

# Real-Time Stereo Vision-Based Lane Detection System

Rui Fan & Naim Dahnoun

Department of Electrical and Electronic Engineering, Merchant Venturers Building,  
University of Bristol, Bristol, BS8 1UB, UK

E-mail: {ranger.fan, naim.dahnoun}@bristol.ac.uk

February 2018

**Abstract.** The detection of multiple curved lane markings on a non-flat road surface is still a challenging task for automotive applications. To make an improvement, the depth information can be used to greatly enhance the robustness of the lane detection systems. The proposed system in this paper is developed from our previous work where the dense vanishing point  $V_p$  is estimated globally to assist the detection of multiple curved lane markings. However, the outliers in the optimal solution may severely affect the accuracy of the least squares fitting when estimating  $V_p$ . Therefore, in this paper we use Random Sample Consensus to update the inliers and outliers iteratively until the fraction of the number of inliers versus the total number exceeds our pre-set threshold. This significantly helps the system to overcome some suddenly changing conditions. Furthermore, we propose a novel lane position validation approach which provides a piecewise weight  $w_g$  based on  $V_p$  and the gradient  $\nabla$  to reduce the gradient magnitude of the non-lane candidates. Then, we compute the energy of each possible solution and select all satisfying lane positions for visualisation. The proposed system is implemented on a heterogeneous system which consists of a Intel Core i7-4720HQ CPU and a NVIDIA GTX 970M GPU. A processing speed of 143 fps has been achieved, which is over 38 times faster than our previous work. Also, in order to evaluate the detection precision, we tested 2495 frames with 5361 lanes from the KITTI database (1637 lanes more than our previous experiment). It is shown that the overall successful detection rate is improved from 98.7% to 99.5%.

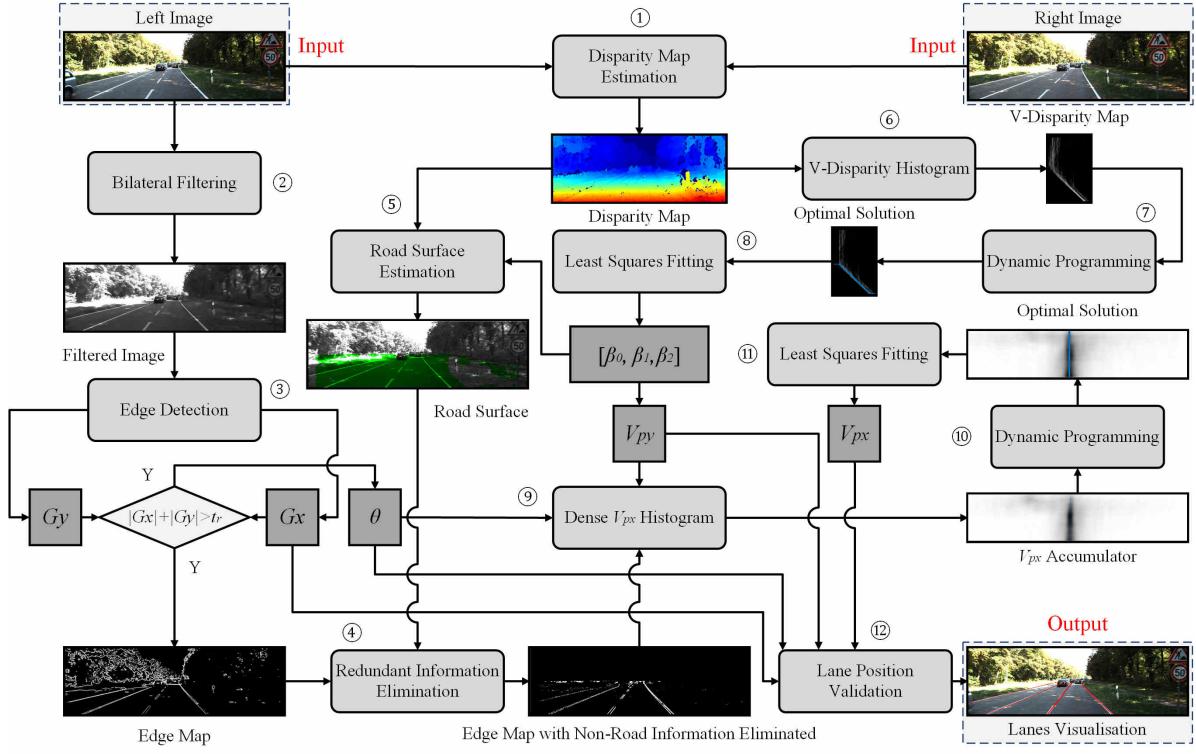
## 1. Introduction

Various prototype vehicle road tests have been conducted by Google in the US since 2012, and its subsidiary X plans to commercialise their autonomous cars as from 2020 [1]. Recently, Volvo has been planning to conduct a series of self-driving experiments involving about 100 cars in China and many companies like Ford and Uber have entered the race to make driver-less taxis a reality. The techniques like lane detection systems in the ADAS (Advanced Driver Assistance Systems) are playing an increasingly crucial role in enhancing driving safety and minimising the possibilities of fatalities.

Current lane detection algorithms can mainly be grouped into two categories: feature-based and model-based [2]. The feature-based algorithms extract the local, meaningful and detectable parts of the image, such as edges, texture and colour, to segment lanes and road boundaries from the background pixel by pixel [3]. On the other hand, the model-based algorithms try to represent the lanes with a mathematical equation based upon some assumptions of the road geometry [4]. The commonly used lane models include: linear, parabolic, linear-parabolic and splines. The linear model works well for the lanes with a low curvature, as demonstrated by the linear lane detection system in our previous work [5, 6]. However, a more flexible road model is inevitable when the lanes with a higher curvature exist. Therefore, some algorithms [7–10] use a parabolic model to represent the lanes with a constant curvature. For some other more complicated cases, Jung et al. propose a linear-parabolic combined lane model, where the nearby lanes are represented as linear models, whereas the far ones are modelled as parabolas [11]. In addition to the models mentioned above, the spline model is another alternative which interpolates the pixels on a lane as an arbitrary shape [4, 12]. However, the more parameters introduced into a flexible model, the higher will be the computational complexities of the algorithm. Therefore, we turn our focus on some additional important properties of 3D imaging techniques instead of being limited to 2D information.

One of the most prevalently used methods is IPM (Inverse Perspective Mapping). With the assumption that two lanes are parallel to each other in the world coordinate system (WCS), IPM is able to map a 3D scenery into a 2D bird's eye view [13]. Furthermore, many researchers [14–17] propose to use the vanishing point  $\mathbf{V}_p = [V_{px}, V_{py}]^\top$  to model lane markings and road boundaries. However, their algorithms work well only if we assume the road surface is flat or the camera parameters are known. Therefore, we pay closer attention to the disparity information provided by either active sensors (radar and laser) or passive sensors (stereo cameras) [18]. Since Labayrade et al. proposed the concept of "v-disparity" [19], disparity information has been widely used to enhance the robustness of the lane detection systems. Our previous work [18] shows a particular instance where the disparity information is successfully combined with a lane detection algorithm to estimate  $V_{py}$  for a non-flat road surface. At the same time, a lot of redundant information can be eliminated from the obstacles by comparing their disparities with the theoretical ones. However, the estimation of  $V_{py}$  suffers from the outliers in an optimal solution, and the lanes are sometimes unsuccessfully detected because the selection of plus-minus peaks is not always effective. Moreover, real-time performance is still a challenging task in [18] because of the intensive computational complexities of the algorithm. Therefore, in this paper we present an improved lane detection system for the problems mentioned above.

The proposed multiple lane detection system is composed of four main components: disparity map estimation, dense  $V_{py}$  estimation, dense  $V_{px}$  estimation and lane position validation. The block diagram of the proposed system is illustrated in Fig. 1. Procedures 1 to 6 run on a GPU because they are more efficient for parallel processing. On the



**Figure 1.** The block diagram of the proposed lane detection system.

contrary, the serially-efficient procedures 7 to 12 execute on a CPU .

Firstly, we estimate the disparity map from a pair of well-calibrated stereo images. The disparity map is mainly used for:

- estimation of dense  $V_{py}$ ,
- road surface detection, and
- elimination of the noise from obstacles.

The disparity information is combined with the lane detection system by analysing a so-called v-disparity histogram. In this paper, we assume the road surface is non-flat and the projection of on-road disparities on the v-disparity map can be represented by a parabola  $f(v) = \beta_0 + \beta_1 v + \beta_2 v^2$ , where  $\beta = [\beta_0, \beta_1, \beta_2]^\top$  is the parameters of the vertical road profile and  $v$  is the row number. Compared with various quadratic pattern detectors, dynamic programming (DP) provides a more efficient way of extracting the best path from the v-disparity map by minimising the energy of an optimisation problem. Then, the extracted path  $M_y$  is fitted to a parabola using the least squares fitting (LSF). However, the outliers  $\mathcal{O}$  in  $M_y$  may affect the accuracy of the LSF significantly. Therefore, we propose to update the parabola function iteratively using RANSAC (Random Sample Consensus) until the capacity of inliers  $\mathcal{I}$  does not change, which greatly helps the system maintain the robustness when estimating  $V_{py}$ . Furthermore, we employ a bilateral filter before approximating the horizontal and vertical derivatives:  $G_x$  and  $G_y$ . This achieves a better performance than the median filter which was used

in our previous work [18] in terms of the edge-preservation and noise elimination.  $V_{py}$  and the orientation of each pixel in the road surface area are then used to estimate  $V_{px}$ , where we apply the same strategy of RANSAC to minimise the influence of outliers  $\mathcal{O}$  on the LSF. An arbitrary lane marking or road boundary can thus be extracted using the information of  $\mathbf{V}_p = [V_{px}, V_{py}]^\top$ . To locate the correct lane markings, we propose a novel lane position validation approach using the information of  $G_x$  (the dark-light transition of the lane markings has a positive and higher  $G_x$  than the non-edge pixels, while  $-G_x$  is positive and higher on the light-dark transition of the lane markings). The angle  $\angle(\mathbf{m}, \mathbf{n})$  between the orientation vector  $\mathbf{m}$ , obtained from the edge detector and the orientation vector  $\mathbf{n}$ , obtained from  $\mathbf{V}_p$ , is used to provide a piecewise weight  $w_g$  to update  $G_x$ , where the value of an edge pixel that does not belong to either lane markings or road boundaries is cut down. After computing the energy of each possible solution, the satisfying local minima are validated as lane positions. The proposed system achieves a real-time execution of 143 fps along with an approximately 99.5% successful detection rate on our heterogeneous system consisting of an Intel Core i7-4720HQ CPU and a NVIDIA GTX 970M GPU.

The remainder of the paper is structured as follows: section 2 describes the proposed lane detection system, section 3 evaluates the experimental results, and section 4 concludes the paper.

## 2. System Description

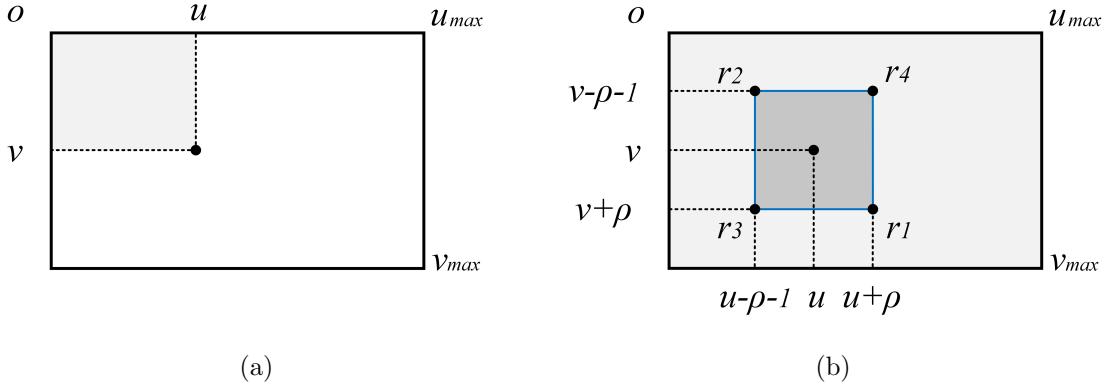
### 2.1. Disparity map estimation

As compared to many other stereo matching algorithms which are aimed at automotive applications, the trade-off between accuracy and speed has been greatly improved in our previous work [20, 21]. Therefore, we employ the algorithm in [21] to acquire the disparity information for the proposed lane detection system.

**2.1.1. Memorisation** Due to the insensitivity to the intensity difference, the Normalised Cross-Correlation (NCC) is utilised as the cost function to measure the similarity between two blocks, as shown in Eq. 1. Each block chosen from the left image is matched with a series of blocks on the same epipolar line in the right image [22]. The block pair with the highest correlation cost is selected as the best correspondence, and the shifting distance between them is defined as the disparity  $d$ .

$$c(u, v, d) = \frac{1}{n\sigma_l\sigma_r} \sum_{i=u-\rho}^{i=u+\rho} \sum_{j=v-\rho}^{j=v+\rho} (I_l(i, j) - \mu_l)(I_r(i - d, j) - \mu_r) \quad (1)$$

where  $c$  is defined as the correlation cost, and  $I_l$  and  $I_r$  represent the pixel intensities in the left and right images, respectively. The centre of the block is  $(u, v)$  and the side length of the block is  $2\rho + 1$ .  $n = (2\rho + 1)^2$  represents the number of pixels within each



**Figure 2.** Integral image processing. (a) original image. (b) integral image.

block.  $\mu_l$  and  $\mu_r$  denote the means of the intensities within the left and right blocks, respectively.  $\sigma_l$  and  $\sigma_r$  represent their corresponding standard deviations [21]:

$$\sigma_l = \sqrt{\sum_{i=u-\rho}^{i=u+\rho} \sum_{j=v-\rho}^{j=v+\rho} (I_l(i, j) - \mu_l)^2 / n} \quad (2)$$

$$\sigma_r = \sqrt{\sum_{i=u-\rho}^{i=u+\rho} \sum_{j=v-\rho}^{j=v+\rho} (I_r(i - d, j) - \mu_r)^2 / n} \quad (3)$$

When the left block is selected, the calculations of  $\mu_l$  and  $\sigma_l$  are always repeated because  $d$  is only used to select the position of the right blocks for stereo matching. Therefore, the four independent parts  $\mu_l$ ,  $\mu_r$ ,  $\sigma_l$  and  $\sigma_r$  can be pre-calculated and stored in static program storage for direct indexing. The integral image algorithm can be used to compute  $\mu_l$  and  $\mu_r$  efficiently [23], which is illustrated in Fig. 2. The algorithm has two steps: integral image initialisation and value indexing from the initialised reference. In the first step, for a discrete image  $I$  whose pixel intensity at  $(u, v)$  is  $I(u, v)$ , its integral image intensity  $In(u, v)$  at the position of  $(u, v)$  is defined as:

$$In(u, v) = \sum_{i \leq u, j \leq v} I(i, j) \quad (4)$$

Algorithm 1 details the implementation of the integral image initialisation, where  $In$  is calculated serially based on its previous neighbouring results to minimise unnecessary computations.

After initialising an integral image, the sum  $s(u, v)$  of pixel intensities within a square block whose side length is  $2\rho + 1$  and centre is  $(u, v)$  can be computed with four references  $r_1 = In(u + \rho, v + \rho)$ ,  $r_2 = In(u - \rho - 1, v - \rho - 1)$ ,  $r_3 = In(u - \rho - 1, v + \rho)$ , and  $r_4 = In(u + \rho, v - \rho - 1)$  as follows:

$$s(u, v) = r_1 + r_2 - r_3 - r_4 \quad (5)$$

**Algorithm 1:** Integral image initialisation

---

**Input** : original image:  $I$   
**Output:** integral image:  $In$

```

1  $In(u_{min}, v_{min}) \leftarrow I(u_{min}, v_{min});$ 
2 for  $u \leftarrow u_{min} + 1$  to  $u_{max}$  do
3   |  $In(u, v_{min}) \leftarrow In(u - 1, v_{min}) + I(u, v_{min});$ 
4 end
5 for  $v \leftarrow v_{min} + 1$  to  $v_{max}$  do
6   |  $In(u_{min}, v) \leftarrow In(u_{min}, v - 1) + I(u_{min}, v);$ 
7 end
8 for  $u \leftarrow u_{min} + 1$  to  $u_{max}$  do
9   | for  $v \leftarrow v_{min} + 1$  to  $v_{max}$  do
10    |   |  $In(u, v) \leftarrow In(u, v - 1) + In(u - 1, v)$ 
11    |   |  $-In(u - 1, v - 1) + I(u, v);$ 
12   | end
13 end

```

---

The mean  $\mu(u, v) = s(u, v)/n$  of the intensities within the selected block is then stored in a static program storage for the computations of  $\sigma$  and  $c$ . To simplify the computations of  $\sigma_l$  and  $\sigma_r$ , we rearrange Eq. 2 and Eq. 3 as shown in Eq. 6 and Eq. 7, respectively.

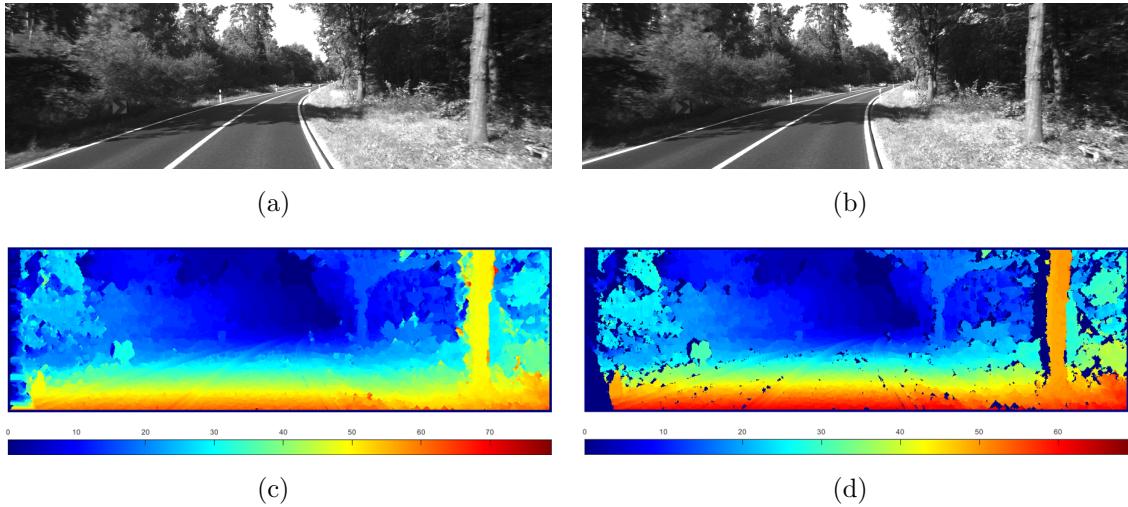
$$\sigma_l = \sqrt{\sum_{i=u-\rho}^{i=u+\rho} \sum_{j=v-\rho}^{j=v+\rho} I_l^2(i, j)/n - \mu_l^2} \quad (6)$$

$$\sigma_r = \sqrt{\sum_{i=u-\rho}^{i=u+\rho} \sum_{j=v-\rho}^{j=v+\rho} I_r^2(i - d, j)/n - \mu_r^2} \quad (7)$$

where  $\sum I_l^2$  and  $\sum I_r^2$  are dot products. Similarly, the computations of  $\sum I_l^2$  and  $\sum I_r^2$  can be accelerated by initialising two integral images  $In_{l2}$  and  $In_{r2}$  as references for indexing. Therefore, the standard deviations  $\sigma_l$  and  $\sigma_r$  can also be calculated and stored in static program storage for the efficient computation of  $c$  as follows:

$$c(u, v, d) = \frac{1}{n\sigma_l\sigma_r} \left[ \sum_{i=u-\rho}^{i=u+\rho} \sum_{j=v-\rho}^{j=v+\rho} I_l(i, j)I_r(i - d, j) - n\mu_l\mu_r \right] \quad (8)$$

From Eq. 8, only  $\sum I_l I_r$  needs to be calculated during the stereo matching. Hence, with the values of  $\mu_l$ ,  $\mu_r$ ,  $\sigma_l$  and  $\sigma_r$  able to be indexed directly, Eq. 1 is simplified as a dot product. The performance improvement achieved by factorising the NCC equation will be discussed section 3.1.



**Figure 3.** Input images and disparity maps. (a) left image. (b) right image. (c) disparity map. (d) disparity map processed with LRC.

**2.1.2. Search Range Propagation (SRP)** In this paper, the disparities are estimated iteratively row by row from row  $v_{max}$  to row  $v_{min}$ . In the first iteration, the stereo matching goes for a full search range  $SR = \{sr | sr \in [d_{min}, d_{max}]\}$ . Then, the search range for stereo matching at the position of  $(u, v)$  is propagated from three estimated neighbouring disparities  $\ell(u - 1, v + 1)$ ,  $\ell(u, v + 1)$  and  $\ell(u + 1, v + 1)$  using Eq. 9 [24], where  $\tau$  is the bound of the search range and set as 1 in our proposed system. The left disparity map is illustrated in Fig. 3(c). More details about the SRP-based disparity estimation are given in algorithm 2. The performance of the SRP-based stereo will be discussed in section 3.1.

$$SR = \bigcup_{k=u-1}^{u+1} \{sr | sr \in [\ell(k, v + 1) - \tau, \ell(k, v + 1) + \tau]\} \quad (9)$$

**2.1.3. Post-Processing** For various disparity map estimation algorithms, the pixels that are only visible in one disparity map are a major source of the matching errors. Due to the uniqueness constraint of the correspondence, for an arbitrary pixel  $(u, v)$  in the left disparity map  $\ell^{lf}$ , there exists at most one correspondence in the right disparity map  $\ell^{rt}$ , namely [21]:

$$\ell^{lf}(u, v) = \ell^{rt}(u - \ell^{lf}(u, v), v) \quad (10)$$

A left-right consistency (LRC) check is performed to remove half-occluded areas from the disparity map. Although the LRC check doubles the computational complexity by re-projecting the computed disparity values from one image to the other one, most of the incorrect half-occluded pixels can be eliminated and an outlier can be found [21]. For  $\ell^{rt}$  estimation, the memorisation of  $\mu_l$ ,  $\mu_r$ ,  $\sigma_l$  and  $\sigma_r$  is unnecessary because they have already been calculated when estimating  $\ell^{lf}$ . The LRC check is detailed in algorithm

---

**Algorithm 2:** SRP-based disparity map estimation

---

**Input :** left image, right image;  
           left mean map, right mean map;  
           left standard deviation map, right standard deviation map;

**Output:** disparity map

```

1 estimate the disparities for row  $v_{max}$ ;
2 for  $v \leftarrow v_{max} - 1$  to  $v_{min}$  do
3   for  $u \leftarrow u_{min}$  to  $u_{max}$  do
4     propagate the search range from row  $v + 1$  using Eq. 9;
5     estimate the disparity for  $(u, v)$ ;
6   end
7 end
```

---

3, where  $tr_{LRC}$  is the threshold and set as 3. The corresponding results can be seen in Fig. 3(d).

---

**Algorithm 3:** LRC check

---

**Input :** left disparity map:  $\ell^f$   
           right disparity map:  $\ell^{rt}$

**Output:** disparity map:  $\ell$

```

1 for  $v \leftarrow v_{min}$  to  $v_{max}$  do
2   for  $u \leftarrow u_{min}$  to  $u_{max}$  do
3     if  $abs(\ell^f(u, v) - \ell^{rt}(u - \ell^f(u, v), v)) > tr_{LRC}$  then
4        $\ell(u, v) \leftarrow 0$ ;
5     else
6        $\ell(u, v) \leftarrow \ell^f(u, v)$ ;
7     end
8   end
9 end
```

---

## 2.2. Dense $V_{py}$ Estimation

Since Labayrade et al. proposed the concept of "v-disparity" in 2002 [19], disparity information has been widely used to assist the detection of either obstacles or lanes. The v-disparity map is created by computing the histograms of each horizontal row of the disparity map. An example of the v-disparity map is shown in Fig. 4(a), which has two axes: disparity  $d$  and row number  $v$ . The value  $m_y(d, v)$  represents the accumulation at the position of  $(d, v)$  in the v-disparity map. In [25], Hu et al. proved that the projection of a flat road on the v-disparity map is a straight line:  $d = f(v) = \alpha_0 + \alpha_1 v$ . The parameters  $\boldsymbol{\alpha} = [\alpha_0, \alpha_1]^\top$  can be obtained by using some linear pattern detectors

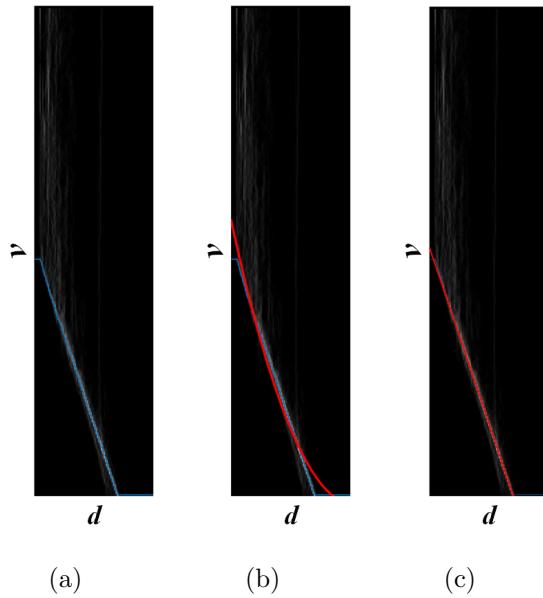
such as the Hough Transform (HT) [26]. In our previous work [18], we used a parabola model  $d = f(v) = \beta_0 + \beta_1 v + \beta_2 v^2$  to depict the projection of a non-flat road surface on the v-disparity map. In this case, DP is more efficient than some quadratic pattern detectors to search for every possible solution and extract the projection path by minimising the energy function in Eq. 11.

$$E = E_{data} + \lambda E_{smooth} \quad (11)$$

Eq. 11 is solved iteratively starting from  $d = d_{max}$  and going to  $d = 0$ . In the first iteration,  $E_{smooth} = 0$  and  $E_{data} = -m_y(d_{max}, v)$ . Then,  $E$  is computed based upon the previous iterations:

$$E(v)_d = -m_y(d, v) + \min_{\tau_y} [E(v - \tau_y)_{d+1} - \lambda_y \tau_y], \text{ s.t. } \tau_y \in [0, 6] \quad (12)$$

In each iteration, the index position of the minimum is saved into a buffer for the solution back-tracing. The buffer has the same size as the v-disparity map. The solution  $\mathbf{M}_y = [\mathbf{d}, \mathbf{v}]^\top \in \mathbb{R}^{k \times 2}$  with the minimal energy is selected as the optima, which is plotted as the blue paths in Fig. 4. The blue path has  $k$  points.  $\mathbf{v} = [v_0, v_1, \dots, v_{k-1}]^\top$  is a column vector recording their row numbers, and  $\mathbf{d} = [d_0, d_1, \dots, d_{k-1}]^\top$  is a column vector storing their disparity values. Then, the parameters  $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2]^\top$  of the vertical road profile can be estimated by solving the least squares problem in Eq. 13. We also plot  $f(v) = \beta_0 + \beta_1 v + \beta_2 v^2$  in red, as shown in Fig. 4(b) and Fig. 4(c).



**Figure 4.** V-disparity map and dynamic programming. (a) v-disparity map. (b) target solution in [18]. (c) target solution in the proposed system. The blue paths are the optimal solution from DP.  $f(v) = \beta_0 + \beta_1 v + \beta_2 v^2$  is plotted in red.

$$\boldsymbol{\beta} = \arg \min_{\boldsymbol{\beta}} \sum_{j=0}^{k-1} (d_j - (\beta_0 + \beta_1 v_j + \beta_2 v_j^2))^2 \quad (13)$$

From Fig. 4(b), we observe that the outliers  $\mathcal{O}$  severely affect the accuracy of the least squares fitting. To improve  $V_{py}$  estimation, we employ RANSAC to update the inliers  $\mathcal{I}$  and  $\boldsymbol{\beta}$  iteratively. This procedure is described in algorithm 4.

---

**Algorithm 4:**  $\boldsymbol{\beta}$  estimation with the assist of RANSAC.

---

**Input :** optimal solution  $\mathbf{M}_y = [\mathbf{d}, \mathbf{v}]^\top$

**Output:**  $\boldsymbol{\beta}$

- ```

1 do
2   select a specified number of candidates  $[d_j, v_j]^\top$  randomly;
3   fit a quadratic polynomial using Eq. 13 to get  $\boldsymbol{\beta}$ ;
4   determine the number of inliers  $\mathcal{I}$  and outliers  $\mathcal{O}$ :  $n_{\mathcal{I}}$  and  $n_{\mathcal{O}}$ , respectively;
5   remove  $\mathcal{O}$  from  $\mathbf{M}_y$ ;
6 while  $n_{\mathcal{I}}/(n_{\mathcal{I}} + n_{\mathcal{O}}) < \epsilon_y$ ;
7 fit  $\mathcal{I}$  into a quadratic polynomial and get  $\boldsymbol{\beta}$ ;

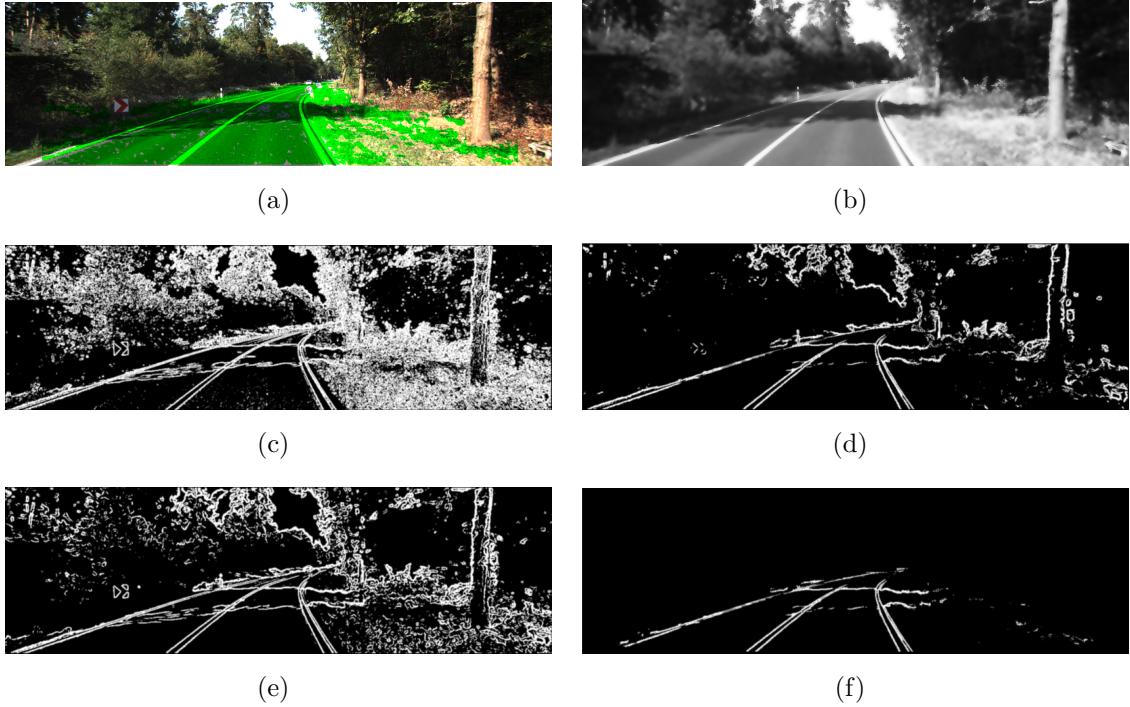
```
- 

Given a candidate  $[d_j, v_j]$ , to decide whether it is an inlier or not, we need to compute the error  $r_j = (d_j - f(v_j))^2$ . If  $r_j$  is smaller than our pre-set tolerance  $tr_y$  ( $tr_y = 4$  in this paper), we mark this candidate as an inlier and update  $\mathcal{I}$ . Otherwise, it will be marked as an outlier and removed from  $\mathbf{M}_y$ . The iteration works until the fraction of inliers versus the number of candidates in the updated  $\mathbf{M}_y$  exceeds our pre-set threshold  $\epsilon_y$  ( $\epsilon_y = 0.99$  in this paper). Finally, the inliers are used to fit a quadratic polynomial to get  $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2]^\top$ . Compared with the parabola obtained in [18], the parabola estimated with the assistance of RANSAC is more reliable and less affected by the outliers (please see Fig. 4(c)).  $V_{py}$  can be computed as follows:

$$V_{py}(v) = v - \frac{\beta_0 + \beta_1 v + \beta_2 v^2}{\beta_1 + 2\beta_2 v} \quad (14)$$

### 2.3. Dense $V_{px}$ Estimation

**2.3.1. Sparse  $V_{px}$  estimation** In order to reduce the redundant information, we set a threshold  $\varpi = 3$  to remove the non-road areas. A pixel at  $(u, v)$  in the disparity map  $\ell$  is identified to be on-road only if it satisfies  $|\ell(u, v) - f(v)| \leq \varpi$  and  $f^{-1}(0) \leq v \leq v_{max}$  ( $f^{-1}$  is the inverse function of  $f(v) = \beta_0 + \beta_1 v + \beta_2 v^2$ ). Otherwise, it is either on the obstacles or in the potholes. The estimated road surface is shown as a green area in Fig. 5(a). For the following procedures, we only focus on the road surface area, which greatly reduces the redundant information used for edge detection and dense  $V_{px}$  estimation.



**Figure 5.** Sparse  $V_{px}$  estimation. (a) road surface estimation. (b) bilateral filtering for Fig. 3(a). (c) edge detection of Fig. 3(a). (d) edge detection of (b). (e) edge detection of the median filtering output. (f) edges in the road surface area. The green area in (a) is the road surface. For the bilateral filter,  $\sigma_s = 300$  and  $\sigma_r = 0.3$ . The window size of the bilateral filter and the median filter is  $11 \times 11$ . The thresholds of the Sobel edge detection in (c), (d), (e) and (f) are 100. For the following procedures, we only focus on the edge pixels in (f).

However, the noise introduced from the input image still makes the edge detectors like Sobel very sensitive to the blobs [27]. Therefore, we use a bilateral filter to further reduce the noise before detecting edges. Compared with the median filter utilised in [18], the bilateral filter is more capable of preserving the edges when smoothing an image. The bilateral filtering [28] is performed as follows:

$$I^{bf}(u, v) = \frac{\sum_{i=u-\rho}^{i=u+\rho} \sum_{j=v-\rho}^{j=v+\rho} w_s(i, j) w_r(i, j) I(i, j)}{\sum_{i=u-\rho}^{i=u+\rho} \sum_{j=v-\rho}^{j=v+\rho} w_s(i, j) w_r(i, j)} \quad (15)$$

where

$$\begin{aligned} w_s(i, j) &= \exp \left\{ -\frac{(i-u)^2 + (j-v)^2}{\sigma_s^2} \right\} \\ w_r(i, j) &= \exp \left\{ -\frac{(I(i, j) - I(u, v))^2}{\sigma_r^2} \right\} \end{aligned} \quad (16)$$

$I(i, j)$  is the intensity of the input image at  $(i, j)$  and  $I^{bf}(u, v)$  is the intensity of the filtered image at  $(u, v)$ . The block size of the filter is  $(2\rho + 1) \times (2\rho + 1)$ , and its centre

is  $(u, v)$ . The coefficient  $w_s(i, j)$  is based on the spatial distance, and the coefficient  $w_r(i, j)$  is based upon the colour similarity.  $\sigma_s$  and  $\sigma_r$  are the parameters for  $w_s$  and  $w_r$ . In our practical experiments,  $\sigma_s$  and  $\sigma_r$  are set to 300 and 0.3, respectively. The output of bilateral filtering is shown in Fig. 5(b). The corresponding edge detection is illustrated as Fig. 5(d). The edge detection of Fig. 3(a) is shown as Fig. 5(c), while the edge detection for the output of the median filtering is depicted in Fig. 5(e). Obviously, although the median filter has removed a lot of redundant edges, the bilateral filter still achieves a better performance in terms of noise elimination and edge preservation.

In the following procedures, we only focus on the pixels on the road surface. The edge map in the road surface area is shown in Fig. 5(f). Based on the gradient  $\nabla(e) = [G_x, G_y]^\top$  approximated by the Sobel edge detector, we can estimate the sparse  $V_{px}$  using Eq. 17 for every pixel  $(u_e, v_e)$  on the edge  $e$ .  $V_{px}^s(u_e, v_e)$  at the position of  $(u_e, v_e)$  is recorded in a sparse  $V_{px}$  map (note: we use the notation  $V_{px}^s$  to represent the sparse  $V_{px}$  in the following content).

$$V_{px}^s(u_e, v_e) = u_e + \frac{v_e - V_{py}(v_e)}{\nabla(e)} \quad (17)$$

Next, we will provide some details about the implementation. In the GPU architecture, a thread is more likely to fetch the memory from the closest addresses that its nearby threads accessed, which makes the use of cache impossible [29]. Therefore, we utilise the texture memory which is read-only and cached on-chip to optimise the caching for 2D spatial locality during the bilateral filtering. Firstly, a 2D texture object is created. Then, the texture object is bound directly to the global memory address of the left image.  $I(i, j)$  used in the bilateral filtering is then fetched from the texture object to reduce the memory requests from the global memory. In addition, as the constant memory is read-only and beneficial for the data that will not change over the course of a kernel execution [29], we create two lookup tables on it to store the values of  $w_s$  and  $w_r$ . So far, the execution of bilateral filtering has been highly accelerated, so we move to the implementation of the Sobel edge detection. The address of  $I^{bf}(u, v)$  is always accessed repeatedly when judging whether the pixel at  $(u + \omega, v + \varrho)$  belongs to an edge or not, where  $\omega, \varrho \in \{-1, 0, 1\}$ . Thus, we load a group of data  $I^{bf}$  into the shared memory for each thread block. All threads within the same thread block will access the shared data instead of fetching them repeatedly from the global memory. In order to avoid the race conditions among different threads which run logically in parallel instead of executing physically concurrently, the threads within the same thread block need to be synchronised after they finish the data loading. Compared with the performance of  $V_{px}^s$  estimation on a Core-i7 4720HQ CPU processing with a single thread, our implementation on a GTX 970M GPU speeds up the execution by over 74 times.

*2.3.2. Dense  $V_{px}$  accumulation* The sparse  $V_{px}$  information is usually less robust than the dense  $V_{px}$  information with regard to validating a correct lane position. Therefore, we propose to vote  $V_{px}^s$  within a band whose longitudinal size is  $2\chi + 1$ . The voting

results are then saved in a 1D histogram. A 2D dense  $V_{px}$  accumulator can be created by shifting an unfixed band from the bottom of the  $V_{px}^s$  map to its top and gathering the 1D histogram from each band. We abbreviate dense  $V_{px}$  as  $V_{px}^d$  in the following content. The details of the  $V_{px}^d$  accumulation are described in algorithm 5.

---

**Algorithm 5:** Dense  $V_{px}$  accumulation.

---

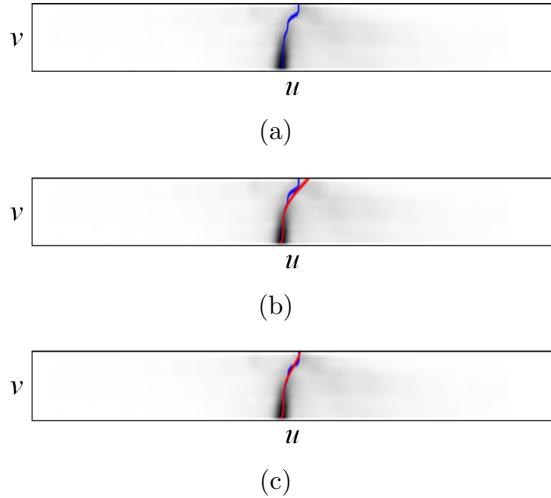
```

Input :  $V_{px}^s$  map
Output:  $V_{px}^d$  accumulator
1 set all elements in  $V_{px}^d$  accumulator as 0;
2 for  $i \leftarrow u_{min}$  to  $u_{max}$  do
3    $| V_{px}^d(V_{px}^s(i, v_{max}), v_{max}) \leftarrow -\varrho$ 
4 end
5 for  $j \leftarrow v_{max} - 1$  to  $f^{-1}(0)$  do
6   if  $j > v_{max} - \chi - 1$  then
7     for  $i \leftarrow u_{min}$  to  $u_{max}$  do
8        $| V_{px}^d(i, j) \leftarrow V_{px}^d(i, j + 1);$ 
9        $| V_{px}^d(V_{px}^s(i, j), j) \leftarrow V_{px}^d(V_{px}^s(i, j), j) - \varrho;$ 
10      end
11    else if  $f^{-1}(0) + \chi \leq j \leq v_{max} - \chi - 1$  then
12      for  $i \leftarrow u_{min}$  to  $u_{max}$  do
13         $| V_{px}^d(i, j) \leftarrow V_{px}^d(i, j + 1);$ 
14         $| V_{px}^d(i, j) \leftarrow V_{px}^d(i, j) + V_{px}^d(i, j + \chi + 1);$ 
15         $| V_{px}^d(V_{px}^s(i, j), j - \chi) \leftarrow V_{px}^d(V_{px}^s(i, j), j - \chi) - \varrho;$ 
16      end
17    else
18      for  $i \leftarrow u_{min}$  to  $u_{max}$  do
19         $| V_{px}^d(i, j) \leftarrow V_{px}^d(i, j + 1);$ 
20         $| V_{px}^d(i, j) \leftarrow V_{px}^d(i, j) + V_{px}^d(i, j + \chi + 1);$ 
21      end
22    end
23 end

```

---

The initial step is to form a 1D  $V_{px}^d$  histogram where each edge pixel  $e$  on row  $v_{max}$  votes  $V_{px}^s$  to the histogram (the parameter  $\varrho$  for each vote is proposed to be 1). Then, the votes of  $V_{px}^s$  on row  $v$  are accumulated with the 1D histogram on row  $v + 1$  to form the accumulation of  $V_{px}^d$  on row  $v$ . This continues until the band is able to reach a  $2\chi + 1$  longitudinal size, where  $\chi = 25$  in this paper. Then, the current band is shifted slightly up to create another 1D  $V_{px}^d$  histogram for statistics. In order to achieve the computational efficiency, a sliding window is used by simply subtracting the votes that appear on row  $v - \chi$  above from the current band and adding the votes that appear on the bottom row of the previous band (Please note, we subtract the votes in order



**Figure 6.** Dense  $V_{px}$  accumulator and dynamic programming. (a) dense  $V_{px}$  accumulator. (b) target solution in [18]. (c) target solution in the proposed system. The blue paths are the optimal solution obtained by DP.  $g(v) = \gamma_0 + \gamma_1 v + \gamma_2 v^2 + \gamma_3 v^3 + \gamma_4 v^4$  is plotted in red.

to ensure consistency with the term "energy minimisation" in DP, and therefore, more votes from the  $V_{px}^s$  map correspond to a negatively higher value in the  $V_{px}^d$  accumulator). Furthermore, due to the higher lane curvature a road may have, a thinner band is more desirable for those top bands [18]. Therefore, only the votes on row  $v + \chi + 1$  are added to them without the subtractions from the current band. The sliding window makes the 1D  $V_{px}^d$  histogram update more efficiently by simply processing the bottom row and the top row. The 2D  $V_{px}^d$  accumulator is illustrated in Fig. 6(a), where the notation  $m_x(u, v)$  represents the votes of  $V_{px} = u$  on row  $v$ .

**2.3.3.  $V_{px}$  estimation** Similarly, the  $V_{px}^d$  accumulator is optimised by minimising the energy function in Eq. 11 using the DP. In the first iteration,  $E_{smooth} = 0$  and  $E_{data} = m_x(u, v_{max})$ . After that,  $E$  is estimated based on the previous iterations:

$$E(u)_v = m_x(u, v) + \min_{\tau_x} [E(V_{px} + \tau_x)_{v+1} + \lambda_x \tau_x] \text{ s.t. } \tau_x \in [-5, 5] \quad (18)$$

The solution  $\mathbf{M}_x = [\mathbf{u}, \mathbf{v}] \in \mathbb{R}^{t \times 2}$  with the minimal energy is selected as the optima, and we plot it as a blue path in Fig. 6. There are  $t$  points on the path.  $\mathbf{u} = [u_0, u_1, \dots, u_{t-1}]^\top$  is a column vector storing their column numbers, and  $\mathbf{v} = [v_0, v_1, \dots, v_{t-1}]^\top$  is a column vector recording their row numbers. The parameters  $\boldsymbol{\gamma} = [\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4]^\top$  can be estimated by solving the least squares problem in Eq. 19. Here, we use the same strategy as the estimation of  $\boldsymbol{\beta}$ .  $\mathcal{I}$  and  $\mathbf{M}_x$  are updated using RANSAC until their capacities do not change any more. Then,  $\boldsymbol{\gamma}$  is estimated from  $\mathcal{I}$ . Algorithm 6 provides more details on the  $\boldsymbol{\gamma}$  estimation.

$$\boldsymbol{\gamma} = \arg \min_{\boldsymbol{\gamma}} \sum_{j=0}^{t-1} (u_j - (\gamma_0 + \gamma_1 v_j + \gamma_2 v_j^2 + \gamma_3 v_j^3 + \gamma_4 v_j^4))^2 \quad (19)$$

---

**Algorithm 6:**  $\boldsymbol{\gamma}$  estimation with the assist of RANSAC.

---

**Input :** The optimal solution  $\mathbf{M}_x$   
**Output:**  $\boldsymbol{\gamma}$

**1 do**

2 select a specified number of candidates  $[u_j, v_j]^\top$  randomly;

3 fit a quadruplicated polynomial using Eq. 19 to get  $\boldsymbol{\gamma}$ ;

4 determine the number of inliers  $\mathcal{I}$  and outliers  $\mathcal{O}$ :  $n_{\mathcal{I}}$  and  $n_{\mathcal{O}}$ ;

5 remove  $\mathcal{O}$  from  $\mathbf{M}_x$ ;

6 **while**  $n_{\mathcal{I}}/(n_{\mathcal{I}} + n_{\mathcal{O}}) < \epsilon_x$ ;

7 interpolate  $\mathcal{I}$  into a quadruplicate polynomial and get  $\boldsymbol{\gamma}$ ;

---

Given an observation data  $[u_j, v_j]^\top$ , to determine whether it is an inlier or outlier, the error  $r_j = (u_j - g(v_j))^2$  will be computed, where the function  $g(v) = \gamma_0 + \gamma_1 v + \gamma_2 v^2 + \gamma_3 v^3 + \gamma_4 v^4$  is the solution of Eq. 19. If  $r_j$  is smaller than our pre-set threshold  $tr_x$  ( $tr_x = 16$  in this paper), the candidate  $[u, v]^\top$  is marked as an inlier and saved in  $\mathcal{I}$ . Otherwise, it will be marked as an outlier  $\mathcal{O}$  and removed from  $\mathbf{M}_x$ . The iteration runs until the fraction of  $\mathcal{I}$  versus the capacity of  $\mathbf{M}_x$  exceeds our pre-set threshold  $\epsilon_x$  ( $\epsilon_x = 0.99$  in this paper). Finally,  $\boldsymbol{\gamma}$  can be estimated from  $\mathcal{I}$  by solving the energy minimisation problem in Eq. 19. Compared with the target solution obtained in [18], as shown in Fig. 6(b), the target solution in the proposed system is less affected by the outliers (please see Fig. 6(c)). In a practical implementation, we rearrange Eq. 19 as Eq. 20 to avoid the data overflow when fitting the quadruplicate polynomial.

$$\boldsymbol{\gamma} = (\kappa \mathbf{P}^\top \mathbf{P})^{-1} (\kappa \mathbf{P}^\top) \mathbf{u} \quad (20)$$

where

$$\mathbf{P} = \begin{bmatrix} 1 & v_1 & \cdots & {v_1}^4 \\ 1 & v_2 & \cdots & {v_2}^4 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & v_t & \cdots & {v_t}^4 \end{bmatrix} \quad (21)$$

$\mathbf{P}$  is a Vandermonde matrix.  $\kappa$  is used to avoid the data overflow problem caused by the higher order polynomials (e.g. when  $v = 375$ ,  $v^8 \approx 4 \times 10^{20}$  which is far beyond the significand range of *long double* type in C language). With  $\boldsymbol{\gamma}$  estimated, the dense  $V_{px}$  can thus be estimated for each row  $v$  with the known parameters  $\boldsymbol{\gamma}$  as follows:

$$V_{px}(v) = \gamma_0 + \gamma_1 v + \gamma_2 v^2 + \gamma_3 v^3 + \gamma_4 v^4 \quad (22)$$

## 2.4. Lane Position Validation

$\mathbf{V}_p$  provides the tangential direction and the curvature information of lanes, which can help us to validate the lane positions. In [18], we form a likelihood function  $V(e) = \nabla(e) \cdot \cos(\theta_e - \theta_{\mathbf{V}_p})$  for each edge point  $e$  to select the plus-minus peak pairs, where  $V(e)$  is the vote from  $e(u_e, v_e)$ .  $\theta_e$  is the angle between the  $u$ -axis and the orientation of  $e$  which is approximated by a Sobel edge detector.  $\theta_{\mathbf{V}_p}$  is the angle between the  $u$ -axis and the radial from an edge pixel  $e$  to  $\mathbf{V}_p(v_e)$ . Although some impressive experimental results have been achieved by the local peak-pair selection algorithm in [18], some inaccurate detections are still inevitable. In this paper, we verify every possible solution efficiently using a global histogram  $\mathcal{H}$  created by analysing the information of  $G_x$  and  $\mathbf{V}_p$ . Then, the favourable local minima are selected from  $\mathcal{H}$  to determine the lane positions. Algorithm 7 provides more details on the proposed approach.

---

**Algorithm 7:** Lane position validation.

---

**Input :**  $\mathbf{V}_p$  and  $G_x$   
**Output:** lane position vector  $\Gamma$

- 1 create two 2D maps  $\mathcal{M}_0$ ,  $\mathcal{M}_1$  with the same size of the input image  $I$ ;
- 2 set all elements in  $\mathcal{M}_0$ ,  $\mathcal{M}_1$  as 0;
- 3 create a 1D histogram  $\mathcal{H}$  of size  $(2\xi + 1)u_{max}$ ;
- 4  $\forall \mathcal{M}_0(i, j) \leftarrow \sum_{x=-\nu}^{x=+\nu} \sum_{y=-\varsigma}^{y=+\varsigma} G_x(i + x, j + y) w_g(i + x, j + y)$ ;
- 5  $\mathcal{M}_1 \leftarrow \mathcal{G} * \mathcal{M}_0$ ;
- 6 **for**  $i \leftarrow -\xi u_{max}$  to  $\xi u_{max}$  **do**
- 7     aggregate  $\mathcal{M}_1$  from row  $v_{max}$  to row  $f^{-1}(0)$  to get the energy  $E$ ;
- 8      $\mathcal{H}(i + \xi u_{max}) \leftarrow E$ ;
- 9 **end**
- 10 **if**  $\mathcal{H}(i) < \min\{\mathcal{H}(i - 1), \mathcal{H}(i + 1)\}$  **then**
- 11     put  $i - \xi u_{max}$  into  $\Gamma$ ;
- 12 **end**
- 13 remove the elements which are smaller than the threshold  $tr_{LPV}$  from  $\Gamma$ ;
- 14 remove the nearby candidates from  $\Gamma$ ;
- 15 multiple lanes visualisation;

---

For a dark-light transition of a lane marking,  $G_x$  is positive and higher than the values at the non-edge positions, while  $-G_x$  is positive and higher on a light-dark transition of the lane markings [18]. Therefore, our aim to validate a lane position turns out to be estimating the position of the middle between two adjacent peaks (dark-light transition and light-dark transition). In order to reduce the noise introduced from non-lane edges, we use a piecewise weighting  $w_g$  to decrease their magnitudes:

$$w_g(u_e, v_e) = \begin{cases} \exp\left(-\frac{|\theta_e - \theta_{\mathbf{V}_p}|}{\sigma_g^2} \cdot \frac{36}{\pi}\right), & |\theta_e - \theta_{\mathbf{V}_p}| \leq \frac{\pi}{6} \\ 0, & \text{otherwise} \end{cases} \quad (23)$$

where we define the step as  $\theta_s = \pi/36$ . The portion  $p = |\theta_e - \theta_{\mathbf{V}_p}|/\theta_s$  is used to provide a Gaussian weight so as to decrease the magnitude of noise pixels, where we set  $\sigma_g = 3.5$ . Then, we accumulate the updated  $G_x$  within a shifting box to further reduce the noise. The box size is  $(2\nu + 1) \times (2\varsigma + 1)$ , where  $\nu = 1$  and  $\varsigma = 3$  in our experiment. The accumulation output is saved in a 2D map  $\mathcal{M}_0$ , where the horizontal gradient from a dark-light peak to a light-dark peak is negative. To approximate the horizontal gradients, a convolution operation between the Sobel kernel  $\mathcal{G} = [1, 2, 1]^\top [1, 0, -1]$  and  $\mathcal{M}_0$  is conducted and the output is stored in  $\mathcal{M}_1$ . Then, we aggregate the cost  $\mathcal{M}_1$  for each possible position from row  $v_{max}$  to row  $f^{-1}(0)$ :

$$E(v)_{u_v} = \mathcal{M}_1(u_v, v) + \lambda_g E(v+1)_{u_{v+1}} \quad (24)$$

where

$$u_v = \frac{V_{px}(v+1) + vu_{v+1} - V_{py}(v+1)u_{v+1}}{v+1 - V_{py}} \quad (25)$$

In order to cover all possible solutions, we set  $\xi = 0.5$ . This implies that  $u$  starts from  $-0.5u_{max}$  and ends at  $1.5u_{max}$ . In the first iteration, we select a possible position  $(u_{v_{max}}, v_{max})$  and the total energy is simply represented as:  $E = \mathcal{M}_1(u_{v_{max}}, v_{max})$ . Then, we use  $\mathbf{V}_p$  to estimate the next position  $(u_{v_{max}-1}, v_{max} - 1)$  on the selected track. The energy  $E$  is updated using Eq. 24, where  $\lambda_g$  is proposed to be 1. The aggregation of  $\mathcal{M}_1$  works until  $v$  reaches  $f^{-1}(0)$ . For each lane starting from  $(u_{v_{max}}, v_{max})$ , its total energy is saved in a 1D histogram  $\mathcal{H}$ . Then, we pick out the local minima which are smaller than our pre-set threshold  $t$  from  $\mathcal{H}$ . At the same time, if two local minima are quite close to each other, we ignore the minima with a larger energy. Finally, the lanes can be visualised by iterating Eq. 25. The lane detection results will be discussed in section 3.2.

### 3. Experimental Results

#### 3.1. Stereo vision evaluation

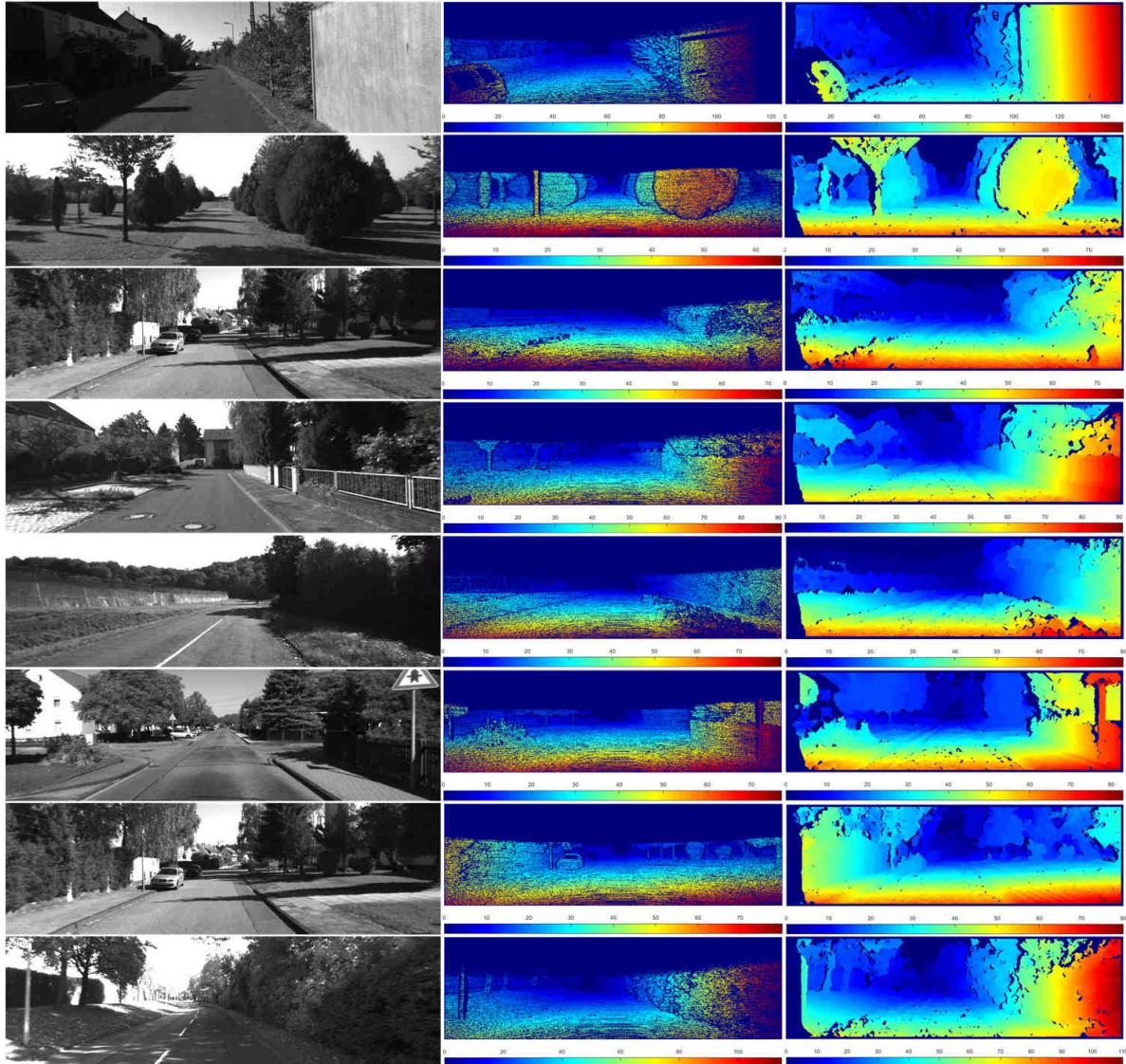
The proposed disparity estimation algorithm is evaluated using the KITTI stereo 2012 dataset [30]. Some experimental examples are shown in Fig. 7, where the first column illustrates the input left image, the second column shows the ground truth of the disparity map, and the third column depicts our experimental results. The overall percentage of the error pixels is approximately 6.82% (error threshold: 2 pixels).

The proposed disparity estimation algorithm is implemented in C language on an i7-4720HQ CPU and a NVIDIA GTX 970M GPU for real-time purpose. Please kindly refer to [21] for more details on the implementation. When  $\rho = 3$  and  $\tau = 1$ , the algorithm execution achieves a speed of 37 fps on GPU. After the memorisation, the values of  $\mu$  and  $\sigma$  are computed and stored in a static program storage for direct indexing in stereo matching, which greatly boosts the algorithm execution. The performance improvement

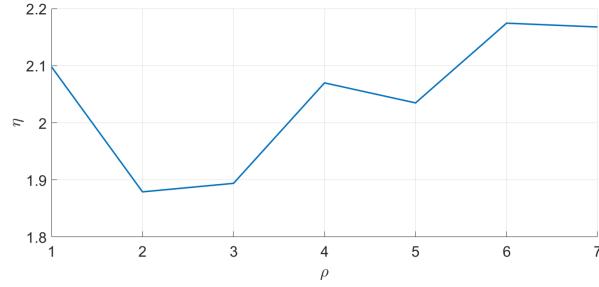
achieved from memorisation is depicted in Fig. 8, where  $\eta$  represents the runtime of prevalent NCC-based stereo versus the runtime of NCC-based stereo with memorisation. In Fig. 8, it can be clearly seen that the memorisation speeds up the algorithm execution by about two times.

### 3.2. Lane detection evaluation

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 [31]. Therefore, many publications for lane detection only focus on the quality of their experimental results. For this reason, we compare the performance of the proposed system with our previous work in [18] in terms of both accuracy and speed. Some successful detection examples are shown in Fig. 9.

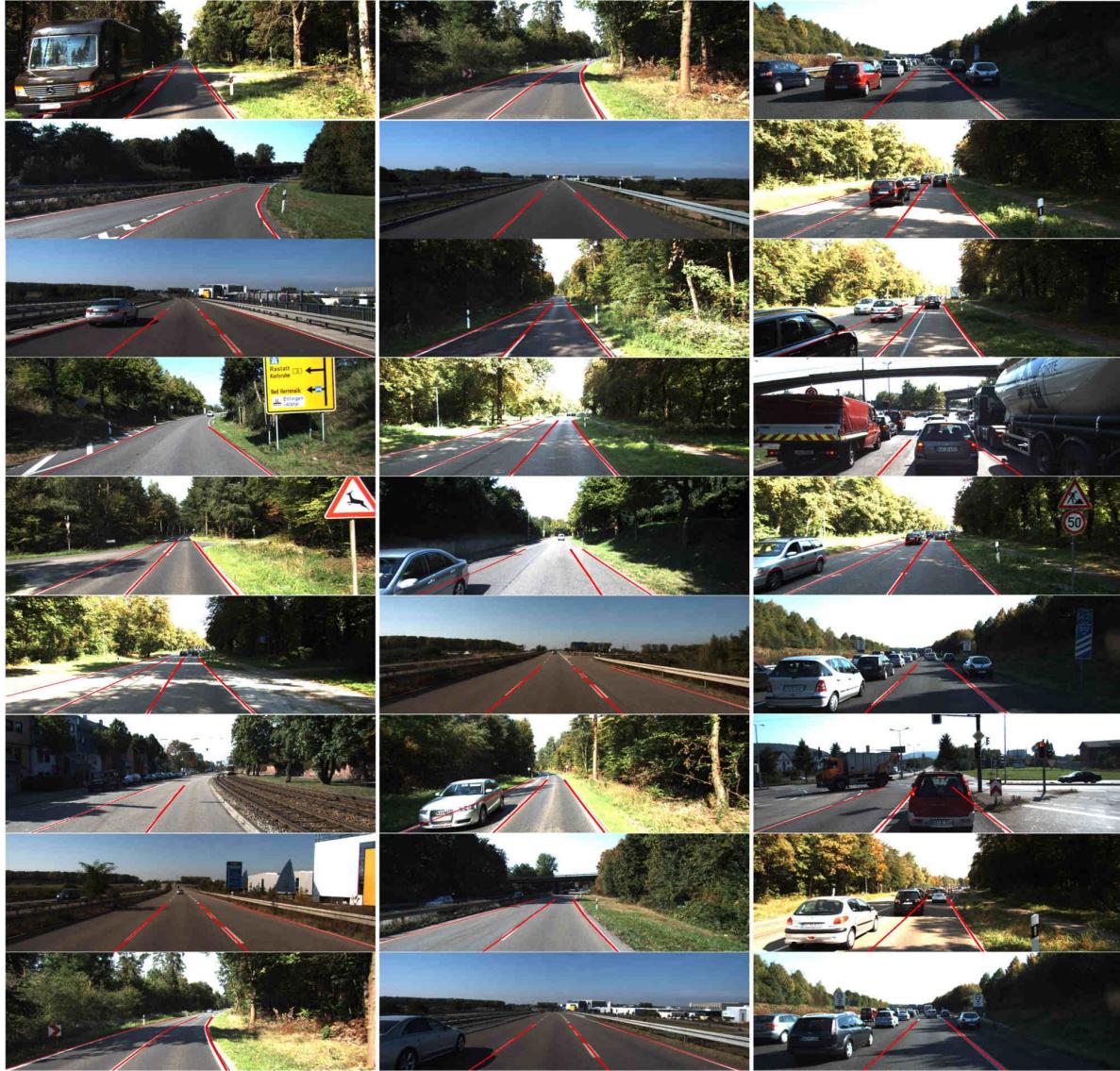


**Figure 7.** Experimental results of the KITTI stereo 2012 dataset.



**Figure 8.** Evaluation of  $\eta$  when using different  $\rho$ .

Firstly, we will discuss the precision performance. The lane detection ratio of the proposed algorithm and our previous work [18] are detailed in tables 1 and 2,



**Figure 9.** Lane detection experimental results. The red curves are the detected lanes.

**Table 1.** 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                   | 0            |
| 5        | 678   | 0                   | 0            |
| 6        | 1060  | 1                   | 2            |
| 7        | 644   | 0                   | 0            |
| 8        | 993   | 18                  | 7            |
| Total    | 5361  | 19                  | 9            |

**Table 2.** Detection results of [18].

| 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           |

respectively. To evaluate the robustness of the proposed algorithm, we tested eight sequences (including two additional sequences): 2495 frames containing 5361 lanes from KITTI database [32] (1637 more lanes than the experiment in [18]). The image resolution is  $1242 \times 375$  in sequences 1 to 6,  $1241 \times 376$  in sequence 7, and  $1238 \times 374$  in sequence 8. From table 1, it can be seen that the proposed algorithm presents a better detection ratio where 99.9% lanes are successfully detected in sequences 1 to 7 (including all the sequences in table 2), while the detection ratio of [18] is 98.7%. The comparison between some failed examples in [18] and their corresponding results of the proposed system is illustrated in Fig. 10.

In Fig. 10(a), we can see that the obstacle areas occupy a larger portion than the road surface area, which severely affects the accuracy of  $V_{py}$ . When we consider both inliers and outliers into the LSF,  $\mathbf{V}_p$  differs too much from the ground truth. This further influences the peak-pair selection and leads to an imprecise detection and a misdetection. When reflecting on Fig. 10(b), it can be seen that the fitting with only inliers increases the precision of the  $V_{py}$  estimation significantly. Moving to the second row, an over-curved lane can be seen in Fig. 10(c). When we fit  $\beta$  and  $\gamma$  with the assist of RANSAC, the improvement can be observed in Fig. 10(d), where a more reasonable lane is detected. In Fig. 10(e), the lane near the left road boundary is misdetected because the low contrast between lane and road surface reduces its magnitude in the



**Figure 10.** Comparison between some failure examples in [18] and the corresponding results in this paper. The green areas in the first column are the road surface. The red curves are the detected lanes. The first column illustrates the failed examples in [18], and the second column shows the corresponding results of the proposed system.

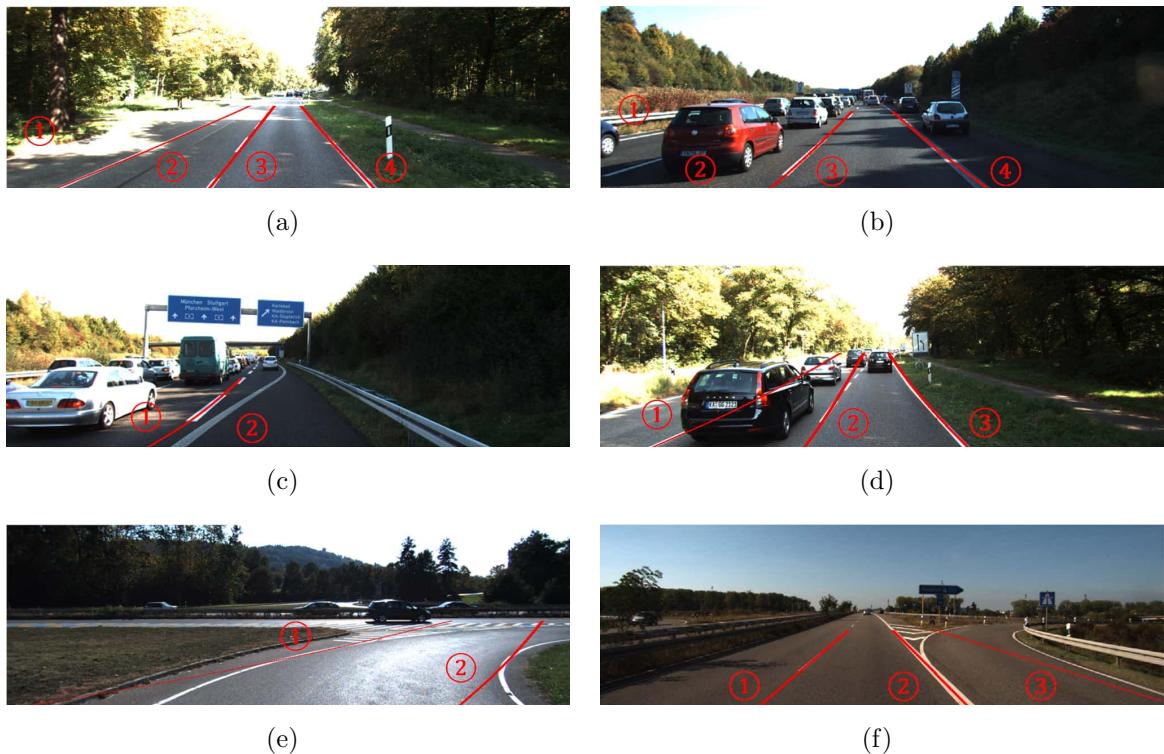
peak-pair selection stage. In Fig. 10(g), an incorrect detection occurs because the magnitude of a plus-minus peak of road markings is higher than the magnitude of lanes. In section 2.4, we proposed a more effective piecewise weighting  $w_g$  to update  $G_x$  for the edge pixels. Then,  $G_x$  of the non-lane edges reduces significantly, which therefore greatly helps us avoid the incorrect detection of some lane markings. Also, we sum  $G_x w_g$  within a shifting box for each position, which increases the magnitude of the lanes which are lowly contrastive to the road surface. The misdetections in Fig. 10(e) and 10(g) are thus detectable, and the failed detection in 10(g) is also corrected. The corresponding results of the proposed system are shown in Fig. 10(f) and Fig. 10(h).

In our experiment, the failed cases consist of misdetections and incorrect detections. The misdetections are mainly caused by: the over-exposure of the image, partially-occluded by the obstacles, forks on the road. The corresponding examples are illustrated in Fig. 11(a), 11(b) and 11(c), respectively. In Fig. 11(a), due to the image over-

exposure, the edges pixels on lane 1 are rare, which leads to its misdetection. In Fig. 11(b), we can see that the vehicles partially occlude lane 1 and lane 2. The occlusion decreases the magnitudes of  $G_x w_g$  when we try to validate the lane positions, which makes lane 1 and lane 2 undetectable. In 11(c), lane 2 forks from lane 1, and thus, has a different curvature information from lane 1, which therefore causes the misdetection.

For the factors leading to the incorrect detections, we group them into three main categories: ambiguous projection of road surface on the v-disparity map;  $\mathbf{V}_p$  that does not exist in the image; different roadways, which are presented in Fig. 11(d), 11(e) and 11(f), respectively. In Fig. 11(d), the obstacles (i.e. vehicles, trees and signageS) take a big portion in the image. Therefore, when  $d$  is around 0,  $m_y$  is mainly accumulated by the pixels on the obstacles and sky, which makes  $V_{py}$  inaccurate and lanes 1 to 3 are slightly above the ground truth when the lane moves to the boundary between the road surface and sky. In Fig. 11(e),  $\mathbf{V}_p$  does not exist, which affects the detection results. In Fig. 11(f), there are two different roadways: roadway between lane 1 and lane 2; roadway between lane 2 and lane 3. The second roadway turns right and therefore has a different  $\mathbf{V}_p$  from the first roadway, which leads to an imprecise detection of lane 3.

Next, we will discuss about the speed performance. The algorithm is implemented on a heterogeneous system consisting of an Intel Core i7-4720HQ CPU and an NVIDIA GTX 970M GPU. The GPU has 10 streaming multiprocessors (SMs) with 128 CUDA cores on each of them. In the sparse  $V_{px}$  estimation stage, the optimisations of the GPU memory allocation make this stage yield a performance speed-up of over 74



**Figure 11.** Examples of the failed detection in the proposed algorithm.

times than the implementation on a CPU with single thread processing. The total runtime of the proposed system is around 7 ms, which is approximately 38 times faster than our previous work where 263 ms was achieved (excluding the runtime of the disparity estimation). The authors believe that the failed cases can be prevented in the future with the addition of a lane tracking algorithm. A sample video is available at <https://www.youtube.com/watch?v=fgrUdy1kv0>.

#### 4. Conclusion

A multiple lane detection system was presented in this paper. The novelties in this paper include: improved dense  $\mathbf{V}_p$  estimation using RANSAC, image pre-processing with the bilateral filtering, lane position validation with  $G_x$  and  $\mathbf{V}_p$ , and a real-time implementation on a CPU-GPU-based heterogeneous system. To evaluate the performance, 5361 lanes from eight datasets were tested. The experimental results illustrated that the proposed algorithm works more accurately and robustly than our previous work with a 99.5% successful detection ratio achieved. By highly exploiting the GPU architecture and allocating different parts on different platforms for execution, a high processing speed of 143 fps is achieved, which is approximately 38 times faster than our previous work in [18].

- [1] Juan Rosenzweig and Michael Bartl, “A review and analysis of literature on autonomous driving,” *E-Journal Making-of Innovation*, 2015.
- [2] Sandipann P Narote, Pradnya N Bhujbal, Abhilasha S Narote, and Dhiraj M Dhane, “A review of recent advances in lane detection and departure warning system,” *Pattern Recognition*, vol. 73, pp. 216–234, 2018.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] Umar Ozgunalp and Naim Dahnoun, “Robust lane detection & tracking based on novel feature extraction and lane categorization,” in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 8129–8133.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.

- [11] 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.
- [12] Yue Wang, Dinggang Shen, and Eam Khwang Teoh, "Lane detection using spline model," *Pattern Recognition Letters*, vol. 21, no. 8, pp. 677–689, 2000.
- [13] 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.
- [14] 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.
- [15] David Hanwell and Majid Mirmehdi, "Detection of lane departure on high-speed roads.," in *ICPRAM (2)*, 2012, pp. 529–536.
- [16] 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.
- [17] Yifei Wang, Naim Dahnoun, and Alin Achim, "A novel system for robust lane detection and tracking," *Signal Processing*, vol. 92, no. 2, pp. 319–334, 2012.
- [18] 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.
- [19] 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*, 2002, vol. 2, pp. 646–651.
- [20] Zhen Zhang, Xiao Ai, and Naim Dahnoun, "Efficient disparity calculation based on stereo vision with ground obstacle assumption," in *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*. IEEE, 2013, pp. 1–5.
- [21] 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.
- [22] Naim Dahnoun, "Stereo vision implementation," *Multicore DSP: From Algorithms to Real-time Implementation on the TMS320C66x SoC*, pp. 604–616.
- [23] John P Lewis, "Fast template matching," in *Vision interface*, 1995, vol. 95, pp. 15–19.
- [24] Rui Fan, Xiao Ai, and Naim Dahnoun, "Road surface 3d reconstruction based on dense subpixel disparity map estimation," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3025–3035, 2018.
- [25] Zhencheng Hu, Francisco Lamosa, and Keiichi 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*. IEEE, 2005, pp. 204–211.
- [26] Dana H Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern recognition*, vol. 13, no. 2, pp. 111–122, 1981.
- [27] C Rafael Gonzalez and Richard Woods, "Digital image processing," *Pearson Education*, 2002.
- [28] Kaiming He, Jian Sun, and Xiaou Tang, "Guided image filtering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.
- [29] NVIDIA, "Cuda c programming guide," September 2017.
- [30] A Andreas, Philip Lenz, and Raquel Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2012.
- [31] 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.
- [32] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun, "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.