2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

Global Correlation Based Ground Plane Estimation Using V-Disparity Image*

Jun Zhao, Jayantha Katupitiya and James Ward
ARC Centre of Excellence for Autonomous Systems
School of Mechanical and Manufacturing Engineering
The University of New South Wales, Sydney, NSW 2052, Australia
J.Katupitiya@unsw.edu.au

***Abstract*— This paper presents the estimation of the position of the ground plane for navigation of on-road or off-road vehicles, in particular for obstacle detection using stereo vision. Ground plane estimation plays an important role in stereo vision based obstacle detection tasks. V-disparity image is widely used for ground plane estimation. However, it heavily relies on distinct road features which may not exist. In here, we introduce a global correlation method to extract the position of the ground plane in V-disparity image even without distinct road features.**

***Index Terms*— V-disparity image, correlation, stereo vision.**

I. INTRODUCTION

Stereo vision is one of the key components in vision-based robot navigation. Today mobile robotics researchers focus on developing navigation in unknown environments where a robot requires to respond to changes in the environment in real-time. Although laser sensors provide refined and easy-to-use information about the surrounding area, they also present some intrinsic limitations to their functioning as mentioned in [1].

Because a vision system provides a large amount of data, extracting refined information sometimes may be complex. In obstacle detection tasks, the purpose is to distinguish the obstacle pixels from the ground pixels in the depth map. Se and Brady [2] quote from Gibson's "ground theory hypothesis"(1950): "there is literally no such thing as a perception of space without the perception of a continuous background surface". In this study, we assume that the ground can be locally represented by a plane [3].

In built environments such as in-door environment, the position of stereo rig relative to the ground is normally fixed, thus the disparity of ground pixels in the disparity map can be determined during calibration stage [4]. However in outdoor environment, The pitch angle between the cameras and the road surface will change due to static and dynamic factors [5], thus the disparity of ground pixels is changing from time to time. Therefore, we need to compute the pitch angle and disparity of ground pixels dynamically. In [5], the authors used four sensors mounted between the chassis and wheels to compute the pitch angle.

*This work is supported in part by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government.

"Plane fitting" is a traditional method for ground estimation and used by different researchers. In [6] , the authors used *RANSAC Plane Fitting* to find the disparity of ground pixels. In [7], pixels (u, v) with a valid value in the depth map are labeled as belonging to the ground plane if the following constraint is satisfied: $d(u, v) \leq au + bv + c + r(d)$. In [8], the authors developed a road detection algorithm utilizing road features called plane fitting errors.

Recently, V-disparity image has become popular for ground plane estimation [1][9][10][11]. In this image, the abscissa axis (w) plots the offset for which the correlation has been computed; the ordinates axis (v) plots the image row number; the intensity value is settled proportional to the measured correlation, or the number of pixels having the corresponding disparity (w) in a certain row (v). Each planar surface in the field of view is mapped in a segment in the V-disparity image [10]. Vertical surfaces in the 3D world are mapped into vertical line segments, while ground plane in the 3D world are mapped into slanted line segment. This line segment, called ground correlation line in this study, contains the information about the cameras pitch angle at the time of acquisition (mixed with the terrain slope information).

Both plane fitting and v-disparity image rely on distinct road features such as lane markings. Without these features, there would not be sufficient ground pixels in the sparse disparity map from which the ground plane can be extracted. In this paper, we first analyze the behavior of ground correlation line in different camera pitch angles in V-disparity image. This is introduced in section II. Based on the behavior, in Section III, we introduce a global correlation method to extract the ground correlation line. In Section IV, we show some experimental results using image pairs without distinct road features. In Section V, we draw some conclusions.

II. GROUND CORRELATION LINE IN V-DISPARITY IMAGE WHEN CHANGING TILT ANGLE

A. Camera placement and geometry

To obtain a real world representation from an image pair, it is necessary to know the cameras placement at the time of acquisition. Consider cameras on an autonomous vehicle that are tilted down an angle θ, as shown in Fig. 1. In this

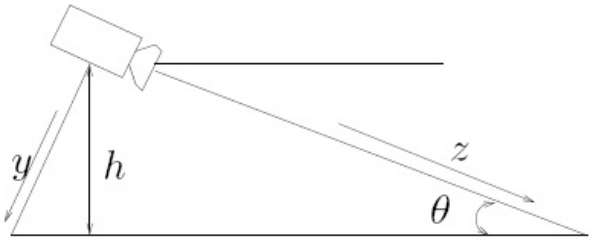


Fig. 1. Camera placement

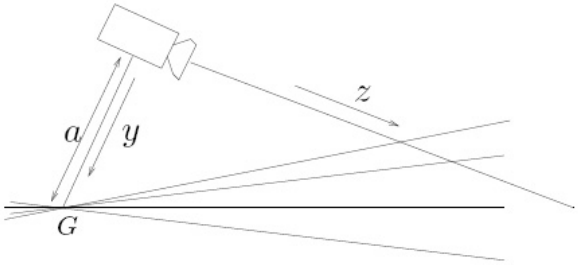


Fig. 2. Conditions for the ground correlation lines to be parallel: a is fixed, all ground planes pass through point G

figure, a camera centered coordinate system, xyz, defines the positions of points in the physical world in front of the cameras. If the cameras are at a distance h above the ground, the ground plane can be represented as

$$z = \frac{h}{\sin\theta} - y\frac{\cos\theta}{\sin\theta} \tag{1}$$

An image coordinate system, uvw, defines the spatial positions of data points on the image plane(uv) and the relative disparity of corresponding points between the left and right images(w). We adopt the pin-hole camera model, then the relation between the world coordinates of a point $P(x, y, z)$ and the coordinates on the image plane (u, v) in the camera is

$$u = x\frac{f}{z}, v = y\frac{f}{z}, w = b\frac{f}{z} \tag{2}$$

where f is the focal distance of the lens and the stereo baseline is b.

From Eq. 1 and 2, we can get that in image plane

$$v = \frac{hw}{b\cos\theta} - f\tan\theta \tag{3}$$

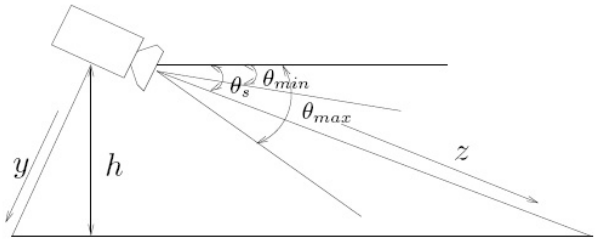


Fig. 3. Pitch variation

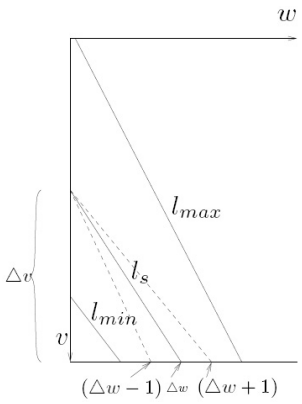


Fig. 4. Slope of ground correlation line during pitch variation. The solid slanted lines represent ground correlation lines. The dashed lines represent the next distinguishable gradient near l_s

In this equation, b and f are constants dependent on the camera geometry and spacing between cameras. In this study, we also assume that the camera height h is also fixed, that is, there only exists pitch oscillation. Thus v is a function of w and θ. At the time of acquisition, θ is also fixed, then the slope g of ground plane in V-disparity image, can be represented as

$$g = \frac{\partial v}{\partial w} = \frac{h}{b\cos\theta} \tag{4}$$

Eq. 4 shows that the slope g becomes greater when θ becomes greater. Thus the ground correlation lines corresponding to different pitch angles would be like those shown in Fig.4.

Actually, for the ground correlation lines to be exactly parallel to each other, it should satisfy the conditions shown in Fig. 2. In this figure, the y axis intersects the ground plane through the point G, the distance(a) between camera and G is fixed, and all different ground planes pass through point G. In this situation, the ground plane can be represented as

$$z = \frac{a}{\tan\theta} - \frac{y}{\tan\theta}$$

together with Eq. 2, we can get

$$v = \frac{a}{b}w - f\tan\theta$$

thus $g = \partial v/\partial w = a/b$ which is a constant. This has been investigated in [12].

At different pitch angles shown in Fig. 3, the slope of ground plane in V-disparity image will be different. Suppose g_s is the slope of l_s using static calibration data when $\theta = \theta_s$, g_{max} is the slope of l_{max} when $\theta = \theta_{max}$, and g_{min} is the slope of l_{min} when $\theta = \theta_{min}$ as shown in Fig. 4. The changing ratio of g

$$\frac{dg}{d\theta} = \frac{h}{b}\sec\theta\tan\theta$$

approaches 0 when θ is small, and

$$\triangle g = \frac{dg}{d\theta}\triangle\theta = \frac{h}{b}\sec\theta\tan\theta\triangle\theta \tag{5}$$

From Eq. 5 we can get

$$g_{max} - g_s \approx \triangle g = \frac{h}{b} \sec\theta_s \tan\theta_s (\theta_{max} - \theta_s) \quad (6)$$

In V-disparity image the resolution of abscissa axis(w) is one pixel. Let $g_s = (h/b)\sec\theta_s = \triangle v/\triangle w$ as shown in Fig. 4. If $\triangle v$ is fixed, from this figure we can see that the next distinguishable gradient bigger than g_s is $\triangle v/(\triangle w - 1)$. Thus for l_s to be parallel with l_{max} in the V-disparity image, it should satisfy

$$g_{max} < \frac{\triangle v}{\triangle w - 1} \quad (7)$$

then we should have

$$\frac{\triangle v}{\triangle w - 1} - \frac{\triangle v}{\triangle w} > \frac{\triangle v}{\triangle w} \tan\theta_s (\theta_{max} - \theta_s) \quad (8)$$

from which we can get

$$\triangle w < \frac{1}{\tan\theta_s (\theta_{max} - \theta_s)} + 1, \quad \theta_s \neq 0 \quad (9)$$

similarly we can get

$$\triangle w < \frac{1}{\tan\theta_s (\theta_s - \theta_{min})} - 1, \quad \theta_s \neq 0 \quad (10)$$

Combining Eq. 9 and 10, we get,

$$\triangle w < min\{\frac{1}{\tan\theta_s (\theta_{max} - \theta_s)} + 1, \frac{1}{\tan\theta_s (\theta_s - \theta_{min})} - 1\}, \quad \theta_s \neq 0 \quad (11)$$

For the ground correlation lines to be parallel with each other at different pitch angles, it only need to satisfy Eq. 11. Thus for a given $\triangle w$ and certain pitch oscillation, we need to reduce θ_s. In [10], the tilt angle is $8.5°$. In [1], the tilt angle is $7.73°$.

Claudio et al [1] have experimentally found out that the behavior of the ground correlation line during a pitch variation is to oscillate, parallel to itself like in Fig. 5. In here, we investigate this characteristic mathematically and give the condition for this characteristic to be valid.

III. GLOBAL CORRELATION METHOD

Originally, Laybarade [10] used Hough transform to extract the ground correlation line from the V-disparity image. It works well for on-road vehicles when there exist distinct road features on the ground. However, when there are no distinct features on the ground, Hough transform is prone to errors. In the previous section, we found that if the tilt angle is small enough, the ground correlation line during a pitch variation is to oscillate, parallel to itself. The slope and position of the ground correlation line under stable condition can be obtained off-line during calibration stage. Then under driving conditions the ground correlation line must be one of those parallel to the ground correlation line under stable conditions and within a certain disparity range as shown in Fig. 5.

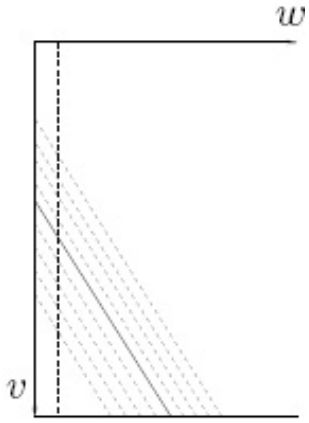


Fig. 5. Ground correlation lines in V-disparity image. The solid slanted line corresponds to the ground correlation line obtained using the static calibration data, while the dashed slanted lines are the ones expected varying the pitch value. The solid vertical line indicates the 0 disparity value, the dashed vertical line indicates the disparity value of points at infinite distance; they do not overlap due to a slight convergence of cameras optical axes.

Claudio [1] estimate the ground correlation line by accumulating the V-disparity image values along each of the candidate lines and choosing the line that accumulated the greatest V-disparity value. This method can reduce the computational load by computing the correlation image values only in correspondence to pixels belonging to the candidate lines. However, it still requests road features. In off-road situations, they use a semi-global method because the weak edges produced by ground texture takes up the largest part of the image, correlation is done by ignoring the absolute value of edges and considering only their phase.

It is well known that, correlation based stereo matching is prone to errors in a uniform grey level area such as that of a road without distinct features because there is not a clear optimum correspondence. A bigger window size can reduce this error. However, correlation assumes that the depth is equal for all pixels of a correlation window. This assumption is violated at depth discontinuities. We assume the ground near the vehicle is clear of obstacles. Then we can stretch the correlation window horizontally like that shown in Fig. 6 in which each scan line can contain only one such window. We then calculate matching cost associated with different disparity for it. In V-disparity image, the disparity of ground is changing according to the ground correlation line, and the slope is determined by the stereo rig geometry and placement, it is also known that during a pitch variation the ground correlation line is to oscillate, parallel to itself. Thus we accumulate the matching cost of the very wide correlation window along each of the candidate lines and choose the one with least accumulated matching cost as ground correlation line. This is like calculating the matching cost for a very wide window. In this window, the disparity is kept constant horizontally, but vertically it is changing according to the

Fig. 6. A very wide matching window(in black box) on the ground

slope of ground correlation line.

In practice, the matching cost is accumulated from the bottom of the V-disparity image to a level lower than the horizon. If the ground plane dominate this area, the quantity of information carried by obstacles is attenuated, because the ground plane contributes most to the matching cost for this area.

IV. EXPERIMENTAL RESULTS

We tested the method on image sequences from [13]. The sequence covers very different traffic and road conditions. The disparity map is computed using normal correlation based algorithm.

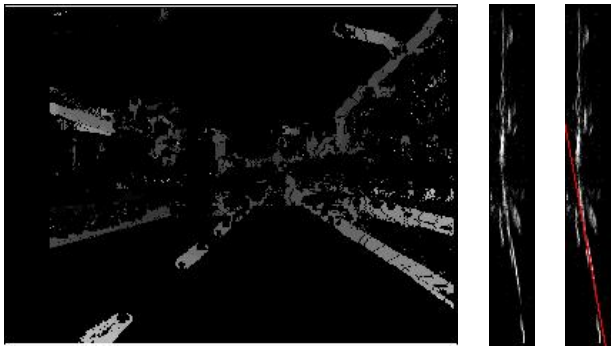
A. Slope and position of ground correlation line in different pitch angles

To investigate the slope of ground ground correlation line under different pitch angles, we first select image pairs with distinct road features in the sequence(one such image pair is shown in Fig. 7). The selected image pairs cover all road and traffic conditions presented in the sequences. We them perform Hough transform on the associated V-disparity images for these image pairs to extract the ground correlation line as shown in Fig. 8. Fig. 8(a) is the disparity image of Fig. 7, Fig. 8(b) is the corresponding V-disparity image. The abscissa axis plots the disparity(w) for which the correlation has been computed; the ordinate axis plots the image row number(v) and the intensity value is settled proportional to the number of pixels having the corresponding disparity (w) in a certain row (v). Ground plane(lane markings) in Fig. 8(a) is mapped into the slanted line segment in Fig. 8(b), the red slanted line in Fig. 8(c) is the extracted ground correlation line using Hough transform.

After Hough transform, we can get the position($\triangle w$) and slope(g) of the ground correlation lines for these image pairs. The result is shown in Fig. 9. The lower plot of this figure shows the position($\triangle w$) of ground correlation lines for different image pairs, which is ranging from 25 to 28. This plot tells us there do exist pitch variation in different conditions. The upper plot shows the slop(g) of ground correlation lines, which ranges from 5.5 to 5.5652. From this Figure we can see that the bigger the slope(g) is, the

(a) (b)

Fig. 7. Test image pair with distinct road features



(a) (b) (c)

Fig. 8. Resulted disparity image of Fig. 7 (a), V-disparity image (b), extracted ground correlation line using Hough Transform (c)

greater $\triangle w$. This confirms the match between Eq. 4 and Fig. 4 in section II. Also, the upper plot shows that the change of slope is very small, the difference between the greatest slope and smallest slope is merely 0.0652, which is 1.17%. Thus the assumption that the ground correlation lines under different pitch angles are parallel, is valid.

B. Ground correlation line extraction using global correlation method

To use global correlation method introduced in Section III, we need to know the slope of ground correlation line beforehand. From the test of section IV-A, we choose the slope to be 5.533. After slope is given, we calculate the matching cost associated with each candidate ground correlation lines shown in Fig. 5. The position of candidate line $\triangle w$ is ranging from 22 to 32. This range is large enough to cover all the possibilities of $\triangle w$ under different pitch angles. The matching cost of each candidate line is accumulated from the bottom of the image to the level in which $v = 150$, or the level in which the disparity of the candidate line reaches 0. Fig. 10(a) shows the plot of matching costs associated with candidate lines using image pair in Fig. 7, the minimum is unique in this plot and the extracted ground correlation line is shown in Fig. 10(b). It is the same as that taken with Hough transform in Fig. 8(c).

Fig. 11 presents an image pair in which there are no distinct road features on the ground. Fig 12(a) is the corresponding disparity image. Fig 12(b) is the V-disparity image built from (a). In Fig. 12(b) there is no clear longitudinal profile of the road. If we use Hough transform

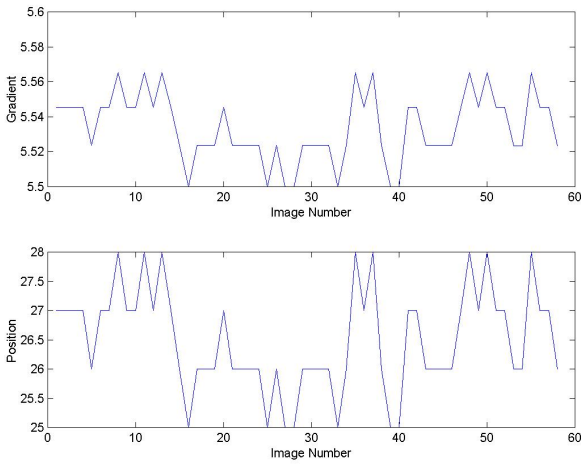


Fig. 9. The slope and position of ground correlation line of image sequence

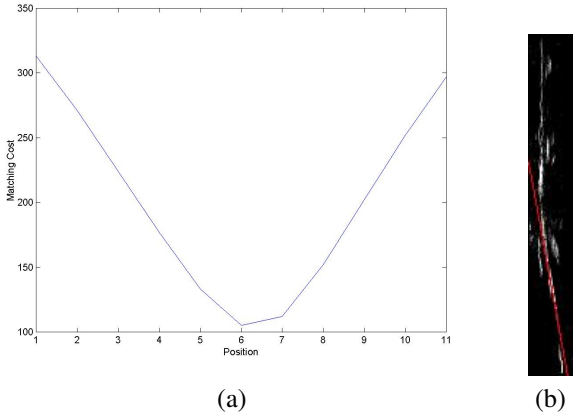


Fig. 10. Plot of matching cost for image pair of Fig. 7 (a), V-disparity image and extracted ground correlation line using global correlation method (b)

Fig. 11. Test image pair without distinct road feartures

to extract a straight line, the result would be that shown in Fig. 12(c). Fig. 13(a) shows the plot of matching costs associated with candidate lines for this image pair, the minimum is still unique in this plot and using the global correlation method the ground correlation line is extracted as shown in Fig. 13(b).

Fig. 14 shows an image pair with a close obstacle. Fig. 15(a) shows the plot of matching costs associated with candidate lines for this image pair, the ground correlation line is extracted as shown in Fig. 15(b) using global correlation method. The disparity map of this image pair is shown in Fig. 16(a). The detected obstacles(red pixels) and horizon(horizontal red line) are shown in Fig. 16(b). In the obstacle detection stage, only the pixels below the horizon in the disparity image are considered, also within this area only the pixels with valid correspondents are checked [1].

The global correlation method is tested on various image pairs, with or without distinct ground features. It can extract the ground correlation line in V-disparity image correctly in both situations. In the plots of Fig. 10, Fig. 13 and Fig. 15, a parabola may be fitted to a local neighborhood (3 or 5 points) of the global minimum to obtain sub-pixel accuracy of $\triangle w$. This accuracy is dependant of the image quality and whether or not the ground pixel dominate this area of the image. We will investigate it in the future. The global correlation method can work in real time because the matching cost has been computed in stereo matching stage.

V. CONCLUSION

In a V-disparity image, the planar ground is represented as a slanted straight line. Previously Hough transform was used to extract this straight line. This method requires local information in the image such as lane markings, road edges belonging to the ground plane. This requirement is caused by the characteristic of correlation based stereo matching. Only strong responses of the correlation function are considered as correspondents in the disparity map. Weak edges produced by ground texture, which take up a large part of the image are rejected. These weak edges belong to the ground plane. In V-disparity image the ground plane is simplified into a slanted straight line which is called ground correlation line here. In this paper, we first analyzed the characteristic of ground correlation line in different angles θ of the ground plane with respect to the reference of the stereo vision system. The slope of ground correlation line can be considered as a constant. Based on this we proposed a global correlation method in which matching cost is

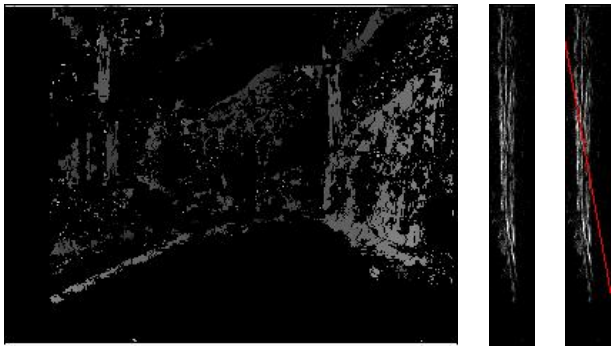
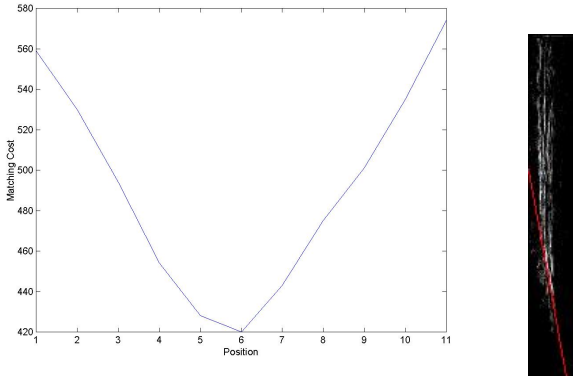


Fig. 12. Resulted disparity image (a), V-disparity image (b), extracted ground correlation line using Hough Transform (c)

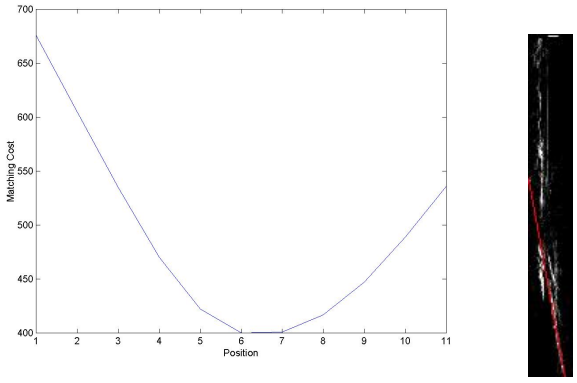


(a) (b)

Fig. 13. Plot of matching cost for image pair of Fig. 11 (a), V-disparity image and extracted ground correlation line using global correlation method (b)

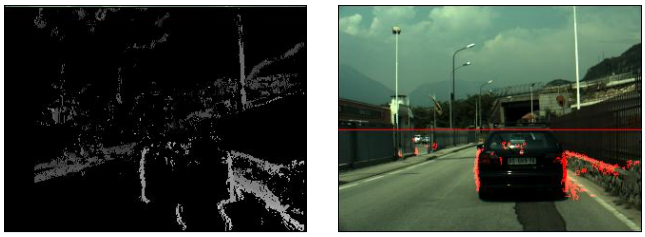
(a) (b)

Fig. 14. Test image pair with close obstacle



(a) (b)

Fig. 15. Plot of matching cost for image pair of Fig. 14 (a), V-disparity image and extracted ground correlation line using global correlation method (b)



(a) (b)

Fig. 16. Disparity image (a) detected obstacles, red pixels belong to obstacles, horizontal red line represents detected horizon(b)

accumulated along the candidate ground correlation lines. Thus the effect of a large amount of weak edges belonging to the ground are taken into consideration. Experimental results for different road conditions show that the minimum of the matching costs associated with candidate lines is outstanding and thus indicates the position of the ground correlation line.

ACKNOWLEDGMENT

The authors would like to thank Paolo Lombardi, Michele Zanin and Stefano Messelodi for making their test sequences available for evaluation, and Claudio Caraffi, Qiang Zhang for valuable discussions.

REFERENCES

[1] C. F. R. G. P. Broggi, A.; Caraffi, "Obstacle detection with stereo vision for off-road vehicle navigation," *Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on*, vol. 3, no. 65- 65, pp. 20–26, June 2005.

[2] S. Se and M. Brady, "Ground plane estimation, error analysis and applications," *Robotics and Autonomous Systems*, vol. 39, no. 2, pp. 59 – 71, 2002.

[3] S. Badal, S. Ravela, B. Draper, and A. Hanson, "Practical obstacle detection and avoidance system," *IEEE Workshop on Applications of Computer Vision - Proceedings*, pp. 97 – 104, 1994.

[4] D. Murray and J. J. Little, "Using real-time stereo vision for mobile robot navigation," *Autonomous Robots*, vol. 8, no. 2, pp. 161 – 171, 2000.

[5] S. Nedevschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, R. Schmidt, and T. Graf, "High accuracy stereo vision system for far distance obstacle detection," *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 292 – 297, 2004.

[6] Q. Yu, H. Araujo, and H. Wang, "A stereovision method for obstacle detection and tracking in non-flat urban environments," *Autonomous Robots*, vol. 19, no. 2, pp. 141 – 157, 2005.

[7] P. Lombardi, M. Zanin, and S. Messelodi, "Unified stereovision for ground, road, and obstacle detection," *IEEE Intelligent Vehicles Symposium, Proceedings*, vol. 2005, pp. 783 – 788, 2005.

[8] X. Li, X. Yao, Y. L. Murphey, R. Karlsen, and G. Gerhart, "A real-time vehicle detection and tracking system in outdoor traffic scenes," *Proceedings - International Conference on Pattern Recognition*, vol. 2, pp. 761 – 764, 2004.

[9] G. Toulminet, M. Bertozzi, S. Mousset, A. Bensrhair, and A. Broggi, "Vehicle detection by means of stereo vision-based obstacles features extraction and monocular pattern analysis," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2364 – 2375, 2006.

[10] D. T. J.-P. Labayrade, R.; Aubert, "Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation," *Vehicle Symposium, 2002. IEEE*, vol. 2, pp. 646–651, June 2002.

[11] J. Rebut, G. Toulminet, and A. Bensrhair, "Road obstacles detection using a self-adaptive stereo vision sensor: A contribution to the arcos french project," *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 738 – 743, 2004.

[12] P. Burt, L. Wixson, and G. Salgian, "Electronically directed 'focal' stereo," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 94 – 101, 1995.

[13] http://tev.itc.it/Research.html.