

Cast Shadow Resistant Ground Plane Detection in Single Image

Xiaoyan Xu¹, Xiaoming Liu¹, Chang Liu^{1*} and Yuan Ling²

¹*School of Information Science and Technology
Dalian Maritime University*

No.1, Linghai Road, Dalian, China

²*Hualu Zhida Technology Co.Ltd*

No.717, Huangpu Road, Dalian, China

xyxhlj@126.com, liuchang@dlnu.edu.cn, lingy@ehualu.com

Abstract

Ground plane detection is useful for vision navigation to robots and autonomous vehicles. Cast shadow on the ground plane is a challenging issue that may cause the detection fail. In this paper, we present a cast shadow resistant ground plane detection approach from a single color image. We first derive an initial ground plane using geometric layout method. We then apply shadow invariant transform on the roughly detected shadow edges to get a gray-scale intrinsic image. The final shadow resistant ground plane result is obtained by employing region growth on the initial seed region in the shadow-free intrinsic image. The approach proposed here does not need priori assumption such as calibration or landmark. Experimental results show the method works for different scenes.

Keywords: *cast shadow, intrinsic image, ground plane detection*

1. Introduction

Ground plane detection is useful for vision-guided obstacle avoidance and navigation, which are the key problems in robot and autonomous vehicles applications. Ground plane detection also is the basis for the detection of objects moving on that plane. Many works have been developed to solve the problem based on a number of different approaches. Variations have included stereo vision, monocular vision, and single image. Some approaches like [1] use a stereo camera setup to detect ground planes from the 3D reconstruction of the scene. While the algorithm can provide reasonable results, this kind of methods has some drawbacks. In many situations, ground plane is textureless and looking for corresponding points is challenges due to lack of texture in this area. Therefore, it may not be suitable to apply techniques of binocular vision or multi-images since it requires matching corresponding points in multiple images. This method also needs camera calibration and is time consuming.

Because of the difficulty as mentioned previously in binocular vision, more recently, researchers have developed monocular approaches. Several methods have been proposed for identifying the ground plane from an image sequence. A widely used approach exploits the homography constraint, i.e., a geometric relationship linking two views of the same planar surface. In [2], a purely vision-based method using a single on-board camera is proposed to estimate the ground plane. The proposed method is based on the reliable estimation of the homography between ground planes by computing matched features in successive images. Conrad et al. [3] proposed a homography-based ground plane detection that uses a modified expectation maximization algorithm to cluster pixels on images as belonging to one of two possible classes: ground and non-ground pixels. Work that use of monocular cameras in this area also includes the use of color information either because of dissimilarity in color [9] or spatial color gradient changes [10] to

distinguish the ground and obstacles. Another technique used in autonomous robotics for ground plane detection is optical flow [4] [11]. However, such methods can be computationally intense. In [5], the author proposed an active vision system to estimate ground planes. It optimizes online the camera motion to maximize the convergence rate of the plane estimation error and obtain higher estimation accuracy in a shorter time.

Humans seldom have any difficulty in inferring the 3D structure of the scene from a single image. This is not only attributed to prior knowledge about the environment but also to using monocular cues to infer depth. For example, Tian et al. [6] apply edge, clarity, color, and pixel position to find the passable regions from a single image for robot navigation. Depth information is very important for a robot to avoid obstacles. Saxena et al [8] used monocular cues and Markov Random Field to recover the depth maps. Literatures [14, 15] recover depth information from defocus. Edge blur is used in [16] as a cue to recover the relative depth. Human can use monocular cues such as texture, interposition, object size (Near object looks more larger than far ones), edge, shading, defocus, etc. In [7], a Markov Random Field based 3D reconstruction is performed to build an approximate depth map of an image. This map is robust against texture variations due to shadows. A texture segmentation algorithm is also applied to determine the ground plane accurately.

A ground plane detection method is usually affected by ground shadows, i.e., detection may failed when shadows exist. In this paper, we propose a shadow resistant ground plane detection method based on intrinsic images. Intrinsic image here is referred to an image insensitive to shadows. In our study, we propose a simple intrinsic method based on roughly detected shadows.

2. Method Description

Given an input image as shown in the top-left in Figure 1, we first label the image into geometric classes using the method proposed by Hoiem et al. [11]. The labeling is shown in the top-right image in Fig.1. It offers us the probability of the ground plane (the cyan color). We found that their result is affected by the shadow regions. To counter that effect, an intrinsic image, as shown in the bottom-left in Figure 1, is generated using the method proposed in section 2.2. The shadows are largely attenuated in the result. Finally, taking Hoiem's ground plan result as initial region, region growth on the intrinsic images is used to extract our final ground plane that is invariant to shadows. As shown in the bottom-right in Figure 1, we can find that ground plane can be detected well without influence of the cast shadows.

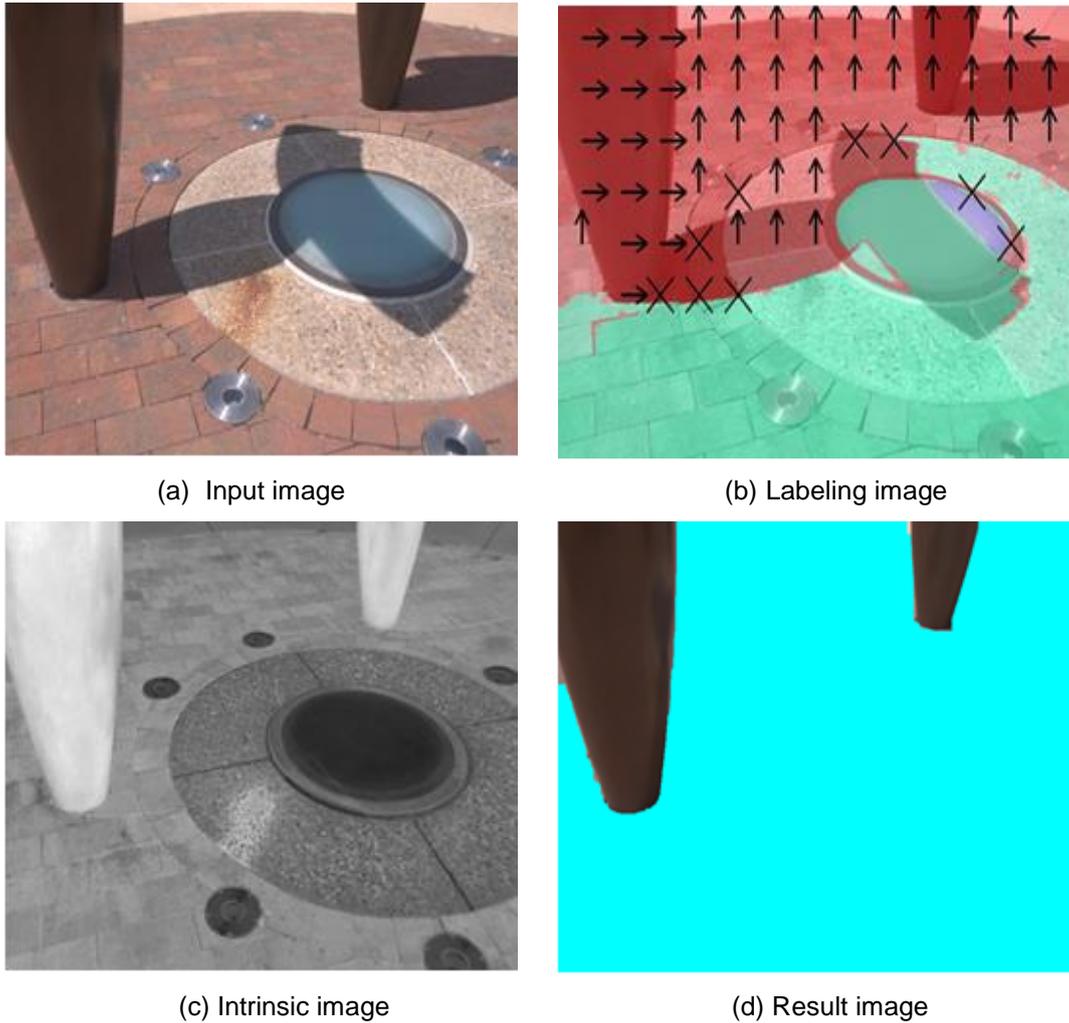


Figure 1. Algorithm Demonstration

Incorporating intrinsic result to scene layout can much improve the ground plane detection result. Top-left (a): original image. Top-right (b): geometric layout by [11], where the cyan denotes detected ground plane. We can see that shadows lead to wrong labels of their method. Bottom-left (c): the intrinsic image. Bottom-right (d): Our detected ground plane by cooperating the geometric label and intrinsic image.

From figure (b) we can see, Hoiem's algorithm is used to detect the ground plane. When shadows exist in the image, the algorithm makes error. It takes shadows as solid obstacles, and eventually leads to the failure of the algorithm. In fact, the shadows are connected region, but Hoiem's algorithm mistaken it as an unconnected region. Taking Hoiem's ground plan result as initial region and using the intrinsic image, we can find that ground plane can be detected well without influence of the cast shadows.

The main method is organized as follows. Method for image labeling is described in section 2.1; A shadow resistant algorithm is proposed in section 2.2; Shadow detection and parameter estimation is described in section 2.3 ; Region growth algorithm is introduced in section 2.4.

2.1. Image Labeling

We first label the image into geometric classes using the method proposed by Hoiem et al. [11]. This work can classify image into different orientation labels (e.g., planar left,

planar center, or planar right) by applying decision-tree Adaboost algorithm to learn a wide variety of appearance-based cues, includes edge, color, texture, and perspective, to contribute to the confidence in each geometric label. In Hoiem's study, they used the logistic regression version of Adaboost, which differs from the original confidence weighted version; they trained separate classifiers to distinguish among the main classes and the subclasses of vertical objects. As shown in Fig.1, on this image, main class labels are indicated by colors (green=support, red=vertical, and blue=sky), and subclass labels are indicated by markings (left/up/right arrows for planar left/center/right, 'O' for porous, and 'X' for solid). The results show that the presence of shadows can easily lead the algorithm fail

Overview of training boosted decision trees:

Input:

- $D_1 \dots D_m$: training data
- $w_1 \dots w_m$: initial weights
- $y_1 \dots y_m \in \{-1, 1\}$: labels
- n_n : number of nodes per decision tree
- n_t : number of weak learner decision trees

For $t = 1 \dots n_t$:

1. Learn n_n -node decision tree T_t based on weighted distribution w_t

2. Assign to each node $T_{t,k}$:
$$f_{t,k} = \frac{1}{2} \log \frac{\sum_{i: y_i=1, D_i \in T_{t,k}} w_{t,i}}{\sum_{i: y_i=-1, D_i \in T_{t,k}} w_{t,i}}$$

3. Update weights:
$$w_{t+1,i} = \frac{1}{1 + \exp(y_i \sum_{t'} f_{t',k_{t'}})}$$
 with $k_{t'} : D_i \in T_{t',k_{t'}}$

4. Normalize weights so that $\sum_i w_{t+1,i} = 1$

Output:

- $T_1 \dots T_{n_t}$: decision trees
- $f_{1,1} \dots f_{n_t, n_n}$: weighted log-ratio for each node of each tree

2.2. Intrinsic Image

Our shadow invariant image is based on work [13], where they proposed a method which uses a Gamma-corrected sRGB image to generate a grayscale image insensitive to shadows. However, they just use the mean spectral power distributions of daylight and skylight to approximate the linear parameters. Different to their work, the method in this paper use inverse Gamma-corrected sRGB (linear sRGB) image to derive grayscale shadow invariant image. This will make the method simpler, and we will estimate the parameters of our algorithm based on a rough shadow detection result. We first apply inverse Gamma correction to convert sRGB to linear RGB in three channels by Equation (1). (Please Ref. to IEC 61966-2.1 for details).

$$\Phi_H = 100 \cdot \left[\left(\frac{\Phi_H^L}{255} + 0.055 \right) / 1.055 \right]^{2.4} \quad (1)$$

Here Φ_H denotes linear and non-Gamma corrected pixel values in H=R, G, B channels respectively and Φ_H^L denotes the correspondingly Gamma corrected pixel values. Literature [13] deduced that the following equation holds.

$$F_H = k_H f_H \quad (2)$$

Here F_H denotes the linear sRGB pixel values in non-shadow area while f_H denotes those in shadow area, k_H denotes the proportional coefficients that represent the ratio of non-shadow pixel and shadow pixel. Taking logarithms of both sides of Equation (2), we have,

$$\log(F_H) = \log(k_H) + \log(f_H) \quad (3)$$

In detail,

$$\begin{cases} \log(R) = \log(k_1) + \log(r) \\ \log(G) = \log(k_2) + \log(g) \\ \log(B) = \log(k_3) + \log(b) \end{cases} \quad (4)$$

From Equation (4), the following equation holds,

$$\log(R) + \log(G) - \beta \cdot \log(b) \quad (5)$$

Here $\beta = \frac{\log(k_1) + \log(k_2)}{\log(k_3)}$.

2.3. Shadow Detection and Parameter β

Prior to calculate β , we should determine k_H first. As illustrated in Figure 2, we estimate the parameters of k_H based on the roughly detected shadow edges. We applied the shadow detection algorithm proposed in [16] to detect shadow edges. For each detected edge, we first apply morphological dilation on the both sides of an edge, and then calculate the mean values of the dilated regions. Then, we divide the mean values of the shadow side from the non-shadow one to get the parameter K_H .

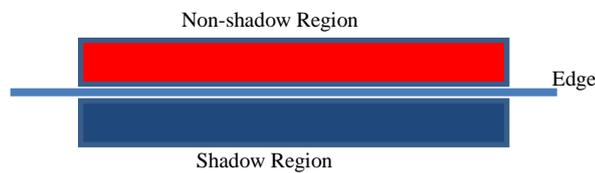


Figure 2. Shadow Edges and Shadow Regions

(The red: Non-shadow region; The blue: Shadow region)

2.4. Region Growth Algorithm

Through the application of the above algorithm, we can get a shadow-free intrinsic image. To get the final shadow resistant ground plane result, we used the region growth algorithm [6]. Their algorithm is summarized as follows: First, they extracted the depth-related features, such as edges, clarity and details, color similarity with lighting invariance, shading, defocus, etc. Then they trained MILN [7] (Multiple-layer In-place

Learning Network) to find passable regions in those untrained images. Finally, they find passable regions from a single still image without image registration. The region growth algorithm [6] divided every training image into 40×40 windows as the feature vectors, which in each training image are marked as passable and non-passable regions manually. MILN combines the feature vectors as inputs and the corresponding passable regions as outputs for supervised learning.

3. Experiment Results

Experiment in this study is implemented by using the MATLAB. The experimental hardware platform is Inter Core i3, 2450M, 2 G, 2.27 GHz, and software environment is the MATLAB R2012b under Windows 7 OS.

In this section, we summarize the steps of our algorithm and show some results:

- Step 1: According to the detected shadow edges, we calculate the K_H parameters.
- Step 2: using the K_H parameters, according to the formula: $\beta = \frac{\log(k_1)}{\log(k_3)} + \frac{\log(k_2)}{\log(k_3)}$, we calculate the β parameter.
- Step 3: According to the shadow invariant algorithm, we get an intrinsic image.
- Step 4: Obtain the initial ground plane results from the geometric layout by [11].
- Step 5: The final shadow resistant ground plane result is obtained by employing region growth algorithm [6] on the initial seed region in the shadow-free intrinsic image.

Figure 3 shows some results of our ground plane detection results on different scenes and different surfaces. We can see our algorithm can detect the ground plane without affected by the cast shadows on the ground.



(a) Original image



(f) Result image



(b) Original image



(g) Result image



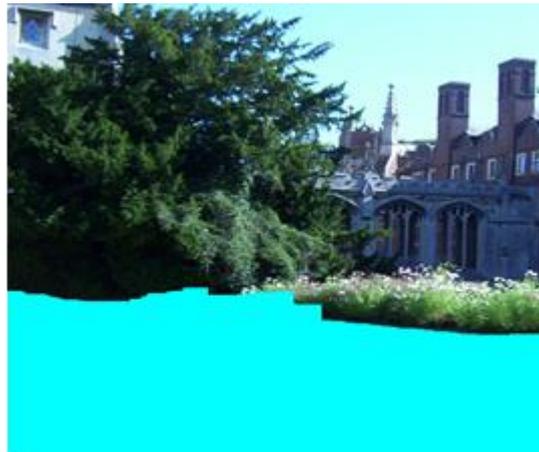
(c) Original image



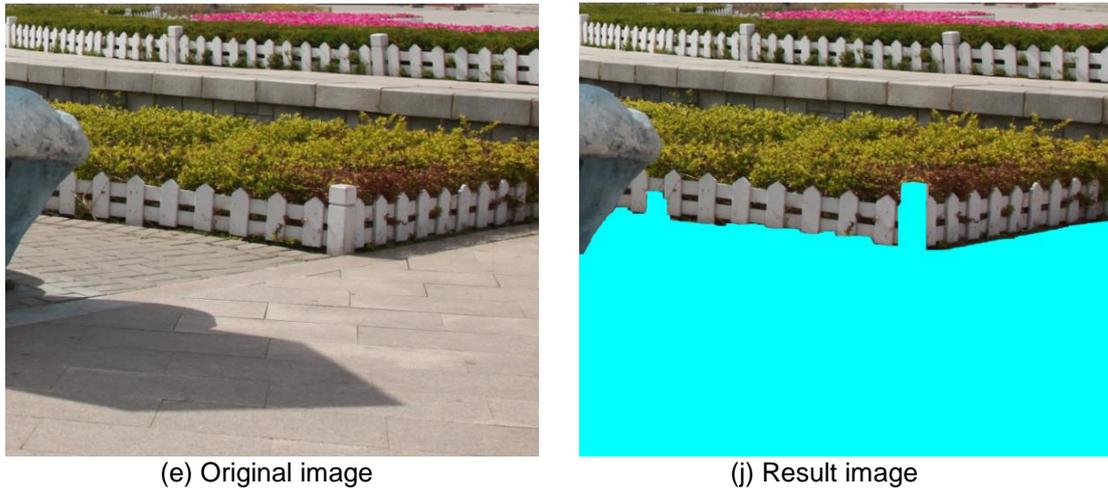
(h) Result image



(d) Original image



(i) Result image



**Figure 3. Results of our Ground Plane Detection Algorithm
(a),(b),(c),(d), (e): Original images; (f),(g),(h),(i),(j): Result images)**

4. Conclusion

Ground plane detection has many applications for autonomously navigating in unknown outdoor environments. In this paper, we proposed an algorithm to detect ground plane from a single color image. We first obtain the initial ground plane using geometric layout. We then obtain the grayscale intrinsic image by shadow invariant transformation on the roughly detected shadow edges. We finally apply the region growth which takes the initial detection as the seed region on the shadow-free intrinsic image. Experimental results show that our algorithm can detect the ground plane and is not sensitive to cast shadows on the ground.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (Grant No. 61301131).

References

- [1] P. Lombardi, M. Zanin and S. Messelodi, "Unified Stereovision for Ground, Road, and Obstacle Detection", Proceedings of IEEE Intelligent Vehicles Symposium, (2005); Las Vegas, United States.
- [2] J. Arr'ospide, L. Salgado, M. Nieto and R. Mohedano, "Homographybased Ground Plane Detection Using a Single On-board Camera", Intelligent Transport Systems, IET, vol. 4, no. 2 (2010), pp. 149–160.
- [3] D. Conrad and G. N. DeSouza, "Homography-based Ground Plane Detection for Mobile Robot Navigation Using a Modified EM Algorithm", Proceedings of IEEE International Conference on Robotics and Automation, (2010); Anchorage, United states.
- [4] V. Grabe, H. H. Bülthoff and P. R. Giordano, "Robust Optical-flow Based Self-motion Estimation for a Quadrotor UAV", Proceedings of IEEE International Conference on Intelligent Robots and Systems, (2012); Vilamoura, Portugal.
- [5] R. Spica, P. R. Giordano and F. Chaumette, "Plane Estimation by Active Vision from Point Features and Image Moments", Proceedings of IEEE International Conference on Robotics and Automation, (2015); Seattle, United States.
- [6] J.-D. Tian, W. Dong and Y.-D. Tang, "Finding Passable Regions from a Single Still Image", International Journal of Robotics and Automation, vol. 24, no. 3, (2009), pp. 228-234.
- [7] J.-Y. Weng, T.-Y. Luwang, H. Lu and X.-Y. Xue, "A Multilayer In-place Learning Network for the Development of General Invariances", International Journal of Humanoid Robotics, vol. 4, no. 2, (2007), pp. 281-320.
- [8] A. Cherian, V. Morellas and N. Papanikolopoulos, "Accurate 3D Ground Plane Estimation from a Single Image", Proceedings of IEEE International Conference on Robotics and Automation, (2009); Kobe, Japan.

- [9] I. Ulrich and I. Nourbakhsh, "Appearance-Based Obstacle Detection with Monocular Color Vision", Proceedings of the AAAI national Conference on Artificial Intelligence, (2000) ; Austin, TX.
- [10] L. M. Lorigo, R. A. Brooks and W. E. L. Crimson, "Visually-guided Obstacle Avoidance in Unstructured Environments", Proceedings of the IEEE International Conference on Intelligent Robots and Systems, (1997); Grenoble, Fr.
- [11] D. Coombs, M. Herman, T.-H. Hong, Nashman and Marilyn, "Real-time Obstacle Avoidance Using Central Flow Divergence, and Peripheral Flow", IEEE Transactions on Robotics and Automation, vol. 14, no. 1, (1998), pp. 49-59.
- [12] D. Hoiem, A. A. Efros and M. Hebert, "Recovering Surface Layout from an Image", International Journal of Computer Vision, vol. 75, no. 1, (2007), pp. 151-172.
- [13] A. Saxena, S. H. Chung, and A. Y. Ng, "3-D depth Reconstruction from a Single Still Image", International Journal of Computer Vision , vol. 76, no. 1, (2007), pp. 53-69.
- [14] W. N. Klarquist and W. S. Geisler, "Maximum Likelihood Depth from Defocus for Active Vision", Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, (1995); Pittsburgh, USA.
- [15] A. N. Rajagopalan, S. Chaudhuri and U. Mudenagudi, "Depth Estimation and Image Restoration Using Defocused Stereo Pairs", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 11, (2004), pp. 1521-1525.
- [16] J. A. Marshall, C. A. Burbeck, A. Dan, J. P. Rolland and K. E. Martin, "Occlusion Edge Blur: A Cue to Relative Visual Depth", Journal of the Optical Society of America A, vol. 13, no. 4, (1996) pp. 681-688.
- [17] J.-D. Tian, X.-J. Qi, L.-Q. Qu and Y.-D. Tang, "New Spectrum Ratio Properties and Features for Shadow Detection", Pattern Recognition, vol. 51, no. 1, (2016), pp. 85-96.

Authors



Xiaoyan Xu, She received her M.S. degree in Signal and information processing from HeiLongjiang University,China, in 2008. She is now a Ph.D. candidate in College of Information Science and Technology, Dalian Maritime University.Her main research interests in image processing and pattern recognition.



Xiaoming Liu, He received his M.S. and Ph.D. degrees in Communication and information system from Dalian Maritime University,China. He is a professor in College of Information Science and Technology, Dalian Maritime University. His research interests include radar signal processing, electronic information system and image processing.

