

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: $d(u, v) \leq au + bv + c + \tau(d)$. 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-DISPARIITY 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

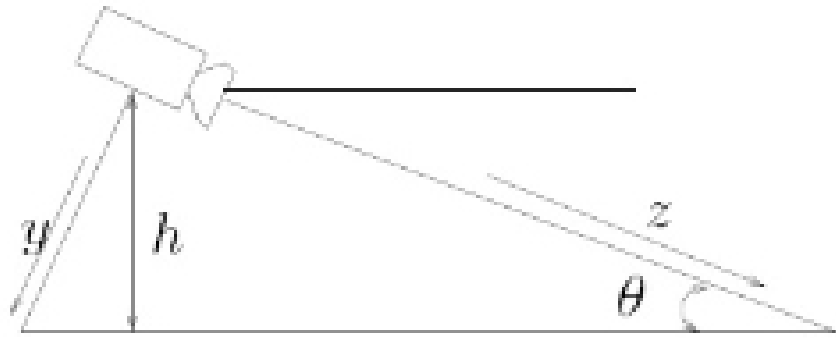


Fig. 1. Camera placement

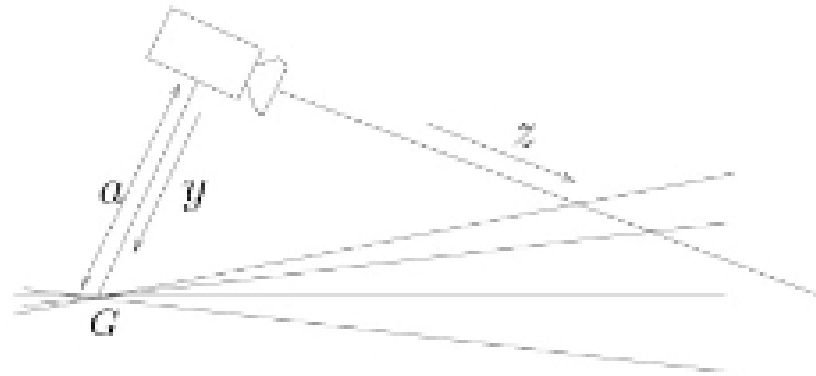


Fig. 2. Conditions for the ground correlation lines to be parallel: a is fixed, all ground planes pass through point G

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

$$z = \frac{h}{\sin \theta} - y \frac{\cos \theta}{\sin \theta} \quad (1)$$

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

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

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

$$v = \frac{hw}{b \cos \theta} - f \tan \theta \quad (3)$$

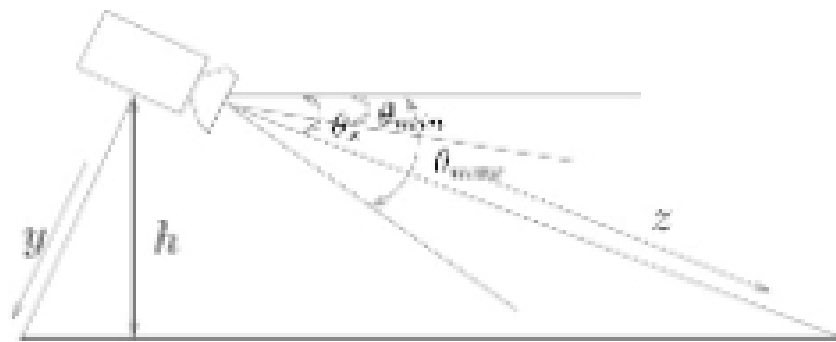


Fig. 3. Pitch variation

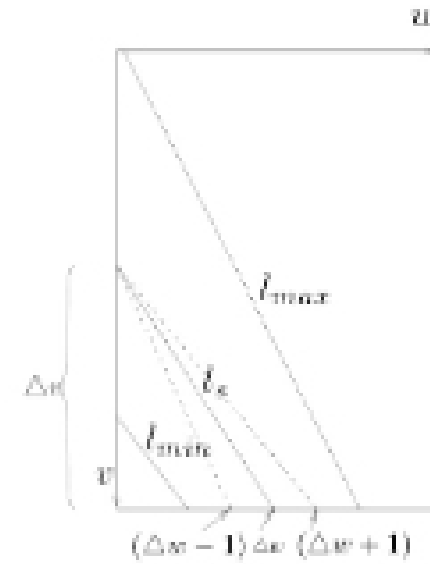


Fig. 4. Slope of ground correlation line during pitch variation. The solid slanted lines represent ground correlation lines. The dashed lines represent the next distinguishable gradient near l_s

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

$$g = \frac{\partial v}{\partial w} = \frac{h}{b \cos \theta} \quad (4)$$

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

$$\Delta g = \frac{dg}{d\theta} \Delta \theta = \frac{h}{b} \sec \theta \tan \theta \Delta \theta \quad (5)$$

From Eq. 5 we can get

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

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

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

then we should have

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

from which we can get

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

similarly we can get

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

Combining Eq. 9 and 10, we get,

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

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 Δw 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.

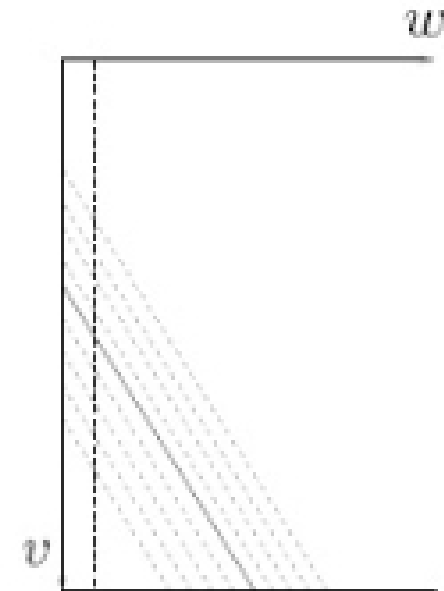


Fig. 5. Ground correlation lines in V-disparity image. The solid slanted line corresponds to the ground correlation line obtained using the static calibration data, while the dashed slanted lines are the ones expected varying the pitch value. The solid vertical line indicates the 0 disparity value, the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