A Vision-Based Lane Detection System Combining Appearance Segmentation and Tracking of Salient Points

Vitor S. Bottazzi
School of Information and
Communication Technology
Griffith University
Gold Coast, Australia
vitor.bottazzi@griffithuni.edu.au

Paulo V. K. Borges
Autonomous Systems Laboratory
CSIRO ICT Centre
Brisbane, Australia
vini@ieee.org

Jun Jo
School of Information and
Communication Technology
Griffith University
Gold Coast, Australia
j.jo@griffith.edu.au

Abstract—Reliable lane detection is a key component of autonomous vehicles supporting navigation in urban environments. This paper introduces the GOLDIE(Geometric Overture for Lane Detection by Intersections Entirety) system, a vision-based software architecture that uses an on-board single camera to determine the position of road lanes with respect to the vehicle. We propose an efficient vision-based lane-detection system that combines an appearance-based analysis with salient point tracking. The appearance-based analysis consists of segmenting high contrast areas that fit inside a Region-Of-Interest(ROI) on the frame. The salient point tracker selects interesting points based in a reference line, that guides a dynamic ROI. The tracking ROI look for paint lane marks close to the last lane reference found, where road marks are likely to emerge, in order to maintain the usability of the salient point tracker. The tracking is performed with the Lucas-Kanade algorithm and the lane points candidates are selected according to a predefined triangular model. Once such lanes points are detected, the vehicle position is estimated based on the intersection of linearised lanes determined through a vanishing point approach. Experiments and comparisons with other algorithms illustrate the applicability of the method.

I. INTRODUCTION

In recent years, significant research has focused on autonomous driving vehicles using lasers [1], [2], and cameras [3]–[5], in an attempt to make autonomous driving in urban environments possible. An important component of urban driving is automated lane detection, which finds application in both autonomous and manned vehicles, playing a significant role in driver assistance systems. Lane detection is commonly employed to localise road boundaries, identify unwanted lane changes, and to help estimate the upcoming geometry of the road. One broadly used approach for lane detection is to employ expensive computer vision techniques to extract specific features from road images in order to identify road marks. In such systems, real time performance is a key requirement when the application targets autonomous vehicle navigation.

In this paper we propose an computationally efficient vision-based lane detection algorithm that combines an appearance-based analysis with salient point tracking. This additional process reduces the probability of missing the lane mark reference due to discontinuity or environmental changes which may disrupt the Hough [6] line detection effort. The appearance-based analysis consists of segmenting high contrast areas within strategic positioned ROIs applying the Hough transform to find extended road-reference lines. In parallel to this process, Lucas-Kanade algorithm is used to track the paint marks selecting the interesting points based on the segmented lane lines position.

Once a match between the two cues are detected, the vehicle's position on the road is estimated using a mathematical model. The directional information is extrapolated from the triangular approximation [7] of the road through model fitting. The distance between two arbitrary intersections detected inside the road visible area, which is bounded by the ROIs, is used to calibrate the triangle's baseline size. This approach provides fast calibration when solid line references are available and reliable localisation for the system to estimate the correct position of the car in relation to the road.

The GOLDIE system presents a low computational complexity and we compare the results with other state of the art algorithms based on inverse perspective mapping [2] and spline models [8], illustrating the applicability of our solution. This paper is organised as follows. The section 2 presents the related work and some state of art lane detection algorithms. Section 3 describes the two cues that compose our solution while section 4 shows the experiment setup with it practical considerations, results and comparison with other approaches. Section 5 conclude the paper and presents the future work.

II. RELATED WORK

A number of works on lane detection have been proposed, focusing on highways [3], [9] or urban environments [2], for single [2], [3], [8], [9] or multiple [4] lanes. Common approaches include detection based on region texture [10], colour [4], [11] or edge points [3], [9]. Several models have been applied to describe borders of a lane, such as piecewise linear segments [12], clothoids [13], parabola [3], hyperbola [9], splines [8], or snakes [6], [12]. Several lane detectors have been implemented and named in literatures, such as GOLD [14], SCARF [11], RALPH [15], MANIAC [16], and LANA [17]. The lane tracking issue is addressed in most of the research articles like Kodagoda [1], Sehestedt [2], Huang [12] and Butdee [18]. Most of the cited articles use extra sensors to assure a valid road boundary reference to start with, including ladar and multiple video cameras.

There are mainly two well defined approaches [8] applied to lane detection using video: feature-based and model-based. The feature based technique aims to localise the lanes using traditional image segmentation [3] searching for lane segments. Those segments are extracted from painted lines, road edges, curbs, etc. The main disadvantage of this method is that it does not impose constraints on lane shapes, which may cause faulty lane detection due to occlusion or noise. The model based approach represents the lane using mathematical models, in order to represent a set of parameters defined by an equation. This approach can easily model straight and parabolic curves on the road based on the lines and splines equations respectively. Our research is focused on the visual information extracted from a single camera. Several common situations were used for a case study to select valid data inputs as shown in Fig 1. This research also employed both the feature-based and model-based approaches.

Fig. 2: Vision pipeline illustrating the parallel lane detection analysis through appearance-based segmentation and salient point tracking.

(a) Ideal situation   (b) Lane occlusion   (c) Abnormal paint

Fig. 1: Driving input samples

III. PROPOSED METHOD

The proposed method combines two independent but simultaneously running cues for lane detection (as illustrated in Fig 2), which consist of the following modules:

- An appearance-based approach for colour based lane segmentation.
- Tracking of feature points and segmenting points that can be clustered according to lane-like characteristics.

Appearance-based detection is executed in parallel with the lane marks tracking to identify the salient point of the road. Both lane references are compared and used to process t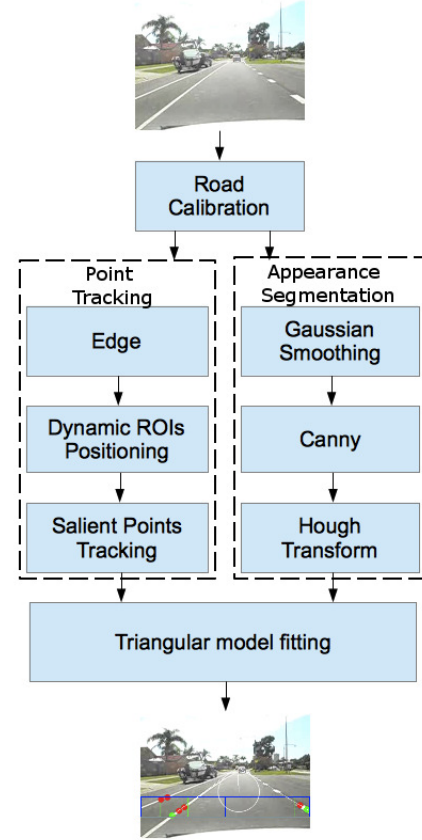he model fitting. The direction and positions of lanes are estimated following a trigonometry approach, which is explained in Section III-D.

A specific road mark pattern (we employ a triangular model) must be identified before applying the trigonometric algorithm to reconstruct missing lane references. This pattern is based on the distance between extended lane segments and the vanishing point [3]. The automatic calibration positions the static segmentation region-of-Interest (SSROI) shown in Fig 3, before the appearance-based analysis is performed. In the following we describe in more detail the components of the system, elaborating the pipeline presented in Fig 2.

A. Automated Road Calibration

B. Appearance-Based Lane Segmentation

The calibration step aims to optimally position the SSROIs using a small number of valid samples, as exemplified in Fig 1a. Firstly, a predefined position at the centre of the bottom half of the image is determined for the SSROIs. Secondly, the image segmentation is applied over the predefined position to calculate the middle distance in between the vanishing point and the bottom part of the image, where the road lanes are detectable, to calibrate the SSROI position.

444

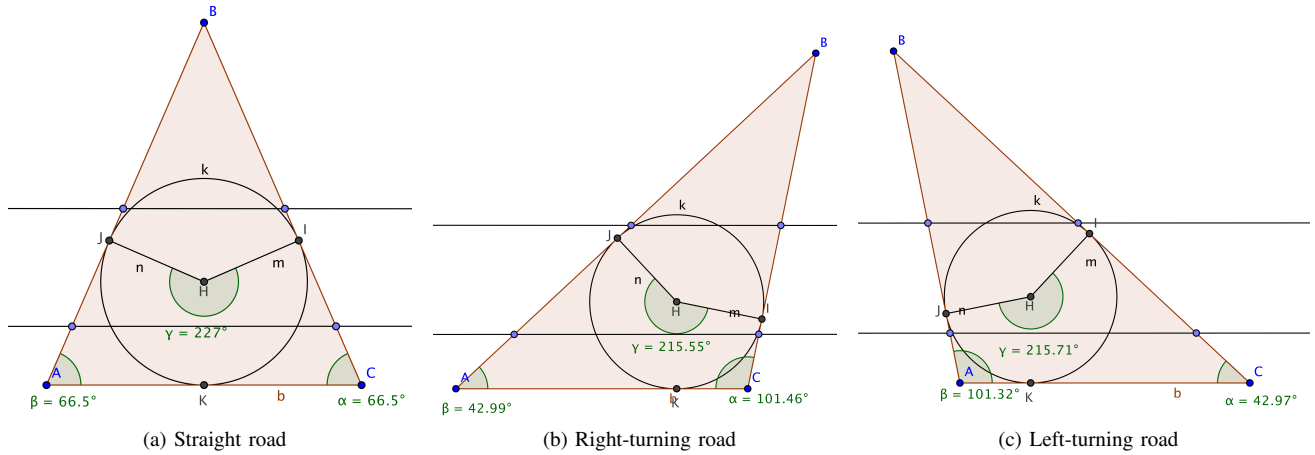| (a) Straight road | (b) Right-turning road | (c) Left-turning road |

Fig. 4: Trigonometry applied to lane tracking.

From this point, the extended lanes segments are extracted from the new SSROIs position and the distance between extended lane segments is calculated. This step aims to approximate the road size, setting up the triangle's baseline ($\overline{AC}$ in Fig 4a) as reference.

The SSROI boundaries are setup by utilising the visible road. The road is captured by a camera that is positioned *approximately* parallel to the ground, aiming to capture a similar perspective of a driver's field of view. This camera scans both sides of the road separately using two SSROI to ideally identify the left and right lane marks, without needing previous classification. After the calibration, the Gaussian smoothing technique is applied to each SSROI, in order to reduce image noise and to improve contour extraction. The SSROIs are then converted to gray scale, to serve as input for the Canny algorithm to extract its contours.

Subsequently, the Hough transform is applied to each road-mark's contour to extract lane line segments. We setup the SSROIs to strategically reduce the amount of "non-interesting" lines found by the Hough transform. This schema reduces the image processing overhead by eliminating objects such as trees, plates and buildings, focusing on extended lane segments creation. The appearance-based lane detection

Fig. 3: Triangular lane model showing the static segmentation ROIs (SSROI) and the dynamic tracking ROIs (DTROI).

approach produces positive results for continuous lane mark and high contrast curbs. However, there are several environment variables that can affect vision-based lane detection such as missing lane marks, road painting information for drivers, road works, wet areas, and surface reflection. The feature tracking algorithm explained in the next section is used to confirm lane mark presence. Furthermore the lane mark references are used in the triangular model fitting.

C. Lane Estimation Based on Tracked Points

One of the main advantages of isolated point tracking is that the system becomes more robust to shadows, different lane colours, illumination changes and lane mark discontinuity. In urban environments lane marks often become discontinuous, disappear or change colour depending on traffic rules (i.e. white to yellow). The Lucas-Kanade [19] algorithm for feature tracking is applied to support the appearance-based method mentioned in Section III-B. We use lane marks tracking to ensure that the correct road lane references are captured by the segmentation stage and also to use it in the model fitting. Additionally, the algorithm considers the relative motion of the lane mark references, within an expected range.

Lane detection systems are expected to keep track of the different lane mark colours on various road sizes. We use two dynamic tracking ROIs (DTROI) moving according to the lane marks distance found in the automatic calibration stage.

1) Dynamic Tracking ROI: The salient point tracking uses two small dynamic regions-of-interest that are strategically positioned according to the triangle's baseline references, shown in Fig 3. The DTROIs are re-positioned every frame, according to the intersections found by segmentation stage and assuming that the calibration stage has detected the correct road size, the tracking DTROIs is placed optimally. The process tracks the last intersection area positioning the DTROIs on each left/right intersection limits detected previously towards the road.
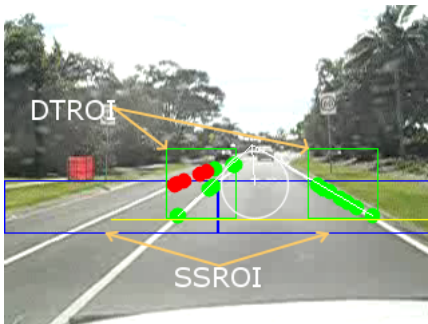
445

The initial position of the DTROIs is defined during calibration stage discussed earlier (see Section III-A). The DTROI then re-position itself according to the past lane reference validated by both segmentation and tracking schema, shown in Fig 3. The distance between the lanes is validated by the triangle's baseline size. If no intersection is detected, the DTROIs search for road marks horizontally up to the maximum expected road size setup at calibration.

2) Tracking: The LK tracker uses a Kanade-Lucas-Tomasi feature tracker, based on the Harris [20] corner detector. Points are tracked across consecutive frames and selected following a linear approximation model based on the determinant of a triangular matrix M (1):

$$M = \begin{bmatrix} X_{vp} & Y_{vp} & 1 \\ X_{lp} & Y_{lp} & 1 \\ X_{tp} & Y_{tp} & 1 \end{bmatrix} \tag{1}$$

where (tp) is the tracked point being classified, (vp) is the vanishing point, and (lp) is the lateral position represented by the left/right intersection. vp and lp are represented by points B and C respectively shown in Fig 4a, whereas P_3 and P_4 (shown in Fig 5) represent the tp set being classified. The determinant D (2) of M (1) is used to classify points:

$$D = \begin{vmatrix} X_{vp} & Y_{vp} & 1 \\ X_{lp} & Y_{lp} & 1 \\ X_{tp} & Y_{tp} & 1 \end{vmatrix} \tag{2}$$
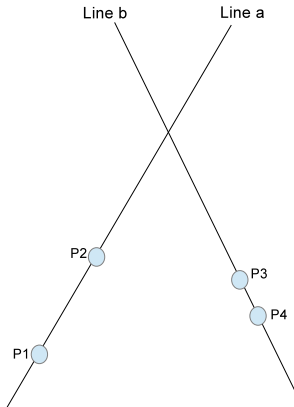

Fig. 5: Extended lane segments.

where $D = 0$ means that the three points are collinear. Additionally, we can test the numerical value of (2) for every $D \neq 0$, searching for thin triangle's areas, see Fig 6a.

This schema guarantees that only points close enough to both sides of the analysed lane line are selected. In our implementation, we setup the selection interval from -500 to 500 pixels as candidate point criteria. This interval corresponds to thin triangle areas with less than 251 pixels. We can discard the triangle area calculation (3) as the determinant already give us enough precision to classify candidates.

$$A = \frac{|D|}{2} \tag{3}$$

The confirmed lane-boundaries are used to maintain the proposed geometric model explained in the next section.

D. Road boundary prediction based on trigonometry

We propose a triangle model to control reference points extracted from the road. One of the most common issues regarding the road lane detection occurs when road mark references disappear. The road reference direction relative to the car is determined by detecting the vanishing point of the road, such as the top intersection shown in Fig 3. The Fig 4a denotes the triangular model concept, where the parallel lines cutting the middle section of the triangle represents the visible part of the road. The two SSROIs are placed within the parallel lines shown in the model to segment each lane separately. The two extended lanes intersect each other above the horizon of the road defining the vanishing point B, as shown in Fig 4.

Lines can be described using two points P_1 and P_2. In a two dimensional representation $P_n(x, y)$, four generic points P_1, P_2, P_3 and P_4 can create an intersection between two lines on the same plane. The intersection of those two line segments is shown in Fig 5. Line a can be described by the equation $P_a = P_1 + u_a(P_2 - P_1)$ and line b is similarly represented by the equation $P_b = P_3 + u_b(P_4 - P_3)$. This is the generic equation system representing two lines intersecting at point P_i for two unknowns (u_a and u_b).

If we consider that $P_i = P_a = P_b$ describe the generic intersection between line a and line b, we can find two new equations from the system of equations above: $x_1 + u_a(x_2 - x_1) = x_3 + u_b(x_4 - x_3)$ and $y_1 + u_a(y_2 - y_1) = y_3 + u_b(y_4 - y_3)$. The solution of this system is given in equations (4) and (5).

$$u_a = \frac{(x_4 - x_3)(y_1 - y_3) - (y_4 - y_3)(x_1 - x_3)}{(y_4 - y_3)(x_2 - x_1) - (x_4 - x_3)(y_2 - y_1)} \tag{4}$$

$$u_b = \frac{(x_2 - x_1)(y_1 - y_3) - (y_2 - y_1)(x_1 - x_3)}{(y_4 - y_3)(x_2 - x_1) - (x_4 - x_3)(y_2 - y_1)} \tag{5}$$

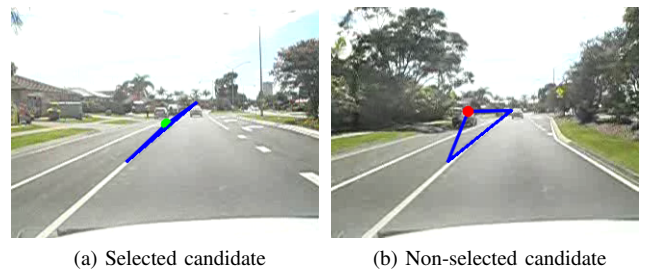(a) Selected candidate   (b) Non-selected candidate

Fig. 6: Salient points classification

Next we calculate u_a using the cartesian coordinates from points P_1, P_2, P_3 and P_4 and apply it into the equations (6) and (7) to determine the vanishing point P_i. The same process is used to detect the other two triangle's vertex, using each one of the extended lane marks plus the triangle's baseline.

446

$$x = x_1 + u_a \left(x_2 - x_1 \right) \qquad (6)$$

$$y = y_1 + u_a \left(y_2 - y_1 \right) \qquad (7)$$

We can also observe in Figs 4b and 4c, the changes in the internal angles in relation with the triangle's configuration. The triangular road model is built using three intersection points. The first intersection, which is the vanishing point, is found at calibration stage and represented by point B in Fig 4a, such that the extension of two linearised lane marks is found. The consecutive two intersections found by segmentation (points A and C) are used in combination with the point tracking to accurately estimate the road lane positions. The model fitting block shown in Fig 2 uses the points A, B and C illustrated in Fig 4 to predict the lateral position of the road marks according to the input images provided by the on-board camera.

IV. EXPERIMENTS

In this section we present results related to the computational complexity and the detection rate of the algorithm. We divide the results in three parts: (i) colour segmentation, (ii) tracked points estimation, and (iii) a combination of both colour segmentation and tracked points estimation. Similarly to other state of art methods, the experiment showed the capability of identifying and reconstructing missing parts of the road lanes using the proposed approach.

A. Practical Considerations

Apart from the recorded videos, we also evaluated the algorithm using a large lane detection database [5] where different day light scenarios were chosen to validate the GOLDIE approach. The expected road width in pixels must be provided, due to the different video resolutions used in the experiment, to guarantee that the calibration stage uses a single lane as reference. This parameter is used to test the lane width, where the car is positioned, avoiding wrong reference calibration. The three different videos selected from [5] were S1C1, S6C2, and S7C2. Their characteristics are respectively:

- a clear day with some clouds and predominant discontinuous road lines;
- a cloudy day with predominant abnormal road painting and heavy traffic; and
- a cloudy day during dawn time with heavy traffic and cars light disturbance.

All those videos used the same resolution 640×480 with 30 frames per second frame rate.

The sample videos used for development were recorded in a clear day with no clouds in real residential roads in Australia. The path was chosen due to predominancy of continuous line painting, the ideal scenario for road calibration purposes, with transitions to discontinuous line painting and shadows in both sides of the road. Additionally, the data sample includes areas where the lane references change

position and direction offering a favourable setup for the development phase. The resolution of the sample video is 320×240 and the frame rate is 30 frames per second. In total, 599 frames containing lanes and a number of road marks were recorded. Some snapshots of the footages are shown in Fig 1.

B. Results

The appearance-based module was able to detect the left, right or both lanes in 78% of all lane instances. Detecting a single lane through pure segmentation may also be regarded as a successful detection as the position of the other lane may be extrapolated using the road size reference determined during calibration stage.

The single point tracking showed positive results under different light conditions, achieving 97.8% of accurate *road-paint tracking*. The 2.2% fail rate was caused by discontinuous lane painting and non-interesting road painting marks. These "non-interesting" marks include speed limits indicators, bicycle lane paintings and turning signs marked on the road. The point tracking had an excellent response to environment variables, however there is a lane classification demand. The tracked point cloud needs a reliable lane model reference to accurately isolate and track lane marks. It is applied when both cues are combined.

According to the experiments, the segmentation algorithm combined with the DTROI provides three reliable points to build a triangular model that accurately estimates the lateral road lane position. The combination of both approaches improved the system reliability to 99.7% based on the local data sample. The 0.3% false positives, similar to the second cue, were caused by discontinuous lane and pedestrian crossing painting. A small sample showing the results can be found at http://www.youtube.com/watch?v=VMIAB2ZCFwU. The statistics were extracted using the indicators from Table I on a *per frame* basis.

Sample data	False pos.	False neg.	Performance
Recorded sample	2%	-	98%
S1C1	8%	1%	91%
S6C2	12%	6%	82%
S7C2	37%	-	63%

TABLE I: Lane detection with no spatiotemporal filtering support.

C. Comparisons

Similar to other segmented based approaches [3], [4], [9], the appearance-based segmentation approach is not reliable enough to deal with large variations of intense light and shadow. False true patterns are mismatched by vehicles changing lanes, abnormal lane paint, road works, non-uniform lanes variation, among others as illustrated in Figs 1b and 1c. Nonetheless, GOLDIE performed efficiently. It can be combined with semantic aspects of roads (i.e. different lane sizes) and be supported by other sensorial modalities to support autonomous driving experience.

447

Additionally, the algorithm was tested using the database sample published in [5], providing 3 different scenarios with challenging light conditions to validate the proposed approach. The results were promising achieving overall average detection and tracking up to 84 %. It is important to mention that the high rate of false positives detected on samples *S6C2* and *S7C2* were mainly caused by long calibration effort at the start of the video caused by discontinuous lane marks, and floating road size during wide curves merging different highways. The sample video *S6C2* starts in the middle of discontinuous lane marks demanding a long time for the road size calibration algorithm. After the calibration is done, the false positive rate drops considerably on sample *S6C2* as shown at http://http://www.youtube.com/watch?v=wZyjE0gcZBU.

The huge number of false positives detected on samples *S7C2* was caused by merging roads with different sizes, car lights arising within the ROIs and on-board car head lights being switched on as shown at http://http://www.youtube.com/watch?v=YslrZKdv1Mg.

Spatiotemporal filtering was not used in this experiment, aiming to extract the best performance possible from the combination of our appearance-based segmentation and point tracking. The computer used to simulate the lane detection is an ordinary Intel Core 2 Duo 1.4 GHz with 4GB of RAM. The GOLDIE system attended to the real-time response demands running a full lane detection cycle in less than 0.1 second.

V. Conclusions

We have proposed a new framework for the detection of road lanes, combining colour based lane segmentation with feature point tracking lane estimation. The traditional colour detection method is supported by the lane estimation, based on points that are tracked across consecutive frames, and selected following an approximated linear model. It does not require excessive computational effort compared to other state of the art methods, as the method does not rely on complex geometric fitting models, presenting higher performance in terms of detection rate.

The main benefit of the GOLDIE is that the system becomes more robust to lighting variations, shadows across the road, breaks on lanes and lane colours changes. The results verify that the method presents good performance under different road sizes and light conditions. However, one of the limitations of the algorithm is that it relies on well painted road marks(or high contrast curbs) in the calibration stage.

Future work includes avoiding false positives caused by vehicles merging to the current lane using vehicle tracking, to get a more accurate estimation regarding lane marks distance through homography, and to implement a spatiotemporal filter to reduce abrupt interferences caused by light sources moving close to the lane marks.

References

[1] W. S. Wijesoma, K. R. S. Kodagoda, and A. P. Balasuriya, "Road-boundary detection and tracking using ladar sensing," *IEEE Transactions on Robotics*, vol. 20, no. 3, pp. 456–464, 2004. [Online]. Available: http://dx.doi.org/10.1109/TRA.2004.825269

[2] S. Sehestedt, S. Kodagoda, A. Alempijevic, and G. Dissanayake, "Efficient lane detection and tracking in urban environments," in *Proc. European Conf. Mobile Robots*, 2007, pp. 126–131.

[3] S. Zhou, Y. Jiang, J. Xi, J. Gong, G. Xiong, and H. Chen, "A novel lane detection based on geometrical model and gabor filter," in *Intelligent Vehicles Symposium*. IEEE, 2010, pp. 59–64. [Online]. Available: http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=5535007

[4] K. Chiu and S. Lin, "Lane detection using color-based segmentation," in *Intelligent Vehicles Symposium, 2005. Proceedings. IEEE*. IEEE, 2005, pp. 706–711.

[5] A. Borkar, M. Hayes, and M. Smith, "A novel lane detection system with efficient ground truth generation," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 13, no. 1, pp. 365–374, 2012.

[6] D. Kang, J. Choi, and I. Kweon, "Finding and tracking road lanes using line-snakes," in *Intelligent Vehicles Symposium, 1996., Proceedings of the 1996 IEEE*. IEEE, 1996, pp. 189–194.

[7] V. Bottazzi and J. Jo, "Trigonometry-based road lane detection," *International Symposium of Robotics for Education - ISRE*, 2010.

[8] Y. Wang, D. G. Shen, and E. K. Teoh, "Lane detection using spline model," *Pattern Recognition Letters*, vol. 21, no. 6-7, pp. 677–689, jun 2000.

[9] A. Assidiq, O. Khalifa, R. Islam, and S. Khan, "Real time lane detection for autonomous vehicles," in *Computer and Communication Engineering, 2008. ICCCE 2008. International Conference on*. IEEE, 2008, pp. 82–88.

[10] J. Zhang and H. Nagel, "Texture-based segmentation of road images," in *Intelligent Vehicles' 94 Symposium, Proceedings of the*. IEEE, 1994, pp. 260–265.

[11] J. Crisman and C. Thorpe, "Color vision for road following," *Vision and Navigation: The Carnegie Mellon Navlab*, 1988.

[12] S.-S. Huang, C.-J. Chen, P.-Y. Hsiao, and L.-C. Fu, "On-board vision system for lane recognition and front-vehicle detection to enhance driver's awareness," in *ICRA*. IEEE, 2004, pp. 2456–2461. [Online]. Available: http://dx.doi.org/10.1109/ROBOT.2004.1307429

[13] R. Danescu, S. Nedevschi, M. Meinecke, and T. To, "Lane geometry estimation in urban environments using a stereovision system," in *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*. IEEE, 2007, pp. 271–276.

[14] M. Bertozzi and A. Broggi, "Real-time lane and obstacle detection on the gold system," in *Intelligent Vehicles Symposium, 1996., Proceedings of the 1996 IEEE*. IEEE, 1996, pp. 213–218.

[15] D. Pomerleau and T. Jochem, "Rapidly adapting machine vision for automated vehicle steering," *IEEE expert*, vol. 11, no. 2, pp. 19–27, 1996.

[16] T. Jochem, D. Pomerleau, and C. Thorpe, "Maniac: A next generation neurally based autonomous road follower," in *Proceedings of the Intl Conf on Intelligent Autonomous Systems: IAS-3 Pittsburgh, Pennsylvania, USA*, 1993.

[17] C. Kreucher and S. Lakshmanan, "Lana: a lane extraction algorithm that uses frequency domain features," *Robotics and Automation, IEEE Transactions on*, vol. 15, no. 2, pp. 343–350, 1999.

[18] S. Butdee and A. Suebsomran, "Automatic guided vehicle control by vision system," in *Industrial Engineering and Engineering Management, 2009. IEEM 2009. IEEE International Conference on*. IEEE, 2009, pp. 694–697.

[19] J. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, 2001.

[20] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey vision conference*, vol. 15. Manchester, UK, 1988, p. 50.