ROBUST GROUND PLANE DETECTION WITH NORMALIZED HOMOGRAPHY IN MONOCULAR SEQUENCES FROM A ROBOT PLATFORM

Jin Zhou and Baoxin Li

Dept. of Computer Science & Engineering
Arizona State University, Tempe, AZ 85287, USA

Abstract

We present a homography-based approach to detect the ground plane from monocular sequences captured by a robot platform. By assuming that the camera is fixed on the robot platform and can at most rotate horizontally, we derive the constraints that the homograph of the ground plane must satisfy and then use these constraints to design algorithms for detecting the ground plane. Due to the reduced degree of freedom, the resultant algorithm is not only more efficient and robust, but also able to avoid false detection due to virtual planes. We present experiments with real data from a robot platform to validate the proposed approaches.

Index Terms – Vision-based navigation, homography

1. INTRODUCTION

In mobile robot navigation, one of the fundamental problems is obstacle avoidance. Among various possibilities, vision-based obstacle avoidance has always been one attractive option. This is even more so nowadays with the availability of low-cost imaging sensors and compact yet high-performance processors as a result of recent technology development. There are different vision-based approaches to obstacle avoidance. In this paper, we focus on the problem of ground plane detection, which entails naturally obstacle detection, since in many applications the mobile robot may be considered to maneuver on a planar surface and thus obstacle detection may be reduced to the problem of ground plane detection: With the ground plane detected, other objects can be viewed as obstacles if they are on the direction of movement and outside of the ground plane.

Various approaches have been proposed to address the problem of ground detection. For example, simple approaches identify the ground floor using color information [2,4,7]. While being simple to implement, these approaches are suitable only for very specific environments. To handle general environments, a few systems attempt to recover the structure of the scene. To this end, different techniques have been used. For example, the work in [3,6] uses stereo vision for this purpose. Presently, monocular vision based approaches

are more popular and attractive since they are typically more cost-effective and also there is no stringent requirement on camera calibration (as in stereoscopic approaches). For example, in [8], optical flow is used in computing the surface normal for different image patches, which are then grouped to detect the ground floor. The optical-flow-based approaches tend to be computationally costly and not robust to unpredictable motion of a mobile platform.

In this paper, aiming at designing a practical system to detect the ground plane with monocular vision, we propose a method that makes use of a sparse set of features detected from a corner detector. The motion of these features is modeled by a homography transformation and in turn used in detecting the ground plane. Further, we utilize the constraints arising from the target application to greatly simplify the homography model, resulting in more efficient and robust algorithms. Moreover, the inherent “virtual plane problem” is solved naturally. Consequently, compared with prior work on the same task using plane homography (such as [5]), the proposed method performs better in finding the real ground plane. Also the proposed method is not restricted to translational motion as in [11].

2. HOMOGRAPHY-BASED GROUND PLANE DETECTION

Our task is to detect the ground plane from a monocular sequence captured by a camera mounted on a robot platform navigating on a planar surface. In theory, in the monocular sequence, points on the same plane share a homography transformation in two views. That is, for a set of point correspondences $\{x_i \leftrightarrow x_i'\}$ in two images, if all of the points are coplanar, then there is a homography matrix H such that

$$x_i' = Hx_i \quad (1)$$

where x represents a homogeneous image coordinate $(x, y, w)^T$, and H is a 3 by 3 matrix. Since x is represented by homogeneous coordinates, Eqn. (1) is true up to an unspecified scale factor. Thus H has only eight degrees of freedom. To determine such an H , four non-degenerated point correspondences are required since each point correspondence provides two independent constraints, although in practice typically

more points are used to improve the accuracy.

Apparently, different planes have different homographies, and thus theoretically if we find a homography that includes at least three points on the ground, it corresponds to the ground plane. Thus Eqn. (1) suggests a way for detecting the ground plane through grouping the detected feature points into coplanar sets, with each set sharing a common homography. If we assume that the ground plane contains the most feature points, then we can detect it by searching for a dominant homography that accounts for the most feature points in two views. We can then use this homography to determine if any other feature point is on the plane and thus achieving the detection of the ground plane.

2.1. The Virtual Plane Problem

Even when the assumption that the ground plane contains the most points is satisfied, there still exists potential problem in the above procedure, which we name as the *virtual plane* problem. That is, we may find a virtual plane, which may contain some ground plane points as well as some other obstacle points. Such a plane may contain more feature points than the actual ground plane does, although this plane does not correspond to a physical planar object (and hence the term “virtual”). This problem can easily occur since the number of feature points from an automatic feature detector is agnostic to the scene objects and thus where the feature points may be is unpredictable. In this paper, we exploit additional constraints arising from the target application to limit the search space for finding the dominant homography. The resultant approach solves the virtual plane problem naturally while lending itself to a more efficient and robust search algorithm, as detailed in subsequent sections.

3. HOMOGRAPHY FOR THE GROUND PLANE

Different planes have different homographies between two given views. For robot navigation on the ground plane, if the camera mounted on a mobile robot is fixed, there is some special pattern for the homography of the ground plane (we assume that the robot is moving via wheels just like a car). We derive this special pattern analytically in the following.

In general, we can set the camera matrices of two different views as:

$$\begin{aligned} P_1 &= K_1 R_1 [I | -C_1] \\ P_2 &= K_2 R_2 [I | -C_2] \end{aligned} \quad (2)$$

where P_i is 3×4 camera matrix, K_i represents the internal camera matrix, R_i the camera rotation matrix, and C_i the camera center coordinates ([9]). For a single camera, we can set $K_1 = K_2 = K$.

Fig 1. Illustration of the coordinate systems in relation to the ground plane.

We define the world coordinates system such that the y-axis is perpendicular to ground plane and the origin is at the same height as the camera, as illustrated in Figure 1. Based on the world coordinates system, we can have $C_i = (x_i, 0, z_i)^T$. The ground plane has coordinates $\pi_0 = (n_0^T, d)^T$ so that for points on the plane we have $n_0^T X + d = 0$, where $n_0 = (0, 1, 0)^T$.

From [9], when $P = K[I | 0]$, $P' = K[R | t]$, this plane correspond to a homography given by

$$H = K'(R - tn^T/d)K^{-1} \quad (3)$$

We can adjust (rotate and translate) the world coordinates system so that it is the same as that of the first camera. Then the new camera matrices become:

$$P_1' = K[I | 0], \quad P_2' = K[R_2 R_1^{-1} | -R_2 \Delta C] \quad (4)$$

where $\Delta C = (C_2 - C_1)$ and $n' = R_1 n$. Thus

$$H = K(R_2 R_1^{-1} + R_2 \Delta C n^T / d) \quad (5)$$

Since $R^{-1} = R^T$, we have

$$H = K R_2 (I + \Delta C n^T / d) R_1^{-1} K^{-1} \quad (6)$$

For a mobile platform moving on the plane, if the camera is fixed on the robot, we can have $R_2 = R_1 \Delta R$ where ΔR indicates the rotation around y-axis. More specifically, if the robot rotate θ degree on the plane, then

$$\Delta R = R_y(\theta) = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \quad (7)$$

From (6), we can compute the **normalized homography** $\hat{H}$ from any H :

$$\hat{H} = (K R_1)^{-1} H (K R_1) = \Delta R (I + \Delta C n^T / d) \quad (8)$$

For the ground plane, since $n_0 = (0, 1, 0)^T$, with $\Delta C = (x_0, 0, z_0)^T$, $\hat{H}$ has the following form:

$$\hat{H} = \begin{bmatrix} \cos(\theta) & x_0/d & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & z_0/d & \cos(\theta) \end{bmatrix} \quad (9)$$

Eqn. (9) shows that the normalized homography of the ground plane has just 3 degree of freedom, namely, θ , x_0/d and z_0/d . However, in order to compute normalized homography, we still need to know K and

R_1 . While K may be obtained by calibration, there is no straightforward way for determining R_1 . Moreover, if we choose different world coordinates system, we will have different R_1 . Here, we prove that if we keep the y-axis of the world coordinate system unchanged (i.e. always perpendicular to the ground plane), different values of R_1 do not change the form of (9).

Proof: Since the y-axis is unchanged, the world coordinate system can only rotate around the y-axis. Suppose R_1 and R_1' represent rotation matrix in different world coordinates system respectively, then we have $R_1' = R_1 \Delta R_1$, where ΔR_1 is rotation around y-axis so

$$\Delta R_1 = R_y(\varphi) = \begin{bmatrix} \cos(\varphi) & 0 & \sin(\varphi) \\ 0 & 1 & 0 \\ -\sin(\varphi) & 0 & \cos(\varphi) \end{bmatrix} \quad (10)$$

Then we have

$$\begin{aligned} \hat{H}' &= (KR_1')^{-1} H (KR_1') = \Delta R_1^{-1} \hat{H} \Delta R_1 \\ &= \Delta R_1^{-1} \Delta R (I + \Delta C n^T / d) \Delta R_1 \\ &= \Delta R \Delta R_1^{-1} (I + \Delta C n^T / d) \Delta R_1 \\ &= \Delta R (\Delta R_1^{-1} I \Delta R_1 + \Delta R_1^{-1} \Delta C n^T \Delta R_1 / d) \\ &= \Delta R (I + \Delta C' n^T / d), \quad \text{with } \Delta C' = \Delta R_1^{-1} \Delta C \end{aligned} \quad (11)$$

Since $\Delta C = (x_0, 0, z_0)^T$, with (10), we have

$$\Delta C' = (x_0', 0, z_0')^T$$

Therefore, the new normalized homography based on R_1' still has the form as (9). Consequently, the above observation provides us with the freedom to choose R_1 . In the next section, we will present a simple technique to determine R_1 as well as techniques to search the ground plane based on the normalized homography.

It is interesting to note that Eqn. (9) also provides the rotation and displacement of the two underlying views. This information can be useful for a robot.

4. GROUND DETECTION BASED ON NORMALIZED HOMOGRAPHY

4.1 Determination of R_1

We design a very simple approach to determine R_1 as follows:

- 1) Take an image with the camera.
- 2) For several (≥ 4) ground points, manually pick their corresponding points on the image.
- 3) Set the 2d coordinates of the ground plane points $\{x_i\}$. With their image coordinates $\{\tilde{x}_i\}$, compute the homography H such that $\tilde{x}_i = Hx_i$.
- 4) Estimate R_1 from H :

$$H' = K^{-1}H = \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix}, \quad \begin{cases} r_1 = p_1 / \|p_1\| \\ r_3 = p_2 / \|p_2\|, \\ r_2 = r_1 \times r_3 \end{cases}$$

$$R_1 = \begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix}$$

The following analysis explains why this approach works. For convenience, we set the origin of world coordinate system on the ground plane (Translation of the world coordinates system does not change R_1). Then for a ground plane point $X = (u, 0, v, 1)^T$, its image point is

$$\begin{aligned} \tilde{x} &= K[R_1 | t]X = K[r_1, r_2, r_3, t](u, 0, v, 1)^T \\ &= K[r_1, r_3, t](u, v, 1)^T = K[r_1, r_3, t]x \end{aligned}$$

Therefore, $H = K[r_1, r_3, t]$.

Note that R_1 needs to be computed only once as long as the camera is fixed on the robot platform or can only rotate around the y-axis.

4.2 Ground Plane Detection

With K and R_1 determined, we can compute the normalized homography from the original homography by Eqn. (8). The ideal normalized homography of the ground plane has the form of Eqn. (9) which has just 3 degree of freedom. Thus searching for the ground plane can be formulated as searching for a dominant normalized homography that has the form of Eqn. (9).

To compute the normalized homography, we first normalize the coordinates of the points: for a point x , its normalized coordinates are given by $\hat{x} = (KR_1)^{-1}x$. Then the homography computed from the normalized points is the normalized homography since:

$$\begin{aligned} x' &= Hx = (KR_1)\hat{H}(KR_1)^{-1}x \\ (KR_1)x' &= \hat{H}(KR_1)^{-1}x \\ \hat{x}' &= \hat{H}\hat{x} \end{aligned}$$

After we obtain a normalized homography, we try to fit it to the normalized homography model for the ground, given in the form of Eqn. (9). To search for a dominant model, we can use the RANSAC scheme. Based on the nature of our problem, we modify the basic RANSAC scheme to obtain the following algorithm.

Algorithm: Loop for N feature points.

- a) Randomly select a point, get its four closest neighbors that are not collinear; Use these 5 points to compute the normalized homography H .
- b) Fit H to Eqn. (9). If it fails to fit, go to (a). Otherwise, use this model to find more inliers and recompute H and fit again to find more inliers until the number of inliers do not increase.

At the termination of the procedure, the model with the most inliers is declared as the ground plane.

5. EXPERIMENTS

The experiments are based on the robot platform of Fig. 2(a). We use a paper with checkerboard patterns to measure R_1 , as shown in Fig. 2(b). We manually pick 7 grid points on the image. With these 7 image points and their corresponding coordinates on paper. We compute the homography and further determine R_1 with methods introduced in Section 4.

With K and R_1 known, we can apply our approach to the image sequences taken by the robot. Fig. 3 illustrates sample results, where we also compare with the approach of searching for the dominant homography directly without using the constraints introduced in this paper. Harris corner detector [1] was used in the experiments in feature detection (green crosses in the images). Feature correspondence was done based on the method reported in our previous paper [reference omitted for anonymity]. Fig 2 (d) is the best results of the direct search method, which include more non-ground points.

Fig 2. (a) The robot platform used in the experiments. (b) The set-up for computing the rotation matrix of the camera on the robot.

6. CONCLUSION

In this paper, we proposed a new method to detect the ground plane from monocular sequences captured by a robot platform. The core idea is to use the fixed camera-ground configuration to impose a strong constraint in searching the dominant homography. We designed complete algorithms and tested with real data from a robot platform. The results demonstrate the advantages of the proposed method.

References

- [1] C. J. Harris and M. Stephens, "A combined corner and edge detector," In *4th Alvey Vision Conference Manchester*, pp 147-151, 1988.
- [2] J. Hoffmann, M. Jungel and M. Lotzsch, "Vision Based System for Goal-Directed Obstacle Avoidance", *8th Int. Workshop on RoboCup 2004*.
- [3] K. Sabe, M. Fukuchi, J.-S. Gutmann, T. Ohashi, K. Kawamoto, and T. Yoshigahara, "Obstacle Avoidance and Path Planning for Humanoid Robots using Stereo Vision," *Proc. Int. Conf. on Robotics and Automation (ICRA'04)*, New Orleans, pp. 592 – 597, Vol.1 April 2004.
- [4] L. M. Lorigo, R. A. Brooks, and W. E. L. Grimsou, "Visually-guided obstacle avoidance in unstructured

environments," *Proc. IEEE/RSJ Int. Conf. Intel. Robot & Systems*, France, vol. 1, pp. 373-379, Sept. 1997.

[5] N. Pears and B. Liang, "Ground plane segmentation for mobile robot visual navigation," In *IROS 2001*, vol. 3, pp. 1513-1518, 2001.

[6] R. Mandelbaum, L. McDowell, L. Bogoni, B. Beich, and M. Hansen, "Real-time stereo processing, obstacle detection, and terrain estimation from vehicle-mounted stereo cameras," In *Proc. 4th IEEE Workshop on Applications of Computer Vision*, Princeton NJ, pp. 288-289, 1998.

[7] S. Lenser and M. Veloso, "Visual sonar fast obstacle avoidance using monocular vision," In *Proc. of the IEEE/RSJ IROS 2003*.

[8] Y. Kim and H. Kim, "Layered ground floor detection for vision-based mobile robot navigation," *Proc. IEEE Int. Conf. on Robotics and Automation*, vol. 1, pp. 13-18, 2004.

[9] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University, Cambridge, 2nd edition, 2003.

[10] J. Zhou and B. Li, "Homography-based Ground Detection for A Mobile Robot Platform Using a Single Camera" In *Proc. ICRA 2006*.

[11] B. Liang, N. Pears and Z. Chen, "Affine Height Landscapes for Monocular Mobile Robot Obstacle Avoidance", In *Proc. Intelligent Autonomous Systems*, pp. 863-872, August 2004.

Fig 3. (a)-(b) illustrate two views with feature points and their correspondence (indicated by red/dark line segments). (c) is the detected ground plane (green/light crosses) with the proposed approach in this paper. (d)-(f) is the results by simply searching for a dominant homography without using the constraint utilized in the proposed approach. (e)-(f) are situations with the virtual planes detected as the ground. In all figures, red/dark crosses indicate non-ground feature points.