

Stereo Vision Based Traversable Region Detection for Mobile Robots Using U-V-Disparity

ZHU Xiaozhou, LU Huimin, *Member, IEEE*, YANG Xingrui, LI Yubo, ZHANG Hui

College of Mechatronics and Automation, National University of Defense Technology, Changsha 410073, China
E-mail: {zhu.xz1011, lyubo1987}@gmail.com, lhmnew@nudt.edu.cn, henry-young@sohu.com, zhanghui_nudt@126.com

Abstract: Accurate and real-time traversable region detection is of great significance to the navigation of mobile robots. In Comparison with lasers and radars, stereo vision can provide richer information and thus has become increasingly popular in this field. In this paper, an efficient traversable region detection algorithm is proposed based on stereo vision by using u-v-disparity. In the algorithm, obstacle disparity image and non-obstacle disparity image are introduced to enhance the performance. The thorough experiments are performed on our NuBot rescue robot in the indoor RoboCup rescue field and the outdoor campus environment and on KITTI Vision Benchmark dataset, and the results show that the traversable region can be detected efficiently with high successful rate by using the proposed algorithm.

Key Words: Traversable Region Detection, Stereo Vision, U-V-Disparity, Mobile Robots

1 Introduction

Reliable traversable region detection is vital to the navigation of mobile robots. It is still full of great challenges because the travel ability can be influenced by both the terrain characteristic and the robot mobility characteristic.

There are two kinds of hazards which would make a region non-traversable: the geometric hazard referring to the object or terrain that can be regarded to be non-traversable owing to its geometric feature, and the non-geometric hazard which can hinder or stop entirely the motion of the robots due to its physical properties such as the density and the cohesion [1]. The former includes obstacles, and the latter includes swamp, drift sand, etc. In this paper, we mainly focus on geometric hazards, so traversable regions can be defined as regions that do not contain geometric obstacles.

In comparison with other sensors such as lasers or radars, vision sensors are more similar to human's perception methods, and can provide the richest information including color, texture, depth, etc. In addition, as a passive sensing means, the interference among robots can be neglected when using vision sensors. Although being sensitive to weather and illumination conditions, vision sensors can still be regarded as ideal sensors for environment perception for mobile robots.

Stereo vision can measure the ranges to objects by calculating disparities between the stereo images. After obtaining the disparity image, there are mainly two ways to detect traversable regions.

The first one is 3D reconstruction according to points cloud derived from disparity map. Then obstacles and traversable regions can be detected using edge detection [2], local safety map [3], plane fitting [4], etc. However, the high computation burden makes 3D reconstruction difficult to meet the real-time requirement in robot navigation.

The second one is u-v-disparity method. V-disparity was originally introduced in [5] by Labayrade aiming at detecting obstacles either on a flat road or on a non-flat road, where "v" represents the coordinate of a pixel in the (u, v)

image coordinate. This method does not need the planar road assumption and the extraction of specific structures like road edges. By accumulating pixels with the same (v, d) in each row, the v-disparity can be built, where d represents the disparity value. Perpendicular obstacles can be mapped to be vertical lines whose pixel intensity represents the width of the obstacles. The road modeled as a succession of planes can be projected as slanted line segment. This line segment, also named as "ground correlation line" in [6], corresponds to the traversable region. Hu [7] extended Labayrade's work and proposed the concept of u-v-disparity. Similar to v-disparity, u-disparity is built by accumulating pixels with the same (u, d) in each column, and perpendicular obstacles can be mapped to be horizontal lines whose pixels intensity represents the height of the obstacles.

Usually, Hough Transform is applied in u-v-disparity images directly in order to extract lines [5, 7, 8]. However, it is inefficient and inaccurate because too many line candidates can be generated by using Hough Transform without removing the noise. In addition, this method can only detect the traversable regions correctly provided that the ground correlation line plays the dominant role in the v-disparity, which often cannot be satisfied, e.g., the traversable regions may only take a small part of the image or large obstacles may appear in the field of view of the camera.

In this paper, we propose an efficient traversable region detection algorithm based on stereo vision by using u-v-disparity for mobile robots. Obstacle disparity image and non-obstacle disparity image are introduced, so the affection caused by the non-traversable regions can be removed in the v-disparity image as much as possible, and thus the detection accuracy of the traversable region can be enhanced even when the traversable regions may only take a small part of the image.

The following parts are organized as follows: Section 2 introduces our NuBot rescue robot system to be used in the experiments briefly; our traversable region detection algorithm is proposed in Section 3; the thorough experiments are performed in Section 4, followed by the conclusion in Section 5.

2 NuBot Rescue Robot System

Our NuBot rescue robot system is shown in Fig. 1. The tracked robot is equipped with four sub-tracks which can greatly enhance its locomotion ability in complex scenarios. The robot is able to deal with slopes up to 45 degrees and climb upstairs and downstairs. The equipped stereo vision system is Bumblebee2 by Point Grey. The cameras are factory-calibrated for lens distortion and camera misalignments, to ensure consistency of calibration across all cameras and eliminate the need for in-field calibration.



Fig. 1. Our NuBot rescue robot system

3 Our Proposed Algorithm

In this paper, we propose an efficient traversable region detection algorithm based on stereo vision by using u-v-disparity. Our algorithm can be divided into three stages as follows:

- (1) Generating disparity image;
- (2) Calculating obstacle disparity image and non-obstacle disparity image;
- (3) Detecting ground correlation line and projecting.

3.1 Generating disparity image

After the left and right image are acquired and denoised, the rectification is performed to ensure that they are aligned. By this means, the correspondences can be restricted to the same line in both images and thus the computation burden in the stereo matching can be reduced.

The disparity image is generated by applying ELAS (Efficient Large-Scale Stereo Matching) algorithm [9], which uses a Bayesian approach and performs quite well even in relatively low-textured images. Firstly, the disparities of a sparse set of support points are computed using relatively strict constraints. The image coordinates of the support points are then used to create a 2D mesh via Delaunay triangulation. Finally, a prior is computed to reduce the matching ambiguities. In particular, this prior is formed by calculating a piecewise linear function induced by the support point disparities and the triangulated mesh.

The left and right rectified image shown in Fig. 2(a) and Fig. 2(b) respectively are a pair of stereo images from KITTI Vision Benchmark dataset [10]. Fig. 3(a) shows the obtained disparity image followed by the same disparity image demonstrated using a pseudo color LUT in Fig. 3(c), where the warmer color indicates the higher value of the disparity and means the smaller distance to the cameras according to equation (1)

$$z(x, y) = \frac{f \cdot b}{d(x, y)} \quad (1)$$

where f , b , $d(x, y)$, $z(x, y)$ represent focal length, base-line length, disparity value and the distance to the cameras

respectively. The dark blue in Fig. 3(c) represents the invalid pixels, mainly caused by low-textured or un-textured regions, e.g., the sky.

What needs to be mentioned is that the leftmost part of the rectified left image does not have correspondent regions in its right counterpart, and thus their disparity values can not be computed by stereo match. Different stereo match algorithms deal with this problem according to different strategies, e.g. replacing with the disparity value corresponding to the neighboring pixels, or by interpolating. As to ELAS, it labels them as invalid pixels.



(a) The rectified left image



(b) The rectified right image

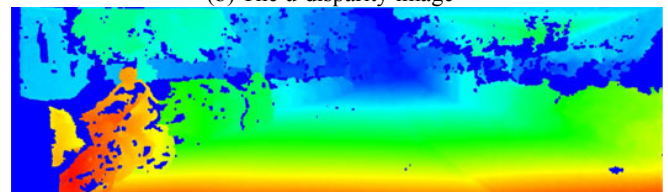
Fig. 2. A pair of rectified stereo images from KITTI Vision Benchmark dataset.



(a) The disparity image



(b) The u-disparity image



(c) The disparity image in LUT



(d) The u-disparity image in LUT

Fig. 3. The disparity image and the u-disparity image computed from the images in Fig. 2.

3.2 Calculating obstacle disparity image and non-obstacle disparity image

The traversable regions can be modeled as a succession of planes, so they can be projected to be slanted line segments in the v-disparity image, i.e. ground correlation lines. As too much noises caused by the non-traversable regions in the v-disparity image would make it difficult to detect the correct ground correlation lines, they should be removed as much as possible before generating the v-disparity image.

We define obstacle disparity image as the regions in the disparity image corresponding to potential obstacles. The

intensity of a pixel (u, d) in the u-disparity image refers to the number of pixels with the same disparity value d in the u th column of the disparity image. If the intensity is higher than a threshold, it means that in a certain column of the image, there are too many points which have the same distance to the camera, and these points correspond to potential obstacles. After taking this into consideration, we perform the following steps to obtain the obstacle disparity image which contains the potential non-traversable regions or obstacle regions:

Step 1 U-disparity is calculated by accumulating pixels with the same disparity d in each column of the disparity image, as shown in Fig. 3(b) followed by the u-disparity image shown using a pseudo color LUT in Fig. 3(d). It is obvious that there are many line segments in the u-disparity image corresponding to obstacles and the obstacle regions should be removed from the original disparity image before calculating the v-disparity image.

Step 2 Binarized obstacle image is built. For every pixel (u, d) in the u-disparity image, if its intensity is greater than a threshold, then in the u th column of the disparity image, those pixels whose disparity equals to d are set to be obstacle pixels, and those pixels whose disparity does not equal to d are set to be potential traversable region pixels. Here the threshold indicates the height of potential obstacles measured in pixels. So the binarized obstacle image can be generated, as shown in Fig. 4(a), where the white regions represent the potential obstacles, and the black regions represent the potential traversable regions.

Step 3 Morphological close operation is performed on the binarized obstacle image to link small spaces between potential obstacles, as shown in Fig. 4(b). Small isolated black regions can also be eliminated directly as noises and changed to be potential obstacle regions, as shown in Fig. 4(c). This is reasonable because as to black regions, too small an area indicates that it cannot be traversable and thus should be treated as potential obstacles.

Step 4 Obstacle disparity image is built according to the binarized obstacle image. For every pixel of the binarized obstacle image, corresponding pixel of the original disparity image is studied. If a pixel of the binarized obstacle image belongs to potential obstacle regions, then the corresponding pixel with the same position in the original disparity image is retained. If a pixel of the binarized obstacle image belongs to potential traversable regions, then the corresponding pixel with the same position in the original disparity image is deleted. So the obstacle disparity image can be generated, as shown in Fig. 4(d) by using a pseudo color LUT.

Step 5 Non-obstacle disparity image can be generated by subtracting the original disparity image with the obstacle disparity image, as shown in Fig. 4(e) by using a pseudo color LUT. From Fig. 4(e), we see that most of the potential obstacle regions or non-traversable regions have been removed from the disparity image.

3.3 Detecting ground correlation line and projecting

By accumulating pixels with the same (v, d) in each row of the non-obstacle disparity image, the v-disparity image can be calculated, as shown in Fig. 5(a) followed by the v-disparity image shown using a pseudo color LUT in Fig.

5(b). Because the affection from the obstacle regions has been removed before constructing the v-disparity image, the ground correlation line can be detected easily by using Hough Transform. The detected line segments are marked as red lines in the v-disparity image, as shown in Fig. 5(c).

All pixels in the v-disparity image that contribute to the ground correlation line can be projected back to the disparity image, so the traversable regions can be found, marked as green regions in Fig. 6. During this projecting process, those regions whose areas being smaller than a threshold are filled or eliminated according to surrounding regions.

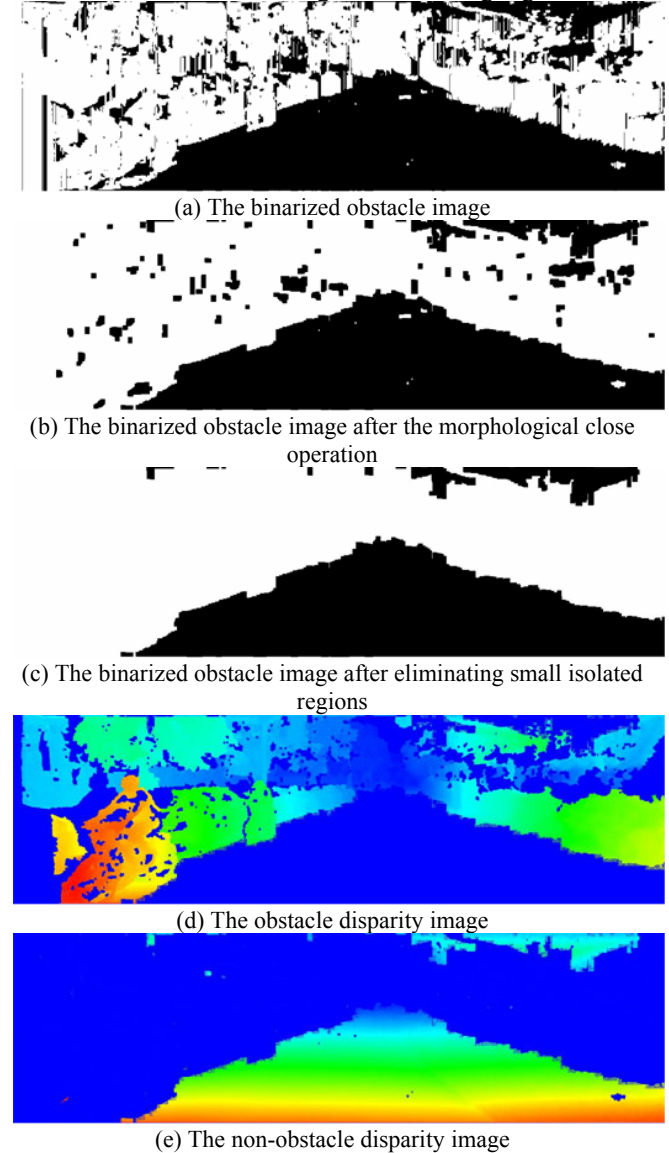


Fig. 4. The process of calculating the obstacle disparity image and the non-obstacle disparity image from the disparity image in Fig. 3.

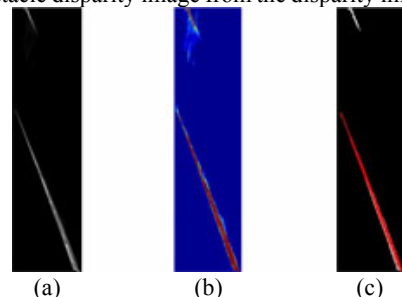


Fig. 5. The v-disparity image obtained from the non-obstacle disparity image in Fig. 4, and the detected ground correlation line.

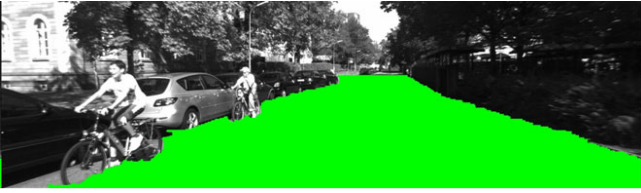


Fig. 6. The detected traversable regions (green regions) by using our algorithm.

4 Experimental Results

In this section, we perform a series of experiments to validate our traversable region detection algorithm. The experiments are conducted by using our NuBot rescue robot in various environments and KITTI Vision Benchmark dataset [10] acquired by a moving vehicle. The performance of our algorithm is evaluated by the successful detection rate and the computation cost.

4.1 The Experiments in indoor rescue environment

In this experiment, we test our algorithm in the indoor RoboCup rescue competition field which is to simulate an environment after disaster. Traversable region detection is a basic ability for the autonomy of the rescue robot. Some typical results are shown in Fig. 7, from which we see that our algorithm works well in this environment, and the non-traversable regions such as the ladder, wood blocks, etc can be excluded from the traversable regions.

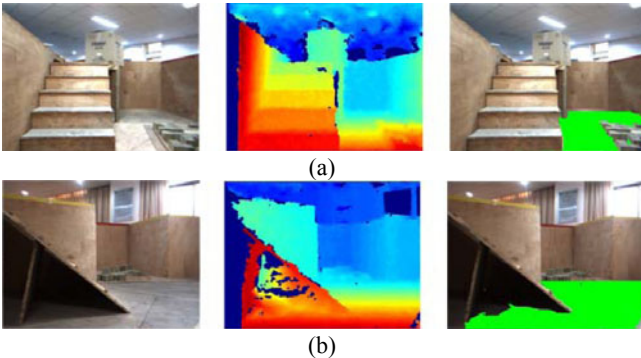


Fig. 7. The experimental results in the indoor rescue environment (from left to right: the rectified left image, the disparity image shown by using a pseudo color LUT, and the detected traversable regions).

4.2 The Experiments in outdoor campus environment

In this experiment, we test our algorithm in several typical environments of the outdoor campus. The results are shown in Fig. 8, from which we see that our algorithm works well in outdoor campus environment, and the non-traversable regions such as sidewalls, road curbs, buildings, trees, bushes, walkers, cars, etc can be excluded from the traversable regions. What needs to be pointed out is that the invalid pixels caused by stereo match in the bottom-left of the images are regarded as non-traversable regions for safety sake.

4.3 The Experiments on KITTI Vision Benchmark dataset

In this experiment, we test our algorithm by using the KITTI Vision Benchmark dataset acquired in the urban environment [10], and the results are shown in Fig. 9. From

the results, we can conclude that our algorithm is also effective in the complex urban environment, especially in Fig. 9(d) where the road takes only a small part of the image.

We process the same images in Fig. 9(d) to compare our algorithm and the traditional method which detects the ground correlation line in the original v-disparity image [5, 7, 8]. The results are shown in Fig. 10. We see that by using our algorithm, most of the affection from the obstacles has been removed in the v-disparity image, and thus the ground correlation line can be detected correctly.

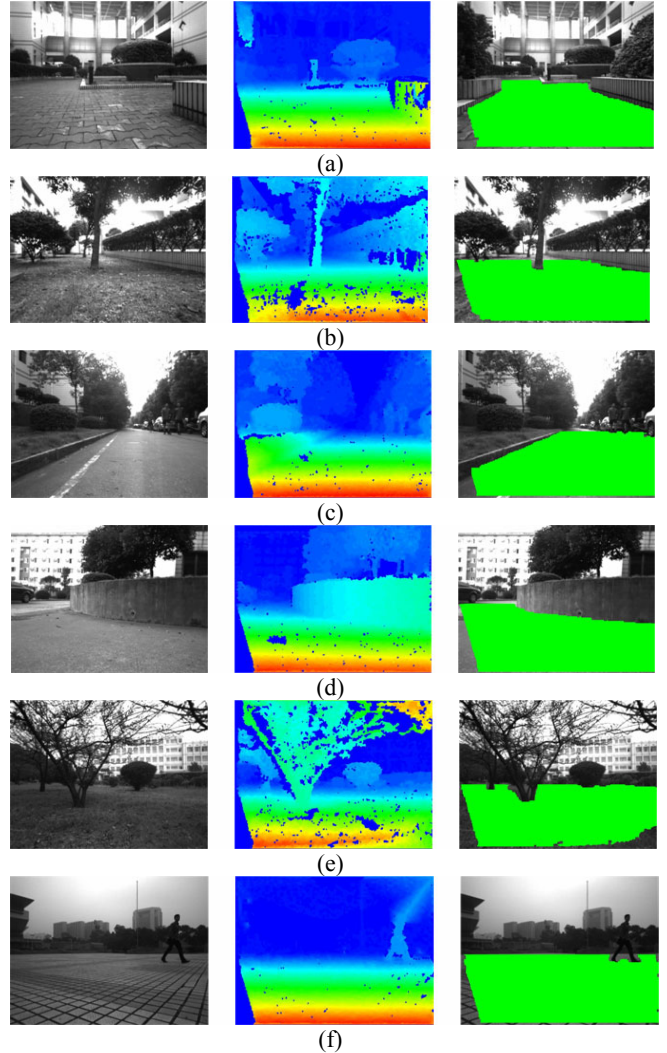


Fig. 8. The experimental results in outdoor campus environment (from left to right: the rectified left image, the disparity image shown by using a pseudo color LUT, and the detected traversable regions).

4.4 Evaluating the performance of our algorithm

In this paper, we evaluate the performance of our algorithm by using successful detection rate. If most of the traversable regions in an image can be detected correctly and no obstacle is detected wrongly as the traversable region, we consider the detection to be successful. We use 14 image sequences from the KITTI Vision Benchmark dataset to obtain the statistical result about successful detection rate, and the average value is 92.3%.

Although high successful rate can be achieved, there are still some deficiencies in our algorithm. When the image becomes very low-textured or even un-textured caused by

terrible environment characteristics like the road sign (Fig. 11(a)), the strong sun light (Fig. 11(b) (d)), or water on the road surface (Fig. 11(c)), etc, large areas of invalid pixels are generated during the process of stereo match, so parts of the traversable region are regarded to be non-traversable, as shown in Fig. 11.

We also test the computation time needed in the whole process of our algorithm. The computer used in the experiments is equipped with an Intel Core i5-3450 3.1GHz CPU and 4GB RAM. The computation time needed to compute the u-v-disparity and detect traversable regions in our algorithm is about 10 ms on the images with the dimension of 512×384 , 17 ms on the images with the dimension of 640×480 , and 23 ms on the images with the dimension of 1344×391 respectively. The details corresponding to the images appeared in the figures of this paper are shown in Table 1. So our traversable region detection algorithm can be run efficiently after the stereo match has been completed. From Table 1, we can also see that the computation cost in the step of stereo match is much higher than that in the step of computing the u-v-disparity and detecting traversable regions. So stereo match should be accelerated in future to enable the whole process of traversable region detection be run in real-time.

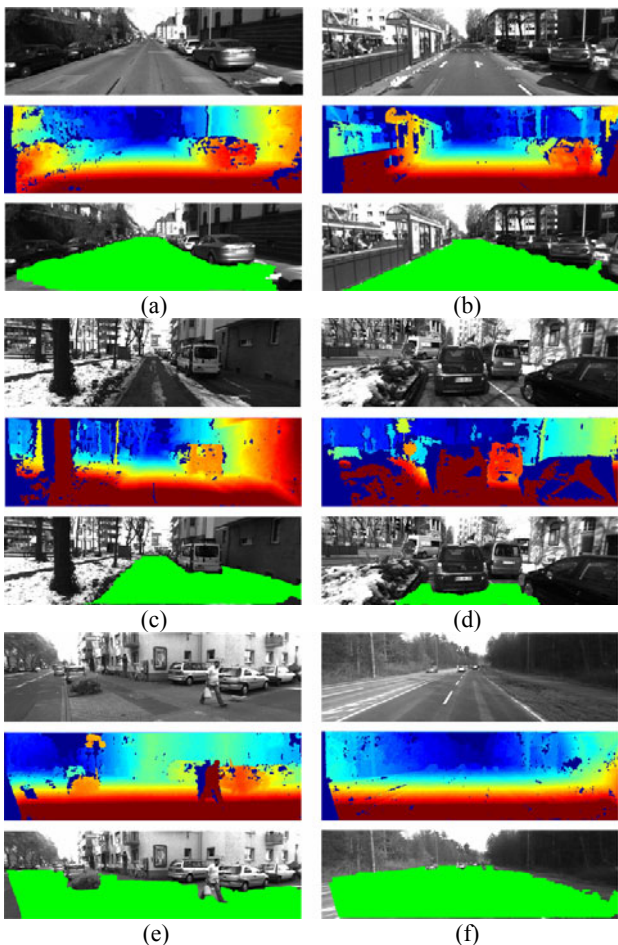


Fig. 9. The experimental results on the KITTI Vision Benchmark dataset (from top to bottom: the rectified left image, the disparity image shown by using a pseudo color LUT, and the detected traversable regions).



Fig. 10. The comparison between the traditional method [5, 7, 8] and our algorithm (from left to right: the v-disparity image shown by using a pseudo color LUT, the detected ground correlation line, and the detected traversable regions).

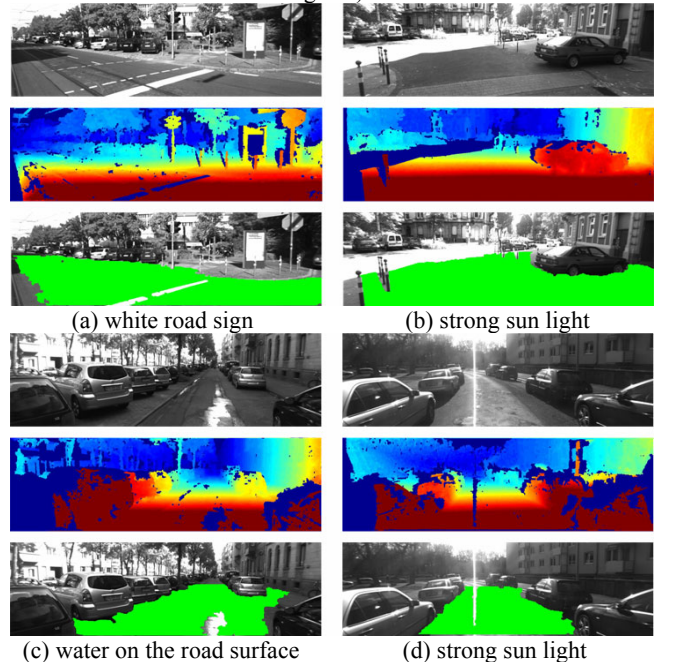


Fig. 11. Parts of the traversable regions are regarded to be non-traversable when the image becomes very low-textured or even un-textured caused by terrible environment characteristics.

Table 1: The needed computation time corresponding to the images appeared in the figures of this paper

	The dimension of the image	The computation time needed in stereo match to compute the disparity image	The computation time needed in traversable region detection by using our algorithm
Fig. 7(a)	512×384	115.5 ms	10.1 ms
Fig. 7(b)	512×384	112.2ms	10.0ms
Fig. 8(a)	640×480	213.3 ms	17.6 ms
Fig. 8(b)	640×480	202.0 ms	16.8 ms
Fig. 8(c)	640×480	190.0 ms	14.6ms
Fig. 8(d)	640×480	184.8 ms	14.8 ms
Fig. 8(e)	640×480	188.9 ms	19.4 ms
Fig. 8(f)	640×480	197.4ms	16.2 ms
Fig. 9(a)	1344×391	213.3 ms	25.4 ms
Fig. 9(b)	1344×391	202.0 ms	23.3 ms
Fig. 9(c)	1344×391	190.0 ms	22.5ms
Fig. 9(d)	1344×391	184.8 ms	23.6 ms
Fig. 9(e)	1344×391	188.9 ms	23.1 ms
Fig. 9(f)	1344×391	197.4ms	23.5ms

5 Conclusion and Future Work

In this paper, an effective traversable region detection algorithm is proposed for mobile robots that employs the concept of the u-v-disparity. Obstacle disparity image and non-obstacle disparity image are introduced, which can remove the affection caused by the non-traversable regions in the v-disparity image as much as possible, and thus enhance the detection accuracy. The thorough experiments on various environments and KITTI Vision Benchmark dataset are performed to test our algorithm, and the results show that our algorithm can detect the traversable regions efficiently with high successful rate by using stereo vision system.

To evaluate our algorithm further, we intend to make more experiments to compare its performance with some other famous algorithms. To increase the robustness of our system, we will try to develop novel algorithms to make stereo match more robust to the environment characteristic. Because the real-time performance of stereo match should be improved, we plan to use GPU to accelerate the process of stereo match.

6 Acknowledgment

The authors would like to thank A. Geiger, P. Lenz, and R. Urtasun for making their vision dataset available for testing, which is very helpful to our work.

References

- [1] D. Helmick, A. Angelova, and L. Matthies, Terrain Adaptive Navigation for Planetary Rovers, *Journal of Field Robotics*, 26(4): 391-410, 2009.
- [2] L. Li, R. Wang, and M. Zhang, Study on Stereo Vision-based Cross-country Obstacle Detection Technology for Intelligent Vehicle, in *Proceedings of the Third International Conference on Natural Computation*, 2007: 719-723.
- [3] A. Murarka, M. Sridharan, and B. Kuipers, Detecting Obstacles and Drop-offs using Stereo and Motion Cues for Safe Local Motion, in *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008: 702-708.
- [4] T. Braun, H. Bitsch, and K. Berns, Visual Terrain Traversability Estimation using a Combined Slope/Elevation Model, in *KI 2008: Advances in Artificial Intelligence, LNCS 5243*, 2008: 177-184.
- [5] R. Labayrade, D. Aubert, and J.-P. Tarel, Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through "V-disparity" Representation, in *Proceedings of IEEE Intelligent Vehicle Symposium*, 2002: 646-651.
- [6] A. Broggi, C. Caraffi, R. I. Fedriga, and P. Grisleri, Obstacle Detection with Stereo Vision for Off-Road Vehicle Navigation, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, 2005.
- [7] Z. Hu and K. Uchimura, U-V-Disparity: An efficient algorithm for Stereovision Based Scene Analysis, in *Proceedings of IEEE Intelligent Vehicle Symposium*, 2005: 48-54.
- [8] G. D. Cubber, D. Doroftei, L. Nalpantidis, G. C. Sirakoulis, and A. Gasteratos, Stereo-based Terrain Traversability Analysis for Robot Navigation, in *Proceedings of IARP/EURON Workshop on Robotics for Risky Interventions and Environmental Surveillance*, 2009.
- [9] A. Geiger, M. Roser, and R. Urtasun, Efficient Large-Scale Stereo Matching, in *Computer vision - ACCV 2010, LNCS 6492*, 2011: 25-38.
- [10] A. Geiger, P. Lenz, and R. Urtasun, Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite, in *Proceedings of the 2012 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012: 3354-3361.