

A Robust Roll Angle Estimation Algorithm Based on Gradient Descent

Rui Fan¹, *Member, IEEE*, Lujia Wang², *Member, IEEE*,
Ming Liu¹, *Senior Member, IEEE*, Ioannis Pitas³, *Fellow, IEEE*,
¹*Robotics and Multi-Perception Laboratory, Robotics Institute,*
the Hong Kong University of Science and Technology, Hong Kong.

²*Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China.*

³*Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece.*
{eeruifan, eelium}@ust.hk, lj.wang1@siat.ac.cn, pitas@aiaa.csd.auth.gr

Abstract—This paper introduces a robust roll angle estimation algorithm, which is developed from our previously published work, where the roll angle was estimated from a dense subpixel disparity map by minimizing a global energy using golden section search algorithm. In this paper, to achieve greater computational efficiency, we utilize gradient descent to optimize the aforementioned global energy. The experimental results illustrate that the presented roll angle estimation method takes fewer iterations to achieve the same precision as the previous method.

I. INTRODUCTION

Binocular vision is a commonly used technology for 3D information extrapolation [1]–[3]. This technology extrapolates the depth of the real-world scenery by comparing the difference between a stereo image pair [4], [5]. This relative positional difference is generally known as disparity [2], [6]. The disparity information enables autonomous systems to segment road regions more effectively [7]. For example, Labayrade et al. [8] introduced an accurate obstacle detection method based on the analysis of a so-called “v-disparity” image, which was created by computing the disparity histogram of each horizontal row. Since then, many algorithms based on v-disparity image analysis have been proposed for road region segmentation [9]–[11]. These algorithms generally hypothesize that the road surface is planar and the baseline of the stereo rig is perfectly parallel to the horizontal road surface [12], [13]. These assumptions make it feasible to represent the road disparity projections using a linear model [14]–[16]. The road regions can, therefore, be extracted by comparing the difference between the actual disparity values and the estimated linear road disparity projection model [17].

However, in practice, the baseline of the stereo rig is not always perfectly parallel to the horizontal road surface. This fact can introduce a non-zero roll angle θ , shown in Fig. 1, into the imaging process, where T_c and h represent the baseline and the mounting height of the stereo rig, respectively; o_l^c and o_r^c are the origins of the left and right camera coordinate systems, respectively. The stereo rig baseline can become perfectly parallel to the horizontal road surface by transforming the original world coordinate system $X^w O^w Y^w$ to a new world coordinate system $X^{w'} O^{w'} Y^{w'}$. The original disparity map

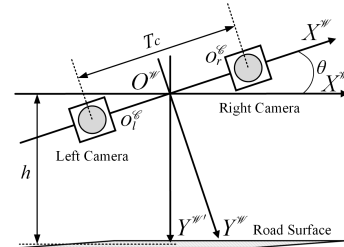


Fig. 1. Roll angle definition in a stereo rig.

is illustrated in Fig. 2, where readers can observe that the disparity values change gradually in the horizontal direction. This makes the way of representing the road disparity projections with a linear model problematic [18]. Therefore, the effects caused by the non-zero roll angle have to be eliminated before performing road region segmentation [18].

In recent years, many researchers have proposed to estimate the roll angle from the disparity map. For example, in [19] and [20], the road surface was assumed to be planar, and the disparities in a selected small area were modeled by a plane $g(u, v) = c_0 + c_1 u + c_2 v$, where u and v are the horizontal and vertical disparity coordinates, respectively. The roll angle was then calculated as $\arctan(-c_1/c_2)$. However, finding a proper small area for plane modeling is always challenging, because such an area may contain an obstacle [18]. This can severely affect the plane modeling precision [21]. In this regard, an iterative roll angle estimation algorithm was proposed in [21], where multiple small areas were selected for plane modeling. The roll angle was then estimated from the optimum selected area. However, the aforementioned algorithms can only work for planar road surfaces. Therefore, in our previous work [18], the projections of road disparities were represented by a parabola, and the roll angle estimation was formulated as an energy minimization problem. To achieve greater processing efficiency, golden section search (GSS) algorithm [22] was utilized to reduce the roll angle search range. However, this method still takes extensive iterations to converge to the minimum. Therefore, the algorithm proposed in this paper solves

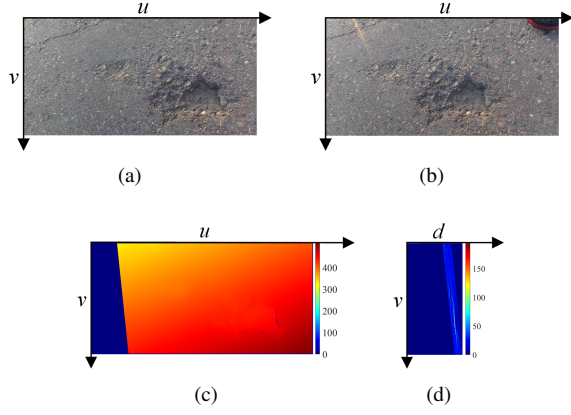


Fig. 2. Disparity map when the roll angle does not equal zero; (a) left stereo image; (b) right stereo image; (c) dense disparity map; (d) v-disparity map.

the energy minimization problem using gradient descent (GD) [23], which greatly reduces the computational complexity of roll angle estimation.

The remainder of this paper is structured as follows: Section II presents the proposed roll angle estimation algorithm. The experimental results for performance evaluation are illustrated in Section III. Finally, Section IV summarizes the paper and provides recommendations for future work.

II. METHODOLOGY

The input of our proposed algorithm is a dense disparity map, which was obtained from a stereo road image pair (see Fig. 2(a) and 2(b)) through our previously published disparity estimation algorithm [20]. The estimated disparity map is shown in Fig. 2(c), and its v-disparity map is shown in Fig. 2(d). To rotate the disparity map around a given angle θ , each set of original coordinates $[u, v]^T$ is transformed to a set of new coordinates $[x(\theta), y(\theta)]^T$ using the following equations:

$$x(\theta) = u \cos \theta + v \sin \theta, \quad (1)$$

$$y(\theta) = v \cos \theta - u \sin \theta. \quad (2)$$

The parabola representing the road disparity projections can, therefore, be written as follows:

$$f(u, v, \theta) = \alpha_0 + \alpha_1 \cdot y(u, v, \theta) + \alpha_2 \cdot y(u, v, \theta)^2. \quad (3)$$

The coefficients of $f(u, v, \theta)$ can, therefore, be estimated by solving an energy minimization problem as follows:

$$\hat{\alpha} = \arg \min_{\alpha} E(\theta) \quad \text{s.t. } \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right], \quad (4)$$

where

$$\begin{aligned} E(\theta) &= \sum_{i=0}^n \left(d_i - f(u_i, v_i, \theta) \right)^2 \\ &= (\mathbf{d} - \mathbf{Y}(\theta)\boldsymbol{\alpha}(\theta))^T (\mathbf{d} - \mathbf{Y}(\theta)\boldsymbol{\alpha}(\theta)), \end{aligned} \quad (5)$$

the column vector $\mathbf{d} = [d_0, d_1, \dots, d_n]^T$ stores the disparity values, the column vector $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \alpha_2]^T$ stores the coefficients of $f(u, v, \theta)$, and

$$\mathbf{Y}(\theta) = \begin{bmatrix} 1 & y_0(\theta) & y_0(\theta)^2 \\ 1 & y_1(\theta) & y_1(\theta)^2 \\ \vdots & \vdots & \vdots \\ 1 & y_n(\theta) & y_n(\theta)^2 \end{bmatrix}. \quad (6)$$

The energy minimization problem in (4) has a closed-form solution [24]:

$$\boldsymbol{\alpha} = \mathbf{J}(\theta)\mathbf{d}, \quad (7)$$

where

$$\mathbf{J}(\theta) = (\mathbf{Y}(\theta)^T \mathbf{Y}(\theta))^{-1} \mathbf{Y}(\theta)^T. \quad (8)$$

The minimum energy E_{\min} can also be obtained by combining (5) and (7):

$$E_{\min}(\theta) = \mathbf{d}^T \mathbf{d} - \mathbf{d}^T \mathbf{Y}(\theta) \mathbf{J}(\theta) \mathbf{d}. \quad (9)$$

According to [18], the desirable roll angle corresponds to the minimum E_{\min} . Therefore, the proposed roll angle estimation algorithm is equivalent to the following energy minimization problem:

$$\hat{\theta} = \arg \min_{\theta} E_{\min}(\theta) \quad \text{s.t. } \theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]. \quad (10)$$

However, deriving the numeric solution of (10) is significantly challenging [18]. Therefore, the proposed algorithm utilizes GD to approximate θ by formulating (10) as an iterative optimization problem as follows [25]:

$$\theta^{(k+1)} = \theta^{(k)} - \lambda \nabla E_{\min}(\theta^{(k)}), \quad k \in \mathbb{N}^0, \quad (11)$$

where

$$\nabla E_{\min}(\theta) = -2\mathbf{d}^T \mathbf{M}(\theta)\mathbf{d}, \quad (12)$$

$$\mathbf{M}(\theta) = \left[\mathbf{I} - \mathbf{Y}(\theta)\mathbf{J}(\theta) \right] \nabla \mathbf{Y}(\theta)\mathbf{J}(\theta), \quad (13)$$

\mathbf{I} is an identity matrix, and λ is the learning rate. If λ is set to a relatively low value, the convergence of (11) may take extensive iterations; If λ is too high, (11) may overshoot the local minimum [22]. Therefore, selecting a proper λ is very important for GD method. Instead of fixing λ with a constant, we utilize backtracking line search algorithm to produce an adaptive λ :

$$\lambda^{(k+1)} = \frac{\lambda^{(k)} \nabla E_{\min}(\theta^{(k)})}{\nabla E_{\min}(\theta^{(k)}) - \nabla E_{\min}(\theta^{(k+1)})}, \quad k \in \mathbb{N}^0. \quad (14)$$

The initial learning rate $\lambda^{(0)}$ is set to 26. We will discuss the selection of $\lambda^{(0)}$ in Section III. The initial approximation $\theta^{(0)}$ is set to 0, because in practical experiments the absolute value of the roll angle is usually not that high. The optimization iterates until the absolute difference between $\theta^{(k)}$ and $\theta^{(k+1)}$ is smaller than a preset threshold δ_{θ} . The performance of the proposed roll angle estimation algorithm will be discussed in Section III.

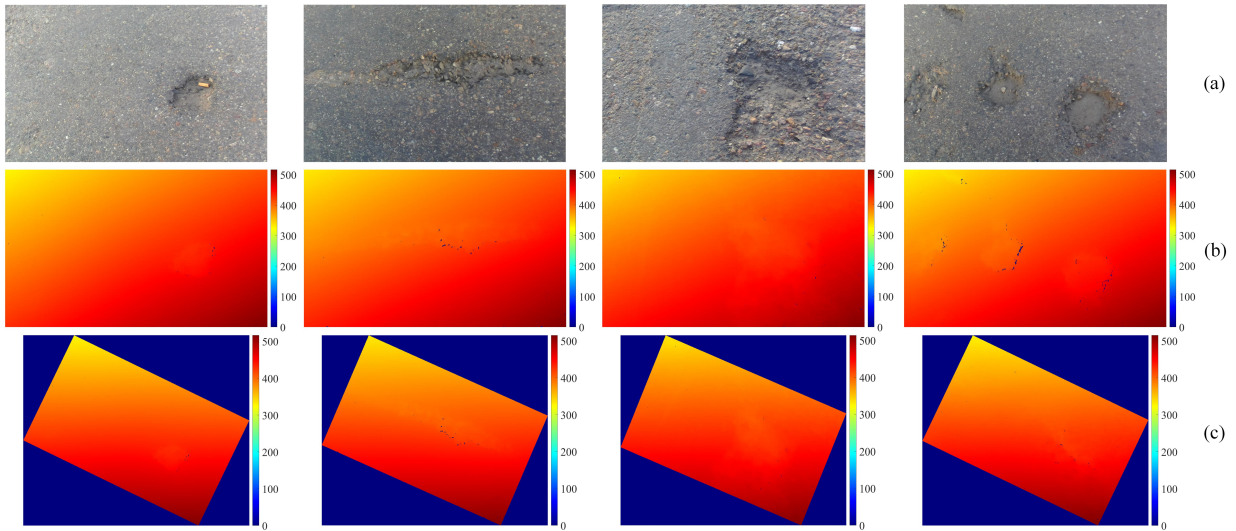


Fig. 3. Examples of the roll angle estimation experimental results; (a) left images; (b) original disparity maps; (c) disparity maps rotated around the estimated roll angles.

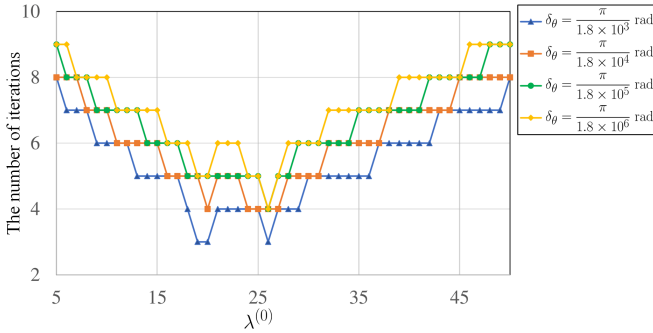


Fig. 4. The number of iterations with respect to different $\lambda^{(0)}$ and δ_θ .

III. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of the proposed roll angle estimation algorithm both qualitatively and quantitatively. The proposed method was implemented in Matlab 2019a on an Intel Core i7-8700K CPU (3.7 GHz) using a single thread.

Firstly, we discuss the selection of $\lambda^{(0)}$. In our experiments, we utilized a ZED stereo camera¹ to capture stereo road image pairs. The disparity maps estimated from the stereo image pairs were then utilized to estimate the roll angles. Some experimental results of our pothole datasets² are shown in Fig. 3, where readers can observe that the disparities on each row in the rotated disparity maps have similar values. Therefore, the proposed algorithm can effectively estimate the roll angle from a given disparity map.

Furthermore, we select a range of $\lambda^{(0)}$ and record the number of iterations that GD takes to converge with respect

TABLE I
COMPARISON BETWEEN GSS AND GD IN TERMS OF THE NUMBER OF ITERATIONS. $\lambda^{(0)} = 26$.

Method	δ_θ (rad)			
	$\frac{\pi}{1.8 \times 10^3}$	$\frac{\pi}{1.8 \times 10^4}$	$\frac{\pi}{1.8 \times 10^5}$	$\frac{\pi}{1.8 \times 10^6}$
GSS	16	21	26	30
GD	3	4	4	4

to different δ_θ , as shown in Fig. 4, where readers can see that $\lambda^{(0)} = 26$ is the optimum value for different δ_θ . When $\delta_\theta = \frac{\pi}{1.8 \times 10^3}$ rad (0.1°), GD takes three iterations to converge, while GD iterates four times, when δ_θ is set to $\frac{\pi}{1.8 \times 10^4}$ rad (0.01°), $\frac{\pi}{1.8 \times 10^5}$ rad (0.001°) or $\frac{\pi}{1.8 \times 10^6}$ rad (0.0001°).

Moreover, we compare the computational efficiency of the proposed algorithm with [18]. As discussed in [18], the computational complexity of GSS is $\mathcal{O}(\log_k \frac{\delta_\theta}{\pi})$, where $k = 0.618$ is the golden section ratio [22]. The comparison of the computational efficiency between GSS and GD is shown in Table I, where we can see that GSS takes more iterations to converge with the decrease of δ_θ , while the performance of GD largely remains the same. Therefore, the proposed roll angle estimation algorithm performs more robustly than [18] with respect to different thresholds δ_θ .

However, GD involves the computations of both ∇E_{\min} and λ , which are more complex than the computation of E . Therefore, we also compare the runtime of GSS and GD with respect to different thresholds δ_θ , as shown in Table II. Readers can observe that the runtime of GSS goes up gradually with the decrease of δ_θ , while the runtime of GD changes slightly with respect to different δ_θ . Overall, the proposed algorithm performs much faster than [18].

To quantify the accuracy of the proposed roll angle estimation algorithm, we utilize a synthesized stereo dataset

¹<https://www.stereolabs.com/>

²www.ruirangerfan.com

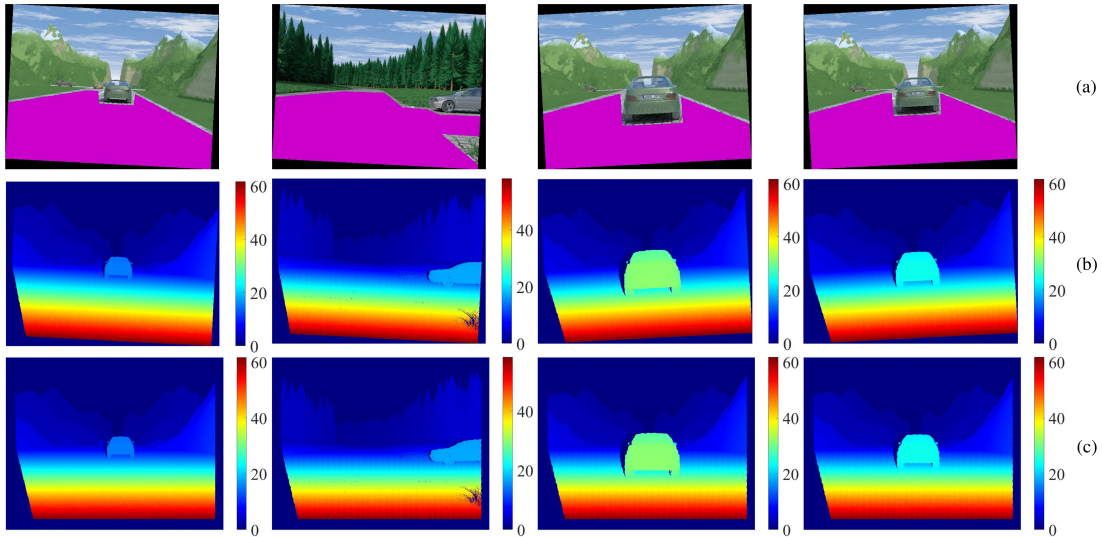


Fig. 5. Experimental results of the EISATS stereo datasets; (a) left stereo images, the areas in magenta are our manually selected road regions; (b) disparity maps; (c) disparity maps rotated around the estimated roll angles.

TABLE II
COMPARISON BETWEEN GSS AND GD IN TERMS OF RUNTIME (IN SECONDS). $\lambda^{(0)} = 26$.

Method	δ_θ (rad)			
	$\frac{\pi}{1.8 \times 10^3}$	$\frac{\pi}{1.8 \times 10^4}$	$\frac{\pi}{1.8 \times 10^5}$	$\frac{\pi}{1.8 \times 10^6}$
GSS	1.551	1.986	2.457	2.810
GD	0.412	0.401	0.511	0.512

from EISATS³ [26], [27], where the roll angle is perfectly zero. The authors manually rotate the disparity maps around a given angle, and then estimate the roll angles from the rotated disparity maps. Examples of the EISATS stereo datasets are shown in Fig. 5, where readers can see that the effects caused by the roll angle are successfully eliminated by performing the proposed algorithm. Furthermore, we compare the accuracy of the estimated roll angles between the proposed method and [18] with respect to different δ_θ . The comparison results are shown in Fig. 6, where $\Delta\theta$ represents the average difference between the actual and estimated roll angles. From Fig. 6, we can find that GD performs more accurately than GSS when δ_θ is set to a lower value. Additionally, the computational efficiency of GD is also greater than that of GSS. Therefore, GD is a better solution than GSS in terms of solving the energy minimization problem in roll angle estimation.

IV. CONCLUSION AND FUTURE WORK

In this paper, a robust GD-based algorithm was proposed to estimate the roll angle from disparity maps. The proposed algorithm was developed based on our previously published work, where roll angle estimation was formulated as an energy minimization problem, which was then solved using GSS. The experimental results demonstrated that the proposed algorithm

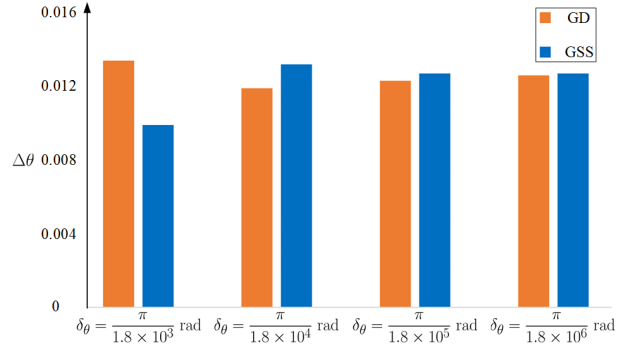


Fig. 6. Comparison between GD and GSS in terms of the roll angle accuracy.

outperforms our previous method in terms of both accuracy and efficiency, and that it can successfully eliminate the effects caused by the non-zero roll angle. In the future, we aim to design a neural network to estimate the roll angle with an unsupervised learning strategy.

ACKNOWLEDGMENT

This work is supported by grants from the Research Grants Council of the Hong Kong SAR Government, China (No. 11210017, No. 16212815, No. 21202816, NSFC U1713211) awarded to Prof. Ming Liu. This work is also supported by grants from the Shenzhen Science, Technology and Innovation Commission (JCYJ20170818153518789), National Natural Science Foundation of China (No. 61603376) and Guangdong Innovation and Technology Fund (No. 2018B050502009) awarded to Dr. Lujia Wang. This work is also supported by funding awarded to Prof. Ioannis Pitas received from the European Union Horizon 2020 research and innovation programme, under grant agreement No. 731667 (Multidrone).

³<https://ccv.wordpress.fos.auckland.ac.nz/>

REFERENCES

- [1] R. Fan, "Real-time computer stereo vision for automotive applications," Ph.D. dissertation, University of Bristol, 2018.
- [2] E. Trucco and A. Verri, *Introductory techniques for 3-D computer vision*. Prentice Hall Englewood Cliffs, 1998, vol. 201.
- [3] R. Fan, Y. Liu, M. J. Bocus, and M. Liu, "Real-time subpixel fast bilateral stereo," *arXiv preprint arXiv:1807.02044*, 2018.
- [4] N. Qian, "Binocular disparity and the perception of depth," *Neuron*, vol. 18, no. 3, pp. 359–368, 1997.
- [5] R. Fan and N. Dahnoun, "Real-time implementation of stereo vision based on optimised normalised cross-correlation and propagated search range on a gpu," in *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*. IEEE, 2017, pp. 1–6.
- [6] R. Fan, Y. Liu, J. Pan, M. J. Bocus, N. Dahnoun, and S. Tancock, "Real-time stereo vision for road surface 3-d reconstruction," in *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*. IEEE, 2018, pp. 1–6.
- [7] R. Fan, J. Jiao, X. Pan, H. Huang, S. Shen, and M. Liu, "Real-time dense stereo embedded in a uav for road inspection," *arXiv preprint arXiv:1904.06017*.
- [8] R. Labayrade, D. 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*, vol. 2. IEEE, 2002, pp. 646–651.
- [9] N. Soquet, D. Aubert, and N. Hautiere, "Road segmentation supervised by an extended v-disparity algorithm for autonomous navigation," in *2007 IEEE Intelligent Vehicles Symposium*. IEEE, 2007, pp. 160–165.
- [10] R. Fan and N. Dahnoun, "Real-time stereo vision-based lane detection system," *Measurement Science and Technology*, vol. 29, no. 7, p. 074005, 2018.
- [11] D. Yiruo, W. Wenjia, and K. Yukihiro, "Complex ground plane detection based on v-disparity map in off-road environment," in *2013 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2013, pp. 1137–1142.
- [12] A. Broggi, C. Caraffi, R. I. Fedriga, and P. Grisleri, "Obstacle detection with stereo vision for off-road vehicle navigation," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, Sep. 2005, p. 65.
- [13] Z. Hu, F. Lamosa, and K. Uchimura, "A complete uv-disparity study for stereovision based 3d driving environment analysis," in *3-D Digital Imaging and Modeling, 2005. 3DIM 2005. Fifth International Conference on*, 2005, pp. 204–211.
- [14] T. Chen, B. Chen, X. Zhang, and Z. Yuan, "Free space detection using stereo confidence metrics and obstacle position probability maps," in *Proc. 14th IEEE Int. Conf. Signal Processing (ICSP)*, Aug. 2018, pp. 1071–1075.
- [15] F. Oniga, E. Sarkozi, and S. Nedeveschi, "Fast obstacle detection using u-disparity maps with stereo vision," in *Proc. IEEE Int. Conf. Intelligent Computer Communication and Processing (ICCP)*, Sep. 2015, pp. 203–207.
- [16] W. Song, M. Fu, Y. Yang, M. Wang, X. Wang, and A. Kornhauser, "Real-time lane detection and forward collision warning system based on stereo vision," in *Proc. IEEE Intelligent Vehicles Symp. (IV)*, Jun. 2017, pp. 493–498.
- [17] H. Ma, Y. Ma, J. Jiao, M. U. M. Bhutta, M. J. Bocus, L. Wang, M. Liu, and R. Fan, "Multiple lane detection algorithm based on optimised dense disparity map estimation," in *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*. IEEE, 2018, pp. 1–5.
- [18] R. Fan, M. J. Bocus, and N. Dahnoun, "A novel disparity transformation algorithm for road segmentation," *Information Processing Letters*, vol. 140, pp. 18–24, 2018.
- [19] U. Ozgunalp, R. Fan, X. Ai, and N. 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.
- [20] R. Fan, X. Ai, and N. Dahnoun, "Road surface 3D reconstruction based on dense subpixel disparity map estimation," *IEEE Transactions on Image Processing*, vol. PP, no. 99, p. 1, 2018.
- [21] M. Evans, R. Fan, and N. Dahnoun, "Iterative roll angle estimation from dense disparity map," in *2018 7th Mediterranean Conference on Embedded Computing (MECO)*. IEEE, 2018.
- [22] P. Pedregal, *Introduction to optimization*. Springer Science & Business Media, 2006, vol. 46.
- [23] J. C. Meza, S. Profile, J. C. Meza, R. A. Oliva, E. O. L. Berkeley, P. D. Hough, and P. J. Williams, "Nonlinear optimization."
- [24] E. W. Weisstein, "Least squares fitting," 2002.
- [25] S. S. Rao and S. S. Rao, *Engineering Optimization: Theory and Practice*. Wiley, 2009.
- [26] T. Vaudrey, C. Rabe, R. Klette, and J. Milburn, "Differences between stereo and motion behaviour on synthetic and real-world stereo sequences," in *Image and Vision Computing New Zealand, 2008. IVCNZ 2008. 23rd International Conference*. IEEE, 2008, pp. 1–6.
- [27] A. Wedel, C. Rabe, T. Vaudrey, T. Brox, U. Franke, and D. Cremers, "Efficient dense scene flow from sparse or dense stereo data," in *European conference on computer vision*. Springer, 2008, pp. 739–751.