

# Obstacle Detection in Outdoor Scenes based on Multi-Valued Stereo Disparity Maps

Qian Ge\* and Edgar Lobaton\*

\*Department of Electrical and Computer Engineering  
North Carolina State University, Raleigh, North Carolina 27695–7911  
Email:qge2@ncsu.edu, edgar.lobaton@ncsu.edu

**Abstract**—In this paper, we propose a methodology for robust obstacle detection in outdoor scenes for autonomous driving applications using a multi-valued stereo disparity approach. Traditionally, disparity maps computed from stereo pairs only provide a single estimated disparity value for each pixel. However, disparity computation suffers heavily from reflections, lack of texture and repetitive patterns of objects. This may lead to wrong estimates, which can introduce some bias on obstacle detection approaches that make use of the disparity map. To overcome this problem, instead of a single-valued disparity estimation, we propose making use of multiple candidates per pixel. The candidates are selected from a statistical analysis that characterizes the performance of the underlying matching cost function based on two metrics: The number of candidates extracted, and the distance from these candidates to the true disparity value. Then, we construct an aggregate occupancy map in  $u$ -disparity space from which obstacle detection is obtained. Experiments show that our approach can recover the correct structure of obstacles on the scene when traditional estimation approaches fail.

## I. INTRODUCTION

Obstacle detection is a fundamental step for autonomous navigation. Correctness is of particular interest in applications areas such as autonomous driving systems [1], [2] for which safety is of big concern. However, obstacle detection often depends on reliable depth estimates. Stereo depth maps estimates have gained some popularity in the transportation area due to the affordable hardware and richness of contextual information. Nevertheless, the outcome from state of the art algorithms is not always reliable due to a number of visual artifacts. The stereo matching step, which requires finding correspondences between pairs of rectified images, is a challenging task due to reflections, lack of texture, weather conditions, and repetitive patterns present in objects. Most approaches [3], [4], [5] rely on the computation of a matching cost from which the minimum is often taken as a possible correspondence between stereo pairs. However, due to the nature of this problem, the true matches may not fall on the minimum of this function but in a local minimum instead. This can lead to incorrect matches. Traditional approaches may also include some indicator function of the correctness or certainty of the match found, which in some cases is used to remove possible mismatches. This leads to depth values being

\*This work was supported by the National Science Foundation under award CNS-1239323.

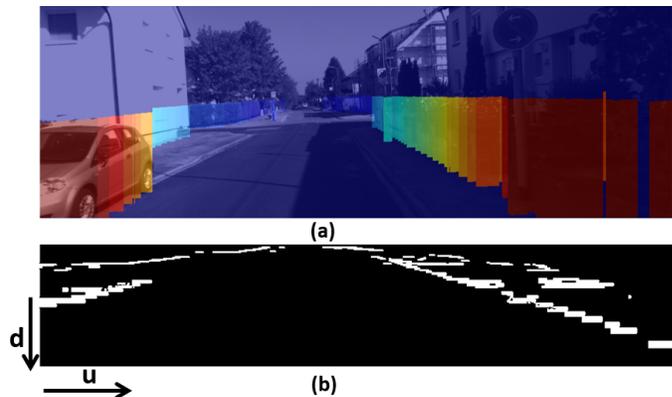


Fig. 1. (a) Environment model. Different colors denote different depth of obstacles detected. Only obstacles above the ground plane and below a maximum-height plane are detected. (b) Corresponding obstacle detection in  $u$ -disparity space. Images used for analysis are from the KITTI dataset [6], [7], [8].

missing, which can also cause a problem in the identification of obstacles.

There are a number of approaches aiming to make disparity computations more robust [4], [9], [5] and extracting obstacle detections from these disparity results in a robust way [10], [11]. Our approach attempts to construct a more reliable occupancy map from an alternative representation of the disparity map, which keeps track of multiple-candidate disparity values. We rely on existing disparity computation algorithms to extract these candidates, and our occupancy map can be used as an input for approaches such as Stixels [10] for robust obstacle detection. In order to capture correct matches on the disparity computations, we make use of multiple-candidates associated with different local minima of the matching cost function instead of relying on a single estimate. These candidates are obtained from an statistical analysis of the performance of the disparity algorithm being used in order to guarantee that we capture the correct structure in our representation with high-confidence. This set of candidates is then used to construct an accumulated occupancy map which provides the probability of occupancy of an obstacle at any given location.

The objective of this paper is to illustrate the more reliable occupancy maps that are obtained by using our multi-candidate approach as oppose to direct computations with the output of a disparity algorithm. As future work, we will integrate our

pipeline with robust obstacle detection approaches (e.g., [10]) in order to quantify the precise impact of our representation.

The remainder of this paper is organized as follows. Section II gives an overview of the current state-of-the-art. Section III gives the problem formulation and overview of our approach. Section IV describes the ground segmentation method used. Section V presents the generation of the multi-valued disparity map. Section VI presents the aggregate occupancy map computation using multi-valued disparity map. Section VII describe a simple obstacle detection approach from the aggregate occupancy map. Results of our proposed method are discussed in section VIII. Finally, section IX summarizes the paper and discusses future work.

## II. RELATED WORK

This section provides a brief overview of the state-of-the-art of different disparity map computation and obstacle detection methods. We focus on the current state-of-the-art of stereo vision based obstacle detection approaches, which make use of disparity maps.

1) *Disparity Map Computation*: In [3], Hirschmuller presents a classical method for disparity map computation. A pixelwise, mutual information based matching cost is used for matching between stereo pairs. This method is able to perform a fast pathwise optimization from all directions. In [12], disparity map is estimated by applying an iconic Kalman filter and known ego-motion. It is able to reduce the variance of disparity of each pixel and increase the density of the disparity map. In [4], to improve the robustness of the correspondence matching, Hermann and Klette introduce an iterative semi-global matching algorithm. The search space for disparity is iteratively reduced by using a pre-evaluated disparity prior. In order to improve fusion of disparity information, Wang *et al.* in [13] propose a post-aggregation method based on the Dempster-Shafer Theory (DST). Recently, Luo *et al.* in [14] propose an efficient deep-learning based stereo matching method, which is able to provide high accurate disparity maps in less than a second of GPU computation.

2) *Obstacle Detection*: Zhang *et al.* in [15] present an on-road obstacle detection algorithm using stereo cameras. Instead of calculating the correspondence for each pixel, they rely on the parameters of stereo cameras to determine the disparity search space for each pixel. Samadi *et al.* in [9] provide a robust obstacle detection method against illumination changes by using a differential image transform algorithm. Khalid *et al.* [16] provide a fast algorithm for obstacle detection based on stereo vision system using Hough transform and morphological processing. Zhang *et al.* [17] propose a method for obstacle detection in unstructured environments based on salient obstacle extraction. First, salient obstacles are extracted through a fast salient obstacle detection method. Then the detection of small obstacles is refined through the continuity and the height constraints among 3D points. Wei *et al.* in [18] presents a new methodology for robust segmentation and detection of obstacles. In their method, a topological persistence analysis is performed on an occupancy map constructed

by using UV disparity methodology. It is able to choose the regions that are most persistent which leads to a more robust segmentation result.

## III. PROBLEM FORMULATION AND APPROACH

In order to simplify the mathematical formulation, we make use of a very simplistic model of the world (see Fig. 1). However, we will show the applicability of the approach to real scenarios throughout this paper. Let the 3D environment be defined by the  $xy$ -plane (referred to as the *ground plane*) and prismatic obstacles erected perpendicularly from the ground to a height of  $h$ . The stereo cameras are parallel to each other with viewing direction along the  $y$ -axis and with displacement along the  $x$ -axis. We will further assume that an obstacle is always present along the  $u$ -axis of the image domain.

The stereo pair of rectified grayscale images  $I_L : \Omega_L \rightarrow \mathbb{R}^+$  and  $I_R : \Omega_R \rightarrow \mathbb{R}^+$  defined over the image domains  $\Omega_L \subset \mathbb{R}^2$  and  $\Omega_R \subset \mathbb{R}^2$  satisfy the relation:

$$I_L(u, v) = I_R(u, v + d(u, v)), \quad (1)$$

where  $(u, v) \in \Omega_L$  and  $d : \Omega_L \rightarrow \mathbb{R}^+$  is the disparity map associated with the distance of an object visible by images  $I_L$  and  $I_R$ . For simplicity, we ignore the effect of occlusions and assume that  $d$  is defined over its entire domain. We let  $\mathcal{O} \subset \Omega_L$  be the set of points in the image domain that correspond to obstacles and not the road, and define the *Obstacle Depth Function* as  $D(u) := d(u, v)$  where  $(u, v) \in \mathcal{O}$ . Note that due to our assumptions about the setup of the camera and the obstacles,  $d(u, v)$  will be the same along the  $v$ -axis as long as the visible point correspond to an obstacle. Our objective is a proper estimation of  $D(u)$ , which is dependent on a good estimate of the disparity map  $d(u, v)$ .

Stereo disparity algorithms often depend on the computation of a matching cost function  $C_p : [a, b] \rightarrow \mathbb{R}^+$  for each pixel  $p = (u, v) \in \Omega_L$ , where  $[a, b]$  denotes the disparity search range, and  $a$  and  $b$  are constants. The matching is challenging due to reflections, lack of texture and repetitive patterns of objects. In these cases, the correct disparity may fall in local minimum of the matching cost curve instead of the global minimum cost. In order to capture the correct disparity values, we make use of multi-candidates per location in the image domain. These candidates are associated with local minima of the cost function, and form our *multi-valued disparity map*. We compute an *aggregate occupancy map* from this set of candidate matches. Fig. 2 shows the pipeline of our proposed approach: first, a ground plane is fitted robustly from the data

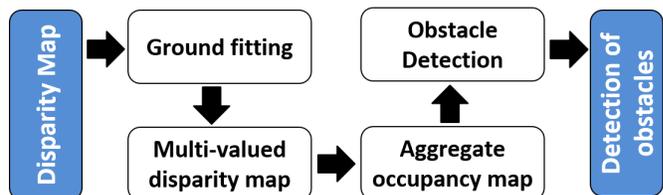


Fig. 2. Overview of proposed approach.

by making use of the disparity values associated with the global minima from the matching cost functions (section IV); then, the multi-valued disparity map is constructed (section V); followed by computation of the aggregate occupancy map (section VI); and finally, an obstacle depth function is estimated (section VII).

For our implementations, we make use of the matching cost defined for the Semi-Global Block Matching approach [3] with  $a = 0$  and  $b = 127$ . However, this can be extended to any other disparity computation algorithms based on a matching cost function.

#### IV. GROUND FITTING

The ground plane is computed by fitting a plane in the  $uvd$ -space, where  $p = (u, v)$  corresponds to the pixel location and  $d$  to the associated disparity value. For the fitting, we make use of a point cloud constructed by selecting the global minimum of the matching cost function  $C_p$  as a corresponding disparity value of the pixel. RANSAC EDIT[19] is used for a robust fitting of the plane.

We parameterize the plane using the form  $v = g_0(u, d)$  since we will be considering bounds on the heights of obstacles in the physical space for their detection. In particular, we will only consider candidate points  $(u, v, d)$  that satisfy the constraint

$$g_0(u, d) + \frac{a_u h_l}{a_v b} \cdot d \leq v \leq g_0(u, d) + \frac{a_u h_u}{a_v b} \cdot d \quad (2)$$

where  $a_u$  and  $a_v$  are the intrinsic focal length parameters of the camera in pixels,  $b$  is the distance of the baseline of the stereo camera system,  $h_l$  and  $h_u$  are the heights above the estimated ground plane used to identify obstacle pixels. In our experiment, we use  $h_l = 200$  and  $h_u = 1700$  for KITTI dataset [6], [7], [8]. Using this setup, the ground plane for ground segmentation is 0.2 meters higher than the fitting plane to make sure we remove all the ground pixels, and we cut every thing above 1.7 meters from the fitting plane to avoid the influence of obstacles above the car. Then the height for obstacle detection is around 1.5 meters in real world space.

#### V. MULTI-VALUED DISPARITY MAP

As mentioned above, finding correspondences in stereo pairs is heavily affected by visual artifacts such as reflections, lack of texture and repetitive patterns. Fig. 3 (a) shows an example of wrong disparity estimation by *Semi Global Block Matching* (SGBM) [3], [20] implemented by OpenCV due to the repetitive structure present in the fence. Fig. 3 (b) shows the cost function  $C_p$  for the point highlighted by a red circle in the image. The cost function applied in OpenCV-SGBM is the Birchfield-Tomasi subpixel metric proposed in [21]. This cost is the measurement of absolute minimum difference of intensities between two pixels in each direction along the epipolar line[3]. The SGBM algorithm select a disparity value close to the global minimum of the function while the groundtruth is at one of the local minima. The periodic pattern in the cost function is due to the repetitive pattern of the fence,

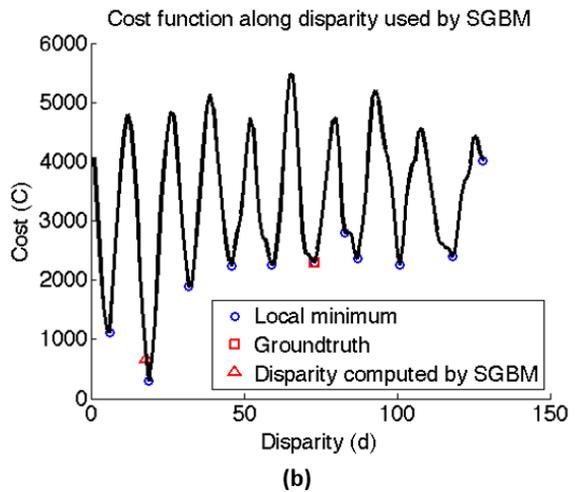
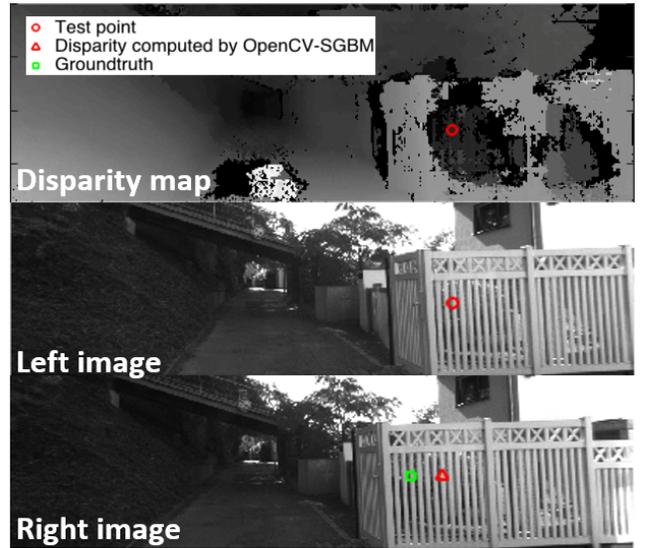


Fig. 3. (a) Example of wrong disparity estimation by SGBM due to the repetitive structure presenting in the fence. Disparity map shows the wrong disparity estimation in some regions of the fence (top). Right image (bottom) shows the estimated correspondence (red triangle) and ground truth (green square) for the test point in left image (middle). (b) Cost function  $C_p$  of the test point highlighted in (a). The estimation is near the global minimum, but the ground truth is near a local minimum. Notice that each local minimum corresponding to a bar of the fence.

which makes it difficult (if not impossible) to determine which one is the correct matching using only local information.

Before further discussion, we first define the local minimum of a cost function. Let  $\bar{d}_{p,i}$  be a local minimum of the cost function  $C_p$  of a pixel  $p$  if it satisfies  $C_p(\bar{d}_{p,i}) \leq C_p(\bar{d}_{p,i} - 1)$  and  $C_p(\bar{d}_{p,i}) \leq C_p(\bar{d}_{p,i} + 1)$ . Note that, although the disparity value is computed to sub-pixel accuracy, the local minimum is computed to pixel accuracy.

One way to capture the correct disparity value is to pick all the local minima,  $\{\bar{d}_{p,i}\}$ , of the cost function  $C_p$  as possible disparity candidates for one pixel in the left image. However, this can lead to too many candidates. In order to select a subset of these candidates, we define the cost ratio of the  $i$ -th local

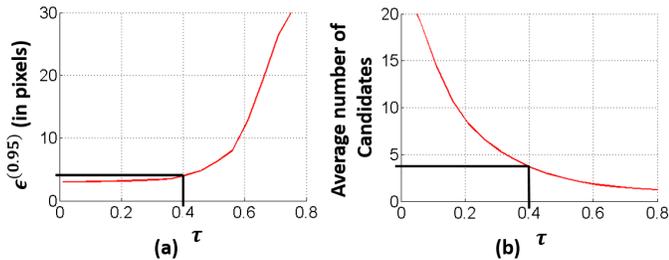


Fig. 4. Threshold choice for disparity candidate selection. (a)  $\epsilon^{(0.95)}$  as a function of parameter  $\tau$ . (b) Average number of candidates as a function of parameter  $\tau$ . A choice of  $\tau = 0.4$  guarantees that  $\epsilon^{(0.95)} < 5$  and that the number of candidates will be below 5 as well. A total of 120 training stereo pairs were used for selecting this threshold.

minimum as the criteria of picking candidates by

$$r_{p,i} = \frac{C_p(\bar{d}_{p,i})}{\min_d C_p(d)}. \quad (3)$$

This ratio was selected as a criteria since the values of the cost functions can vary greatly between different pixel locations, which makes it is unrealistic to pick a global threshold for the entire image based on cost alone. Other than cost ratio, we also tested the prominence ratio as a criteria, which is defined by the ratio of maximal prominence among all local minimum and prominence of a certain local minimum. The experiment showed that cost ratio was a better choice. As part of our future work, we will test the performance of using confidence scores discussed in [22] as our another candidate selection criteria. A subset  $\{d_{p,j}\}$  of candidate disparity values is selected by keeping the local minima with cost ratio  $r_{p,j}$  below a threshold  $\tau$  and for which the point  $(p, d_{p,j})$  satisfy the ground plane condition in Eqn. 2.

The value of  $\tau$  is selected statistically from a set of 120 stereo pairs from the KITTI dataset (which includes groundtruth depth estimates) in order to make sure that the candidate points includes a point near the true disparity with high confidence, while minimizing the overall number of candidates. This is done by enforcing that 95% of the candidate sets have a candidate within a distance  $\Delta$  of the true disparity. The value of  $\Delta$  is empirically selected to be 5 pixels as it offers a good compromise between the number of candidates and the distance from the true disparity.

In order to perform the analysis described above, we compute for each pixel  $p$  the minimum distance to the true disparity value

$$\epsilon_p(\tau) = \min_j |d_p^{true} - d_{p,j}|, \quad (4)$$

where  $d_p^{true}$  is the true disparity value for the corresponding pixel, and  $j$  is the index for the candidate disparity values. We also keep track of the number of candidates  $n_p$  for corresponding pixel. These quantities are accumulated over all pixels and images in the dataset. Additionally, we compute  $\epsilon^{(0.95)}(\tau)$  as the distance associated with the 95% percentile of all  $\epsilon_p$  values. We ensure that at least one candidate is within  $\Delta$  from the true value for 95% of all the samples by picking

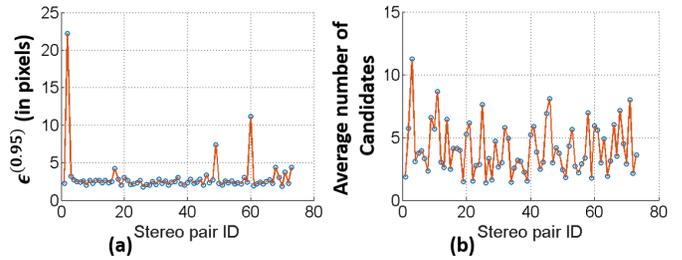


Fig. 5. (a)  $\epsilon^{(0.95)}(0.4)$  of each stereo pair in test set. (b) Average number of candidates per pixel of each stereo pair in test set. For most of pairs,  $\epsilon^{(0.95)}(0.4) < 5$  and average number of candidates per pixel is less than 6.

a value of  $\tau$  that satisfies  $\epsilon^{(0.95)}(\tau) \leq \Delta$ . Fig. 4 illustrates the curves obtained from  $\epsilon^{(0.95)}(\tau)$  and the average number of candidates as a function of  $\tau$ . Since the number of candidates increases monotonically with the value of  $\tau$ , then it is sufficient to pick the largest value of  $\tau$  for which  $\epsilon^{(0.95)}(\tau) \leq \Delta$ . From our analysis, this corresponds to  $\tau = 0.4$ , which gives an average number of candidates less than 5.

As discussed above, the value of  $\tau$  was selected using a set of 120 stereo pairs for training. A different set of 73 stereo pairs are used for testing the performance of this threshold. Fig. 5 (a) and (b) illustrate  $\epsilon^{(0.95)}$  and the average number of candidates for each test stereo pair for  $\tau = 0.4$ . It is observed that  $\epsilon^{(0.95)}(0.4)$  is less than 5 for most of the stereo pairs and the average number of candidates is less than 9 for all except one stereo pair. That is, for most pixels in each stereo pair, there exists at least one candidate disparity within 5 pixel distance of the ground truth and with average number of candidates less than 9.

Notice that this threshold may change if we use a dataset with a very distinct type of images or a different disparity computation algorithms. However, the value of  $\tau$  can be tuned following this same process. This parameter choice is how our approach takes into account some statistical information about the environment and the performance of the cost function used for matching. Also notice that capturing the correct disparity values by multiple disparity values requires that the cost function reflects the correct matching cost between stereo pair. That is, it is likely to find the true matched point near a local minimum. There are no guarantees for the case of reflections since this artifacts yield cost function with misleading local minima.

## VI. AGGREGATE OCCUPANCY MAP COMPUTATION

The aggregate occupancy map is computed using a variation of the visibility based approach introduced in [23], in which a probability of occupancy is obtained for points in the so-called  $u$ -disparity space (or simply  $ud$  space). The main difference for our approach is that we consider multiple candidate points per each pixel due to our multi-valued disparity map instead of a single point estimate.

In the original approach, a point  $s = (u, d)$  in the  $u$ -disparity space is assigned a probability of occupancy by counting the number of visible and observed obstacle points falling within

the site. Since our disparity map is multi-valued, a single pixel  $p$  in the image domain may have multiple points in the point cloud. In order to compensate for this, we assign a weight to each candidate point equal to the  $1/n_p$ , where  $n_p$  is the number of candidate points for that pixel. As our definition of local minimum in section V, all the points within the flat region of a cost function are considered as local minimum as well. However, the contribution of this pixel to the occupancy grid computation is normalized by the number of candidates for this pixel. The end result is that pixels with large flat regions will not influence the occupancy grid. Please refer to [23], [18] for more details on occupancy computation.

The second row of Fig. 6, 7, 8, 9 and 10 illustrate the aggregate occupancy map computed from multi-valued disparity map (left) and OpenCV-SGBM disparity map (right). Compare the occupancy map from the two different approaches, aggregate occupancy map is able to capture the structure of obstacles as good as OpenCV-SGBM disparity map for the regions where OpenCV-SGBM disparity map does well, and for the regions where OpenCV-SGBM provides wrong estimation which highlighted by the black rectangle, the aggregate occupancy map is able to avoid it through the normalization process during occupancy computation. Because pixels with good disparity estimation have small number of candidates since the global minimum cost of those pixels is much smaller than other local minimum, and most of the bad disparity estimations are the outcome of several local minimum closed to the global minimum.

## VII. OBSTACLE DETECTION

A simple approach for obstacle detection from the aggregate occupancy map is to apply a threshold  $\tau_p$  to the probability of occupancy [23]. Detections in the image domain are obtained by displaying a vertical segment for each location  $(u, d)$  in which there was a detection in the occupancy map. The vertical segments are drawn between the lower and upper planes defined during the ground fitting procedure. We make use of a value of  $\tau_p = 0.02$  for detections using our aggregate occupancy map, and a threshold of  $\tau_p = 0.15$  for detections using the traditional occupancy map using SGBM. Simple threshold has the disadvantage of being very sensitive to the choice of the parameter value. Hence, as future work, we hope to make use of the approach introduced in [18] which uses topological persistence to make obstacle detection more robust.

## VIII. RESULT

The disparity map and correspondence matching cost computation is implemented in C++ using OpenCV, and all the other processes are implemented in MATLAB. We use stereo pairs from the KITTI dataset [6], [7], [8] to test our approach. We tested our approach on a complete dataset from the KITTI benchmark suite, and observed either comparable or improved results on the occupancy map between our approach and OpenCV-SGBM. In this section, we only highlight the most extreme cases. Figs. 6, 7, 8, 9 and 10 show five examples of the

obstacle detection using the proposed multi-valued disparity map and the OpenCV-SGBM disparity map. Fig. 6 illustrates a case in which both approaches give good results. We noticed that since most images in the datasets are of good quality and feature well textured obstacles on the road, then disparity computation results are often correct. Our approach produces very similar results in these cases as well. Figs. 7, 8, 9 and 10 show four examples in which the multi-valued disparity map provides better obstacle detection than OpenCV-SGBM. Obstacle detections shown in the image domain are obtained by finding the obstacle with highest disparity value for each column in the  $u$ -disparity domain. That is, the closest obstacle among all detected obstacles at each  $u$ . We map these detection back to the image domain based on the height of the ground and sky plane at position  $(u, d)$  obtained during the ground fitting process mentioned in Section IV.

In Fig. 6, both approaches detect the obstacle correctly. We can see little difference between the two occupancy maps. This is due to the fact that, as discussed in Section VI, pixels with good disparity estimation have small number of candidates.

For Fig. 7, 8 and 9, it is observed that the OpenCV-SGBM disparity maps are bad in some regions due to the brightness or shadow on the ground. Because of it, the occupancy maps obtained using these disparity maps are noisy in the corresponding regions in  $u$ -disparity domain, which are highlighted by the black rectangle. This leads to the wrong obstacle detection results on the ground. For Fig. 10, the bad disparity is due to the repetitive pattern of the fence. The occupancy map in (d) shows some wrong detection in the fence area highlighted by the black rectangle and the obstacle detection (h) shows most parts of the fence are detected as farther obstacle. The occupancy map obtained by our approach in (c) shows several detections including the correction detection in the fence of region and in (g), it shows our approach obtains much better result than (h). Our method is able to mitigate these problem. This can be explained by noting that regions with wrong disparity values often have multiple candidates, and their contribution to the occupancy map is distributed over all candidate matches.

## IX. CONCLUSION

In this paper, we propose a methodology for obstacle detection in outdoor scenes using multi-valued stereo disparity maps. The multi-valued stereo disparity map is obtained by selecting the local minimum of the matching cost function using statistics from a training set of stereo pairs. An aggregate occupancy map is computed from this multi-valued disparity map. Finally, obstacles are detected by thresholding the aggregated occupancy map. The experiment shows that the proposed approach is able to provide more reliable occupancy maps for obstacle detection than traditional single-valued stereo disparity computations.

In the future, we plan to integrate this approach with topological persistent segmentation [18] and other representation such as Stixels [10] to make obstacle detection more robust.

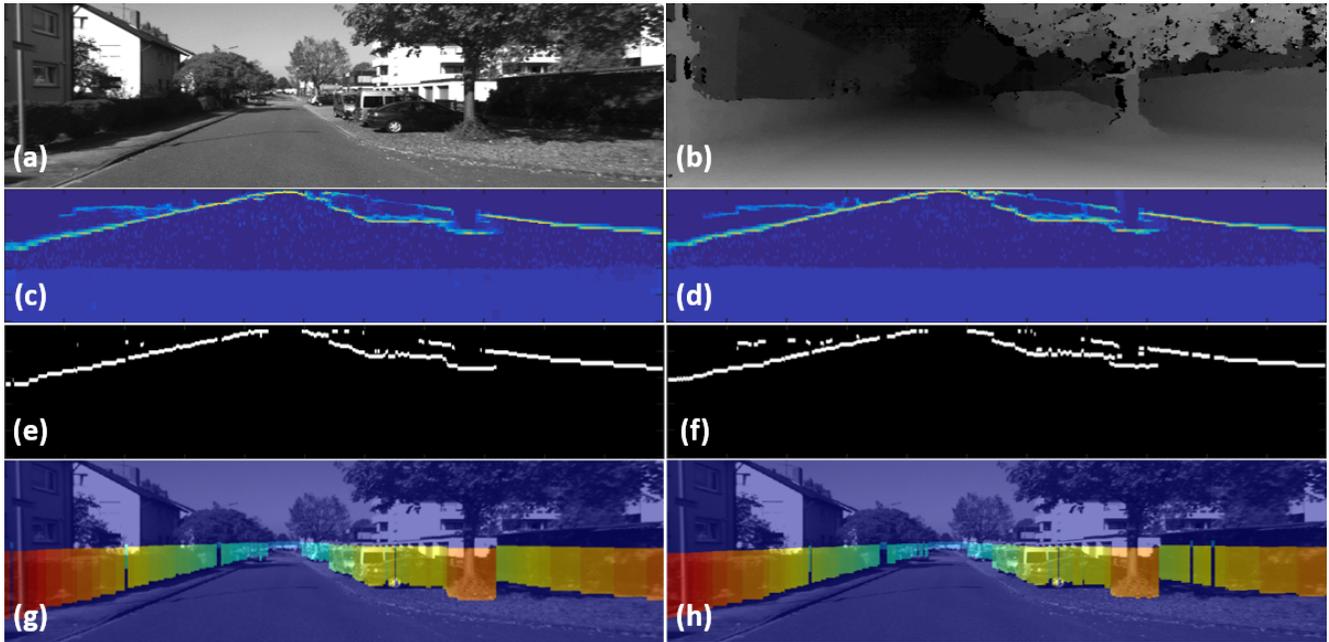


Fig. 6. (a) Left image of the stereo pair. (b) Disparity map from OpenCV-SGBM. (c) Aggregate occupancy map from multi-valued disparity map. (d) Occupancy map from OpenCV-SGBM disparity map. (e) Obstacle detection using aggregate occupancy map in  $u$ -disparity domain. (f) Obstacle detection using OpenCV-SGBM occupancy map in  $u$ -disparity domain. (g) Obstacle detection using aggregate occupancy map in image domain. (h) Obstacle detection using OpenCV-SGBM occupancy map in image domain. In this case, both approaches give good results. (c) and (d) show similar occupancy grid maps, since pixels with good disparity estimation have small number of candidates.

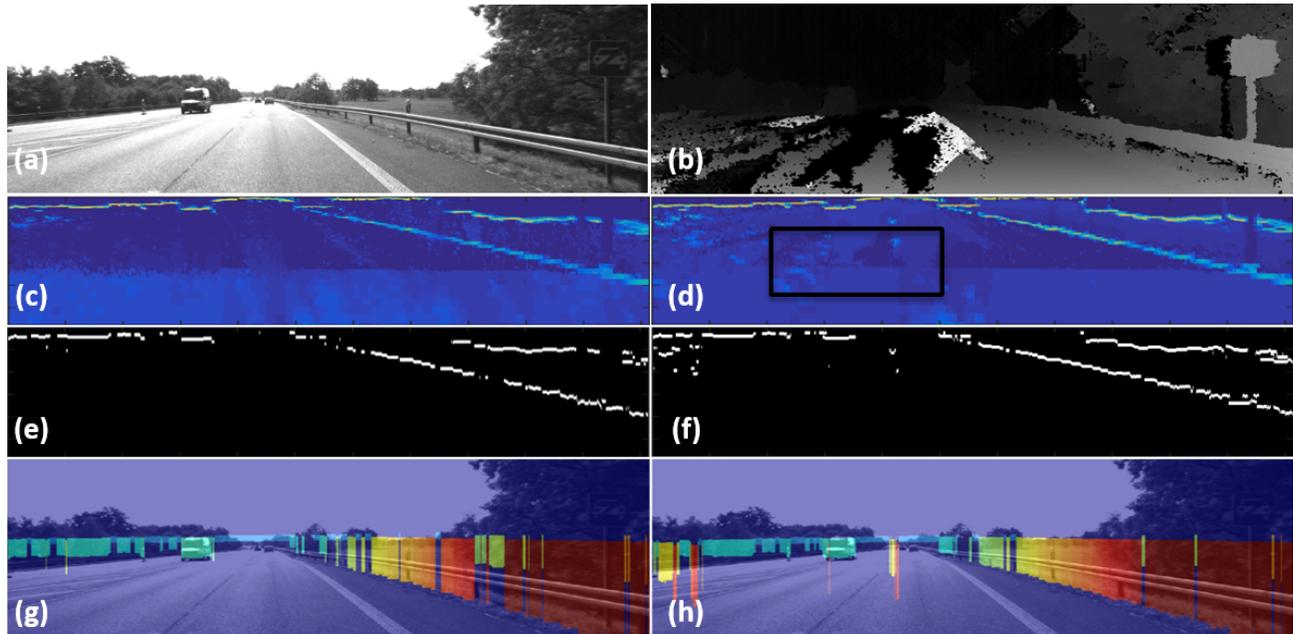


Fig. 7. (a)-(h) are the same as Fig. 6. In this case, the disparity map in (b) is wrong in some regions of ground due to the brightness. The occupancy map computed using this disparity map is noisy in the corresponding regions highlighted by the black rectangle shown in (d). This leads to the wrong obstacle detection on the ground shown in (h). (g) shows our approach is able to avoid those false negatives.

## REFERENCES

- [1] J. Duan, L. Shi, J. Yao, D. Liu, and Q. Tian, "Obstacle detection research based on four-line laser radar in vehicle," in *Robotics and Biomimetics (ROBIO), 2013 IEEE International Conference on*, Dec 2013, pp. 2452–2457.
- [2] A. Broggi, M. Buzzoni, M. Felisa, and P. Zani, "Stereo obstacle detection in challenging environments: The viac experience," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, Sept 2011, pp. 1599–1604.
- [3] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *Pattern Analysis and Machine Intelligence, IEEE Trans-*

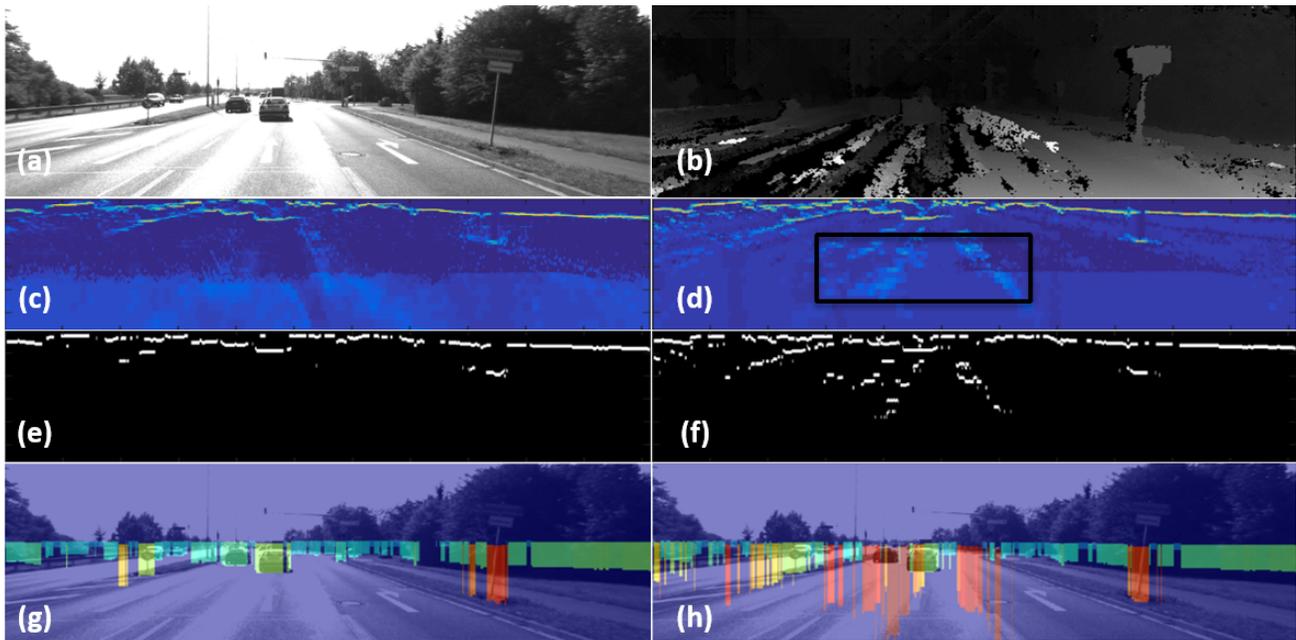


Fig. 8. (a)-(h) are the same as Fig. 6. In this case, the disparity map in (b) is wrong in some regions of ground due to the brightness. The occupancy map computed using this disparity map is noisy in the corresponding regions highlighted by the black rectangle shown in (d). This leads to the wrong obstacle detection on the ground shown in (h). (g) shows our approach is able to avoid those false negatives.

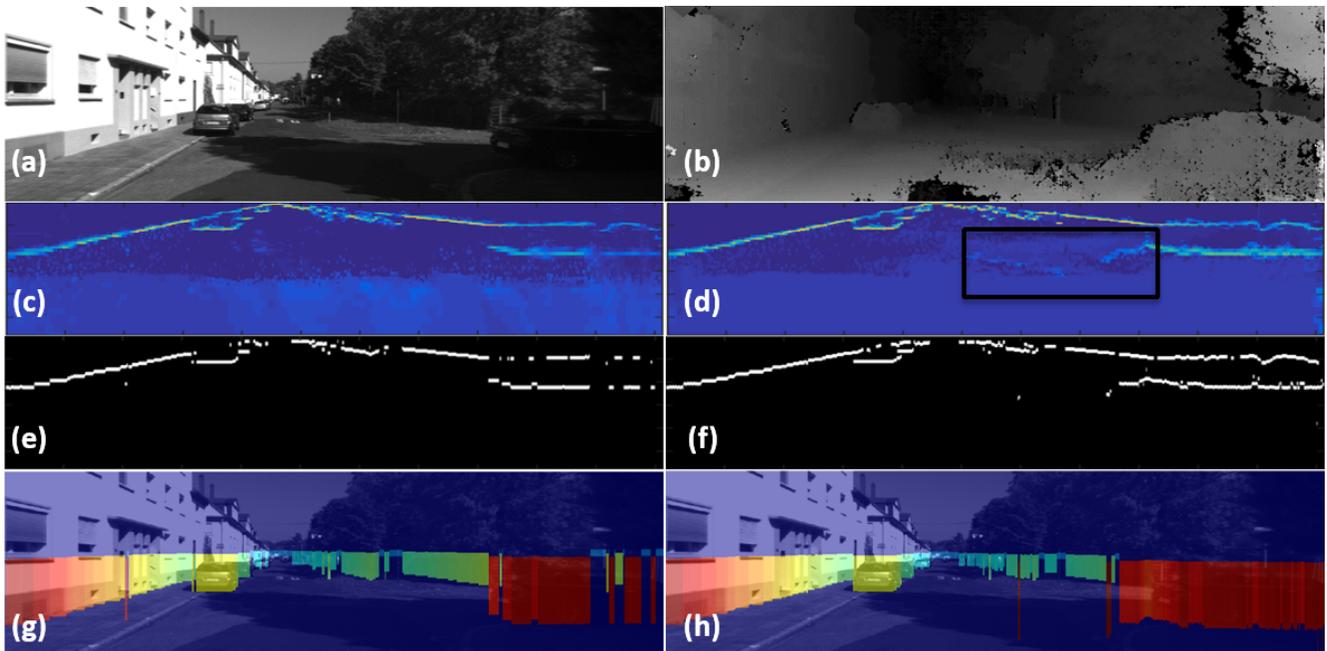


Fig. 9. (a)-(h) are the same as Fig. 6. In this case, the disparity map in (b) are wrong in the right region of ground due to the shadow. The occupancy map computed using this disparity map has wrong detection in the corresponding regions highlighted by the black rectangle closed to the black car shown in (d). This leads to the larger detected obstacle in the right part of the image shown in (h). (g) shows our approach is able to avoid those false negatives and gives the correct detection of the car.

actions on, vol. 30, no. 2, pp. 328–341, 2008.

- [4] S. Hermann and R. Klette, “Iterative semi-global matching for robust driver assistance systems,” in *Computer Vision–ACCV 2012*. Springer, 2013, pp. 465–478.
- [5] K. Ito, M. Sasaki, T. Aoki, T. Ishigami, and A. Nishimura, “Generating robust and stable disparity map using phase-based correspondence matching from stereo video sequence,” in *Pattern Recognition (ACPR), 2013 2nd IAPR Asian Conference on*, Nov 2013, pp. 586–590.
- [6] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? the kitti vision benchmark suite,” in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
- [7] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, “Vision meets robotics:

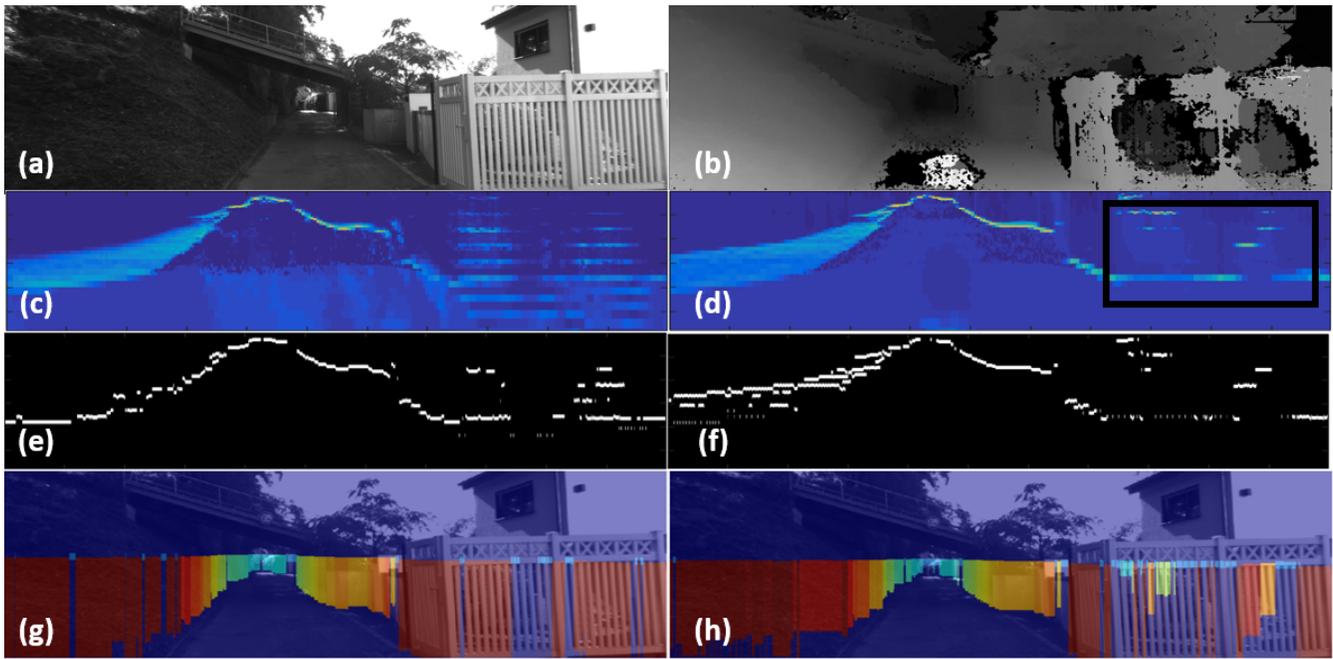


Fig. 10. (a)-(h) are the same as Fig. 6. In this case, the disparity map in (b) is wrong in some regions of the fence due to the repetitive pattern of the fence. The occupancy map shown in (d) has wrong detection in the region of fence highlighted by the black rectangle and the obstacle detection in (h) shows most parts of the fence are detected as farther obstacle. The occupancy map obtained by our approach in (c) shows several detections in the region of fence including the correct detection, since the pixels with bad disparity estimation have large number of candidates. (g) shows although our approach does not obtain the completely correct detection of the fence, it is able to obtain much better result than OpenCV-SGBM.

The kitti dataset,” *International Journal of Robotics Research (IJRR)*, 2013.

- [8] J. Fritsch, T. Kuehnl, and A. Geiger, “A new performance measure and evaluation benchmark for road detection algorithms,” in *International Conference on Intelligent Transportation Systems (ITSC)*, 2013.
- [9] M. Samadi, M. F. Othman, and S. H. Amin, “Stereo vision based robots: Fast and robust obstacle detection method,” in *Control Conference (ASCC), 2013 9th Asian*. IEEE, 2013, pp. 1–5.
- [10] D. Pfeiffer, S. Gehrig, and N. Schneider, “Exploiting the power of stereo confidences,” in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, June 2013, pp. 297–304.
- [11] M. Zhang, P. Liu, X. Zhao, X. Zhao, and Y. Zhang, “An obstacle detection algorithm based on u-v disparity map analysis,” in *Information Theory and Information Security (ICITIS), 2010 IEEE International Conference on*, Dec 2010, pp. 763–766.
- [12] C. Højlund, T. B. Moeslund, C. B. Madsen, and M. M. Trivedi, “Improving stereo camera depth measurements and benefiting from intermediate results,” in *Intelligent Vehicles Symposium (IV), 2010 IEEE*. IEEE, 2010, pp. 935–940.
- [13] F. Wang, A. Miron, S. Ainouz, and A. Bensrhair, “Post-aggregation stereo matching method using dempster-shafer theory,” in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 3783–3787.
- [14] W. Luo, A. G. Schwing, and R. Urtasun, “Efficient deep learning for stereo matching,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 5695–5703.
- [15] Z. Zhang, Y. Wang, J. Brand, and N. Dahnoun, “Real-time obstacle detection based on stereo vision for automotive applications,” in *Education and Research Conference (EDERC), 2012 5th European DSP*. IEEE, 2012, pp. 281–285.
- [16] Z. Khalid, E.-A. Mohamed, and M. Abdenbi, “Stereo vision-based road obstacles detection,” in *Intelligent Systems: Theories and Applications (SITA), 2013 8th International Conference on*. IEEE, 2013, pp. 1–6.
- [17] Y. Zhang, X. Xu, H. Lu, and Y. Dai, “Two-stage obstacle detection based on stereo vision in unstructured environment,” in *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2014 Sixth International Conference on*, vol. 1. IEEE, 2014, pp. 168–172.
- [18] C. Wei, Q. Ge, S. Chattopadhyay, and E. Lobaton, “Robust obstacle segmentation based on topological persistence in outdoor traffic scenes,” in *Computational Intelligence in Vehicles and Transportation Systems (CIVTS), 2014 IEEE Symposium on*. IEEE, 2014, pp. 92–99.
- [19] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981. [Online]. Available: <http://doi.acm.org/10.1145/358669.358692>
- [20] “Open Source Computer Vision (OpenCV) Library,” <http://opencv.org/>, 2015.
- [21] S. Birchfield and C. Tomasi, “Depth discontinuities by pixel-to-pixel stereo,” in *Computer Vision, 1998. Sixth International Conference on*, Jan 1998, pp. 1073–1080.
- [22] X. Hu and P. Mordohai, “A quantitative evaluation of confidence measures for stereo vision,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 11, pp. 2121–2133, Nov 2012.
- [23] M. Perrollaz, J.-D. Yoder, A. Negre, A. Spalanzani, and C. Laugier, “A visibility-based approach for occupancy grid computation in disparity space,” *Intelligent Transportation Systems, IEEE Transactions on*, vol. 13, no. 3, pp. 1383–1393, Sept 2012.