

Multiple Lane Detection Algorithm Based on Optimised Dense Disparity Map Estimation

Han Ma, Yixin Ma, Jianhao Jiao, M Usman Maqbool Bhutta,
Mohammud Junaid Bocus, Lujia Wang, Ming Liu, Rui Fan*

Abstract—Lane detection is very important for self-driving vehicles. In recent years, computer stereo vision has been prevalently used to enhance the accuracy of the lane detection systems. This paper mainly presents a multiple lane detection algorithm developed based on optimised dense disparity map estimation, where the disparity information obtained at time t_n is utilised to optimise the process of disparity estimation at time t_{n+1} ($n \geq 0$). This is achieved by estimating the road model at time t_n and then controlling the search range for the disparity estimation at time t_{n+1} . The lanes are then detected using our previously published algorithm, where the vanishing point information is used to model the lanes. The experimental results illustrate that the runtime of the disparity estimation is reduced by around 37% and the accuracy of the lane detection is about 99%.

Index Terms—lane detection, self-driving vehicles, stereo vision, disparity estimation, vanishing point.

I. INTRODUCTION

THE deployment of autonomous vehicles has been increasing rapidly since Google first launched their self-driving car project in 2009 [1]. In recent years, with a number of technology breakthroughs being witnessed in the world where science fiction inventions are now becoming a reality, the competition to commercialise driver-less vehicles by companies like GM, Waymo and Daimler-Bosch is fiercer than ever [2]. For instance, Volvo conducted a series of self-driving experiments involving around 100 cars in China [3]. The 5G network is also utilised in self-driving vehicles to help them communicate with each other [2]. Furthermore, the

H. Ma is with the Department of Precision Instrument, School of Mechanical Engineering, Tsinghua University, Beijing, China. (e-mail: mh15@mails.tsinghua.edu.cn)

Y. Ma is with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China. (e-mail: y.ma@sjtu.edu.cn)

J. Jiao is with the Robotics and Multi-Perception Laboratory, Robotics Institute, The Hong Kong University of Science and Technology, Hong Kong SAR, China. (e-mail: jjiao@connect.ust.hk)

M. U. M. Bhutta is with the Robotics and Multi-Perception Laboratory, Robotics Institute, The Hong Kong University of Science and Technology, Hong Kong SAR, China. (e-mail: mumbhutta@connect.ust.hk)

M. J. Bocus is with the Visual Information Group, the University of Bristol, BS8 1UB, UK. (e-mail: junaid.bocus@bristol.ac.uk)

L. Wang is with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. (e-mail: lj.wang1@siat.ac.cn)

M. Liu is with the Robotics and Multi-Perception Laboratory, Robotics Institute, The Hong Kong University of Science and Technology, Hong Kong SAR, China. (e-mail: eelium@ust.hk)

R. Fan is with the Robotics and Multi-Perception Laboratory, Robotics Institute, The Hong Kong University of Science and Technology, Hong Kong SAR, China. (e-mail: rui.fan@ieee.org)

*Corresponding author: Rui Fan.

computer stereo vision technique has also been prevalently used in prototype vehicle road tests to provide the depth information for various automotive applications, e.g., road condition assessment [4], [5], lane and obstacle detection [3], [6], and vehicle state estimation [7], [8].

The state-of-the-art lane detection algorithms can mainly be classified as feature-based and model-based [6]. The former extracts local and useful information of an image, such as edges, texture and colour, for the segmentation of lanes and road boundaries [9]. On the other hand, the model-based algorithms aim to represent the lanes with mathematical equations, e.g., linear, parabolic, linear-parabolic and spline, based upon some common road geometry assumptions [10]. The linear model works well for lanes with a low curvature [3], but a more flexible road model is inevitable when lanes with a higher curvature are considered. Therefore, some algorithms [11]–[14] use a parabolic model to represent the lanes with a constant curvature. To address some more complex cases, Jung et al. proposed a combination of linear and parabolic lane models, where the far lanes are modelled as parabolas whereas the nearby lanes are represented as linear models [15]. In addition to the models mentioned above, the spline model is an alternative method whereby the lane pixels are interpolated into an arbitrary shape [10], [16]. However, the more parameters introduced into a flexible model, the higher will be the computational complexity of the algorithm. Therefore, some authors resort to additional important properties of 3-D imaging techniques, such as disparity estimation [6], instead of being limited to only 2-D information.

One of the most prevalently used 3-D based methods is Inverse Perspective Mapping (IPM). With the assumption that two lanes are parallel to each other in the World Coordinate System (WCS), IPM can map a 3-D scenery into a 2-D bird's eye view [17]. In addition, some authors [18]–[21] proposed to use the vanishing point $p_{vp} = [u_{vp}, v_{vp}]^\top$ to model lane markings and road boundaries, where u_{vp} and v_{vp} denote the vertical and horizontal coordinates of the vanishing point, respectively. An example of vanishing point is illustrated in Fig. 1, where l_1 and l_2 are two parallel straight lines. Two 3-D points p_1^w and p_2^w in the WCS are projected on the image plane π and their corresponding points in the Image Coordinate System (ICS) are p_1 and p_2 , respectively. Therefore, the projections of l_1 and l_2 in the image are two straight lines and they intersect at the vanishing point p_{vp} . However, their algorithms work well only if the road surface is assumed to be flat or the camera

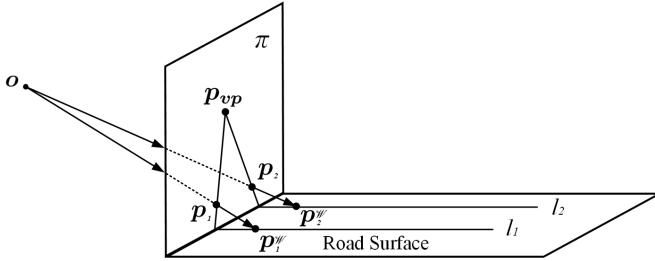


Fig. 1. Vanishing point.

parameters are known. Hence, some researchers pay closer attention to the disparity information which can be obtained using either active sensors, e.g., radar and laser, or passive sensors, e.g., stereo cameras [21]. Since Labayrade et al. proposed the concept of “v-disparity” in 2002 [22], disparity information has been widely used to enhance the robustness of the lane detection systems. The work presented in [21] shows a particular instance where the disparity information is successfully combined with a lane detection algorithm to estimate p_{vp} for both flat and non-flat road surfaces. At the same time, the obstacles containing a lot of redundant information can be eliminated by comparing the actual and fitted disparity values. However, the estimation of p_{vp} is greatly affected by the outliers when performing the Least Squares Fitting (LSF), and the lanes are sometimes wrongly detected because the selection of plus-minus peaks is not always effective. Moreover, achieving real-time performance is still challenging in [21] because of the intensive computational complexity of the algorithm. Hence in this work, we propose a more efficient lane detection algorithm where the disparity information acquired at time t_{n-1} is utilised to improve the process of the disparity estimation at time t_n . The flow chart of the proposed lane detection algorithm is depicted in Fig. 2.

Firstly, we create a vector $\alpha = [\alpha_0, \alpha_1, \alpha_2]^\top$ to store the parameters of the road model. The keypoints $P = [p_0, p_1, \dots, p_n]^\top$ and $Q = [q_0, q_1, \dots, q_n]^\top$ are then extracted and matched using Binary Robust Invariant Scalable Keypoints (BRISK), where $p_n = [u_{pn}, v_{pn}]^\top$ and $q_n = [u_{qn}, v_{qn}]^\top$ denote the matched keypoints in the left and right images, respectively. P and Q are then utilised to create a sparse v-disparity histogram from which the road model $f(v) = \alpha_0 + \alpha_1 v + \alpha_2 v^2$ can be estimated. The search range for the disparity estimation at time t_0 is subsequently controlled according to $f(v)$, where the search range at row v is limited to $[f(v) - \tau, f(v) + \tau]$. The lanes are then detected using our previously published method in [21]. As for the disparity estimation at time t_{n+1} ($n \geq 0$), α is estimated by creating a dense v-disparity map using the disparity information acquired at time t_n ($n \geq 0$) and then fitting a parabola to the best path extracted from the dense v-disparity map. This not only improves the accuracy of the disparity estimation at time t_{n+1} but also boosts its processing speed.

The remainder of the paper is structured as follows: Section II presents the proposed lane detection algorithm. In Section

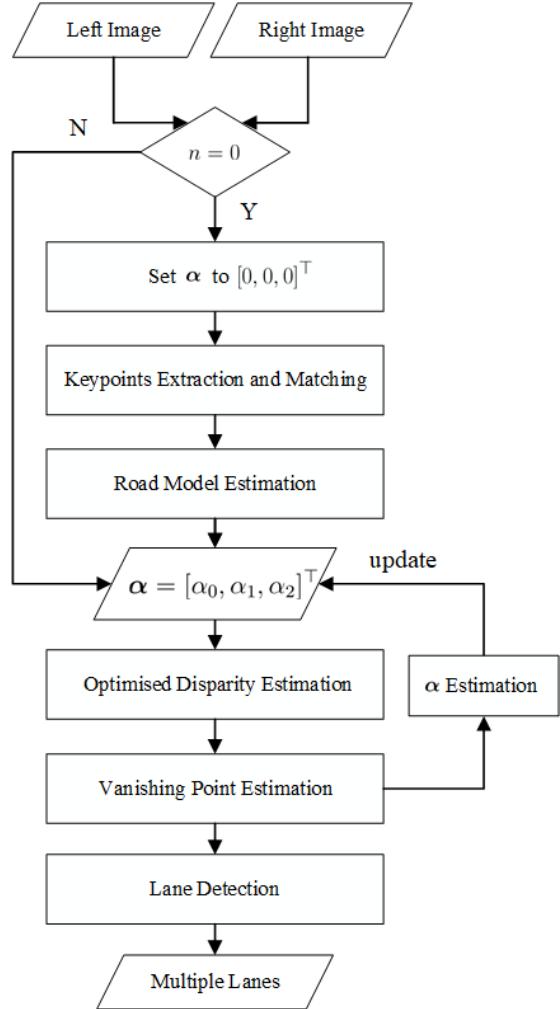


Fig. 2. Flow chart of the proposed algorithm.

III, the experimental results are illustrated and the evaluation of the proposed system is carried out. Finally, Section IV summarises the paper and provides some recommendations for future work.

II. ALGORITHM DESCRIPTION

A. Disparity Estimation

The proposed disparity estimation approach in this paper is developed based on our previously published algorithm [2], where the search range at the position of (u, v) is propagated from three estimated disparities at the positions of $(u - 1, v)$, (u, v) and $(u + 1, v)$. In this subsection, we optimise the process of the disparity estimation by utilising the disparity information acquired at time t_{n-1} to control the search range at time t_n , where n is a positive integer.

As for the stereo matching at time t_0 , the parameter vector $\alpha = [\alpha_0, \alpha_1, \alpha_2]^\top$ is first initialised as $[0, 0, 0]^\top$. Then, α is estimated by solving a least squares problem with a set of reliable correspondences $P = [p_0, p_1, \dots, p_n]^\top$ and $Q = [q_0, q_1, \dots, q_n]^\top$. In this paper, we use BRISK to detect and match P and Q . It allows a faster execution to

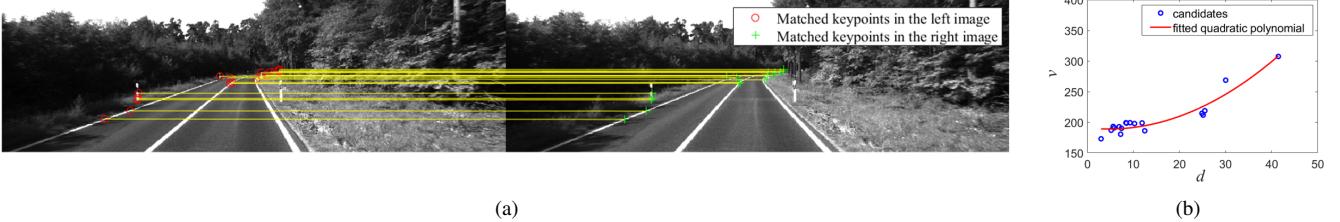


Fig. 3. Road model estimation for the initial stereo image pair. (a) BRISK-based on-road keypoints extraction and matching between the left and right images. (b) the corresponding sparse v-disparity histogram of Fig. 3(a).

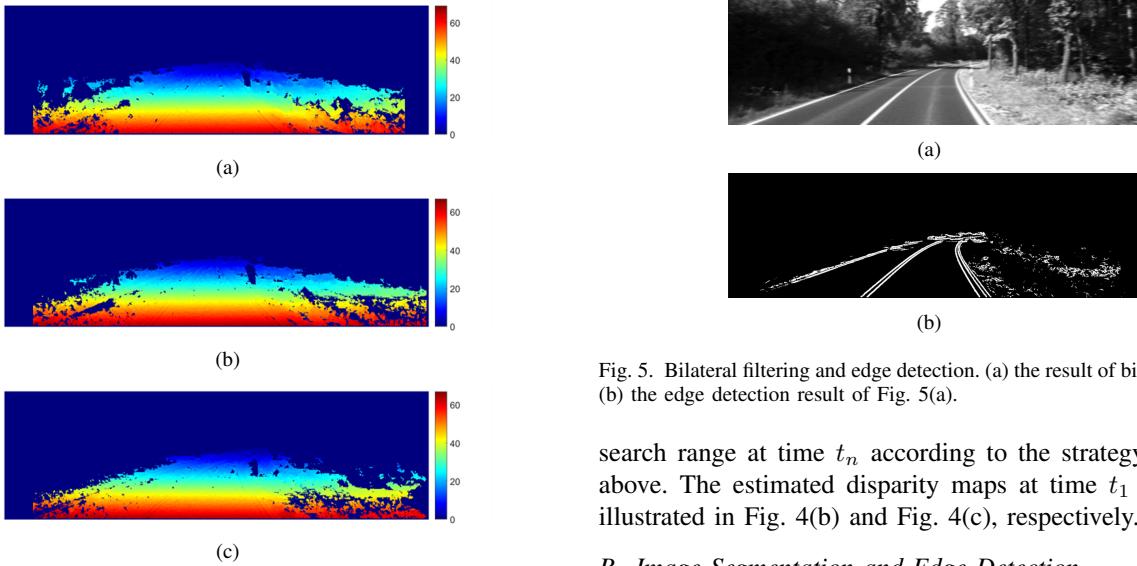


Fig. 4. Disparity map estimation at time t_n . (a) $n = 0$. (b) $n = 1$. (c) $n = 38$.

achieve approximately the same number of correspondences as SURF (Speeded-Up Robust Features) and SIFT (Scale-Invariant Feature Transform) [23]. An example of keypoints detection and matching is shown in Fig. 3(a). The correspondences are then utilised to create a sparse v-disparity histogram as shown in Fig. 3(b). The values of α_0 , α_1 and α_2 can then be obtained by fitting a parabola to the non-zero candidates in the sparse v-disparity map. Since the outliers can severely affect the accuracy of the LSF, Random Sample Consensus (RANSAC) is utilised to iteratively update α until the ratio $r = n_{\mathcal{I}}/(n_{\mathcal{O}} + n_{\mathcal{I}})$ reaches a maximum value, where $n_{\mathcal{I}}$ and $n_{\mathcal{O}}$ represent the number of inliers and outliers, respectively. Then, the search range at time t_n according to the strategy mentioned above. The estimated disparity maps at time t_1 and t_{38} are illustrated in Fig. 4(b) and Fig. 4(c), respectively.

B. Image Segmentation and Edge Detection

1) *Bilateral Filtering*: The input image is always noisy, making various edge detectors such as Sobel and Canny very sensitive to the blobs [6]. Therefore, we first perform bilateral filtering on the input image to reduce redundant information while still preserving the edges. The filtered image is shown in Fig. 5(a).

2) *Road Model Estimation*: As discussed in subsection II-C, the boundary between the road and sky $v = v_{vp}|_{min}$ can be obtained by solving a quadratic function $f(v) = 0$. Furthermore, the road area can be extracted by computing the difference between the actual and fitted disparity values, i.e., \tilde{d} and $f(v)$, and then finding the pixels that meet the requirement of $d(u, v) \in [\max\{0, f(v) - \mu\}, f(v) + \mu]$, where μ is a threshold set to remove the pixels on the obstacles [6].

3) *Edge Detection*: In this subsection, we use Sobel operator to detect the edge information by convolving the horizontal and vertical Sobel operators with the original image [3]. The corresponding edge detection result is shown in Fig. 5(b).

C. Lane Detection

DP is an optimisation method which aims at solving a complicated problem by breaking it down into a series of simpler sub-problems [6]. According to our previously published algorithm in [21], DP is performed to minimise the energy function in Eq. 1, where E represents the total energy and the path with the minimum E is selected as the optimum solution.

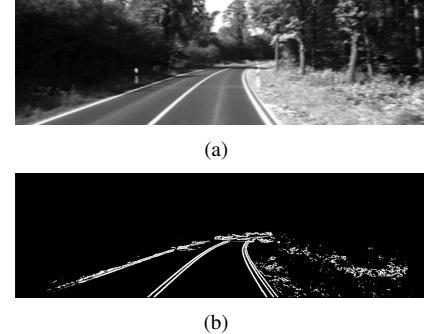


Fig. 5. Bilateral filtering and edge detection. (a) the result of bilateral filtering. (b) the edge detection result of Fig. 5(a).

search range at time t_n according to the strategy mentioned above. The estimated disparity maps at time t_1 and t_{38} are illustrated in Fig. 4(b) and Fig. 4(c), respectively.

Afterwards, the disparity information acquired at time t_{n-1} ($n > 0$) is utilised to create a dense v-disparity map. The latter is then optimised using Dynamic Programming (DP) and the optimum solution corresponds to the vertical road profile. More details are provided in subsection II-C. The optimum solution is then interpolated into a quadratic polynomial and the values of α_0 , α_1 and α_2 at time t_{n-1} can be obtained. The parameter vector α at time t_{n-1} is then used to control the

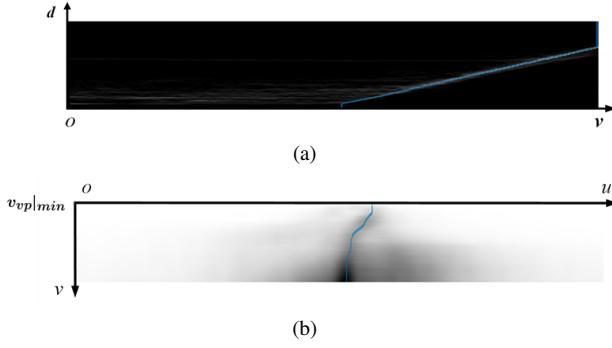


Fig. 6. Dense p_{vp} estimation. (a) dense v-disparity map. The path in blue represents $f(v)$. (b) dense u_{vp} map. The path in blue represents $g(v)$.

E_{data} depends on the total votes each accumulator gets and E_{smooth} with a smoothness constant λ , penalises the changes in p_{vp} .

$$E = E_{data} + \lambda E_{smooth} \quad (1)$$

In this section, DP is first utilised to optimise the v-disparity map by minimising the energy function in Eq. 2. The blue path in the v-disparity map is the optimum solution, as shown in Fig. 6(a). The optimum solution is then interpolated into a quadratic polynomial: $f(v) = \alpha_0 + \alpha_1 v + \alpha_2 v^2$, and v_{vp} with respect to each row can thus be computed.

$$\begin{aligned} E(v)_d &= -m_v(d, v) \\ &+ \min_{\tau_v} [E(v - \tau_v)_{d+1} - \lambda_v \tau_v], \text{ s.t. } \tau_v \in [0, 6] \end{aligned} \quad (2)$$

Then, DP is used to optimise the dense u_{vp} accumulator by minimising the energy function in Eq. 3. The blue path in the dense u_{vp} accumulator map is the optimum solution, as shown in Fig 6(b). u_{vp} with respect to each row can then be estimated by interpolating the optimum solution into a quadruplicate polynomial: $g(v) = \beta_0 + \beta_1 v + \beta_2 v^2 + \beta_3 v^3 + \beta_4 v^4$.

$$\begin{aligned} E(u)_v &= m_u(u, v) \\ &+ \min_{\tau_u} [E(u_{vp} + \tau_u)_{v+1} + \lambda_u \tau_u], \text{ s.t. } \tau_u \in [-5, 5] \end{aligned} \quad (3)$$

p_{vp} provides the tangential direction and the curvature information of lanes, which can help to validate the lane positions. In our previous paper [21], a likelihood function $V(p_e) = \nabla(p_e) \cdot \cos(\theta_{p_e} - \theta_{p_{vp}})$ is formed for each edge point p_e and the plus-minus peak pairs are selected for lane visualisation, where θ_{p_e} is the angle between the u -axis and the orientation of the edge point p_e , and $\theta_{p_{vp}}$ is the angle between the u -axis and the radial from an edge pixel p_e to $p_{vp}(v_e)$. More details are provided in [21], [24]. The lanes can thus be visualised using $f(v)$, $g(v)$ and $V(p_e)$. Some examples of lane detection results are shown in Fig. 7.

III. EXPERIMENTAL RESULTS

In this section, we use the KITTI dataset to quantify the robustness of the proposed algorithm. Some experimental results are shown in Fig. 7, where the green regions are the estimated road areas and the lines in red are the detected lanes.

TABLE I
DETECTION RESULTS OF THE PROPOSED ALGORITHM.

Sequence	Lanes	Incorrect detection	Misdetection
1	860	0	0
2	594	0	0
3	376	0	0
4	156	0	3
5	678	0	9
6	1060	10	2
Total	3724	10	14

TABLE II
DETECTION RESULTS OF [21].

Sequence	Lanes	Incorrect detection	Misdetection
1	860	0	0
2	594	0	0
3	376	0	0
4	156	0	9
5	678	0	17
6	1060	14	7
Total	3724	14	33

Currently, it is impossible to access a satisfying ground truth dataset for the evaluation of lane detection algorithms because accepted test protocols do not usually exist [25]. Therefore, many publications related to lane detection only compare the quality of their experimental results with the ones obtained using some other published algorithms [21]. For this reason, we compare the performance of the proposed system with [21]. Table I and Table II details the successful detection rates of the proposed algorithm and the algorithm presented in [21], respectively. The overall successful detection rate is around 99% and it is increased by around 0.6%.

In addition, the proposed algorithm is implemented in C programming language. The runtime of the algorithm for a single frame with a resolution of 1242×375 is 0.21 seconds (using a single thread of Intel Core i7-8700K CPU). By controlling the search range at time t_n using the disparity information acquired at time t_{n-1} , the runtime of the disparity estimation is reduced by around 37%. Although the algorithm does not perform in real time, we believe that the execution speed of the implementation can be further accelerated by highly exploiting the parallel computing architecture.

IV. CONCLUSION AND FUTURE WORK

This paper presented a multiple lane detection algorithm. The process of the disparity estimation is optimised by using the disparity information acquired at time t_n to suggest the search range at time t_{n+1} ($n \geq 0$). This not only improved the accuracy of the estimated disparities but also reduced the runtime of the algorithm. The runtime of the disparity estimation is reduced by approximately 37%. The lanes were detected using our previously published algorithm in [6], [21]. The experimental results illustrated that the proposed system works robustly and precisely for both highway and urban

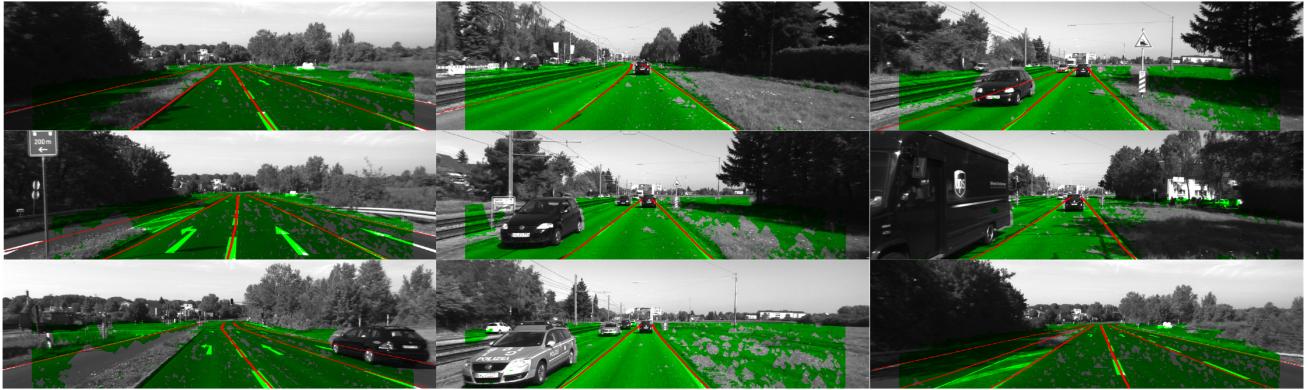


Fig. 7. Experimental results. The regions in green are estimated road surface areas. The lines in red are detected lanes.

scenes and a 99% successful detection rate was achieved when processing the KITTI dataset.

However, some actual road conditions may result in failed detections. Therefore, we plan to train a deep neural network for dense vanishing point estimation. Furthermore, we plan to implement the proposed algorithm on some state-of-the-art embedded systems, such as Jetson TX2 for real-time purposes.

V. ACKNOWLEDGEMENTS

This work is supported by the Research Grant Council of Hong Kong SAR Government, China, under Project No. 11210017, No. 16212815 and No. 21202816, NSFC U1713211 and Shenzhen Science, Technology and Innovation Commission (SZSTI) JCYJ20160428154842603, awarded to Prof. Ming Liu; Shenzhen Science, Technology and Innovation Commission (SZSTI) JCYJ20170818153518789 and National Natural Science Foundation of China No. 61603376, awarded to Dr. Lujia Wang.

REFERENCES

- [1] James A Brink, Ronald L Arenson, Thomas M Grist, Jonathan S Lewin, and Dieter Enzmann, “Bits and bytes: the future of radiology lies in informatics and information technology,” *European Radiology*, pp. 1–5, 2017.
- [2] Rui Fan and Naim Dahnoun, “Real-time implementation of stereo vision based on optimised normalised cross-correlation and propagated search range on a gpu,” in *Imaging Systems and Techniques (IST), 2017 IEEE International Conference on*. IEEE, 2017, pp. 241–246.
- [3] Rui Fan, Victor Prokhorov, and Naim Dahnoun, “Faster-than-real-time linear lane detection implementation using soc dsp tms320c6678,” in *Imaging Systems and Techniques (IST), 2016 IEEE International Conference on*. IEEE, 2016, pp. 306–311.
- [4] R. Fan, X. Ai, and N. Dahnoun, “Road surface 3D reconstruction based on dense subpixel disparity map estimation,” *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3025–3035, June 2018.
- [5] Rui Fan, Yanan Liu, Xingrui Yang, Mohammad Junaid Bocus, Naim Dahnoun, and Scott Tancock, “Real-time stereo vision for road surface 3-d reconstruction,” *arXiv preprint arXiv:1807.07433*.
- [6] Rui Fan and Naim Dahnoun, “Real-time stereo vision-based lane detection system,” *Measurement Science and Technology*, vol. 29, no. 7, 2018.
- [7] Meghan Evans, Rui Fan, and Naim Dahnoun, “Iterative roll angle estimation from dense disparity map,” in *2018 7th Mediterranean Conference on Embedded Computing (MECO)*. IEEE, 2018.
- [8] Rui Fan, Mohammad Junaid Bocus, and Naim Dahnoun, “A novel disparity transformation algorithm for road segmentation,” *Information Processing Letters*, vol. 140, pp. 18–24, 2018.
- [9] Massimo Bertozzi and Alberto Broggi, “Gold: A parallel real-time stereo vision system for generic obstacle and lane detection,” *IEEE transactions on image processing*, vol. 7, no. 1, pp. 62–81, 1998.
- [10] Yue Wang, Eam Khwang Teoh, and Dinggang Shen, “Lane detection and tracking using b-snake,” *Image and Vision computing*, vol. 22, no. 4, pp. 269–280, 2004.
- [11] Karl Kluge and Sridhar Lakshmanan, “A deformable-template approach to lane detection,” in *Intelligent Vehicles' 95 Symposium., Proceedings of the*. IEEE, 1995, pp. 54–59.
- [12] Yan Wang, Li Bai, and Michael Fairhurst, “Robust road modeling and tracking using condensation,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 4, pp. 570–579, 2008.
- [13] Yong Zhou, Rong Xu, Xiaofeng Hu, and Qingtai Ye, “A robust lane detection and tracking method based on computer vision,” *Measurement science and technology*, vol. 17, no. 4, pp. 736, 2006.
- [14] Chris Kreucher and Sridhar Lakshmanan, “Lana: a lane extraction algorithm that uses frequency domain features,” *IEEE Transactions on Robotics and automation*, vol. 15, no. 2, pp. 343–350, 1999.
- [15] Cláudio Rosito Jung and Christian Roberto Kelber, “An improved linear-parabolic model for lane following and curve detection,” in *Computer Graphics and Image Processing, 2005. SIBGRAPI 2005. 18th Brazilian Symposium on*. IEEE, 2005, pp. 131–138.
- [16] Yue Wang, Dinggang Shen, and Eam Khwang Teoh, “Lane detection using spline model,” *Pattern Recognition Letters*, vol. 21, no. 8, pp. 677–689, 2000.
- [17] Marcos Nieto, Luis Salgado, Fernando Jaureguizar, and Julian Cabrera, “Stabilization of inverse perspective mapping images based on robust vanishing point estimation,” in *Intelligent Vehicles Symposium, 2007 IEEE*. IEEE, 2007, pp. 315–320.
- [18] David Schreiber, Bram Alefs, and Markus Clabian, “Single camera lane detection and tracking,” in *Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE*. IEEE, 2005, pp. 302–307.
- [19] David Hanwell and Majid Mirmehdi, “Detection of lane departure on high-speed roads,” in *ICPRAM (2)*, 2012, pp. 529–536.
- [20] Basel Fardi and Gerd Wanielik, “Hough transformation based approach for road border detection in infrared images,” in *Intelligent Vehicles Symposium, 2004 IEEE*. IEEE, 2004, pp. 549–554.
- [21] Umar Ozgunalp, Rui Fan, Xiao Ai, and Naim Dahnoun, “Multiple lane detection algorithm based on novel dense vanishing point estimation,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 621–632, 2017.
- [22] Raphael Labayrade, Didier Aubert, and J-P Tarel, “Real time obstacle detection in stereovision on non flat road geometry through “v-disparity” representation,” in *Intelligent Vehicle Symposium, 2002. IEEE*. IEEE, 2002, vol. 2, pp. 646–651.
- [23] Stefan Leutenegger, Margarita Chli, and Roland Y Siegwart, “Brisk: Binary robust invariant scalable keypoints,” in *2011 International conference on computer vision*. IEEE, 2011, pp. 2548–2555.
- [24] Rui Fan, Yanan Liu, Mohammad Junaid Bocus, Lujia Wang, and Ming Liu, “Real-time subpixel fast bilateral stereo,” *arXiv preprint arXiv:1807.02044*.
- [25] Aharon Bar Hillel, Ronen Lerner, Dan Levi, and Guy Raz, “Recent progress in road and lane detection: a survey,” *Machine vision and applications*, vol. 25, no. 3, pp. 727–745, 2014.