Ground Plane Stereo for Obstacle Detection

Kwang Hee Won and Soon Ki Jung
School of Computer Science and Engineering
Kyungpook National University
1370 Sankyuk-dong, Buk-gu, Daegu, 702-701, South Korea
Email: khwon@vr.knu.ac.kr, skjung@knu.ac.kr

***Abstract*—We propose a dense stereo matching algorithm for obstacle detection by using the ground plane assumption in which the scene includes objects standing on the ground plane only. The algorithm produces dense disparity map and its zero-valued pixels indicate the ground plane of the input image. The algorithm refines disparity map by filtering out the patterns of disparity values, which violate the ground plane assumption. Our algorithm can produce accurate disparity map for a given scene than the normal stereo algorithms with limited disparity budget and the disparity map is directly applicable to the obstacle detection task.**

I. Introduction

The "Smart Car" became an important keyword not only in the vehicle industry but also in the computer vision field. It is because the scene needs to be analyzed properly and precisely to notify any event or to augment additional information on the driver's view for safe driving. So, cameras or range sensors are equipped on the vehicle to enlarge the view of drivers and to detect obstacles near the car.

Among the sensors, the stereo vision sensor is the representative passive range sensor. It is useful because the stereo produces 2.5D range data and images simultaneously. The resulting range data or the disparity map can be exploited for tasks such as obstacle detection or path planning.

The typical stereo matching algorithms assume the zero-disparity plane is parallel to one of the image planes. However, this assumption may not be effective under the observation that the vehicles and other objects necessarily stand or move around on a certain ground plane. The ground plane can be a good alternative as the zero-disparity plane for several reasons. First, it produces better (more discriminative) matching costs for horizontal surfaces [1]. Second, the ground plane based disparity can represent the longer distance on the road with fixed disparity range as shown in Fig. 1. For the next, many applications compute the ground plane from range image to detect obstacles and to compute the height of those objects but ground plane based disparity can be utilized directly to those applications. The bad pixels can easily be removed by evaluating whether the pixel is a part of the ground-standing object or not.

In this paper, we propose a dense stereo matching algorithm based on the ground plane. The algorithm produces the disparity map which has zero-disparity on the ground plane of a given scene. The proposed algorithm effectively filters out the matching failures using the ground plane scene assumption.

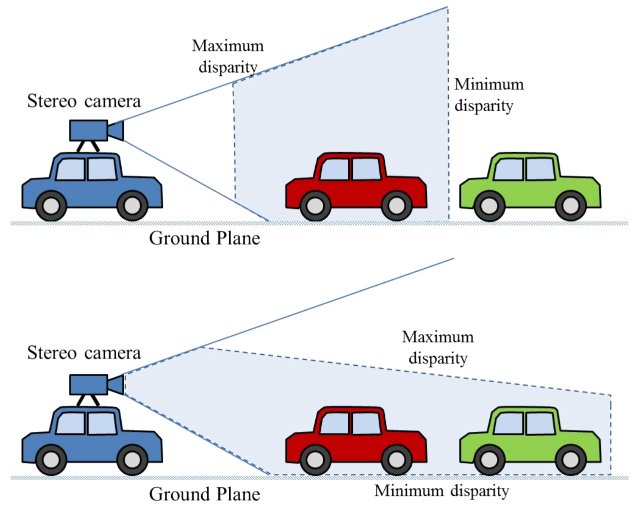


Fig. 1. Depth perception with limited disparity budget: (first row) conventional stereo matching algorithm, and (second row) ground plane stereo matching algorithm.

This paper consists as follows. In section 2, we introduce the previous researches about the ground plane based stereo algorithms. In section 3, we describe the ground plane based stereo matching algorithm. In section 4, we show some experimental results, and in section 5, we conclude this paper.

II. Related Works

Many literatures about object detection and unmanned automatic vehicle researches introduced the utilization of ground plane homography.

Arróspide et al. made use of ground plane homography of single-view consecutive images to detect moving objects on the road [2]. They obtained the homography by feature matching with prior knowledge. Se and Brady extracted edges from stereo images and matched them to detect objects on the ground [3]. They employed Kalman Filter to trace the edges, and estimated the ground plane using RANSAC. Similarly, Xie extracted objects by classifying the overlapped contours obtained by applying ground plane homography to one of stereo images [4].

We try to find ground plane based dense disparity map instead of feature correspondences. Our algorithm employs the Semi-Global Matching algorithm [5] to aggregate matching costs and finds globally optimal disparity values. The dense

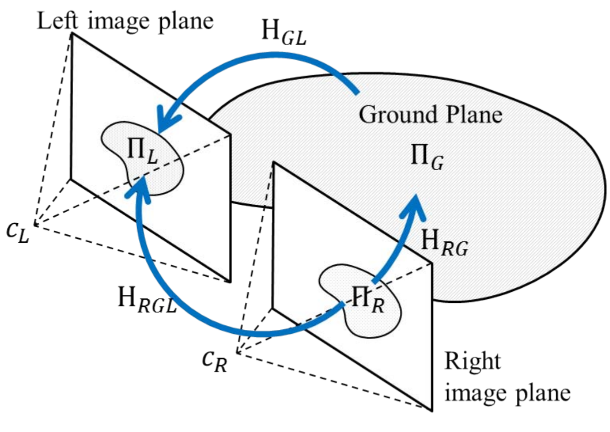


Fig. 2. Ground plane homography.

disparity map is useful because some patterns (such as the increasing disparity from the ground plane) can be applicable to detect objects.

Williamson and Thorpe compared the matching costs of horizontal and vertical stereo matching costs and extracted objects from them [1]. They made groups of pixels that have the similar depth values and location, and filtered out a small sized groups. Our algorithm discards bad pixels by validating whether the pixel is a part of standing objects or not. After that, we detect each connected component on the disparity map as objects.

III. Ground Plane Stereo Matching

Fig. 2 shows the relationship between the stereo camera system installed on a vehicle and the ground plane Π_G. Π_L and Π_R represent the projected plane of Π_G on each image plane, respectibly. Then the relation between Π_L and Π_R respect to the Π_G can be denoted by a single homography, H_{RGL} as denoted in equation (1).

$$H_{RGL} = H_{GL} \times H_{RG}, \tag{1}$$

where H_{RG} and H_{GL} are the homography from Π_R to Π_G, and from Π_G to Π_L, respectively. Fig. 3 shows an overlapped example of a left image and its corresponding H_{RGL}-applied right image. In this ground plane stereo geometry, the matching cost of a pixel p with disparity value d is represented as follows,

$$C(p, d) = |I_L(p) - I_R(H_{RGL}(p) + d)|, \tag{2}$$

where I_L and I_R are the intensity values of the given pixel.

The matching costs are computed for all disparity range and aggregated by the equation from the Semi-Global Matching algorithm (in short, SGM) [5]. SGM aggregates matching cost along various directions instead of 2D global optimizations, using the following equation [6], [7],

$$\begin{aligned} L_r(p, d) &= C(p, d) \\ &+ \min\Big(L_r(p_{-1}, d), L_r(p_{-1}, d \pm 1) + P_1, \\ &\min_i L_r(p_{-1}, i) + P_2\Big), \end{aligned} \tag{3}$$

where $L_r(p, d)$ is an aggregated matching cost of pixel p and disparity d. p_{-1} represents the precedence pixel along the direction of aggregation. P_1 and P_2 are the penalty for the small (± 1) and large $(1 <)$ disparity changes. The P_2 is determined adaptively according to the difference of pixel values from P_1 to P_2.

The optimal disparity is then determined as the disparity value that has the smallest sum of aggregated costs for all directions. After the initial disparity map is computed, the median filter is applied to remove the flicking noise. For the next step, we filter out the patterns of disparity values that do not fit the ground plane assumption. In outdoor environment, artifacts can be produced for various reasons such as the different exposure, motion blur (especially the closer part of the scene), the occlusion, asynchronous acquisition and so on.

The filtering rules are simple: (1)if the disparity value of a certain pixel is less than ϵ, then the pixel is accepted (pixels on the ground plane), and (2)if the disparity value is larger than ϵ, there must be any neighbor that is connected to the pixels of (1). These two rules can be represented by the recursive definition of δ function as follows,

$$\delta(p) = \begin{cases} \text{if } D(p) < \epsilon, & 1, \\ \text{if } D(p) \geq \epsilon \ \& N_p = \phi, & 0, \\ \text{otherwise} & \max_i\{\delta(q_i \in N_p)\}, \end{cases} \tag{4}$$

where $D(p)$ is the determined disparity of pixel p and N_p is a set of non-visited neighboring pixels within radius r and acceptable difference of disparity value t. We show the example result of refinement using ground plane constraint in Fig. 5. The equation 4 is represented in a recursive form, but we implemented it using multiple scans and merge for the computational efficiency. The obtained disparity map can

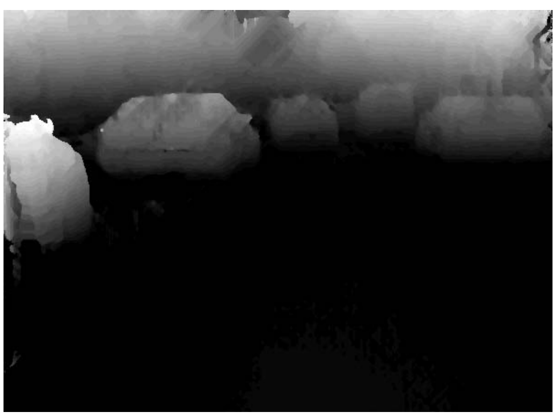
Fig. 3. An overlayed left image and H_{RGL}-applied right image.

be transformed to the traditional vertical plane based disparity map by inversely applying the homography relation in equation (2). Otherwise, it can be used directly by finding components using the discontinuity of disparity values on the object boundary.

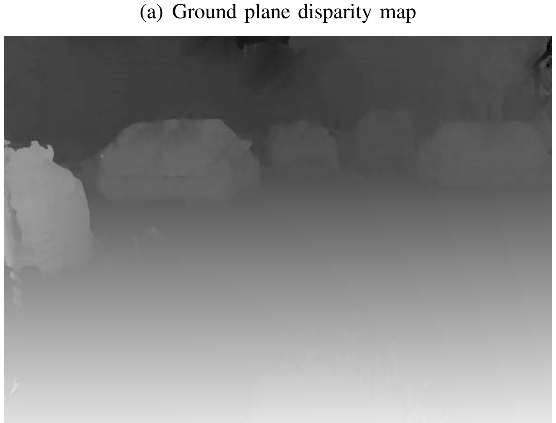
IV. Experimental Results

We used a stereo camera system which includes two CCD cameras, with 6 mm C-mount lens on a 200 mm-horizontal rig. Then, we obtained stereo video from a moving vehicle and computed its disparity maps. The resolution of the input is 640-width and 480-height and the 20-pixel boundaries of each side are discarded.

We first compared the disparity map of ground plane stereo and that of vertical plane based stereo matching. For the experiment, we implemented our method and the SGM. The ground plane homography is obtained manually by assigning the correspondances of pixels on the ground plane. The automated acquisition of the precise ground plane homography is also an important issue but it is out of scope for this paper. The left-right consistency check is omitted in both implementations. The penalty parameters P_1 and P_2 are adjusted experimentally to produce the best disparity map. Fig. 4 shows an example of disparity map resulted from the ground plane based and vertical plane based methods.



(a) Ground plane disparity map



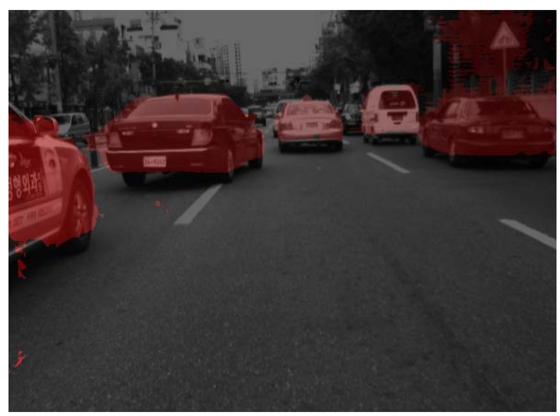
(b) Transformed disparity map

Fig. 6. Ground plane based dispariy map and transformed disparity map.

The 4(c) and 4(d) show that the SGM requires more wide disparity ranges for the same input scene. The ground plane stereo matching produces clear ground plane with objects which are distinguishable from the ground and other backgrounds. Fig. 6 shows the transformed disparity map from the ground plane based one. The recovered disparity values are about 90-levels. Fig. 7 shows more results with various ground plane scenes with vehicles, pedestrian, and other obstacles. The same levels of disparity range with Fig. 4 are used.

Finally, we illustrate the detected objects by removing the ground plane and trim off the background objects which are far enough in Fig. 8.

(a) Detected object regions.



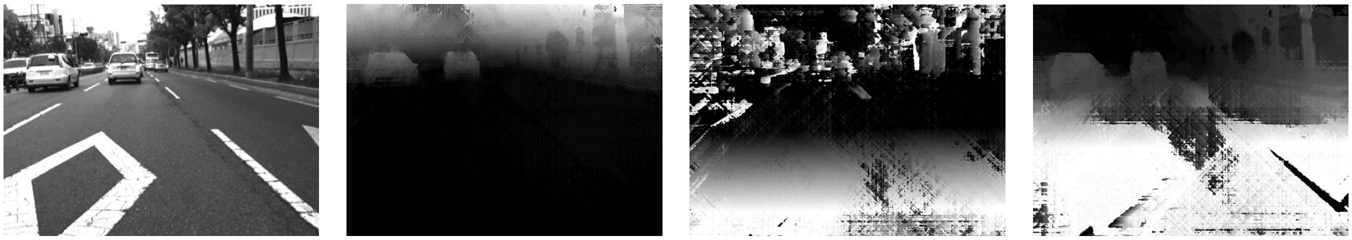
(b) Overlayed objects on the left image.

Fig. 8. Object detection using ground plane disparity map.

V. Conclusion

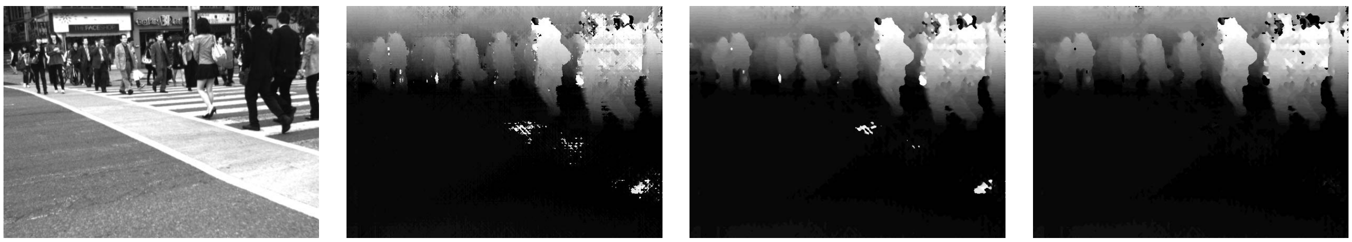
In this paper, we proposed a dense stereo algorithm to produce ground plane based disparity map. We also proposed a filtering method for the abnomal disparity patterns on the ground plane scene. The proposed algorithm produces more abundant disparity on a typical vehicle scenes and the disparity can directly be used for an object detection task.

The ground plane based disparity map provides good estimation of the ground plane, but it produces rather weak clues for the vertical surfaces than SGM. In the future, we will combine the strong points of both methods to produce



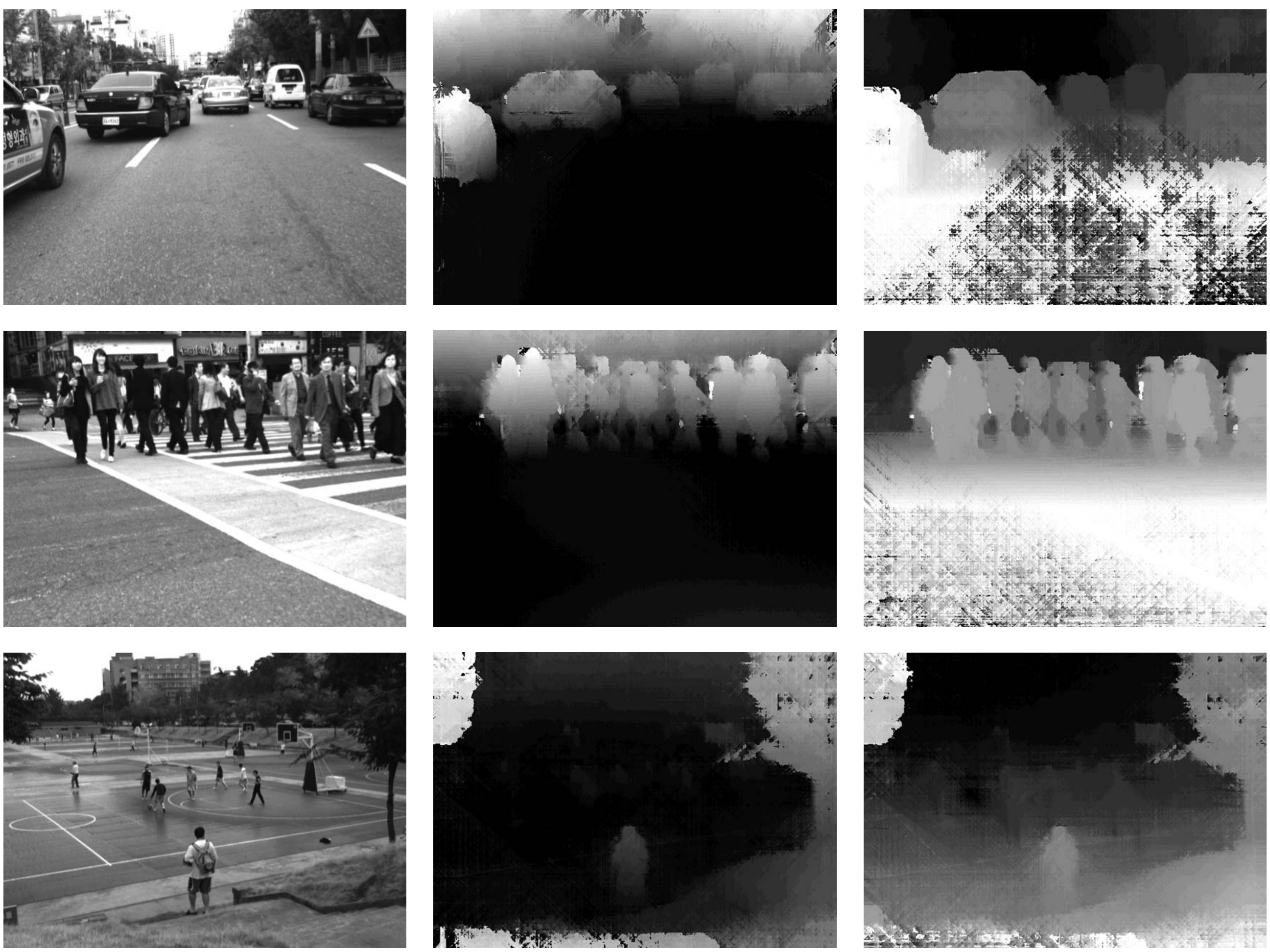
(a) Left image. (b) Ground plane based stereo matching (32-level disparity). (c) SGM (32-level disparity, large disparity values). (d) SGM (32-level disparity, medium disparity values).

Fig. 4. Ground plane stereo matching: 4(a) is the input image. 4(b) is the ground plane based disparity map which is obtained using 32-level disparity values. 4(c) and 4(d) show the 32-level disparity range cannot represents the whole scene with the traditional representation of disparity map.



(a) Left image. (b) Initial disparity map. (c) Median filter (5×5) applied. (d) Equation (4) applied.

Fig. 5. Disparity map refinement: bad pixels that violate the ground plane constraint are removed.



(g) Left image. (h) Ground plane stereo matching. (i) SGM with 32-level disparity.

Fig. 7. Ground plane scene with various objects and the corresponding disparity maps.

better disparity map, which facilitates as important evidence for object detection.

References

[1] T. Williamson and C. Thorpe, "Detection of small obstacles at long range using multibaseline stereo," in *Proceedings of the IEEE International Conference on Intelligent Vehicles*, 1998.

[2] J. Arróspide, L. Salgado, M. Nieto, and R. Mohedano, "Homography-based ground plane detection using a single on-board camera," *IET Intell. Syst.*, vol. 4, no. 2, pp. 149–160, 2010.

[3] S. Se and M. Brady, "Stereo vision-based obstacle detection for partially sighted people," in *In Proc. Third Asian Conference on Computer Vision*, 1998.

[4] M. Xie, "Matching free stereovision for detecting obstacles on a ground plane," *Machine Vision and Applications*, vol. 9, pp. 9–13, 1996.

[5] H. Hirschmüller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Proc. CVRP*. IEEE Computer Society, 2005, pp. 807–814.

[6] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," in *in Proc. ICCV*, vol. 2, 2001, pp. 508 –515.

[7] J. Sun, Y. Li, S. Bing, and K. H. yeung Shum, "Symmetric stereo matching for occlusion handling," in *CVPR*, 2005, pp. 399–406.