown in various studies such as [35] to understand reasons for crashes. A haphazard vehicle trajectory will be reflected in a random vehicle trajectory curvature, whereas a vehicle moving straight will show a near zero values for $C$ over a period of time. Similarly, if the vehicle trajectory is curving, it will be reflected in $C$ being positive or negative over an interval of time during the drive.

![Fig. 8. 65-km route map on highways on which the test drives were conducted.](image)

In this paper, we limit the scope of the use of vehicle trajectory curvature to determine the road curvature. Given that the proposed study in this paper is related to highway scenarios, the curvature of the vehicle trajectory can be related to the road curvature itself in most cases. This is because of the more streamlined vehicle movement in highway scenarios as compared to urban or city-limit roadways. Fig. 7 shows the vehicle trajectory curvature (in degrees/m) that was covered during a particular drive. It can be seen from the thumbnails at $t_2$ and $t_3$ in Fig. 7 that if $C$ is greater than a positive threshold $T_+$, then the road is also curving to the right. Similarly if

B. Vehicle Trajectory Curvature

In this section, we will determine the curvature of the vehicle trajectory. The curvature of the vehicle trajectory
Fig. 9. Drive related semantics extracted from (A) drive 1, and (B) drive 2. Legend of curves (top to bottom in (A) and (B)): 1. Lane positions and lane change events. Left lane change in \(\textcolor{green}{\text{spikes}}\), right lane change in \(\textcolor{red}{\text{spikes}}\), 2. Ego-vehicle position (in meters) with respect to the center of the lane, 3. Lane marker type information - left lane marker in \(\textcolor{green}{\|y\| = 0.5}\) and right lane marker in \(\textcolor{red}{\|y\| = 1}\), 4. Vehicle speed information in miles/hr, 5. Vehicle trajectory curvature in (degrees/meter).

\[ C < T_− \text{ where } T_− < 0, \text{ it indicates a left curve. Therefore, the curvature of the vehicle trajectory can be directly related to road curvature in highway drive analysis. However, while } t_2 \text{ and } t_3 \text{ in Fig. 7 indicate the correct detection of left and right curvatures of the road based on the formulation above, } t_1 \text{ represents a right lane change and not a right road curvature. This is possible because the yaw angle varies similarly during the vehicle trajectories of both a right lane change event and a right turn. Therefore, using the vehicle dynamics alone from the IMU and CAN bus may result in false positives (such as } t_1 \text{ in Fig. 7). Combining the curvature information with the lane change events allows to remove such false positives. Also, the visual information about the lanes can also be used to detect the presence of road curvatures more accurately. For example, the curvature subtended by the lane features extracted from the farther scan bands in LASER can be used to detect curved lanes ahead of the ego-vehicle. Such vision-based techniques can be explored further for detection of road curvature in conjunction with the vehicle trajectory curvature. Additionally, digital maps if available could be used as a complementary source of information in addition to the in-vehicle sensors to determine the curvature of the road.]

VII. DRIVE ANALYSIS BY COMBINING VEHICLE DYNAMICS & LANE SEMANTICS

As shown in Table I, the different mid-level information from speed, curvature and lanes can be combined to perform a detailed drive analysis of naturalistic driving data. In this section, we will illustrate the generation of drive analysis reports on two test drives (drive 1 and drive 2) that were conducted by two different drivers on the 65-km test route on highways. As indicated previously, being one of the first contributions in this specific area of research on NDS, the scope of this study is limited to highways in this paper. The techniques extracting mid-level information based on lanes and vehicle dynamics both need to be further extended to cater to urban and city-limit roads. This is because, as compared to highways, city-limit roads are lesser constrained in terms of lane geometry, infrastructural requirements, and regulations on vehicle movement and speed. Therefore, analyzing drives on such roads requires additional requirements, which are out of scope of this paper.

Fig. 8 shows the 65-km test route on highways for the drives 1 and 2. For each drive, Fig. 9 plots the 5 main components of drive analysis information that are extracted from the sensors, i.e. lane position and lane change information, vehicle localization, lane marker types, speed and curvature, which are all time synchronized with each other. This information is further used to determine the different drive analysis semantics that were listed in Table I. Firstly, the lane change event detection was performed to generate the ROCs shown in Fig.
10. Nearly 90% of the lane changes were detected accurately in both drives at a false positive rate of 1 false positive in every 1000 frames. The lane marker positions detected by LASeR algorithm were then used to determine the ego-vehicle localization, i.e. the mean deviation of the ego-vehicle from the center of the lane during the drive. The information about the lanes is fused with speed of the vehicle and curvature of the road to determine 23 different semantics shown in Table II for the two drives.

It can be seen from Table II that drive 2 involved more lane changes than drive 1. The ego-vehicle is deviated from the center of the lane in a similar way in both drives. Both drivers spent majority of the time on the middle lanes (i.e. left and right lane markers are both dashed). However, driver 2 seems to have spent more time on the left most lane as compared to driver 1. From the speed profiles, it can also be established that both drivers exceeded the speed limit of 65 mph for most part of their drives by about 5 mph. Combining the lane marker type information with the speed information shows that driver 2 drove at less than 50 mph when he was on the right most lane. This is in contrast to driver 1 who recorded a speed of 65.6 mph when he was on the right most lane. It can also be seen that the average vehicle deviation is higher in both drivers when they are on the rightmost lanes as compared to leftmost lanes. Another semantic of interest is the vehicle speed and deviation during the different curvatures. It can be seen that driver 2 is more cautious than driver 1 because he records lower speeds during the curvatures. It can also be seen that the average vehicle deviation during such curvatures is lesser for driver 2 as compared to driver 1. This can be attributed to the lower speeds of driver 2. Additionally, the speed profiles can also be used to determine that 12.5% of drive 2 experienced high traffic, when the speed was reduced to 33 mph.

VIII. CONCLUSIONS & FUTURE DIRECTIONS

In this paper, we investigated the need for automated drive analysis so as to reduce the time and effort needed for data reduction in NDS. The proposed drive analysis techniques are shown to synergistically fuse data from a forward looking camera, in-vehicle CAN bus, IMU and GPS to extract lane positions, vehicle localization in the lane, types of lane markers, lane change events, speed information and vehicle trajectory curvature information. These mid-level semantics were further fused in different ways to automatically determine a set of 23 semantics about the drive. The drive analysis process is demonstrated on naturalistic driving data involving multiple drivers over a 65-km test route involving more than 150,000 frames of data and accompanying vehicle dynamics. However, such NDS involve a number of semantics that are of interest to specific behavioral studies. Therefore, there is a need to generate more techniques such as those presented in this paper to cater to the different data needs and requirements of NDS. Future work [37] includes developing more such techniques and toolboxes to extract mid-level semantics, and also analyze it further for behavioral investigations.

ACKNOWLEDGMENTS

The authors would like to thank the associate editor and the reviewers for their constructive suggestions and comments. We would also like to thank our sponsors, Toyota’s Collaborative Safety Research Center (CSRC), particularly Dr. Pujitha Gunaratne, and UC Discovery Program. Finally, we thank our colleagues Ms. Sujitha Martin, Mr. Minh Van Ly, and Mr. Greg Dawe from Laboratory for Intelligent and Safe Automobiles (LISA).

REFERENCES


Ravi Kumar Satzoda received his B.Eng. (with First Class Honors) degree in electronics from Nanyang Technological University (NTU), Singapore in 2004 under Accelerated Bachelors Program. He completed his M.Eng. (by Research) from NTU in 2007 with a Master’s dissertation on VLSI architectures for cryptosystems. He received his PhD in 2013 from NTU with a dissertation on efficient computing architectures for Advanced Driver Assistance Systems (ADAS). He is currently a Post Doctoral Fellow at University of California San Diego (UCSD), and associated with Laboratory for Intelligent and Safe Automobiles. His research interests include computer vision, embedded vision systems and intelligent vehicles. He is the recipient of the prestigious SIA-NOL (Singapore) scholarship for his under-graduation (2000-2004). He is currently serving as a reviewer for various IEEE journals and conferences including IEEE Transactions on Computers, IEEE Transactions on Intelligent Transportation Systems, and IEEE Transactions on Industrial Electronics.

Mohan Manubhai Trivedi (F’08) received the B.E. (with honors) degree in electronics from Birla Institute of Technology and Science, Pilani, India, in 1974 and the M.S. and Ph.D. degrees in electrical engineering from Utah State University, Logan, UT, USA, in 1976 and 1979, respectively. He is currently a Professor of electrical and computer engineering and he is the Founding Director of the Computer Vision and Robotics Research Laboratory, University of California San Diego (UCSD), La Jolla, CA, USA. He has also established the Laboratory for Intelligent and Safe Automobiles, Computer Vision and Robotics Research Laboratory, UCSD, where he and his team are currently pursuing research in machine and human perception, machine learning, human-centered multimodal interfaces, intelligent transportation, driver assistance, active safety systems and Naturalistic Driving Study (NDS) analytics. He regularly serves as a Consultant to industry and government agencies in the United States, Europe, and Asia. He has given over 70 Keynote/Plenary talks at major conferences. He has coauthored a number of papers winning Best Paper awards.

Prof. Trivedi is a Fellow of the International Association of Pattern Recognition (for contributions to vision systems for situational awareness and human-centered vehicle safety) and the Society for Optical Engineering (for contributions to the field of optical engineering). He received the IEEE Intelligent Transportation Systems Society’s highest honor, Outstanding Research Award in 2013, the Pioneer Award (Technical Activities) and the Meritorious Service Award of the IEEE Computer Society, and the Distinguished Alumni Award from Utah State University, Logan, UT, USA. He serves on the Board of Governors of IEEE ITS Society and on the Editorial advisory board of the IEEE Trans on Intelligent Transportation Systems.