Drivable-Region and Lane-Detection System for Autonomous Vehicle Navigation

Aravinth R[1], Arunkumar R[2], Kaviyazhagan G[3], Baskar A[4]
Department of Computer science and Engineering
[1,2,3,4] E.G.S.Pilay Engineering College,Nagapattinam.

Abstract- *Modern road vehicles are employing features of driver assistance systems (DAS) to improve drivability performance, comfort, and safety. In the future perspective, the advances in this field will lead these systems to the level of autonomous and cooperative driving, based on sensors networks and sensor fusion. This paper aims to present the readers a novel strategy for lane detection and tracking, which fits as a functional requirement to deploy DAS features like Lane Departure Warning and Lane Keeping Assist. To achieve the presented results, the digital image processing was divided into three levels. At the low-level, the input image dimensionality is reduced from three to one layer, the sharpness is improved, and region of interest is defined based on the minimum safe distance from the vehicle ahead. The feature extractor for lane edges detection design is part of the mid-level processing. The lane tracking strategy development is discussed in the high-level stage; Hough Transform and a shape-preserving spline interpolation are used to achieve a smooth lane fitting. The experimental outcomes were qualitatively and quantitatively evaluated using a ground truth comparison. The strategy shows good accuracy levels, including scenarios with shadows, curves, and road slope.*

Keywords- Driver assistance systems, image processing, lane detection, lane tracking, monocular camera.

I. INTRODUCTION

MOBILITY is an important social and economic factor for humanity as it provides quality of life for individuals and it is the backbone of commercial trading and services. For the outcomes, the inhabitants of industrialized countries have achieved a high degree of mobility due to the mass production of vehicles and road infrastructure investments.

However, the popularization of automobiles has caused some logistical problems, such as traffic congestion and the increase of the risk of accidents [1], [2]. According to data from the National Highway Traffic Safety Administration (NHTSA), 94% of the critical flaws in the chain of events preceding an accident are assigned to the driver, which justifies the investment in support systems [3].

Driver Assistance Systems (DAS) offer solutions to reduce the effects caused by the above-listed problems. Bengler et al. [4] have been analyzing over the past three decades the future perspective of DAS development, and structuring it according to the technological point of view:

From the late 1970s to the mid-1990s: The early systems were based on proprioceptive sensors to stabilize vehicle dynamics, such as ABS (Anti-lock Brake System), TCS (Traction Control System) and ESC (Electronic Stability Control);

From the early 1990s to the late 2000s: Exteroceptive sensors based systems with information, alarm and comfort functions such as LDW (Lane Departure Warning), ACC (Adaptive Cruise Control) and Park Assist;

From the middle of the years 2000 to 2030: The goal is to reach autonomous and cooperative direction level through networks of sensors.

II. RELATED WORKS

In the last decades, several functionalities have been gradually added in the vehicles towards autonomous driving. For camera-based vision systems, there are the monocular vision and stereo vision modes. For front view based applications such as ACC, monocular cameras with conventional lens achieve satisfactory accuracy level over relevant distances [9]. Side and rear-view systems need wide-angle (or fisheye) lenses because of their greater field of view coverage (conventional lenses cover about 45°, whereas wide-angle lenses can cover up to 180°) [10], [11]. The use of stereo vision is advantageous in applications related to s. the reconstruction of three-dimensional scenarios, objects detection, and tracking [12], [13].

Sensor data fusion strategies using proprioceptive (usually taken from the vehicle serial communication bus) and exteroceptive sensors are highly suitable for DAS. For both, feature extraction and decision-making. The sensor data fusion is an interesting approach to support a wide-range of applications, like lane detection [14], to choose the optimum

path for vehicle traffic [15], for obstacles [16] and pedestrian detection [17].

DAS are functionalities that provide the driver assistance and LDW and LKA are functionalities, which essentially consists in monitoring vehicle's state about the nearest lanes and issuing an alarm in case an undesired lane departure occurs. The research of McCall and Trivedi [18] and Hillel et al. [19] complement each other to provide an overview of the main techniques applied in the development of these systems in recent years.

In the lane detection algorithm implementation, Jung and Kelber [20] have proposed a linear-parabolic highway model. By dividing ROI into two sections, resulting in a near and a far field, the lane boundaries of the near field were approximated by a linear interpolation function and the edges of the far field by a parabolic interpolation function. The LDW metric is like a function of the symmetry between the left and right tracks,

An alternative to IPM transformation is the spatiotemporal images. They are built by accumulating the pixels on a scanline along the time axis and then aligning the scanline. As a result, the trajectory of lane points appears smooth and forms a straight line [32]. In comparison with IPM, it demonstrates a superior computational efficiency and improved detection rates.

Usually, the algorithm's performance evaluation must take into account the accuracy (for selected scenarios) or robustness (accuracy across different scenarios) [33]. Therefore, it is necessary to get a manually marked ground-truth for each

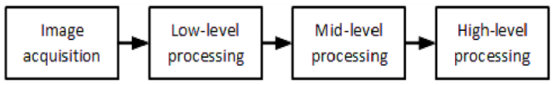


Fig. 1.The general workflow for the proposed strategy

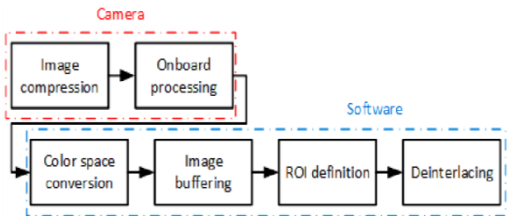


Fig. 2.Block diagram representing the steps of low-level processing.

of the samples to compare with the proposed lane tracking strategy. In [18] three different metrics were evaluated; mean absolute error in position, the standard deviation of error in position and standard of error in rate of change of lateral position. While [34] considers the entire image, to classify a frame if most frames are in the inner part limited by the threshold the tracking is correct. If the tracking area overcomes the ground-truth, it is classified as false-positive.

The next section presents this work strategy according to the division in three processing levels. Low-level processes involve primitive operations such as image preprocessing to reduce noise, contrast enhancement, and image sharpening. Mid-level processing involves tasks such as edges and contours extraction. Finally, high-level processes involve the interpretation of the segmented objects, normally using cognitive functions associated with vision [35].

III. ROAD LANE DETECTION AND TRACKING STRATEGY

Fig. 1 presents the workflow for the proposed strategy, divided into four processes: Image acquisition; Low-level processing; Mid-level processing and High-level processing.

The image acquisition process uses a monocular digital camera in vehicle's windshield. The image output size is 480x640 at the frame rate of 30fps (frames per second), with H.264 compression standard, which provides high-quality video at considerably, lower bit rates [36].

The development steps for low-level processing are presented in the next section.

A. Low-Level Processing

Fig. 2 depicts the low-level processing related operations, which includes camera embedded image processing and the application of software-developed techniques.

The compression/decompression (CODEC) encoding process can be handled on the CPU or through the camera directly. The use of an H.264 encoding camera is more efficient in relation to a regular camera regarding low power consumption, CPU utilization and image quality [36]. Onboard processing option for autofocus is disabled to ensure uniformity in the samples and to avoid environment related incidents, for example, loss of focus in night periods due to other vehicles headlamp.

TABLE I DISTANCE TRAVELED UP TO THE MEAN REACTION TIME FOR BRAKING [41]

Condition	Reaction time [ms]	S.D.[a] [ms]	Traveled distance [m]
Control	392	33	12.99
Radio	408	31	13.41
Conversation with passenger	453	47	15.28
Conversation using handheld phone	464	41	15.43
Conversation using hands-free phone	465	51	15.43

[a]Standard deviation.

RGB (Red Green Blue) to grayscale conversion reduces input image dimensionality from three to one layer, facilitating thus the digital image manipulation in subsequent processing steps. A popular method to perform this conversion is the NTSC CCIR 601 standard [37], using this approach Gamma correction is also achieved [38].

In the image buffering process, a recursive sum of n image frames weighted by Kn gain. The result is then normalized to maintain matrix values in a predefined range. The adoption of this strategy emphasizes the contrast between the track and lane markings, as the frames overlapping reinforce the intensity of the pixels relative to lane markings thus improving contrast. The buffering also contributes to the elimination of transient noise and the identification of weak and broken line markings.

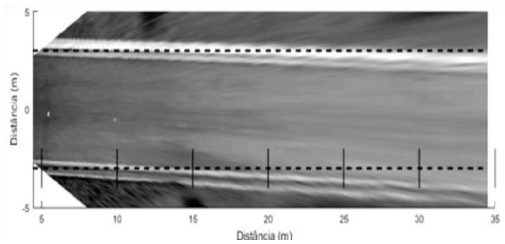


Fig. 3. Inverse perspective mapping with an ROI of 150 lines.

TABLE II : RANGE/HEIGHT OF ROI

Height [number of rows]	Range [m]	Range/Height [m/pixel]	ΔRange/ΔHeight [m/pixel]
100	10.38	0.1037	-
110	12.05	0.1095	0.1669
120	14.37	0.1198	0.2320
130	17.82	0.1372	0.3448
140	23.49	0.1678	0.5682
150	34.54	0.2304	1.1111

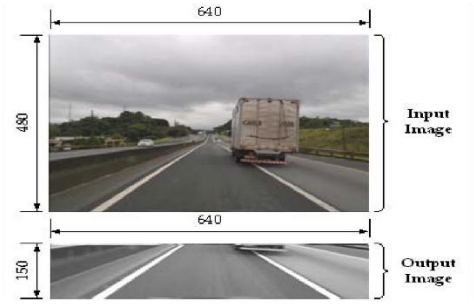


Fig. 4. Output from the low-level processing step.

be similar to Fig. 4. With the reduced frame size, defined by the ROI and reduced dimensionality, from three layers (RGB) to one (intensity).

This output is propagated to the mid-level processing steps as described in the sequence.

B. Mid-Level Processing

The mid-level processing holds a strategy to extract edges from the closest lane markings with the highest signal-to-noise ratio (snr), as depicted in Fig. 5. In this process, another color space conversion occurs, from intensity (grayscale) to binary.

The first stage of the process is the primary edge detector in the ROI; the canny filter was applied due to its characteristics of good detection, location, and singularity of the response for each edge [44]. However, there is no angular restriction in the

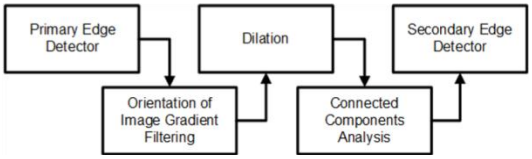


Fig. 5. Block diagram representative of mid-level processing.

Fig. 6. Canny edge detector.

Fig. 7. Changes in the angular orientation of the lanes in different scenarios.algorithm, so the object extraction is made regardless of the shape format.

Fig. 6 illustrates the results of feature extraction over the output of the low-level processing shown in Fig. 4. The excess of information extracted by the algorithm may trigger the appearance of undesirable events in the later processing stages, such as false positive marking.

Taking the assumption that, from the perspective of a camera fixed to the central region of the vehicle windshield, the orientation of the lanes fall within a predictable range of angles. It is possible to implement a filter to remove edges with the undesirable orientation of image gradient components.

However, in curved roads, the lane orientation undergoes considerable changes. The variation of the internal angle of the lanes is shown in Fig. 7 for the following scenarios, straight, left turn and right turn. This observation is relevant since research such as those of [20] use a measure of symmetry, defined experimentally, between these angles to determine the lane departure warning.

The yaw rate signal is directly related to the direction of movement of the vehicle [45]. Thus, the algorithm parameters were defined according to the experimental analysis of the signal behavior.

During overtaking and lane changes, the LDW must be deactivated because it characterizes an intentional act of

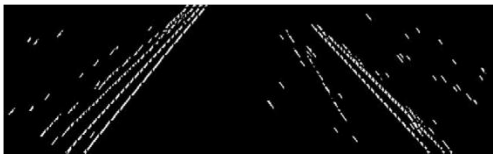


Fig. 8. Filtering on the image gradient orientation.

Fig. 9. Application of dilation to connect points of interest.

the driver. The drifting issue is a critical situation which can be solved with the application of other systems, like Electronic Stability Control.

Fig. 8 illustrates the result of the orientation of gradient filtering on the image of Fig. 6. The output presents the smaller number of white pixels in comparison to the output of the Canny filter, maintaining features of the edges in closes lanes the region. The remaining noise can be attributed to two main factors. Firstly, some pixels form edges with the same orientation as the highway lanes, and filtering is therefore insensitive. Secondly, the use of a safety range for the angles of interest to avoid excessive filtering of the image, keeping the extraction of features robust and functional regardless of the position of the vehicle.

Two disconnected pixels were analyzed at the third line of the figure.

By analyzing the histogram of the labels, it is possible to obtain the area of each set. The proposed algorithm maintains labels with an area greater than a threshold.

The selection of labels process output is shown in Fig. 11(a). Finally, the edge-thinning step provided in the Canny filter algorithm is applied to obtain the result of Fig. 11(b). The application of dilation increases the probability of the occurrence of connections between the pixels related to lanes and noisy-pixels, as depicted in some regions of the figure.

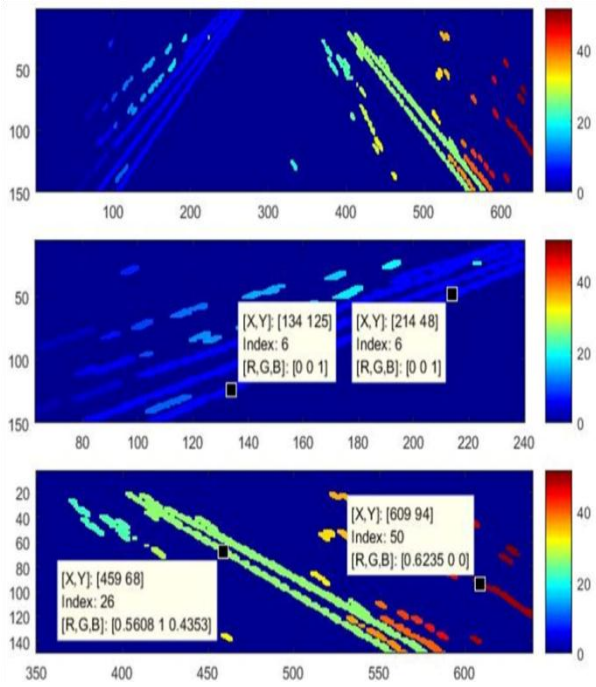


Fig. 10. Dilated image output labeling.

Fig. 11. Output of mid-level processing.

However, it is a tradeoff for reducing the computational efforts in the analysis.

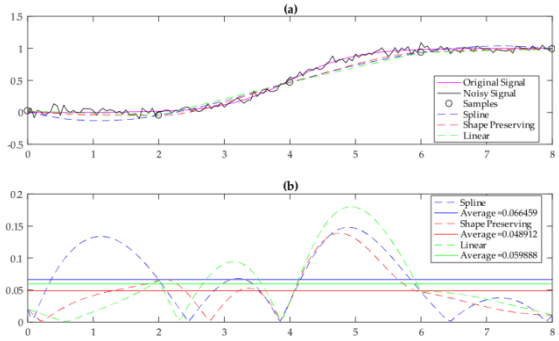
The output is then propagated to the high-level processing step.

C. High-Level Processing

The objects extracted in the mid-level processing step are input for the Hough transform (HT) algorithm to extract possible line segments from the parameter space. Based on [27], the ROI was divided into sections to accommodate the lane shape adequately. The height of each section is gradually reduced from the bottom to the top of the image. By analyzing the lower part of the image, it is assumed that the lane can be approximated by a line. In the distant field, this definition is only acceptable if the area analyzed is relatively small [20].

Fig. 12 illustrates this strategy applying the ROI division and the linear interpolation between the points of the four most relevant pairs of coordinates extracted from HT's accumulator matrix.

The analysis of a high number of coordinates increases algorithm robustness in the following scenarios:



If the most voted pairs of coordinates are inappropriate, due to a non-effective feature extraction;

To standardize the algorithm output tracking only the intern edges of the lane markings.

Fig. 12. Division of ROI sections and detected coordinates.

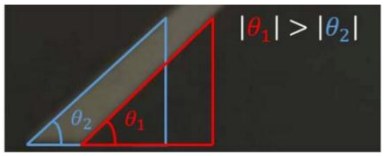


Fig. 13. Comparison between the angles formed by the inner and outer edges of the strip.

Thus, a coordinate selection step turns out to be mandatory. Our selection criterion is based on the line orientation and its evolution considering the neighboring sections. The coordinates extracted from the first section are the basis for the analysis of the subsequent sections. This choice is justified by the fact that the lower section has the largest area and is less affected by the perspective effect; consequently, it is the region with the highest integrity of the lanes.

The following steps presents, the line selection in the lower section (Fig. 12 item 1):

- Calculate the absolute value of the intern angle formed by the lower and upper points of the section;
- Check if the angles fit within the minimum and maximum orientation limits;
- Check if the horizontal distance from the lower point of the line fits within the specified range;
- Select the valid points referring to the line with the greater internal angle, if there are no valid points keep those from the previous iteration through n iterations.

The described process is valid for both lane markings. In summary, besides discarding invalid coordinates, the process favors the selection of the line with the largest internal angle, which represents the inner edge of the lane marking, as shown in Fig. 13. In this way, the system prioritizes the tracing of the edges closest to the vehicle.

If by applying the Hough transform, it is not possible to extract valid points by a large period, the system will no longer display lane tracking until a successive sequence of valid coordinates. Such hysteresis is necessary to avoid periods of intermittent system operation. The operation of the system is also interrupted when the driver intentionally changes lanes, so vehicle's status is monitored from the CAN bus [46].

For the subsequent sections of the image, the coordinate selection process can be similar. However, the third step is

Fig. 14. Analysis of the behavior of the interpolation functions.replaced by checking over the variation in the direction of the line concerning the coordinates selected in the previous section. If the difference between the angle of the same lane in the analyzed section and the preceding section is greater than the proposed threshold, the coordinate under analysis is invalidated.

In the process of checking the coordinates of the last section a fifth step, responsible for activating an indicator

when none of the coordinates analyzed have valid points in the iteration is added. The indicator's function is to deactivate lane tracking visual representation in this section which as a result of the camera perspective effect is a region where each pixel represents a larger area in the image in relation to the near field

Fig. 15 shows the interpolation of the selected points as the conclusion of the lane-tracking process. As expected it is possible to note that due to the point verification and selection steps the tracing occurs on the inner edge of the lanes. Thus, the completion of the low, medium and high-level processing stages completes the development of the detection and tracking strategy proposed in this work.

Fig. 15. The algorithm's output in a straight line.

\

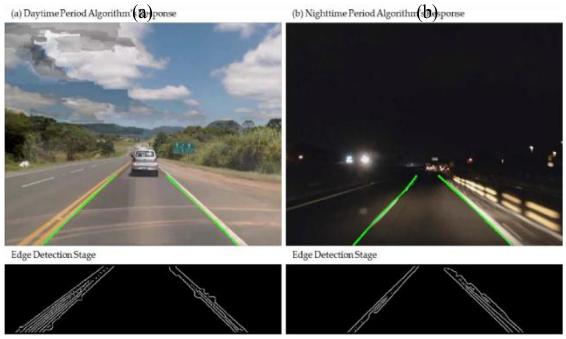


Fig. 16. The algorithm's response in (a) daytime and (b) nighttime periods.

The application of this strategy in different scenarios is the base for the experimental outcomes discussed in the next section.

IV. EXPERIMENTAL OUTCOMES

This section begins with the presentation of the results and their discussion from the qualitative point of view. This analysis contributes to the critique and understanding of the reasons that led the algorithm to present the observed behavior in order to expose the right decisions and the flaws of the proposed strategy.

A. Qualitative Analysis of Results

The algorithm should operate in daytime and nighttime periods, as shown in Fig. 16. However, it is known that some situations may compromise the image acquisition by the camera, such as glare caused by direct sunlight, the reflectivity of the asphalt, the visibility of the lane markings and the presence of obstacles or dirt on the road.

In ordinary conditions, the algorithm presents an equivalent response in both periods. Fig. 17(a) depicts a scenario where the tracking suffers a position deviation in relation the real lane marking due to opposite vehicles headlight incidence. It becomes evident in edge detection stage since the highlighted region has a signal-to-noise ratio much smaller than the right lane.

Detail 1 of Fig. 17(b) exposes the effect of implementing, where tracing was maintained even with the presence of

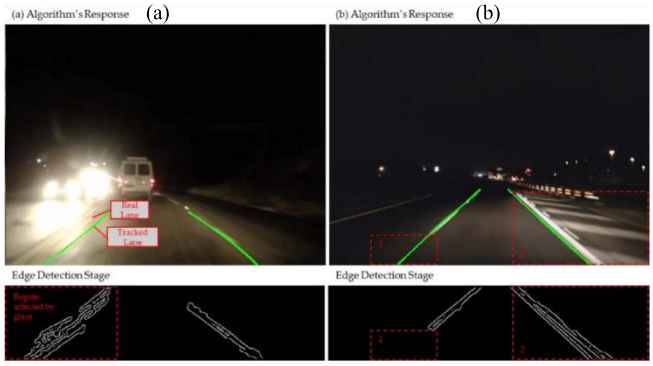


Fig. 17. (a) Obfuscation caused by the headlamp of the vehicle in the opposite direction and (b) maintenance of the last valid coordinates

gap located in the first section. The detail 2 shows the filtering based on the orientation of the image gradient, in this scenario the edge detection extracts only the features inside the delimitation range.

The visual representation of the road far-field tracking is disabled if there are no valid coordinates in the set of points. This algorithm's activation feature occurs especially when the preprocessing is not able to extract the edges at the top of the ROI, which is the case shown in both scenarios of Fig. 18.

In (a), the non-detection of the upper part of the lanes is due to the segmentation in the marking of the left lane. In (b), nondetection may be associated with the reduced contrast between the yellow marking and the asphalt caused by shadow.

Fig. 19(a) exposes a case where lane tracking was improper due to the detection of contrast-related edges on asphalt. As the contrast occurs is in the same orientation of the lanes, it was not possible to filter its components.

The road lanes always must be recognized and only the application function should not generate requests. In Fig. 19(b) driver intention is displayed in detail 1.

B. Quantitative Analysis of Results

The quantitative results were obtained through the use of the RoadMarker in-house software, developed based on the [47] Lane Position Deviation (LPD) technique, in combination with the Root Mean Squared Deviation (RMSD). The software allows the ground truth definition, which consists of a

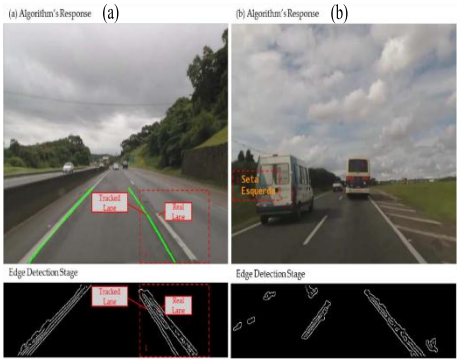


Fig. 19. Detection of edges related to asphalt contrast.

Fig. 20. The process of ground truth marking.

frame-to-frame lane position manual marking. Fig. 20 illustrates the result of a ground truth marking for a frame.

The evaluation process takes both, ground truth and algorithm output files, as inputs. It is necessary to define two input parameters: "ROI n° of rows" and "Deviation threshold." The first one defines the region of interest (in pixels) of the image evaluated by the software. Any coordinate that lies beyond the ROI is disregarded. The "Deviation threshold" is used to establish the algorithm accuracy, based on the LPD result, which will be described in the sequence.

$$= \frac{\delta LPD}{(1)\ y\max - y\min\ _{y=y\min}} \delta x$$

=

The equation (1) describes the LPD calculus. The y variable defines the size of the ROI, which is given by the entrance parameter, "ROI n° of rows." Thus, the ymin represents the basis of the image, and the ymax represents the value defined by the parameter. The δx represents the quadratic x-variation, from the ground truth lane position to the algorithm estimated position. The results for each frame are then compared to the "Deviation threshold" parameter, where if the modulus of result value is equal or below the parameter, it is considered a true positive (T P), otherwise a false positive (FP). If there is no algorithm output for an evaluated line (within the ROI), it is also considered a false positive. Having this values, it is possible to calculate the accuracy (ACC) percentage of the algorithm regarding the analyzed video sequence, as shown in

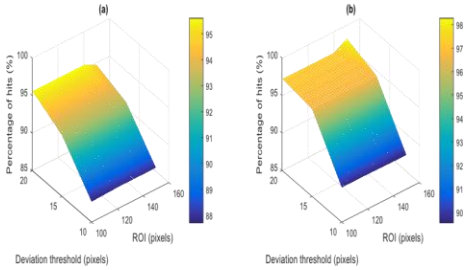


Fig. 21. The general mean of accuracy index.
the equation (2).

$$= ACC \frac{TP}{TP + FP} \quad (2)$$

The results presented by RoadMarker are composed by the calculated ACC, the absolute mean of the RMSD (which consists of the square root of the LPD), and standard deviation. The analysis was performed over six samples segmented from the local GSA database. The selection criteria were the diversity of the scenarios, which includes the lane marking type, the daytime period and the presence of contrast, shadow or glare in the sample. The performance indexes were calculated for each combination of ROI length (100, 120, 140 and 150) and deviation threshold (10, 15 and 20).

We provide the outcomes of road testing on the GSA portal (http://pg.utfpr.edu.br/gsa) in the section "Publications". It includes the complete statistical analysis, tables, and screenshots from the used video samples scenarios.

Fig. 21 shows the surface plot for the accuracy average considering the six samples. In general and as expected, the tracking accuracy grows up with higher deviation threshold values. Therefore an analysis of the maximum allowed deviation should not be neglected to deploy this strategy in a commercial vehicle. The variation of the ROI does not have the same impact on the results; instead, it remains at the same level which proves that the algorithm presents similar accuracy at the near and far fields of the input image.

The results concentrate in the range of 87∼98% of accuracy. Considering the diversity of the test scenarios, the strategy shows promising results. Although it will be necessary to go further in the research to improve results in certain critical conditions, such as reflexive roads, camera's saturation, and poor road infrastructure conditions.

V. CONCLUSIONS

The presented outcomes have been acquired based on the samples from real-world scenarios of the road environment as of the road testing, which the images and vehicle signals have been acquired from a monocular camera installed in a commercial vehicle windshield and instrumentation of the data bus.

The use of IPM algorithm allows the range determination, of the tracking for several ROI sizes. The smaller ROI (100 lines) analyzed covers a range of 10.4m ahead of the vehicle; the largest (150 lines) reaches a coverage of 34.5m. Considering the average reaction time for braking presented in the work of [41], the minimum ROI to avoid collision with a static obstacle at the front of the vehicle before the driver reaction time should have 120 lines, which covers about 14.4m.

From the qualitative analysis of the results of the algorithm, it was possible to detect the scenarios where the tracking of the algorithm is more assertive, as well as the highway patterns or environmental conditions where the tracking is unstable. It was observed that the visibility of the lane markings could be compromised by several factors, such as the reflectivity of the lane, the camera's glare, the presence of shadows and the deterioration of the paint.

The highest assertiveness indices were observed in tracks with continuous lane markings; the lowest performance indices were observed double marked with the inner lane dashed. Due to the fluctuation of the tracing between the internal and external part of the marking, this scenario presented the largest absolute mean error among the samples.

REFERENCES

[1] E. Holden Achieving Sustainable Mobility: Everyday and Leisure-Time Travel in the EU. Farnham, U.K.: Ashgate, 2012.

[2] D. Schrank, B. Eisele, and T. Lomax, "TTI's 2012 urban mobility report," Texas A&M Transp. Inst., Texas A&M Univ. Syst., Austin, TX, USA, Tech. Rep., 2012.

[3] "Critical reasons for crashes investigated in the national motor vehicle crash causation survey," NHTSA, Washington, DC, USA, Tech. Rep. DOT HS 812 115, Feb. 2015, p. 2.

[4] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner, "Three decades of driver assistance systems: Review and future perspectives," IEEE Intell. Transp. Syst. Mag., vol. 6, no. 4, pp. 6–22, Oct. 2014. [Online]. Available: https://ieeexplore.ieee.org/ document/6936444/

[5] Volvo Trucks. (2013). European Accident Research and Safety Report. [Online]. Available: http://www.volvotrucks.com/ iteCollectionDocuments/VTC/Corporate/Values/ART%20Report% 202013_150dpi.pdf

[6] FHWA. (Aug. 26, 2016). Roadway Departure Safety. Accessed: Nov. 7, 2016. [Online]. Available: http://safety.fhwa.dot.gov/roadway_ dept/

[7] N. Naylor, "NHTSA announces final rule requiring rear visibility technology," Tech. Rep., 2014. [Online]. Available: https://www. transportation.gov/briefing-room/nhtsa-announces-final-rule-requiringrear-visibility-technology

[8] Application of Star Rating. 2015 European New Car Assessment Programme. Version 1.2, Jun. 2015. [Online]. Availabe: https://cdn. euroncap.com/media/27428/euro-cap-application-of-star-ratingsprotocol-v14.pdf