Fast Nonparametric Road Disparity Estimation and Gradient Constrained Obstacle Detection for UGV Navigation

Zhihao Zeng, Tingbo Hu, Xiangjing An
College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha, Hunan 410073, China
zengzzz@126.com

Abstract. Road estimation and obstacle detection are two key techniques for autonomous driving vehicles, requiring high accuracy, efficiency and robustness. In this paper, we propose an efficient nonparametric road disparity estimation algorithm that makes full use of the characteristics of the road disparity. It can adapt to both planar and nonplanar road. In addition, to satisfy the real-time requirements of the application, we exploit a new texture gradient based constraint to compute the free space and detect the obstacle, which will also help to obtain more accurate results. Experimental results with real stereo sequences show that the proposed algorithm outperforms the standard methods both in terms of detection accuracy and runtime performance, which can run over 200 fps with normal PC.

1. Introduction
The road estimation task is the key step toward the advanced driver assistance systems (ADAS) or unmanned ground vehicle (UGV). The free space refers to the world regions without any collision for the application of UGV navigation. To compute the free space faces many challenges in realistic road scenarios because the road can be planar or nonplanar (e.g., uphills, downhills, and undulating roads). In binocular system, we usually assume the road model first, then fit the model parameters to obtain the road disparity, and compute the free space using the disparity image. Existing road estimation methods generally fit the road surface into rigid body model (e.g., planar [1], clothoid [2], or B-Spline [3]). However, specific model cannot adapt to a variety of complex reality environment. Then it will seriously affect the subsequent free space computation and obstacle detection. Reference [4] proposed the first nonparametric road disparity estimation method to adapt to a variety of different roads but the process was too complicated.

This paper makes full use of the characteristics of road disparity, and presents a more effective nonparametric method to obtain the road disparity, which does better both in estimation accuracy and runtime performance.

In many obstacle detection algorithms, the stixel [5] based method has received great attention. Stixel world is composed of the ground plane and obstacles like vertical sticks standing in the ground, called stixel (stick pixel). Therefore, the stixel based method mainly contains two computational procedures. First is to compute the free space, second is to extract the stixel. The free space can be computed using the probabilistic occupancy map [5][6] or directly from the disparity image [7][8]. The drawback of the original stixel world framework is that it depends on the disparity image of a certain quality too much. Wrong disparities in cases of occlusions, reflections or image regions with little texture information will result to many false obstacles. Therefore, researchers tried to fuse the color information into the original framework [9][10].
Rather than using the color information, this paper exploits the gradient information from the gray image to constrain the free space computation and the obstacle detection, which reduces the computation time significantly. An obstacle detection result in highway situation is shown in Figure 1. The colors contain the distance information, red means close while blue means distant. Fast nonparametric road estimation ensures the quality of the road disparity. Then gradient constrained method extracts the contours of the road and the obstacles accurately, even in the distance.

This paper is organized as follows. Section II describes the stereo geometry, the characteristics of road disparity and the proposed fast nonparametric road disparity estimation algorithm. Section III introduces the proposed gradient constrained free space computation and the obstacle detection. Evaluation and conclusions are discussed in Sections IV and V, respectively.

Figure 1  Obstacle detection results

2. Fast Nonparametric road disparity estimation
In this section, several important characteristics of road disparity will be discussed. Based on these characteristics, the proposed fast nonparametric road estimation algorithm will be introduced in detail.

2.1. Road Disparity Characteristics
In the camera coordinate system, the position of a scene point $P(X,Y,Z)$ in the image plane is given by its coordinates $(u,v)$. Assume that $\theta$ is the cameras’ pitch angle; $h$ is the height of the cameras above the ground; $b$ is the distance between the cameras (i.e. the baseline); $f$ is the focal length measured in pixels; $(u_0,v_0)$ is the image coordinate of the projection of the camera optical center; $i$ indicates left camera or right camera, and $\epsilon_i = -1$, $\epsilon_r = -1$. Then the coordinate $p(u,v)$ is given as

\[ u_i = \frac{fX - \epsilon_i \frac{fb}{2}}{(Y+h)\sin \theta + Z \cos \theta} + u_0 \]

\[ v = \frac{f(Y+h)\cos \theta - fZ \sin \theta}{(Y+h)\sin \theta + Z \cos \theta} + v_0 \]

The disparity value $d = u_i - u_r$ is given in

\[ d = \frac{fb}{(Y+h)\sin \theta + Z \cos \theta} \]

For the planar road under the vehicle, it can be given as $Y = 0$, then using (1) and (2), a relationship between $d$ and $v$ is given as
\[ d = \frac{b(v-v_o) \cos \theta + fb \sin \theta}{h} \]  

(3)

Reference [1] proposed the concepts of v-disparity image and u-disparity image first. For the v-disparity image, the value on coordinate \((d, v)\) is obtained by accumulating the points with the same disparity \(d\) in the scan line \(v\) of the disparity map. Similarly, the u-disparity image is obtained by accumulating the points with the same disparity in a column-wise manner.

Equation (3) is a straight-line equation. That means the road given as \(Y = 0\) is an oblique line in v-disparity image. Its slope and intercept are denoted as \(k_l\) and \(b_l\), then the cameras’ pitch angle \(\theta\) and height \(h\) can be computed by (3). By observing the road in the image and analyzing (3), we can infer the following road characteristics in v-disparity image:

1. The farther from the camera the road, the smaller the disparity. That means \(d\) decreases with the decrease of \(v\) no matter whether it is a planar or nonplanar road. As for the planar road, the decrement per row derived by (3) is denoted as

\[ \Delta d = k_l \]  

(4)

2. On each row, the road disparity is not larger than any other objects except for the holes because the road extends forward while the objects grow upward. Therefore, on the ideal road, on each row in v-disparity image, the disparity (i.e. the column coordinate) of first non-zero point usually belongs to the road.

3. On normal cases, especially at the bottom of the image, the road surface is the major component of the scene. Also, in the road environments, the road disparity values usually are concentrated while other objects have different disparity values. Therefore, on each row in v-disparity image, the disparity of maximum value point usually belongs to the road.

2.2. Fast Nonparametric Road Estimation

The strategy of our algorithm is to find out the best disparity for the road based on the above characteristics. We utilize the third characteristic to choose the candidate disparity of the road. Then we utilize the second characteristic to confirm whether the candidate disparity belongs to the road or not. Once we have confirmed the road disparity on one row, then the road disparity above the row can be looked for based on the first characteristic. For subpixel precision, each disparity value will be interpolated.

The detailed fast nonparametric road estimation algorithm is shown as below:

1. First step, choose the candidate road disparity in the whole row. Start from the last row of the v-disparity image to up, find the maximum value of the current row, and the corresponding disparity is regarded as the road disparity because the disparity of maximum value point usually belongs to the road, based on the third characteristic. Then we also need to check whether the maximum value is also far larger than its all left values (i.e. larger than 10 times) or not, except for its left neighbors to resist noise. Once it satisfies the conditions, the road disparity is confirmed because the road disparity is not larger than any other objects, based on the second characteristic. Then, record the disparity and its repeat times \(N_{rep}\) is one.

2. Second step, look for the road disparity in local area. Before this step, we need to obtain a parameter in advance to help to look for the rest of the road disparity. Compute \(\Delta d\) in the v-disparity image using Hough transform. Then, obtain the maximum number of repeated rows in v-disparity image, which is only needed to compute in the first frame.

\[ N_{\text{max}} = \frac{k}{\Delta d} \]  

(5)
\( N_{\text{max}} \) is used to constrain the road disparity. Although the original disparity value contains a decimal, we regard them as an integer by rounding to generate v-disparity image. As a result, the value of road disparity may repeat in successive rows. The inverse of \( \Delta d \) from (4) means maximum number of repeated rows under normal circumstances. We raise this parameter by a multiplication factor \( k \) to adapt to undulating road conditions. When \( k \) equals 1.5, it is enough to adapt to most cases. If we don’t limit the number of repeated rows, the same disparity may repeat many times. That means we may mistake an obstacle as the road because the obstacle in v-disparity image is a vertical line, especially in the distance where the obstacles occupy more pixels than the road.

Now, start the second step, go to the next row to look for the rest of the road disparity based on the first characteristic. If \( N_{\text{rep}} \) is smaller than \( N_{\text{max}} \), find the maximum value among the position of the confirmed disparity (i.e. the road disparity of last row) and its two left neighbors. If \( N_{\text{rep}} \) is not smaller than \( N_{\text{max}} \), it means we may encounter an obstacle so we need to find the maximum value among the three left neighbors. Then, the corresponding disparity of the maximum value point is regarded as the confirmed road disparity of current row. If the new disparity is the same as the last one, \( N_{\text{rep}} \) adds one. If not, set \( N_{\text{rep}} \) as one. Keep doing this step until the maximum value becomes zero or the road disparity becomes zero, which means we meet the horizon.

Because the road disparity will not be sharp declines, we search for the road disparity on new row only in three near positions in case we mistake the deep holes or low-lying places even cliffs as the road.

3. Third step, subpixel refinement. We use linear interpolation given as below to obtain subpixel precision.

\[
ml = \max(0, m-l) \\
mr = \begin{cases} 
\max(0, m-r), & \text{if } N_{\text{rep}} < N_{\text{max}} \\
m, & \text{others}
\end{cases}
\]

\[
d_{\text{new}} = \begin{cases} 
\frac{d - 0.5 + ml}{2mr}, & \text{if } ml < mr \\
\frac{d + 0.5 - mr}{2ml}, & \text{if } ml > mr \\
d, & \text{others}
\end{cases}
\]

In the above equation, \( m, ml \) and \( mr \) indicates the maximum value, the value of its left neighbor and right neighbor respectively. \( d \) and \( d_{\text{new}} \) is the confirmed road disparity and new subpixel disparity respectively.

Reference [4] proposed the similar nonparametric algorithm. The main difference is that we simplify and optimize the disparity extraction in the v-disparity image so that we don’t need to remove the obvious obstacle through the u-disparity image, which saves lots of time. Besides, we use subpixel refinement to get higher accuracy.

3. Gradient Constraint Obstacle Detection
Stixel based obstacle detection method mainly contains two computational procedures: free space computation and stixel extraction [5][7]. The first is used to find the base point (or foot point) of the obstacle (i.e. the stixel) on the ground, the second is aimed to find the stixel’s top point. These two steps both need to calculate the cost of every pixel in the disparity image and accumulate them column-wise, then solve a dynamic programming problem to get a more accurate and smooth result. It
provided a computing framework, but the cost function is not so accurate and efficient. Besides, it considered all the candidate points and caused too much computation.

In this section, we will first introduce how to exploit the gradient information of the candidate points to accelerate the algorithm, and then propose new evidence functions for the two procedures.

3.1. Gradient Constraint

As we know, edge is the most attractive visual information. It is usually used to detect the object contours and distinguish the road, foreground and background. Similarly, the obstacle’s base point on the ground and its top point usually have a significant edge, especially horizontal edge.

Therefore, when detecting obstacle’s base point or top point, we can consider the point whose vertical gradient is larger than its neighbors’ in the vertical direction as a candidate. In this paper, the disparity image is divided into several horizontal stripes and inside each stripe, for each image column, the pixel with the maximal vertical gradient is selected as the candidate pixel \( p_{u,s} \) for the base point or the top point, the subscript \( u \) indicates column number, and \( s \) indicates stripe number. Besides, because the candidate pixel has maximum vertical gradient in local area rather than whole area, there is still a great possibility of belonging to the road. Therefore, we also fuse the gradient information into the cost function to measure the possibility of being a basepoint.

Reference [11] had used this technique to speed up the free space computation. We extend it to the obstacle detection and more importantly, we fuse the gradient information into the cost function to get more accurate results.

3.2. Free Space Computation

We obtain the base point scores directly from the disparity image. For all selected candidate point \( p_{u,s} \), a score \( c_{u,s} \) is computed, containing three sub-scores: the first for road evidence \( c_r \), the second for obstacle evidence \( c_o \) with penalty coefficient \( \alpha \), the last for gradient constraint \( c_g \) with penalty coefficient \( \beta \) and gradient threshold \( \gamma \).

\[
C_{u,s} = C_r + \alpha C_o + \beta C_g
\]

\[
C_r = \sum_{v=v_b}^{V} | d_r(v) - d_v |
\]

\[
C_o = \sum_{v=v_b-h_0}^{V} | d_o(v) - d_v |
\]

\[
C_g = \max(0, gra(v_b) - \gamma)
\]

In the above equation, \( v \) indicates the row coordinate; \( v_b \) is the row coordinate of the candidate point; \( V \) is the height of the image; \( d_v \) is the disparity value of current point; \( d_r(v) \) is the road disparity value of current row; \( gra(v_b) \) is the vertical gradient of the candidate point; \( h_0 \) corresponds to the height in pixels of an assumed obstacle standing on the ground (i.e. row \( v_b \)) with a minimum limited value, computed by stereo geometry each row.

Compared with [7], we add a gradient constraint term. It can effectively reduce the possibility of mistaking the road as an obstacle.

These scores are then used in a dynamic programming scheme to extract the optimal free space path cutting the image from left to right with a smoothness constraint. The goal is to find the optimal
horizontal stripe for each stixel.

\[ s^*(u) = \min_{i(u)} \left( \sum_u C_{u,v} + \sum_{u-1,u} S(p_{u-1,s}, p_{u,s}) \right) \]  

(8)

In (8), \( S(p_{u-1,s}, p_{u,s}) \) corresponds to the smoothness constraint between the last base point \( p_{u-1,s} \) and the current base point \( p_{u,s} \). It can be the punishment for the jump between disparities or pixels locations.

3.3. Stixel Extraction

This step is used to determine the top points of all stixels. For all points above the base point, its membership \( M_{u,v} \) to the foreground object is computed.

\[ \Delta D = d_{v_{1}} + f_{d}(Z_{v_{1}} + \Delta Z) \]

\[ M_{u,v} = \max(-1,1 - \text{abs}(d_{v} - d_{v_{1}}) / \Delta D) \]  

(9)

In (9), \( d_{v_{1}} \) indicates the disparity of base point. Function \( f_{d} \) is the mapping of depth onto disparity using stereo geometry. \( Z_{v_{1}} \) corresponds to the depth of the base point and \( \Delta Z \) is the depth difference threshold.

Compared with [5], we modify the membership function to have a more efficient calculation while maintaining its performance.

Then, membership values are used to compute the cost \( C_{u,s} \) for all candidate points:

\[ C_{u,s} = \sum_{v=0}^{v_{t}-1} M_{u,v} - \sum_{v=v_{t}}^{v_{t}} M_{u,v} \]  

(10)

In (10), \( v_{t} \) indicates the row coordinate of candidate point. These costs are then used in a dynamic programming scheme to find the optimal horizontal stripe for each stixel like in (8). The smoothness term is defined as [5].

4. Evaluation

We have evaluated our algorithms on two stereo datasets compare with [4] and [7]. The first dataset is 333*1024 highway sequences with ground truth, some with heavy rain [12], marked with “highway”. The second dataset is the challenging 500*1000 campus sequences with undulating roads and complex environment, marked with “campus”. The datasets will be computed using SGM [13] to obtain the disparity images first. We use an Intel Core i7-5500U CPU and multithread technology for part of the algorithms, also for the compared algorithms.

4.1. Road Estimation Experiment

The first experiment is to evaluate the proposed fast nonparametric road disparity estimation (FNRDE), compared with the Hough transform (HT) method and algorithm in [4] (NT). We define the average road disparity error (ARDE) as the evaluation index.

\[ ARDE = \frac{1}{V - v_{h}} \sum_{v=v_{h}}^{V} (d_{e}(v) - d_{RGT}(v)) \]  

(11)

In (11), \( v_{h} \) denotes the row coordinate of the horizon (i.e. the end of the road). \( d_{e} \) denotes the estimated disparity while \( d_{RGT} \) denotes the ground truth disparity.
As shown in Table 1, the proposed FNRDE outperforms two standard methods in detection accuracy and runtime performance whether it is in highway or campus. It runs much faster than NT because of the more efficient procedure, without removing the obvious obstacle through the u-disparity image. Although, HT is designed for planar road, it doesn’t show better performance in highway, and worse in complex campus. Because the seemingly planar road may be nonplanar in disparity image especially when the light condition is poor or the stereo algorithm is not enough robust. Therefore, the nonparametric method is more common in use.

4.2. Free Space Computation Experiment
The second experiment is to evaluate the free space accuracy of the proposed gradient constrained stixel method (GCS), compared with the standard stixel algorithm in [7] (SS). The main parameters are set as below: the stripe height is 9 for the base point candidates, and 7 for the top point candidates. We use the F1-Measure value as the evaluation index.

\[
F1 = 2 \times \frac{PRE \times REC}{PRE + REC} \\
PRE = \frac{TP}{TP + FP} \\
REC = \frac{TP}{TP + FN}
\]  

(12)

In (12), F1 denotes the F1-Measure value; PRE is the precision rate; REC is the recall rate; TP, FP and FN is the true positive rate, false positive rate and false negative rate, respectively. The index is only counted in highway datasets.

As seen in Table 2, the FNRDE method and GCS method both can significantly reduce the running time and improve the performance compared with the HT method and the SS algorithm. The running time here contains three parts: the processing time of road disparity estimation, free space computation and stixel extraction. Fig. 2 shows the free space computation results in highway situations. The GCS method is more competent to approach the bottom of obstacles than the SS method but it is also more likely to produce image aliasing near the guard bars since the candidate points between columns are discrete and jumping. In addition, the FNRDE method can extract the road disparity more accurately, especially in the distance than the HT method. Fig. 3 shows the free space computation result in undulating campus environment. In the undulating crossroads, only the FNRDE method can accurately detect the free space.

4.3. Stixel Extraction Experiment
The third experiment is to evaluate the stixel extraction accuracy of the proposed gradient constrained stixel method (GCS), compared with the standard stixel algorithm in [7] (SS). We evaluate the extraction results using the annotated bounding boxes of the stixel, only in highway datasets.

A result of obstacle detection in highway situation is shown in Fig. 1. The stixels bounding box error is shown in Fig. 4. The GCS method also does better in the bounding box accuracy than the SS method. It proves the beneficial of the gradient constraint method, which can both improve the running time and detection accuracy. Besides, the FNRDE method outperforms the HT method because it can estimate the road disparity more accurately.

| Table 1 | Road Estimation Results |
|---------|-------------------------|
|         | Campus | Highway |
|         | HT     | NT     | Prop. FNRDE | HT     | NT     | Prop. FNRDE |
| ARDE    | 0.51   | 0.27   | 0.12       | 0.40   | 0.25   | 0.14       |
| Runtime (ms) | 7.62  | 9.11   | 2.05       | 6.19   | 7.20   | 1.66       |
Table 2  Free Space Estimation Results

|          | HT+SS | HT+GCS | FNRDE+GCS |
|----------|-------|--------|-----------|
| PRE (%)  | 89.86 | 89.58  | 89.93     |
| REC (%)  | 97.42 | 98.50  | 98.58     |
| F1 (%)   | 93.46 | 93.81  | 94.04     |
| Runtime  | 23.59 | 10.76  | 4.82      |

Figure 2  Free Space Computation Experiment Results In Highway

Figure 3  Free Space Computation Experiment Results In Campus
5. Conclusion

In this paper, we have proposed a new approach for road estimation. It can be applied to both planar and nonplanar road, with more accurate results and fast running speed. As for the closely related obstacle detection, we have also improved the original algorithm both in accuracy and running time, supported by our experiments results.

The frame rate of the entire system including road estimation, free space computation and obstacle detection is up to more than 200 frames per second, which leaves enough time for other applications and is more likely to be integrated in embedded systems.

The future work is to transplant the algorithms to FPGA based embedded systems for the real-time applications of ADAS and autonomous driving vehicles.

References

[1] R. Labayrade, D. Aubert, and J. P. Tarel, “Real time obstacle detection in stereovision on non flat road geometry through ‘v-disparity’ representation,” in Proc. IEEE Intell. Veh. Symp., 2002, vol. 2, pp. 646–651.

[2] S. Nedevschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, T. Graf, and R. Schmidt, “High accuracy stereovision approach for obstacle detection on non-planar roads,” in Proc. IEEE INES, Cluj Napoca, Romania, 2004, pp. 211–216.

[3] A. Wedel, H. Badino, C. Rabe, H. Loose, U. Franke, and D. Cremers, “B-spline modeling of road surfaces with an application to freespace estimation,” Transactions on Intelligent Transportation Systems, vol. 10, pp. 572 – 583, 2009.

[4] M. Wu, S.-K. Lam, and T. Srikanthan, “Nonparametric Technique Based High-Speed Road Surface Detection,” IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 2, pp. 874–884, 2015.

[5] H. Badino, U. Franke, and D. Pfeiffer. The stixel world - a compact medium level representation of the 3d-world. In German Association for Pattern Recognition (DAGM), pages 51–60, Jena, Germany, September 2009.

[6] Badino, H., Franke, U., Mester, R.: Free space computation using stochastic occupancy grids and dynamic programming. In: Workshop on Dynamical Vision, ICCV, Rio de Janeiro, Brazil (October 2007).

[7] D. Pfeiffer and U. Franke, “Efficient representation of traffic scenes by means of dynamic stixels,” in IV, 2010, pp. 217–224.

[8] S. Kubota, T. Nakano, and Y. Okamoto, “A global optimization algorithm for real-time on-board stereo obstacle detection systems,” in Intelligent Vehicles Symposium, IEEE, 2007.

[9] W. P. Sanberg, G. Dubbelman, and P. H. de With, “Extending the stixel world with online self-supervised color modeling for roadversus-obstacle segmentation,” in IEEE Conference
on Intelligent Transportation Systems (ITSC), 2014.

[10] W. P. Sanberg, G. Dubbelman, and P. H. N. de With, “Color-based free-space segmentation using online disparity-supervised learning.” in ITSC, 2015.

[11] R. Benenson, M. Mathias, R. Timofte, and L. v. Gool. Fast stixel computation for fast pedestrian detection. In ECCV Workshops, 2012.

[12] David Pfeiffer, Stefan K. Gehrig, Nicolai Schneider: "Exploiting the Power of Stereo Confidences", Proceedings of the CVPR 2013, Portland, OR, USA.

[13] H. Hirschmueller. Stereo processing by semiglobal matching and mutual information. PAMI, 30:328–41, 2008.