ステレオカメラによる平坦ではない路面の形状検出システムとその応用

A Stereovision-based Non-flat Surface Area Detection Method and Its Application

游　贛梅*　Ganmei YOU
陳　超*　Chao CHEN
師　忠超*　Zhongchao SHI
魯　耀傑*　Yaojie LU
王　剛*　Gang WANG

要　旨

自動クルーズコントロールを目的とする先行車の検出や，衝突回避のための歩行者の認識など，道路上の目標検出技術は，運転支援システムを実現する上で非常に重要である．現在，道路が平坦であれば，その路面形状は精度よく検出できているが，上り下りのある坂道やカーブなどの複雑な路面形状は平坦な場合よりも精度が劣る．本研究では，実際の状況に近い平坦ではない道路における距離情報を利用した路面形状検出方式を提案するものである．この方式では，まず入力画像イメージから複数の視差情報画像を生成し，次にこの視差情報画像から道路の凹凸や断面情報が算出され，更に平坦ではない路面形状が抽出される．実際の路上にて本方法を適用した結果，視差情報画像を利用することで平坦ではない複雑な路面の状況をうまく抽出することができた．

ABSTRACT

In driving assistance systems, on-road object detection is very important for many applications, for example, detecting the preceding car for auto cruise control, recognizing pedestrians to avoid collisions, etc. Existing road surface detection methods could detect flat road areas very well, but in the case of upwards or downwards or sloping or curved road surfaces, the performance is degraded due to the complex nature of non-flat road surfaces.Therefore, in this paper, we propose a general non-flat surface area detection method using stereovision technology since it provides points' distance data that is useful in distinguishing objects. The method is composed of three steps. Firstly, multiple V-disparity maps are generated from the input disparity image, and then the profile and the cross-section points of the surface are estimated from the V-disparity maps. Finally, the non-flat object surface is extracted. We applied this method in road surface detection. Tests show our method can detect complex non-flat surface especially in sparse disparity images.

*　リコーソフトウエア研究所（北京）有限公司　視覚処理技術ラボ
Vison processing laboratories, Ricoh Software Research Center Beijing Co., Ltd.

1. Introduction

Accurate road surface detection is very useful for many fields in advanced driver assistance systems (ADAS), such as off-road warning, roll-over warning, etc. Also, it is the foundation of obstacle detection, object recognition, guardrail detection, etc.

In the real driving environment, sometimes complex road surfaces, such as arching/slanting, convex, concave surfaces and upward/downward surface make accurate surface detection hard. The stereo vision method is a promising way to extract surfaces because we can deduce the surface points using the distance and height data of image pixels that the method provides.

Traditional stereo vision based surface detection methods are “plane fitting” methods and V-disparity map based methods. “Plane fitting” methods adopt a plane equation to fit surfaces. For example, RANSAC plane fitting has been used to detect road surface pixels[1)]; a plane equation has been used to distinguish road pixels from obstacle pixels[2)]. However, it is time consuming to find planes. The V-disparity map based methods exploit the nature of the road V-disparity map to reduce the number of road pixels[3-7)] that fit the slanting line representing the road in the V-disparity map. However, these methods assume the ground is flat. In one method, the non-flat road surface is processed[8)]; however, it does not refer to how to deal with the slanting road cross section. In summary, these methods do not handle the complex non-flat road surface accurately in real-time for driving assistance.

In this paper, we propose a stereo vision based surface detection method that firstly divides the disparity image into regions, then extracts the surface profile piecewise, which indicates surface depth direction in the sub-V-disparity map, and finally fits the region surface plane using the extracted surface cross section and the surface profile of the region. Since the surface may consist of several planes, its profile consists of piecewise lines. Then, the method can detect the horizontal non-flat (slanting/convex/concave) surface and the depth direction non-flat (upward/downward) surface. This also makes the detected surface fine-grained and accurate, that is, the points of detected surface are just those of the realistic surface, not noises or points of other objects. An accurate surface is useful for accurate calculation of the height to the surface for disparity pixels. Our experiments show that the method is effective in real time in different non-flat road scenarios.

The rest of this paper is organized as follows. Section 2 is devoted to system introduction. In section 3, we describe algorithms applied in the system in detail. Section 4 is dedicated to experimental comparison. Finally, we draw our conclusion in section 5.

2. System Overview

The block diagram in Figure 1 shows the main components of the system. The system input consists of the disparity image and the gray image, and the output is the surface in the disparity image.

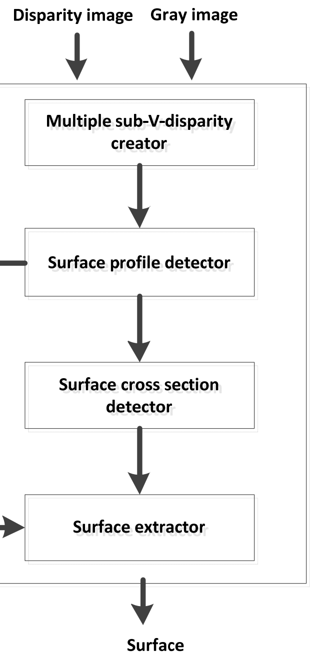

Fig. 1 Block diagram of surface detection.

Disparity is the difference between the column coordinates of the pixel locations of the corresponding pixels in the left and right image of a stereo pair. It is inversely proportional to the distance from the pixel location to the plane of the cameras. We can deduce the distance from the disparity. A disparity image is an image where one pixel has a disparity value. The sub-disparity image is part of the disparity image. A gray image is an image in which the value of each pixel is the intensity information of light. A V-disparity map/image is an image that accumulates points with the same disparity for each horizontal line of the disparity image. A sub-V-disparity map/image is the V-disparity map of part of the disparity image.

The multiple sub-V-disparity creator firstly divides the disparity image into sub-disparity images and then constructs a sub-V-disparity map for each sub-disparity image. Multiple sub-V-disparities and sub-disparity images are the outputs. This module enables the following detection of surface profile and the surface depth direction of fine-grained and accurate surfaces. The surface profile detector fits the surface profile in each sub-V-disparity map piecewise so that in each piece the surface profile line is a straight line. The disparity pixels in the corresponding sub-disparity image whose distances are in the range and whose height to the surface profile lines are in a threshold form a surface segment. Therefore, there is only one surface depth direction in one surface segment. Finally, the surface profile lines in the sub-V-disparities are output to denote surface depth directions and surface segments. The surface profile lines also show the surface position in the depth direction. In this way, an upward/downward surface can be detected accurately. The surface cross-section detector firstly divides each surface segment into such a sub-segment/region so that the surface is straight in each sub-segment/region. The disparity image in the sub-segment/region is called the sub-segment/region disparity image. Secondly, it transforms the sub-segment/region disparity image into a real world 3-dimensional image and projects the sub-segment/region along its depth direction to a plane Z to get its cross section. Thirdly, it extracts the surface projection and fits its cross-section. Finally, the regions and the surface cross-sections are output. In this way, we can detect the slanting/arching/concave/convex cross-section surface. The surface extractor constructs the surface by fitting a plane in each sub-segment/region based on surface profile and surface cross-section.

3. Stereovision-based Non-flat Surface Area Detection

In a disparity image, each pixel is represented by (u, v, d); (u, v) is an image coordinate, and d is the disparity value. Since the sparse disparity images calculated from the left and right images of the stereo cameras can increase the whole system speed by reducing the disparity calculation and object detection, our system uses sparse disparity images. Therefore in this section, we will use them to show our results.

3-1 Multiple sub-V-disparity maps creator

Firstly, the disparity image is divided into sub-disparity images horizontally. If there are some markings on the surface such as lanes, zebra crossings, etc., we can detect them and then divide the surface based on them. If there is no marking, we simply divide the surface based on the horizontal distance. For example, the disparity in Fig. 2 can be divided into two sub-disparity images. Secondly, a sub-V-disparity map is generated for each sub-disparity image. Figure 2 shows the process for the road surface. As shown in figure 2, firstly, the lane markings are detected in the gray image. Secondly, the trapezoid areas of interest (ROIs) in the disparity image based on the land marking position are determined. Thirdly, a sub-V-disparity map for each ROI is created.

3-2 Surface profile detector

In this part, the surface profile in each sub-V-disparity map is detected piecewise. Surface segments and surface profile piecewise lines are output to indicate surface depth direction for each sub-V-disparity map.

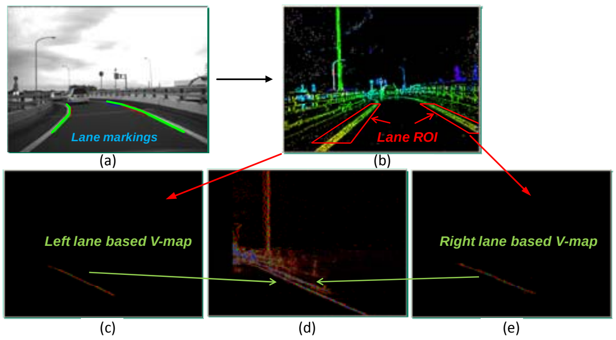

Fig. 2 Multiple sub-V-disparity maps creator for road surface, (a) detected lane markings (green lines) in gray image, (b) ROI in disparity image, (c) sub-V-disparity map of left sub-disparity image, (d) V-disparity map of whole image and lines corresponding to those in (c) and (e), (e) sub-V-disparity map of right sub-disparity image.

> Surface profile piecewise line estimation algorithm:
>
> a) Select the nearest points in k meter distance,
>
> b) Fit a line for points from a), get line $L1$ $v = k_v \cdot d + b_v$, and remove the nearest points,
>
> c) Calculate the v/vertical distance of the remaining points to line $L1$,
>
> d) Extend $L1$ to the distance of the adjacent points whose v/vertical distance to line $L1$ is smaller than the pre-defined distance,
>
> e) Remove fitted points from the sub-V-disparity map.
>
> f) Repeat a) – e) until there are no points left.

The surface profile lines indicate the surface depth direction. However, the surface projection in each sub-V-disparity map may not be just a line. It may include points around the line since the surface cross-section may be slanting/arching, concave, or convex. The surface cross-section detection will be discussed later.

Take the gray image and its disparity image in Figure 3 for example. The left lane marking is detected, and the corresponding ROI is extracted. Then, the sub-V-disparity of the ROI is created as shown in Figure 4 (b). If the surface is upward/downward, the surface profile is piecewise. The algorithm of fitting such a Z/depth non-flat surface is as follows.

In step a), the number of selected points should be larger than the specified number, and the density (point number per meter) of the selected points should be larger than a predefined threshold so that the following line fitting is reliable.

Figure 4 shows the process of the algorithm. (c) shows the selected points of step a); the fitted first segment of the surface line in (d) is the result of step b); (f) extended surface line based on the distance of the point to the line (red line) is the output of step c) and d); (g) the fitted second segment of the surface line (red line) repeats step b); (h) surface profile piecewise lines (red lines) and end of each piece (green lines) is the surface profile, which is the final result of the algorithm.

Each surface segment corresponds to a surface plane, which may be horizontal or upward or downward in the Z/depth direction. In each surface segment of the sub-V-disparity map, the surface profile line is $v = k_v \cdot d + b_v$.

Fig. 3 Gray image (upper) and its disparity image (lower).

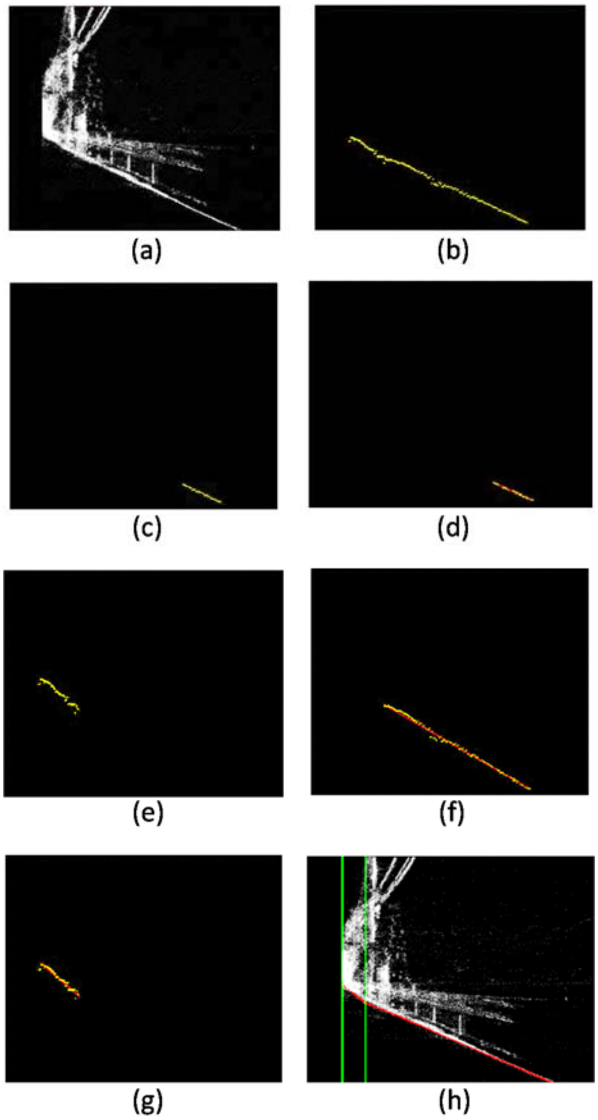

Fig. 4 Process of fitting surface line piecewise, (a) V-disparity map, (b) sub-V-disparity map, (c) points selected based on depth distance, (d) fitted first segment of surface line, (e) remaining points, (f) extended surface line based on distance of point to line (red line) (g) the fitted second segment of surface line (red line), and (h) surface profile piecewise lines (red lines) and end of each piece (green lines).

3-3 Surface cross-section detector

For each surface segment obtained from the sub-V-disparity map, we will detect its cross-section.

In the real world, the surface in each surface segment gotten from the surface profile detector may be non-flat in the horizontal direction: arching/slanting, convex, or concave on surface cross section. Therefore, in this section we will discuss how to get the surface cross-section in each surface segment. The algorithm is as follows.

Surface segment cross-section detection algorithm:

a) Divide the surface segment into straight-lane sub-segments/regions.
b) Compute the surface vanishing point (SVP) of the region.
c) Move the horizontal coordination (u) of the SVP to the image center, transform the position of the disparity points in the region according to its relative position in relation to the new position of the SVP, and get surface orientation aligned disparity region.
d) Transform the surface orientation aligned disparity region into a real world 3D image.
e) Compute real world direction of each segment.
f) Project the real world image of each sub-segment to a plane Z = max_D along the segment real world direction and get real world surface cross-section projection of region.
g) Fit real world surface cross-section projection of each region.
h) Compute surface cross-section in disparity image of each region.

In step a), in each surface segment, we divide the surface in the Z/depth direction into sub-segments so that the surface in each sub-segment is straight and then get the straight surface sub-segment.

In step b), we calculate the sub-segment's surface vanishing point (SVP) in each sub-segment/region. For example, in a road scenario, we extend the detected parallel left and right lanes to get their intersection point in the disparity image, which is the vanishing point SVP (u_{vp}, v_{vp}, 0) of the sub-segment.

In step c), the surface vanishing point is moved to SVPC (w/2, v_{vp}, 0), and the coordinates of all the disparity points in this region are changed according to their relative distance to the SVP and surface orientation aligned disparity region. The surface orientation aligned disparity region enables proper object projection to the plane Z = max_D (max_D is the maximum distance in the region) in step f).

In step d), for each disparity pixel *(u', v')* in the surface orientation aligned disparity region, we calculate its horizontal, vertical, and depth distance to the vanishing point and get its 3-dimensional coordinate *(x', y', z')* in the real world base on equation (1), (2), and (3). Then we transform the disparity image of each straight-lane surface sub-segment into the real world image.

$$x' = \frac{(u - u_{vp}) \cdot b}{d} \quad (1)$$

$$y' = \frac{(v - v_{vp}) \cdot b}{d} \quad (2)$$

$$z' = \frac{b \cdot f}{d} \quad (3)$$

b is the pixel distance between two lens, d is pixels' disparity value, and f is the camera focal value.

In step e), we convert the sub-segment surface depth direction in the V-disparity map ($v = k_v \cdot d + b_v$, d is disparity) to that in the real world ($y = k_r \cdot D + b_r$, D is real distance).

In step f), we project pixels of the sub-segment image along the real world surface depth direction to plane Z = max_D, where max_D is the maximum distance in the sub-segment to get the real world cross-section projection. For a real world pixel *(x', y', d')*, we get its projection *(x_p, y_p,* max_D*)* according to equation (4) and (5).

$$x_p = x' \quad (4)$$

$$y_p = y' - k_r (b \cdot f / d' - \max\_D) \quad (5)$$

In step g), we firstly cluster pixels in the real world cross-section based on horizontal distance. Secondly, we extract the surface part based on height and width. If the cluster's height is higher than a predefined threshold (*H*) or its width is larger than a predefined threshold (*W*), the cluster is not a surface cluster. Finally, we fit the surface cluster points and the lowest points of the non-surface cluster to get the whole surface cross-section projection.

In step h), we firstly trace the surface pixels in the real world image based on the relationship in (4) and (5). Secondly, we trace the surface pixels in the disparity image based on the relationship in (1), (2), and (3). Finally, we get the surface cross-section at the plane Z = *max_D*. Equation (6) and (7) represent this.

$$v = f_{road}(u, d) \quad (6)$$

$$d = md \quad (7)$$

md is the disparity of distance *max_D*.

The relationship between the surface in the real world and the surface profile in the V-disparity is shown in Fig. 5. The schematic diagram of the surface segment cross-section detection algorithm is shown in Fig. 6. An example of surface cross-section detection is shown in Fig. 7.

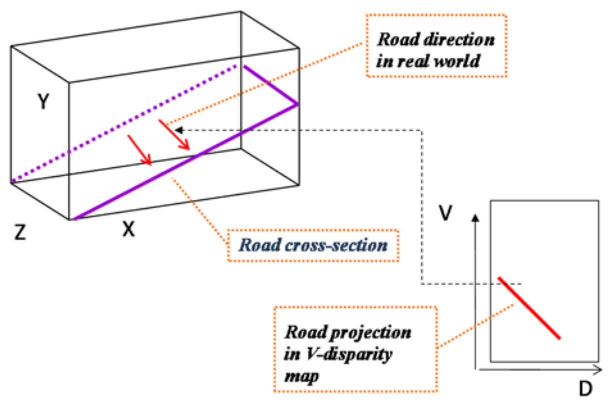

Fig. 5 Relationship between surface in real world and surface profile in V-disparity in road scenario.

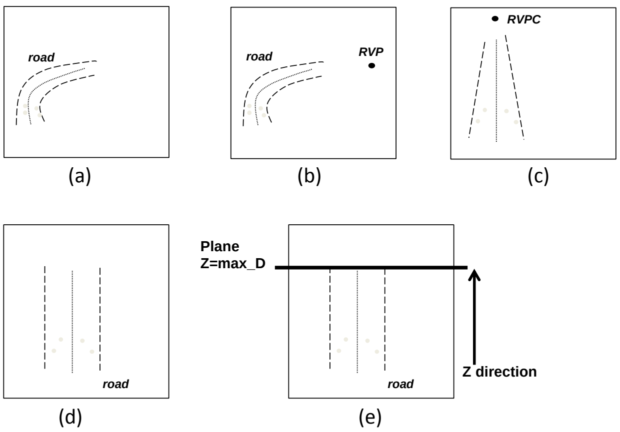

Fig. 6 Schematic diagram of surface cross-section detection in road scenario, (a) disparity image, (b) SVP in disparity image, (c) surface orientation aligned disparity image, (d) top view of real world image, and (e) projection to the plane Z = max_D.

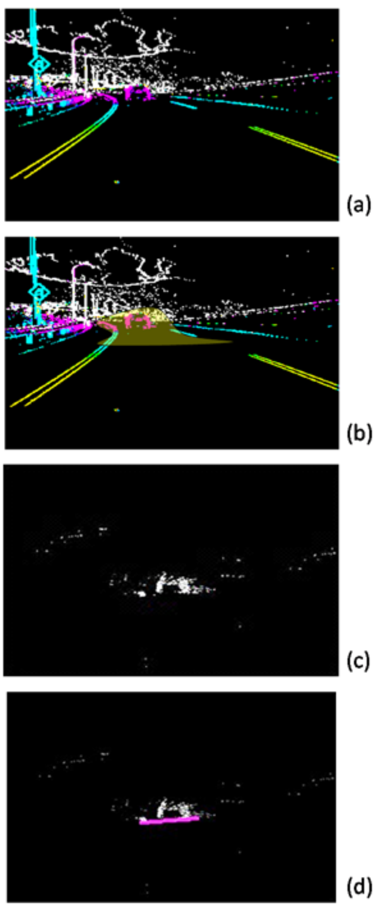

Fig. 7 Example of surface cross-section detection in road scenario, (a) disparity image, (b) sub-segment of surface segment of disparity image, (c) projection of sub-segment, and (d) fit surface projection.

3-4 Surface extractor

In each region, based on surface line fitting in sub-V-disparity map (Z/depth direction) and cross section fitting in each sub-segment (U/horizontal direction), we fit a surface plane for each region as follows.

$$v = k_v \cdot d + b_v + \Delta v \tag{8}$$

$$\Delta v = f_{road}(u, md) - (k_v \cdot md + b_v) \tag{9}$$

Δv is vertical distance of surface point to the surface profile line, which is the same for all the points at different distance in the region. Based on equation (8) and (9), we deduce equation (10) which shows the plane for the surface in the region.

$$v = k_v \cdot d + b_v + f_{road}(u, md) - (k_v \cdot md + b_v) \tag{10}$$

In this way, we get all region surface planes.

3-5 Height to surface calculation of each disparity point

For each disparity point p' (u', v', d') in the disparity image, its height to surface is calculated as (11).

$$h = (v_{road} - v') \cdot b / d' \tag{11}$$

Based on (10) and (11), we deduce (12). Then, we can calculate the disparity point's height to surface according to (12).

$$h = (k_v \cdot d + b_v + f_{road}(u', md) - (k_v \cdot md + b_v - v')) \cdot b/d' \tag{12}$$

b is the pixel distance between two lens, and *b/d* means at disparity *d* how much one pixel represents in real distance.

When the height to surface of the point is almost zero, the point is a surface point.

The height to surface of each disparity point is very useful in detecting on-surface objects such as vehicles and pedestrians.

4. Experimental Results

The environment of the implementation is shown in Table 1. The images are captured by a 3D camera with depth information. Each image is 1280 × 960. We test on an AM3517 EVM board with 600 MHz and PC with P8600@2.4G CPU, 1.89G memory, and Windows XP SP3 OS. We test our method using different samples among the dataset, including flat road, upward/downward road, and slanting road. We compare the detection results of our method with those of the classic V-disparity map method. Some samples are shown in Fig. 8. The experiment results show that 1) when the road surface is flat, the detection results by both methods are almost the same and 2) when the road surface is slanting in the horizontal direction, our method can extract more road surface disparity pixels than the classic V-disparity map method, therefore our method is superior. The speed test result in Table 1 shows that our method is fast and can process in real time.

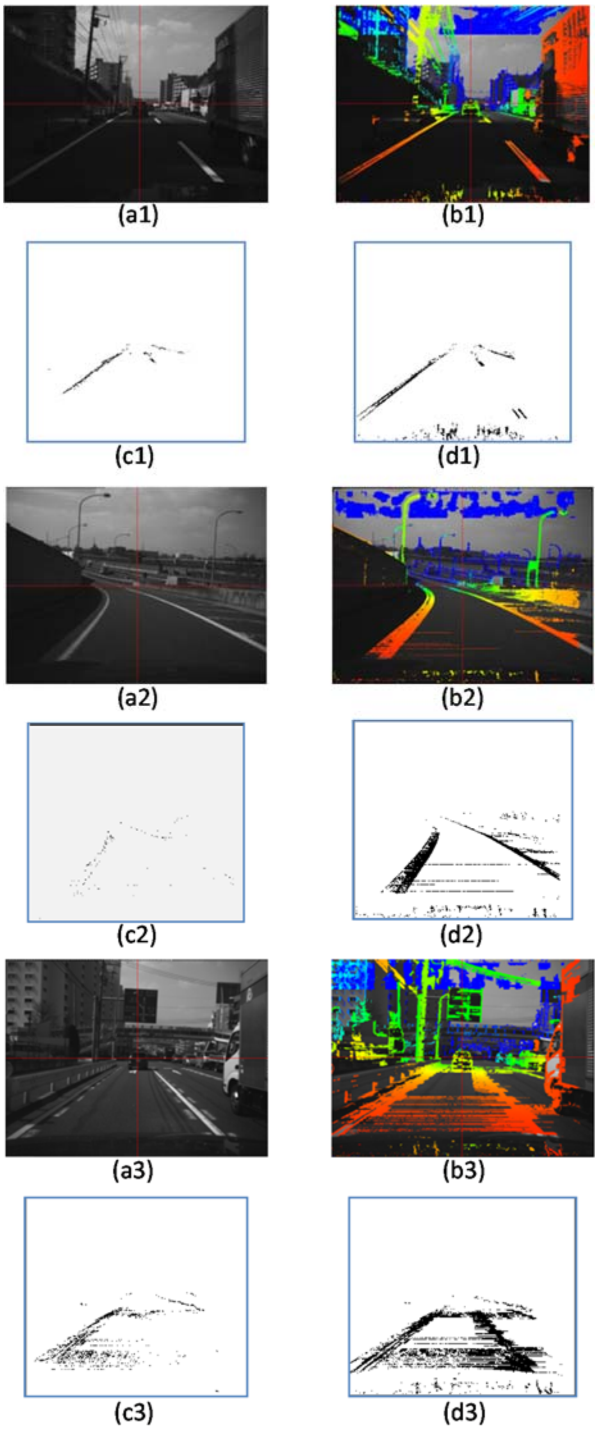

Fig. 8 (c1), (c2), and (c3) show road surfaces extracted by traditional V-disparity method; (d1), (d2), and (d3) are road surfaces extracted by our method. (a1)-(d1) are flat road surface detection, (a1) gray image of flat road surface, (b1) sparse disparity image shown on gray image, colored pixels are disparity pixels, different colors indicate different distances, (a2)-(d2) are slanting in the horizontal and upward road surface detection, and (a3)-(d3) are downward road surface detection.

Table 1 Test speed.

Road detection	PC			Board		
	Average	Max	Min	Average	Max	Min
Processing time/ms	6.77	12.13	5.00	17.25	19.10	15.43

5. Conclusion

We have proposed a stereovision based method for the non-flat surface detection. The method divides the disparity image in the horizontal direction. Then it detects surface profiles in each sub-V-disparity image and surface cross section in sub-disparity map to extract non-flat surfaces. The experiments show that our method is effective in extracting complex road surface especially from sparse disparity images in real time. In the future, we will do further research in more complex surface situations such as convex/concave surfaces.

References

1) Q. Yu, H. Araujo, 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).

2) P. Lombardi, M. Zanin, S. Messelodi: Unified stereovision for ground, road, and obstacle detection, *IEEE Intelligent Vehicles Symposium, Proceedings*, Vol. 2005, pp. 783-788 (2005).

3) G. Toulminet et al.: 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).

4) R. Labayrade, D. Aubert, J. Tarel: Real time obstacle detection in stereovision on non-flat road geometry through “V-disparity” representation, *Vehicle Symposium, 2002 IEEE*, Vol. 2, pp. 646-651 (2002).

5) J. Rebut, G. Toulminet, 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).

6) GB. Vitor et al.: A 2D/3D Vision Based Approach Applied to Road Detection in Urban Environments, *Intelligent Vehicles Symposium (IV), 2013 IEEE*, Australia, pp. 952-957 (2013).

7) C. Oh, B. Kim, K. Sohn: Automatic Illumination Invariant Road Detection with stereo vision, *7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2012, pp. 889-893 (2012).

8) J Zhao, M Whitty, J Katupitiya: Detection of non-flat ground surfaces using V-Disparity images, *IROS 2009*, pp. 4584-4589 (2009).