Safe Maneuverability Zones & Metrics for Data Reduction in Naturalistic Driving Studies

Ravi Kumar Satzoda, and Mohan M. Trivedi

Abstract—Naturalistic driving studies (NDSs) capture drive data from multiple sensor modalities over long periods of time and under varying road conditions. NDS data reduction dictionaries list a range of events that are directly related to the conflicts and threat posed by the dynamics of the surrounding vehicles on the ego-vehicle. Manual reduction of such large scale data for events/conflicts related to dynamics of multiple vehicles is inefficient and prone to errors. In this paper, we present drive analysis techniques for automated NDS data reduction that can be deployed to identify, quantify and visualize threats posed to the ego-vehicle. In this regard, we propose safe maneuver zones (SMZs) that are derived based on the dynamics of surrounding vehicles with respect to the ego-vehicle. A set of metrics are formulated using the SMZs to quantify the threat posed by surrounding vehicles on the ego-vehicle. A detailed drive analysis of naturalistic driving data comprising more than 500,000 frames of data from over 5 hours of highway driving is presented. The resulting drive analysis reports characterize 7 different drives using the different metrics from the SMZs.

I. INTRODUCTION

According to National Highway Traffic Administration (NHSTA), highway deaths claim more than 30,000 lives each year in the USA [1]. In order to determine the various factors leading to crashes and near-crashes, naturalistic driving studies (NDSs) have become an important part of research studies in the area of intelligent transportation systems [2], [3]. NDSs capture data from day-to-day driving scenarios of road users who are not subjected to controlled and simulated driving experiments. NDSs usually involve large number of vehicles ranging from hundreds to thousands, which are driven by multiple drivers. These vehicles are equipped with a variety of sensors such as multiple cameras, in-vehicle sensors, global positioning systems (GPS) etc. which collect large volumes of data. Additionally, NDSs are usually carried out over multiple years. Fig. 1 shows the magnitude and complexity of data that is collected in NDSs.

Data reduction is one of the key steps in NDSs and is an effective way to manage such large volumes of naturalistic driving data. This step conventionally entails manual extraction and analysis of events by trained employees called data reductionists [2]. The higher level semantics and events that lead to crashes or near-crashes are extracted from raw data and catalogued, which are made available for studies by different transportation research and accident prevention groups. Some of these semantics and events include driving styles and behaviors of the drivers, distraction events (e.g. use of mobile phones during driving), impact of external events such as overtaking vehicles, tailgating etc. [2], [4].

Recent studies such as [2] aim at employing vision and learning based techniques towards automating the data reduction process, which is also termed as drive analysis. While distraction events due to hand and head movements of the driver were extracted automatically and analyzed in [5], lane related semantics were extracted from visual data from outside looking cameras in [2]. However, methodologies and techniques such as those presented in [2], [5] address a small percentage of NDS events among a list of few hundred events and categories that are listed in data reductionist dictionary [4].

In our first work on drive analysis in [2], lane related events such as lane changes, lane types and lane curvatures were extracted from NDS data. In this paper, we present techniques that characterize the ego-vehicle maneuvers in the context of the threat posed by the dynamics of the surrounding vehicles. In order to perform this analysis, information about lanes, dynamics of the ego-vehicle and the surrounding vehicles, and ego-vehicle kinematics are combined to determine safe maneuver zones (SMZs). Metrics are derived using SMZs that are used to quantify the threat posed by surrounding vehicles on the ego-vehicle in the SMZs. The proposed drive analysis is demonstrated using more than half a million frames of visual data and corresponding vehicle kinematics data that is collected during multiple hours of
naturalistic driving.

<table>
<thead>
<tr>
<th>NDS</th>
<th>Duration</th>
<th>Sensors</th>
<th>Data complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 Car Study [1], 2006</td>
<td>12-13 months</td>
<td>5 cameras, vehicle kinematic sensors, radar</td>
<td>2 million vehicle miles, 43,000 hours of driving data, 241 drivers, 100 cars</td>
</tr>
<tr>
<td>UDRIVE [6], 2012</td>
<td>4 years</td>
<td>Cameras, vehicle kinematic sensors, GPS, radar</td>
<td>240 cars, 150 trucks, 80 motorcycles, 11 countries</td>
</tr>
<tr>
<td>SHRP2 [7], 2012</td>
<td>3 years</td>
<td>5 cameras, vehicle kinematic sensors, GPS, cellphone records, radar</td>
<td>3,900 data years, More than 4 petabytes of data, 2.5 million trip files, 3,100 primary drivers, all vehicle types, 6 sites</td>
</tr>
</tbody>
</table>

II. RECENT STUDIES & OBSERVATIONS

A. Naturalistic Driving Studies

NDSs have been initiated by multiple organizations around the world [6]. Table I lists three notable studies which involve data collection in large scale. The 100-car study is the first NDS initiative by Virginia Tech Transportation Institute in USA [1], which involved 100 passenger cars with nearly 50,000 hours of driving data. A variety of behavioral studies related to extraction of driver behavior, driver inattention, vehicle maneuvers etc. were reported using the data from 100-car study.

The 100-car study paved way for two new and ongoing NDSs. The UDRIVE project is being conducted in collaboration with 11 different countries in Europe involving different kinds of vehicles [6]. Similar behavioral studies but in different geography and demography will be conducted as part of UDRIVE. The SHRP2 or Strategic High Research Program 2 is direct extension of 100-car study. More than 4 petabytes of data is expected to be collected during 3 years of data collection phase in SHRP2. Initial studies based on SHRP2 data are presented in [2], [5].

B. Role of Surrounding Vehicles in NDS

Considering that this paper particularly deals with safe ego-vehicle maneuvers in the presence of surrounding vehicles, we present some recent studies that discuss the role of surrounding vehicles in NDSs. In one of the earliest works on NDS data in [1], the 100-car study reports the definitions of different kinds of crashes and incidents, which are defined as results occurring due to conflicts between the maneuvers of the ego-vehicle and the surrounding vehicles. One of the recent most studies from SHRP2 NDS [7] describes different types of collisions between the ego-vehicle and surrounding vehicles (e.g., rear-end collisions, forward collisions etc.). This report also describes the role of collision avoidance systems in avoiding crashes/near-crashes. It was reported that the effectiveness of such systems rely on a number of factors including the detection of the surrounding vehicles, and the dynamics between the ego-vehicle and the surrounding vehicles.

[8] lists different types of conflicts such as conflict with lead vehicle, following vehicle, oncoming traffic, vehicle in adjacent lane, merging vehicle etc. Another term that is often related to NDS is precipitating event, which is defined as the action of a driver that begins the chain of events leading up to a crash, near-crash, or incident. Some key questions that are asked in this regard are related to the state of the vehicles surrounding the ego-vehicle before and after the crashes.

The brief survey presented above shows that most works related to NDS have identified the significance of dynamics of the surrounding vehicles for analyzing the safe maneuvers that can be made by the ego-vehicle. However, the conventional data reduction process in NDS is still largely a manual process conducted by trained human reductionists. Such manual reduction can be a time taking effort and also prone to errors, especially when it involves a manual inspection of dynamics of multiple vehicles around the ego-vehicle. This motivates the drive analysis presented in this paper.

Fig. 2. Different components and dynamic cues in the proposed drive analysis presented in this paper.

III. SCOPE & OVERVIEW OF PROPOSED STUDY

NDSs involve multiple sensing modalities resulting in the study of hundreds of factors and scenarios that could lead to crashes/near-crashes. Therefore, we present an overview of the proposed study, which also sets the scope of this paper. The drive analysis process in this work is particularly aimed at reducing naturalistic driving data based on the threat that is posed by analyzing the dynamics of the surrounding vehicles on the ego-vehicle. In order to do this, we consider the following as inputs to the proposed analysis: (a) Relative positions of the vehicles in two perspectives, i.e. from front and rear of the ego-vehicle, (b) Ego-vehicle kinematics: ego-vehicle velocity, and frequency of capturing ego-vehicle kinematics. The relative positions of the surrounding vehicles with respect to the ego-vehicle (Item 1) are assumed to be determined using either the front and rear cameras or the radars, both of which are usually part of NDSs such as SHRP2 [7]. Both the sensor modalities give the relative longitudinal distances between the ego-vehicle and the surrounding vehicles, and it is assumed that these are accurate to acceptable levels. Next, the ego-vehicle kinematics are recorded using in-vehicle sensors through CAN (controller area network) databus. The CAN bus gives a range of vehicle...
kinematics data. In this study, we use the ego-vehicle speed \( v_e \), and acceleration \( \dot{v}_e \).

Fig. 2 illustrates the overview of the proposed drive analysis in this paper. Given the sensor data described above, the visual data is first analyzed to estimate the lanes. The positions of the vehicles in the rear and forward views of the ego-vehicle (that are obtained using either the visual data or the radar data) are then localized within the lanes, i.e., the positions of the vehicles are localized to one of the three lanes - ego-lane, left and right adjacent lanes. The dynamics of the surrounding vehicles relative to the ego-vehicle are determined using the positions of the surrounding vehicles. All this information, i.e., the lane positions, the ego-vehicle localization, the dynamics of the surrounding vehicles and ego-vehicle kinematics are then used to determine the safe maneuver zones (SMZ) for the ego-vehicle. SMZs quantify the threat posed by the surrounding vehicles on the ego-vehicle. A detailed drive analysis using SMZs is presented in the forthcoming sections.

It is to be noted that we limit the scope of the analysis to constant velocity and straight road models. The proposed metrics and analysis can be extended for constant acceleration and curved road models also. Additionally, the techniques and analysis are particularly catered for NDSs involving highway scenarios (such as SHRP2) where the roads are usually marked and have specific geometries.

![Diagram of surrounding vehicle localization using vehicle positions and lane positions.](image)

**IV. Analyzing Dynamics using Safe Maneuver Zones**

**A. Detection & Localization of Lanes and Vehicles**

As shown in Fig. 2, detection of lanes and vehicles plays a vital role in the proposed drive analysis. However, it is to be noted that the objective of this paper is not to detect these objects. Instead, given the positions of the lanes and the vehicles, the proposed techniques are used to analyze safety critical parameters of the drive. For the completeness of the paper, we briefly describe the techniques used to detect the lanes and vehicles.

The LASeR algorithm (lane analysis using selective regions) [2] is used to detect the lanes in which the ego-vehicle is moving. The positions of the lane markings in both the rear and forward views of the ego-vehicle are detected using LASeR. Next, LASeR is used to determine the type of the left and right lane marking of the ego-lane. Additionally, LASeR is also used to detect drifts of the ego-vehicle as described in [9]. Similarly, detection of vehicles is extensively studied in works such as [10], where inverse perspective mapping (IPM) is used to determine longitudinal positions of the detected vehicles from visual images. The following two sets of coordinates are computed: \( D^F = \{ (x^F_1, y^F_1), (x^F_2, y^F_2), \ldots, (x^F_n, y^F_n) \} \) and \( D^R = \{ (x^R_1, y^R_1), (x^R_2, y^R_2), \ldots, (x^R_n, y^R_n) \} \). \( D^F \) and \( D^R \) refer to the positions of vehicles in front and rear of the ego-vehicle respectively. The lane positions, \( D^F \) and \( D^R \), and the lane type information are then used to localize the detected vehicles into 6 regions \( r_0 \) to \( r_5 \) as shown in Fig. 3. A vector \( d \in \mathbb{R}^{5 \times 1} \) is defined where \( d_i \in d \) is set to the distance of the vehicle in \( r_i \)-th region. If there are multiple vehicles in the same region, e.g. vehicles at \( y^F_1 \) and \( y^F_2 \) in \( r_5 \) in Fig. 3, then the vehicle closest to the ego-vehicle is considered for further analysis. If the lane type is shown as solid or if there is no vehicle in the region, then a null is assigned to \( d_i \)'s that are present on the other side of the solid lane. Referring to the case shown in Fig. 3, \( d = [y_3, \phi, y_4, y_5, \phi, y_1]^T \).

![Diagram of surrounding vehicles with threat functions.](image)

**B. Safe Maneuver Zone (SMZ) Estimation**

The conflict between the ego-vehicle and a surrounding vehicle is quantified and characterized using safe maneuver zone (SMZ). For example, in Fig. 3, there is a smaller SMZ behind the ego-vehicle as compared to the vehicle ins front of it in the ego-lane. This denotes that the rear approaching vehicle poses higher threat than the leading vehicle in ego-lane. However, proximity alone does not describe the SMZ completely.

Fig. 4 is used to illustrate the derivation of an SMZ. In Fig. 4(a) we show two scenarios of SMZ between the ego-vehicle (black) and the leading vehicle (red). The ovals represent the
SMZs. The threat posed by the leading vehicle on the ego-vehicle at different distances is indicated by the gradation of color in the oval (blue being safe and red being unsafe). Although the leading vehicle is at the same distance from the ego-vehicle (as seen in Fig. 4(a)), the threat posed by the same leading vehicle on the ego-vehicle can be different. The threat in the SMZ is denoted by the function \( T(y) \).

\( T(y) \) is a function of the following two parameters: (a) Relative proximity \( d_j \) of the surrounding vehicle from the ego-vehicle, (b) Relative velocity vector \( \vec{v}_r \) between the surrounding vehicle and the ego-vehicle. Relative proximity \( d_j \) directly influences the threat posed to the ego-vehicle by the surrounding vehicle in the lane. However, if the relative velocity between the two vehicles is high and if they are approaching each other, the threat is higher as compared to the scenario when the two vehicles are receding away from each other. Therefore, we include the sign and magnitude of the relative velocity \( \vec{v}_r \) in the formulation of \( T(y) \). In the proposed model for \( T(y) \), we assume constant velocity in \( \delta t \) time instance where \( \delta t \) is defined by the frequency of video frame capture set to 15~25 frames per second. In such short interval of time, such an assumption can be considered as valid. However, the model can be extended if acceleration is also considered. Also, it is to be noted that \( d_j \) and \( v_r \) are the longitudinal distance and velocity between the ego-vehicle and the surrounding vehicle along the direction of the lane. Different orientations of the surrounding vehicles due to maneuvers made by them (e.g. lane changes and drifts) with respect to the ego-vehicle are not included as part of this study.

Given the above observations, the different parameters of \( T(y) \) are derived using \( d_j \) and \( v_r \). First, we normalize the \( d_j \) to a value \( p_j \) such that \( 0 \leq p_j \leq 1 \) in the following way:

\[
p_j^i = \begin{cases} \frac{d_j^i}{D} & \text{if vehicle present within } D \\ 1 & \text{otherwise} \end{cases} \tag{1}
\]

where \( D \) is the maximum range (distance) of detection. The superscript \( t_i \) represents the time instance, i.e. \( p_j \) is computed for time instance \( t_i \). Therefore \( p_j \) is closer to 1 if the leading vehicle is away from the ego-vehicle. If the vehicle is not present within \( D \), then we set \( p_j \) to 1 resulting in an SMZ covering the entire lane till \( D \). For the remainder of this section, we drop the subscript \( j \) (denoting the \( j \)-th SMZ in Fig.3) in the explanation. All the variables henceforth in this section are derived for one SMZ.

Next, the relative velocity \( v_j^i \) between the ego-vehicle and a surrounding vehicle is computed using the relative longitudinal distances at time instances \( t_i \) and \( t_{i-1} \) in the following way: \( v_j^i = v_j^i - v_j^i = \frac{d_j^i - d_j^{i-1}}{\Delta} \). Therefore, if \( v_j^i < 0 \), the vehicles are approaching each other, resulting in higher risk of collision.

The relative velocity \( v_j^i \) and the normalized proximity \( p_j^i \) are used to define the Gaussian threat function \( T(y) \) in the following way. We compute the mean \( \mu_j^i \) and standard deviation \( \sigma_j^i \) for the Gaussian function for \( T(y) \) using the following equations: \( \mu_j^i = p_j^i \) and \( \sigma_j^i = \frac{v_j^i}{V_{\text{max}}} \), where \( k_y = 1 \) if \( v_j^i > 0 \) otherwise \( k_y = -1 \), and \( V_{\text{max}} \) is the maximum relative velocity. Therefore, for a given \( y \) (distance between the ego-vehicle and the leading vehicle) we compute the threat posed by the leading vehicle on the ego-vehicle \( \beta \) using:

\[
\beta_j = \frac{1}{\text{height}(y) + \text{width}(\vec{v}_r)} \quad \text{for } 0 \leq y \leq p_j^i \tag{2}
\]

It is to be noted that the longitudinal distance between the surrounding vehicle and the ego-vehicle \( y \) is scaled between 0 and \( p_j^i \). The threat values in (2) are then used to define the safe maneuver zone (denoted by \( A_{\text{SMZ}} \)) as a one sided ellipse centered at \((0,0)\) (center of the ego-vehicle) with the following parameters:

\[
\text{major axis } a = \frac{W}{2} \quad ; \quad \text{minor axis } b = p_j^i \tag{3}
\]

where \( W \) is the width of the lane. At each \( y \) position in SMZ, i.e. \( A_{\text{SMZ}}(x,y) \), we have an associated threat factor that is computed using (2). \( A_{\text{SMZ}} \) is used to determine the following two parameters:

1) \( \beta_{\text{min}} \)-Minimum threat factor in \( A_{\text{SMZ}} \): Having a higher \( \beta_{\text{min}} \) in \( A_{\text{SMZ}} \) indicates that the SMZ is more risky to enter for the ego-vehicle.

2) \( \bar{\beta} \)-Mean weighted area of \( A_{\text{SMZ}} \): This metric is computed using the following equation:

\[
\bar{\beta} = \frac{\sum A_{\text{SMZ}}(x,y)}{\text{Area}(A_{\text{SMZ}})} \tag{4}
\]

A higher \( \bar{\beta} \) implies the entire SMZ is more risky for the ego-vehicle to enter.

Setting thresholds on the above parameters will result in segmenting the drive into sections where the ego-vehicle faced different levels of conflict from the surrounding vehicles. The above formulations when applied on the different regions \( r_0 \) to \( r_5 \) in Fig. 3 result in six SMZs denoted by \( A_{\text{SMZ}}^0 \) to \( A_{\text{SMZ}}^5 \) respectively.

<table>
<thead>
<tr>
<th>TABLE II</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANALYSIS OF SMZS AT THE THREE TIME INSTANCES SHOWN IN FIG. 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>Region</th>
<th>Forward view</th>
<th>Rear view</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t = 1 )</td>
<td>( b )</td>
<td>0.580</td>
<td>0.753</td>
</tr>
<tr>
<td></td>
<td>( \sigma )</td>
<td>0.285</td>
<td>0.181</td>
</tr>
<tr>
<td></td>
<td>( \beta_{\text{min}} )</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>( \beta )</td>
<td>0.678</td>
<td>0.643</td>
</tr>
<tr>
<td>( t = 17 )</td>
<td>( b )</td>
<td>0.679</td>
<td>0.929</td>
</tr>
<tr>
<td></td>
<td>( \sigma )</td>
<td>0.604</td>
<td>0.741</td>
</tr>
<tr>
<td></td>
<td>( \beta_{\text{min}} )</td>
<td>0.665</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>( \beta )</td>
<td>0.443</td>
<td>0.551</td>
</tr>
<tr>
<td>( t = 70 )</td>
<td>( b )</td>
<td>0.542</td>
<td>0.850</td>
</tr>
<tr>
<td></td>
<td>( \sigma )</td>
<td>0.839</td>
<td>0.504</td>
</tr>
<tr>
<td></td>
<td>( \beta_{\text{min}} )</td>
<td>0.244</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>( \beta )</td>
<td>0.616</td>
<td>0.354</td>
</tr>
</tbody>
</table>

V. EVALUATIONS ON SAMPLE NDS DATA

In this section, we use smaller segments of the NDS data to evaluate and demonstrate SMZs and the associated metrics. The inputs for the analysis are the video data captured from
the forward and rear cameras and ego-vehicle speed. Vehicles are detected using the vehicle detection algorithm in [10] on the forward and rear camera data. The longitudinal distances are obtained by applying camera calibration and inverse perspective mapping (IPM) to the forward and rear input video frames. In this section, we present the detailed visualization of the SMZs first. This is important because visualizations of the SMZs are also an aid for data reductionists in NDS. Then, we show how the safe maneuver zones from the drives can be analyzed to reduce NDS data to specific instances of conflict segments.

A. Safe Maneuver Zone Analysis

In Fig. 5, we show a sample visualization of the safe maneuver zones for three time instances $t_i = 1, 17$ and 70 during a drive. For each $t_i$, we show the top view of the forward and rear views of the ego-vehicle that are obtained by applying IPM on the image frames from forward and rear cameras. The IPM images are scaled such that the left adjacent lane, ego lane and right adjacent lane in the forward and rear images are aligned. For each time instance, we show the SMZs for the six regions $r_0$ to $r_5$ (as shown in Fig. 3) that are derived using the different parameters listed in the previous sections. The lane type is used to visualize the lane markings, i.e., dashed lanes are plotted to show that there are left and right adjacent lanes. If the lane type is determined as solid, then the lane next to it (left of left lane and right of right lane) will not have any SMZs. The center horizontal white line represents the position of the ego-vehicle and $y$-axis (vertical axis) represents the scaled version of the longitudinal distance from the ego-vehicle.

The composite SMZ visualization shown for each time instance is also scaled for ease of visualization. This visualization aids the data reductionists to see instances of conflicts that the ego-vehicle faces in a visually perceivable way.

Table II lists the different metrics that are derived from the SMZs for the three time instances in Fig. 5. Drive analysis can be performed in a quantitative manner using these parameters. Analyzing $b$ aids to find how close the surrounding vehicles are. For example, it can be seen that the zones in right lane, both forward and rear zones, will be selected as high conflict zones if a threshold on $b$ is set to less than 0.25. Similar analysis can be applied on all the regions. Next, the value of $\beta_{min}$ quantifies the minimum starting threat that the ego-vehicle would face in an SMZ. Three different scenarios can be inferred from the forward left lane and ego-lane (regions $r_5$ and $r_0$) in the three instances shown in Fig. 5 and their corresponding $\beta_{min}$s in Table II. At $t = 1$ it can be seen that $\beta_{min} = 0$ for the left and ego lanes for different $b$ values. This can be visualized by the dark blue regions in Fig. 5. In other words although the vehicle in the left lane ($V^{r_5}$) is closer to the ego-vehicle (longitudinally) as compared to the vehicle in ego-lane ($V^{r_0}$), the $\beta_{min}$s show that both pose equivalent threat if the vehicle enters into the corresponding zones. However, at $t = 17$, we see that the $V^{r_7}$ is farther than $V^{r_5}$, but the immediate threat posed by $V^{r_7}$ is higher than $V^{r_5}$, i.e., $\beta^{r_7}_{min} = 0.065$ and $\beta^{r_0}_{min} = 0.163$. This indicates that although a vehicle may be farther away from the ego-vehicle, it need not necessarily pose a lesser threat as compared to a closer vehicle closer. At $t = 70$, we see the more obvious scenario with $V^{r_5}$ and $V^{r_0}$, i.e. a vehicle that is farther away $V^{r_0}$ poses lesser threat compared to the vehicle that is closer $V^{r_5}$.

While $\beta_{min}$ quantifies the immediate threat spatially, the mean weighted area of SMZ $\overline{B}$ is Table II assesses the threat over the entire SMZ. This is illustrated for the forward SMZ in ego-lane, i.e., $r_0$ for $t = 1$ and $t = 70$. It can be seen that while $\beta_{min} \approx 0$ in both cases, $\overline{B}^{r_0} = 0.083$, which is 4 times lesser as compared to $\overline{B}^{r_7} = 0.354$. This shows that although in both cases, the immediate threat is low, in the case of $t = 70$, there is more sustained risk in the SMZ.

We plot the three different parameters of the SMZs for the drive in Fig. 6. The first and second columns in Fig. 6 correspond to the plots for the forward and rear views of the ego-vehicle. Fig. 6(a), (b) and (c) plot $b$, $\beta_{min}$ and $\overline{B}$ respectively for the SMZs in left lane (in red), ego lane (in blue) and right lane (in green). When a vehicle is not found within the maximum bounds, the values are forced to -0.5.

The quantitative analysis presented above can be used to generate different kinds of drive analysis reports for NDS on the threat posed on the ego-vehicle by the surrounding vehicles during the drive.
VI. DATA REDUCTION USING SMZ ON EXTENSIVE NATURALISTIC DRIVING DATA

In this section, we automatically generate drive analysis reports for 7 different drives - Drive 1 to Drive 7, with a total visual data of 509,832 frames. Based on the survey of related studies, this kind of analysis is the most extensive in terms of the amount of data that is processed using automated tools for assessing threat from surrounding vehicles.

<table>
<thead>
<tr>
<th>Drive Analysis Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drive Number</td>
</tr>
<tr>
<td>Traffic density</td>
</tr>
<tr>
<td>Threat from overtaking vehicles on right</td>
</tr>
<tr>
<td>Threat from overtaking vehicles on left</td>
</tr>
<tr>
<td>Rear-end collision threat from leading vehicle</td>
</tr>
<tr>
<td>Rear-end collision threat from approaching vehicles</td>
</tr>
<tr>
<td>Sudden cutting in occurrences</td>
</tr>
<tr>
<td>High threat drifts</td>
</tr>
</tbody>
</table>

SMZs are derived for all the drives, which are used to compute minimum threat $\beta_{min}$ and mean weighted area of SMZ $\bar{\beta}$. Thresholds are applied on these semantics to derive characteristics about the different drives in the drive analysis reports shown in Table III.

Seven different characteristics are listed for each drive in Table III. Traffic density and threat from overtaking vehicles from left/right are indicated by low (L), medium (M) or high (H). Traffic density is determined by checking $\beta_{min}$ for the six regions. For a given frame, traffic density is set according to the following threshold settings:

$$
\text{Traffic density} = \begin{cases} 
H & \text{if 4+ regions have } \beta_{min} \geq 0.6 \\
M & \text{if 4+ regions have } 0.4 \leq \beta_{min} < 0.6 \\
L & \text{otherwise}
\end{cases}
$$

Similar thresholds are used on the left and right adjacent regions to indicate the threat posed by overtaking vehicles from left and right adjacent lanes as L/M/H. The above
conditions are applied on each frame and an average value is used to determine the drive characteristic for the entire drive in Table III. In some drives, the threat levels change during the drive, which is indicated by $\rightarrow$ in Table III. It can be seen from Table III that Drives 2 & 6 had higher traffic density compared to the remaining drives.

Instances of rear-end collision threat to the ego-vehicle from leading vehicle in front and approaching vehicle from rear are determined by finding if $b$ in ego-lanes drops to less than 0.1, i.e. the leading/approaching vehicles are in close proximity to the ego-vehicle spatially. The number of such instances in each drive are indicated in Table III. It can be seen that the drive analysis reports for Drives 2 and 6 show higher number of instances of rear-end collisions, which can be explained by higher traffic densities in these drives.

The drive analysis also reports number of cutting in instances. This is determined by analyzing the weighted mean of the threat, i.e. $\bar{\beta}$. Rapid change in $\bar{\beta}$ from low to high in the ego-lane SMZ, and a similar rapid drop in $\bar{\beta}$ in the adjacent lane’s SMZ indicate that there is sudden cutting in from adjacent lane to the ego-lane. For the analysis presented in Table III, the drive analysis looks for a change in $|\bar{\beta}|$ of 0.4 or more. It can be seen that Drive 1 witnessed the most number of cutting-in instances.

Finally, drive analysis reports the number of high threat drifts. This is determined by finding the threat posed by vehicles in the lanes towards which the ego-vehicle is drifting. If the threat was greater than 0.5, then the ego-vehicle drift was tagged as high threat drift.

Fig. 7 shows some sample results from the automated drive analysis process on the drives that were analyzed in this section. High and low traffic scenarios (Fig. 7(a) & (b)), side sweeping/overtaking event scenario (Fig. 7(c)) and a high risk lane drifting case (Fig. 7(d)) are illustrated. The details of the different SMZ parameters for each of the cases are listed in the figure caption. The visual data was also manually analyzed to check the correctness of the drive analysis reports in Table III. To our best knowledge based on the current literature, metrics to quantify the different drive characteristics in Table III have not been concretely formulated. The proposed metrics and analysis can act as reference for future studies.

VII. CONCLUDING REMARKS & FUTURE DIRECTIONS

In this paper, we presented safe maneuver zones (SMZs) which provide a visual tool and quantitative metrics to analyze the dynamics of surrounding vehicles for the data reduction process in NDS. Extensive NDS driving data involving multiple drivers and over 500,000 frames of data was used to demonstrate and evaluate the proposed techniques and metrics. The resulting drive analysis reports reduced such large amounts of data to meaningful events and semantics related to threat assessment.

This study is one of the first of its kind in automated NDS data reduction following our recent work in [2]. While [2] catered to event detection, this study relates the dynamics of ego-vehicle with the surrounding vehicles. However, this is still a preliminary study. For example, we did not consider lane change and lane drift maneuvers made by the surrounding vehicles in this paper. Determining the orientation of the surrounding vehicles, and their impact on the threat and SMZs is a possible future direction. Similarly, city limit and urban conditions require different models to define the dynamics of vehicles. There are many such possible directions, considering that NDSs involve a variety of factors and events.

REFERENCES