2018 15th International Conference on
Control, Automation, Robotics and Vision (ICARCV)
Singapore, November 18-21, 2018

Off-Road Lane Detection Using Superpixel Clustering And RANSAC Curve Fitting

Sanskar Agrawal*1 Indu Kant Deo*2 Siddhant Haldar*3 G Rahul Kranti Kiran*4 Vaibhav Lodhi5
Debashish Chakravarty6

Abstract— **Lane detection is the most important issue to be resolved for successful locomotion of Intelligent Ground Vehicles (IGV). Problems in lane detection often occur in an external setting mainly due to glare or shadow defects. A robust and real-time approach to off-road lane marker detection for IGVs is being presented here. A novel model fitting based lane detection algorithm has been developed. Linear combination of image planes is used which removes the background and uncovers the white lanes. Simple Linear Iterative Clustering is applied to the processed frame and essential thresholding is performed for noise reduction. Two operations namely a novel approach for lane model identification and estimation of chosen lane mode using RANSAC are followed in sequence on the obtained image.The proposed image processing pipeline has been successfully validated in outdoor field conditions.**

I. INTRODUCTION

Autonomous road vehicles are becoming increasingly important in recent times and lane identification is an extremely important task in the context of autonomous driving. With the increase in computation power, the past decade witnessed an increase in numerous image processing based lane detection algorithms which are critical to the development of driver assistance systems. Lane detection is the problem of detecting road boundaries and hence determining the safely driveable regions. The proposed lane marker detection algorithm performs background filtering based on thresholding a linear combination of the BGR color space as a preprocessing step followed by the use of Inverse Perspective Transformation for top view generation. The filtered pixels are then clustered and further thresholded to obtain reliable estimates of the possible lanes upon which curves are fit to get the lane. This algorithm was implemented and pruned for solving the problem statement of Intelligent Ground Vehicle Competition 2018. The Intelligent Ground Vehicle Competition is an international collegiate level competition being held annually at the Oakland University by the Association for Unmanned Vehicles Systems International (AUVSI) since 1993. The competition challenges the multidisciplinary teams to build and program an autonomous robot. In AUTO-NAV challenge of IGV competition, an autonomous ground vehicle is required to negotiate the outdoor course, navigating through the obstacles while staying in between the lanes. Leaving the lane causes termination of run in the AUTO-NAV challenge. Therefore, a system that could provide a robust lane detection, potentially, can win the challenge. Some design choices in the algorithm have been specifically altered to improve performance of the robot in this competition. However, the overall idea from this paper would be applicable in any off-road driving scenario to help in the detection of lanes.

1 sanskaragrawal@iitkgp.ac.in
2 indukantdeo@iitkgp.ac.in
3 siddhanthaldar14@iitkgp.ac.in
4 krantikiran.68@iitkgp.ac.in
5 vaibhav.lodhi@iitkgp.ac.in
6 dc@iitkgp.ac.in
* Authors have equal contribution

II. RELATED WORK

A lot of research is being done on lane detection and tracking with respect to autonomous driving in recent times. The approaches adopted in [4] and [6] search for bright horizontal lines on a top view image of the ground planes obtained using inverse perspective transformation. In [9], detection and tracking of the lanes are done in the Hough space. Adaptive thresholding on brightness transitions is used to detect segments followed by the use of a simple temporal averaging for tracking. Some lane detection systems combine the detection with a tracking and estimation step. A commonly used approach for this combination is the use of particle filters which allow fully integrated detection and tracking [11]–[13], [16]. The use of splines for lane fitting as in [15] is also done in practice. The method of the mixed channel [2] is used to distinguish the lane from grass background in an off-road driving scenario. Several lane detection algorithms as in [1], [5] and [14] use the inverse perspective mapping to remove the perspective effect and exploit the homogeneous information content. In [7], road segmentation is performed by extracting superpixels from the original image using the Simple Linear Iterative Clustering (SLIC) algorithm. A Random Sample Consensus algorithm [8] can be used to estimate the lane model parameters. For this paper, a lane estimation approach with multiple lane model identification followed by the appropriate lane curve fitting is proposed and implemented.

The paper is divided into multiple sections. Section III describes the proposed methodology. The experimental results have been described in Section IV and Section V describes the improvements that were made with respect to other approaches. The conclusion to this technical investigative work is given in Section VI.

III. PROPOSED METHODOLOGY

The proposed pipeline acquires the RGB image from the camera in front view. The input image is transformed into the top view using Inverse Perspective Transformation. This is followed by subtraction of the green background by using a linear combination of color planes. Simple Linear Iterative Clustering (SLIC) with appropriate thresholding is applied for outlier reduction. The resulting image is subjected to lane model identification and a RANSAC based lane model estimation algorithm.

Finally, the output image is utilized for the generation of the local waypoint. The pipeline has been shown in Fig.1 and each step has been described in detail in the paper.

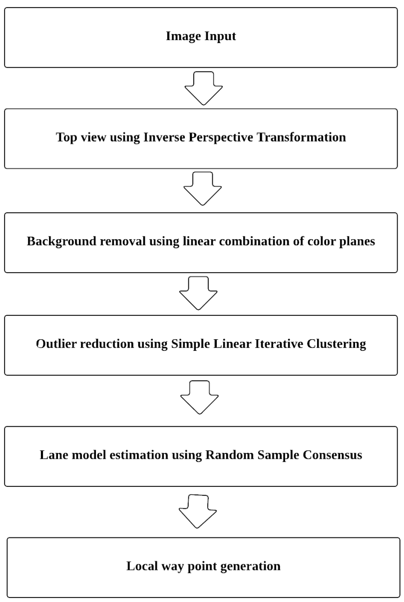

Fig. 1: Proposed Pipeline

A. Top View Generation

Top view generation has a couple of advantages:

- The lanes in the image that appear to converge at the horizon now become vertical and parallel.
- It reduces the region of interest in the input image where we want to detect the lanes, thus reducing the computation time.

Inverse Perspective Transformation (IPT) is used to obtain the top view of an image with respect to the camera center. IPT allows removing the perspective effect from the acquired image, remapping it into a new 2-dimensional domain in which the information content is homogeneously distributed among all pixels. IPT can be of use in structured environments, where, for example, the camera is mounted in a fixed position. To remove the perspective effect, it is necessary to know the specific acquisition conditions (focal length, optical center, pitch angle, yaw angle, and height above ground) and the scene represented in the image (the ground, which is assumed to be flat). Now, a point on the image can be obtained from a point in the world frame using the equation

$$\begin{pmatrix}\tilde{u}\\ \tilde{v}\\ \tilde{w}\end{pmatrix} = \underbrace{\underbrace{\begin{pmatrix}\frac{1}{\rho_u} & 0 & u_0\\ 0 & \frac{1}{\rho_v} & v_0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}f & 0 & 0 & 0\\ 0 & f & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}}_{K}\begin{pmatrix}R & t\\ 0_{1\times 3} & 1\end{pmatrix}^{-1}}_{C}\begin{pmatrix}X\\ Y\\ Z\\ 1\end{pmatrix}$$

Here f denotes the focal length of the camera. It represents the scale of the image, that is, if the focal length is big, then the image is going to be zoomed in and if it is small, then a wide angle image is obtained. The variables u_0 and v_0 represent the position coordinates of the center of the image. The height of the pixels is represented by ρ_u and ρ_v which help convert the image into the pixel scale. The matrices denoted by K are the intrinsic parameters of the camera. The matrix containing R and t is said to be the extrinsic parameters of the camera and they tell the whereabouts of the camera in the world. R gives the orientation of the camera i.e. which way it is pointing and t provides the location of the camera in the world. All these matrices together are called the camera matrix C and it transforms a point $[X, Y, Z]$ in the world to the homogeneous coordinates of a pixel in the image.

B. Background Subtraction and Thresholding

The next processing step in our implementation is background removal which is mainly green coloured regions appearing due to grass. The image acquired from the camera is in BGR color space. This image is split into individual B, G, and R color components. Each color component of the image is referred to as an image plane or channel. Different combinations of image channels were tested. In our experiment, the difference of twice the blue channel and green channel was ultimately used. Reducing the three channels into a single channel image also reduces the computational time of the algorithm. This single channel image is stored in a separate memory for future reference and is called a mixed channel image. To differentiate the white pixels of the lane from the green background, a novel method is proposed and implemented which compares the intensity values of the grayscale, B,G and R channels with a selected threshold value. The selected threshold value is a hyper parameter that is a function of the local illumination and background properties. The grayscale image is generated using the weighted average of three color channels according to the NTSC television standard which uses the formula $Y = 0.299R + 0.587G + 0.114B$ to convert the RGB color space into a single intensity value. A comparison of each pixel intensity value of grayscale channel, B, G, and R channels with respect to a threshold value is performed and if all the channel values are less than the threshold, the accepted pixels of the mixed channel image are converted to the maximum value. This creates a binary image from the mixed channel image.

C. Superpixel Clustering for outlier reduction

The thresholded image still contains some outliers, which might have an adverse effect on curve fitting. The outliers are generally characterized by a unique property, which can be exploited for carrying out outlier reduction. Outliers are sparse and are not grouped. A common approach for outlier reduction can be erosion of the given image. However, it reduces the number of lane pixels. We have devised an efficient way for outlier reduction which does not affect the lanes. Simple Linear Iterative Clustering (SLIC) is initially applied to the image which divides it into very small and fixed number of clusters. In our case, we used 600 clusters. Then the image is thresholded with the following conditions :

- The clusters containing number of pixels less than $(img.rows \times img.cols)/12000$ pixels are rejected.
- For every cluster, we have count_pix which is the number of pixels in that cluster and count_col which is the number of white pixels in that cluster. Clusters with the value of $(count\_col/count\_pix)$ less than 0.7 are rejected. This step is significant as it removes nearly all of the rare outliers.

These threshold values are purely experimental. The clusters passing above thresholds are retained in the image. The outliers which are still left, are taken care by RANSAC curve fitting algorithm described later in the paper.

D. *Lane model identification*

Modeling of lanes for straight and curve roads is usually done using splines and snakes. The problem arises when the lane markers appear nearly horizontal on reaching lane intersections or due to the presence of stop lines. The proposed algorithm uses two different lane models, line to handle horizontal lanes and parabola to handle straight and curved lanes. The algorithm optimally selects between a line and quadratic polynomial for the lane model. All models of lines present in the clustered image are determined using the Hough algorithm. From all the estimated models of line, one with the maximum number of inliers is used for gradient calculation. The computed gradient is compared with a threshold value to determine the lane model. If the estimated gradient of the detected line is within a specific range, which in our case was observed to be lying between 20 degrees and -20 degrees, the lane model is estimated using a line. If not, a parabola based model estimation is selected. The problem in off-road lane detection is that the presence of a single lane in the camera frame leads to loss of information regarding whether the visible lane is the left or the right lane. Subsequently, there is a loss in the orientation of the vehicle which may cause the vehicle to move in the backward direction. While transitioning from the heading direction to backward direction, the lane appears as a horizontal line in one of the acquired frames. In such a scenario, Hough line based lane estimation helps in preventing the backward motion of the vehicle.

E. *Lane Model Estimation*

1) RANSAC Curve Fitting: For RANSAC curve fitting, the origin is assumed to be at the bottom left corner of the image with y representing the rows and x representing the columns in the image. Two parabolas are simultaneously estimated as a curve in the RANSAC algorithm. The model of parabolas used are:

$$\mathrm{y}^2 = \lambda_1.(x-a)$$

and

$$\mathrm{y}^2 = \lambda_2.(x-a-w)$$

The choice of horizontal parabola and parameters w, λ_1 and λ_2 is made in such a way that the shift between the vertex of the two parabolas is related to the actual width of the lane w. λ_1 and λ_2 represent the curvature of the two polynomials. This selection helps in eliminating the wrong estimation of curves when outliers are large in number. Curves with w less than a certain threshold and λ_1 and λ_2 having small values with opposite signs are eliminated. The RANSAC confidence simultaneously takes into account the model of both the lanes. So even in case of occlusion, the confidence score is high. This makes the model quite robust to occlusion. In case of single lane, the other lane is predicted by contemplating whether the detected lane is the left or the right lane. The other lane is then assumed to be of the same curvature and the width is considered to be the same as that obtained in the last frame comprising two lanes. This model has one limitation. It fails to fit in some cases where the lanes are nearly horizontal. This case arises, at the intersection of lanes. We use Hough Line estimation for dealing with the horizontal lanes.

2) Hough Line Detection: In the case of nearly horizontal lanes, Hough Line detection [3] is used to obtain a reliable estimate of the single lane visible in the camera frame. Hough Line transform determines the best fit line by estimating multiple lines and then considering the number of inliers lying on each of them. The line with the maximum number of inliers is considered to be the best fit line. Experimentally, it has been observed that Hough Line detection provides a reliable estimate of the lane parameters when it comes to nearly horizontal lanes.

F. *Local Waypoint Determination*

After determining the equations of lanes, we pass the local waypoint, between the lanes, to the planner. Waypoint is calculated at a distance of 3 meters ahead and the heading direction is calculated as the average of gradients over all the points on the lane. This point, along with the orientation, is then passed into the planner.

IV. EXPERIMENT AND RESULTS

This section presents the performance of the proposed pipeline on a simulated off-road arena with and without obstacles. The algorithm manages to detect and extract the left and right lanes markers. The proposed algorithm has been tested on video images from a video obtained from an onboard camera placed on the vehicle. $Figure 2$ shows a series of images, each row demonstrating a different scenario. From left to right, each row depicts the image acquired from the camera, the top view image, the preprocessed image,

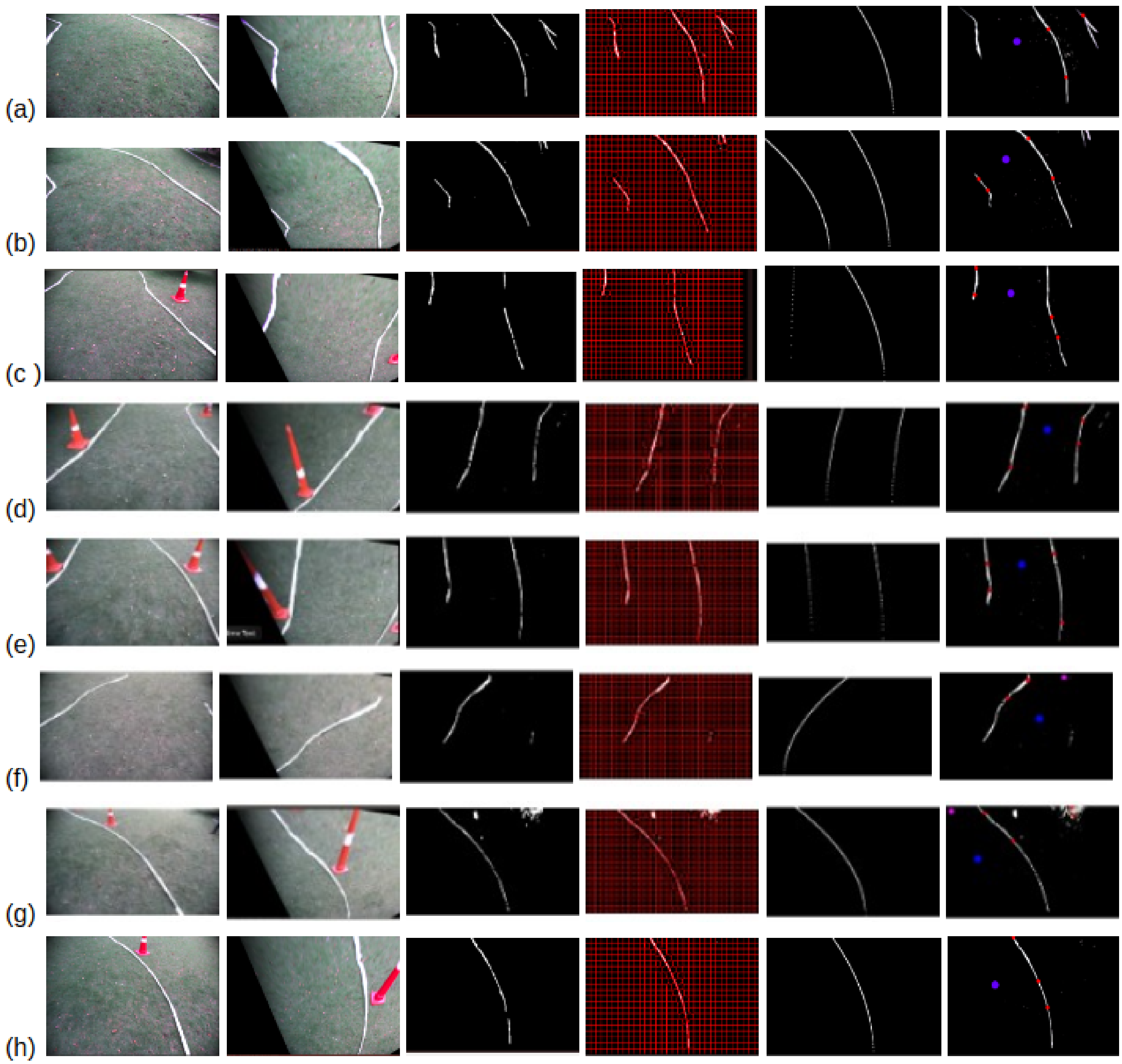

Fig. 2: From left to right (a)Acquired image from the camera (b) Top view image (c) Preprocessed image (d) Clustered image (e) Estimated lane model (f) Estimated heading point. The results clearly demonstrate that objects appearing close to the lanes do not have any adverse effect on the performance of the algorithm.

the result for SLIC, the estimated lane model followed by the estimated heading point. $2(a)$ and $2(b)$ depict successful estimation of the lanes markers when both the lanes are visible in the camera frame. $2(c)$ to $2(e)$ show examples containing obstacles having white markers painted on them. These examples demonstrate the robustness of the algorithm when it comes to detecting lanes in the presence of obstacles. $2(f)$ illustrates how the pipeline succeeds in lane estimation even when only one lane is visible in the camera frame. $2(g)$ and $2(h)$ exhibit the robustness of the proposed algorithm by successfully estimating single lanes even in the presence of obstacles. Therefore, it is demonstrated experimentally that this algorithm is robust to outliers and enables successful lane estimation.

V. IMPROVEMENTS OVER OTHER APPROACHES

Many approaches view lane fitting as fitting a single lane twice after removing the inliers for the first lane. The number of outliers, in this case, are usually large and the probability of poor fitting is increased. The problem in this model arises, often, in the cases where only a single lane is visible and middle portion of that lane is occluded. This model may try to fit two different lanes for the upper and lower part, but since our model optimizes over both the lanes simultaneously, it knows that the other lane must be some distance apart. Hence, our model works even in such non-favourable conditions.

Stability of RANSAC depends mostly on the ratio of inliers and outliers. While fitting the first curve in a single lane model, the other lane acts as an outlier. Also, the actual outlier points are counted twice. In our model, both the curves are inliers, and hence the ratio of inliers to the outliers is more. Hence RANSAC is more stable in our case.

VI. CONCLUSION

This paper describes a customized lane detection algorithm to be used with a mounted monocular camera for off-road lane detection scenarios. For validation and testing of the algorithm, Blackfly 2.3 MP Mono GigE PoE camera was mounted on the three wheeled differential drive robot. The proposed lane detection algorithm is a RANSAC-based curve fitting algorithm involving the use of Simple Linear Iterative Clustering (SLIC) algorithm. gSLIC [10], a GPU based real-time implementation of SLIC was used. The proposed algorithm is robust to effects of either sunlight or shade. The use of SLIC helps in outlier reduction, thus, enabling robust lane detection. The lane detector was tested on Intel i5 7th generation processor with Nvidia GTX-1050 GPU. The camera‘s frame rate was set at 15 fps to enable the proper actuation of motors. The algorithm works at a frequency of 15 Hz on the above-mentioned system. Our entry using the proposed algorithm received the 2nd prize in the Autonomous Navigation challenge at the Intelligent Ground Vehicle Competition 2018, thus validating satisfactory performance of the proposed method in external controlled environments. The output in these cases were found to be satisfactory, although there are several modifications for future work so that the robustness may be increased.

ACKNOWLEDGEMENT

We would like to thank team Autonomous Ground Vehicle, IIT Kharagpur and Sponsored Research & Industrial Consultancy, IIT Kharagpur for providing us with the resources and funding for this project. We would also like to acknowledge the efforts put in by Apoorve Singhal, Arvind Jha, Siddhant Agarwal, Shrey Shrivastava, Shreyas Kowshik and Deepank Agrawal during this task.

REFERENCES

[1] Mohamed Aly. Real time detection of lane markers in urban streets. *2008 IEEE Intelligent Vehicles Symposium*, pages 7–12, 2008.

[2] Andrew Reed Bacha. *Line detection and lane following for an autonomous mobile robot*. PhD thesis, Virginia Tech, 2005.

[3] Dana H Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern recognition*, 13(2):111–122, 1981.

[4] M. Bertozzi and A. Broggi. Real-time lane and obstacle detection on the gold system. In *Proceedings of Conference on Intelligent Vehicles*, pages 213–218, Sept 1996.

[5] M. Bertozzi and A. Broggi. Gold: a parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Transactions on Image Processing*, 7(1):62–81, Jan 1998.

[6] M. Bertozzi, A. Broggi, G. Conte, and A. Fascioli. Obstacle and lane detection on argo. In *Proceedings of Conference on Intelligent Transportation Systems*, pages 1010–1015, Nov 1997.

[7] Patricio Córdova and Yuqing Du. Autonomous driving: Road-estimation.

[8] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[9] John B. McDonald and John B. Mc Donald. Application of the hough transform to lane detection and following on high speed roads. In *in Motorway Driving Scenarios, in Proceeding of Irish Signals and Systems Conference*, 2001.

[10] Carl Yuheng Ren and Ian Reid. gslic: a real-time implementation of slic superpixel segmentation. *University of Oxford, Department of Engineering Science*, 6, 2011.

[11] Stephan Sehestedt, Sarath Kodagoda, Alen Alempijevic, and Gamini Dissanayake. Efficient lane detection and tracking in urban environments. In *EMCR*, 2007.

[12] Yuan Shu and Zheng Tan. Vision based lane detection in autonomous vehicle. In *Fifth World Congress on Intelligent Control and Automation (IEEE Cat. No.04EX788)*, volume 6, pages 5258–5260 Vol.6, June 2004.

[13] S. Vacek, S. Bergmann, U. Mohr, and R. Dillmann. Rule-based tracking of multiple lanes using particle filters. In *2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 203–208, Sept 2006.

[14] Jun Wang, Tao Mei, Bin Kong, and Hu Wei. An approach of lane detection based on inverse perspective mapping. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pages 35–38, Oct 2014.

[15] Yue Wang, Eam Khwang Teoh, and Dinggang Shen. Lane detection and tracking using b-snake. 22:269–280, 04 2004.

[16] W. Zhu, F. Liu, Z. Li, X. Wang, and S. Zhang. A vision based lane detection and tracking algorithm in automatic drive. In *2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, volume 1, pages 799–803, Dec 2008.