Obstacle detection by ground homograph estimation during autonomous navigation

Yuan Gao, Jiayi Ma, Wenjuan Jiang, Jinwen Tian
State Key Laboratory for Multi-spectral Information Processing Technologies, IPRAI,
Huazhong University of Science and Technology, Wuhan 430074, China

ABSTRACT

In this paper, a simple and flexible framework based on the duality of obstacle detection and ground homograph estimation is proposed for obstacle detecting during autonomous navigation. The virtue of which is that the obstacle can be detected and localized without requiring the parameters of the camera. The iterative frame is implemented in our method to avoid the influence of the obstacles when estimating the ground homography. We initially estimate a homograph for the ground using matching points to register the image pairs; then, the two images are differenced so that the homograph disparity image (HDI) can be generated; after that, we segment the HDI to obtain a coarse location of obstacles; and the matching points out of the obstacle regions will be considered as input for the latter loop. This process is iterated until the results converge. The feasibility of this scheme is analyzed, and the experimental results demonstrate it can gain robust results for obstacle detection.

Keywords: obstacle detection, ground homograph estimation, iterative frame

1. INTRODUCTION

Autonomous navigation has become a hot issue in most developed countries these years. The application of autonomous navigation spans lots of areas. For instance, the research about that could lead to results of driver assistance systems which can increase road safety and comfort when driving [1-3]. Moreover, it can also be implemented in military industry to manufacture unmanned aerial vehicle and guided missiles. It is concerned that the US Department of Defense and NASA has been working on this topic for many years.

The fundamental problem of autonomous navigation is obstacle detection. Researchers have made much progress in obstacle detection such as DARPA Grand Challenge [4]. Labayrade et al. [5] use histogram analysis to find the disparities that belong to the ground surface. Regions located above this surface are considered to be obstacles. An improved version of this approach for obstacle detection can be used in rough terrain [6]. These methods usually using a precise mathematical model to give analytical solutions. However, for the lack of certainty model, they are difficult to guarantee robustness.

Our approach is based on ground homograph estimation. Which uses image pairs that can be obtained by stereo camera or monocular camera without requirement of the camera parameters. The experiment shows it can perform a robust result.

This paper is organized as follows: In section 2, the stages of the method are proposed. Section 3 presents experiments on real images as well as the analysis of the results. Finally, we make concluding remarks.

2. THE OBSTACLES DETECTION ALGORITHM

In this section, we describe the basic steps of our approach. Firstly, we introduce feature points extraction, image registration and evaluate the homograph. After that, the segmentation algorithm is applied, which can be used to localize the obstacles. Finally, an overall iterative algorithm is proposed for detecting and localizing the exact position of obstacles.

2.1 Homograph Evaluation

Image of points on a plane is related to corresponding image point in a second view by a (planar) homograph. This is a projective relation since it depends only on the intersections of planes with lines. It is said that the plane induces a

MIPPR 2011: Automatic Target Recognition and Image Analysis, edited by Tianxu Zhang, Nong Sang,
Proc. of SPIE Vol. 8003, 80030K · © 2011 SPIE · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.901206

homograph between the views. The homograph map transfers points from one view to another if they were images of points on the plane [7].

As shown in Figure 1, the ray corresponding to a point x is extended to meet the point x_π in the world plane π ; this point is projected to a point x' in the other image. The map from x to x' is the homograph induced by the plane π . There is perceptivity $x = H_{1\pi} x_\pi$ between the world plane π and the first image plane; and perceptivity $x' = H_{2\pi} x_\pi$ between the world plane and second image plane. The composition of the two perceptivities is a homograph $x' = H_{2\pi} H_{1\pi}^{-1} x = Hx$ between the image planes.

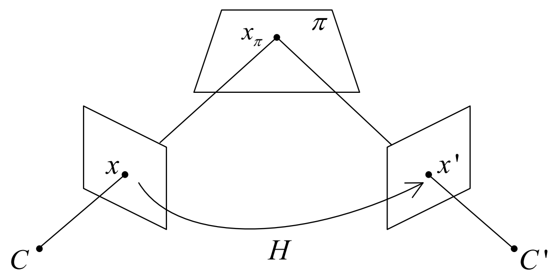

Fig. 1. The homograph induced by a plane

The search for discrete image correspondences can be divided into three main steps. First, feature points are selected at distinctive locations in the image, such as corners, blobs, and T-junctions. Next, a distinctive and robust descriptor is used to describe the interest points. Finally, the descriptor vectors are calculated to match between the corresponding images. The matching vectors are often measured by a distance between them, e.g. the Mahanalobis or Euclidean distance. The Scale Invariant Feature Transform (SIFT) [8] descriptor is chosen in our approach, since it is invariant to image scaling and rotation, and partially invariant to change in illumination and 3D camera viewpoint. Large numbers of features can be extracted from typical images with efficient algorithms.

Once we have the putative matches, we can use geometric alignment to verify the inliers as well as outliers. For example, if we expect the whole image to be translated or rotated in the matching view, we can fit a global geometric transform and keep only those feature matches that are sufficiently close to this estimated transformation. Here, the state-of-art algorithm in which the mismatch can be removed known as Victor Field Consensus (VFC) of Zhao and Ma [9] is applied. After that, the matching points we gain are highly likely to be inliers. And then, the homograph can be estimated using these matches.

2.2 Obstacle segmentation and localization

After estimating the homograph, we use it to match the image pair. Then calculate the absolute difference of the two matched images, which is named as homograph disparity image (HDI). The ground region has nearly zero value in HDI since it approximately fits the homograph. As a result, the obstacle region is highlighted. Then we apply a segmentation algorithm to segment the HDI into some homogeneous regions.

In this paper, Felzenszwalb's method [10] is implemented to segment the HDI. They start with a pixel-to-pixel dissimilarity measure $w(e)$ such as intensity differences between N_8 neighbors. For any region R, its internal difference is defined as the largest edge weight in the region's minimum spanning tree, as formula (1) shows:

$$Int(R) = \max_{e \in MST(R)} w(e) \tag{1}$$

For any two adjacent regions with at least one edge connecting their vertices, the difference between these regions is defined as the minimum weight edge connecting the two regions, which is indicated as formula (2):

$$Dif(R_1, R_2) = \min_{e=(v_1, v_2) | v_1 \in R_1, v_2 \in R_2} w(e) \tag{2}$$

And their algorithm merges any two adjacent regions whose difference is smaller than the minimum internal difference of these two regions, as denoted in formula (3):

$$MInt(R_1, R_2) = \min(Int(R_1) + \tau(R_1), Int(R_2) + \tau(R_2)) \quad (3)$$

Where $\tau(R)$ is a heuristic region penalty they set to $k/|R|$, and $|R|$ denotes the size of R, and k is some constant parameter. In practice k sets a scale of observation, in that a larger k causes a preference for larger components.

Eventually, the regions corresponding to large value in HDI are localized, which is considered to be obstacles.

2.3 An iterative obstacle detection algorithm

There is cyclic dependency between the previous two steps. In order to obtain an accurate homograph, we need to know exact obstacle regions. On the other hand, an ideal localization of obstacles requires an accurate homograph. We make use of the duality of obstacle detection and homograph estimation, and solve this problem by using an overall iterative framework similar to Expectation Maximum algorithm [11] to integrate the previous two steps, as shown in Figure 2.

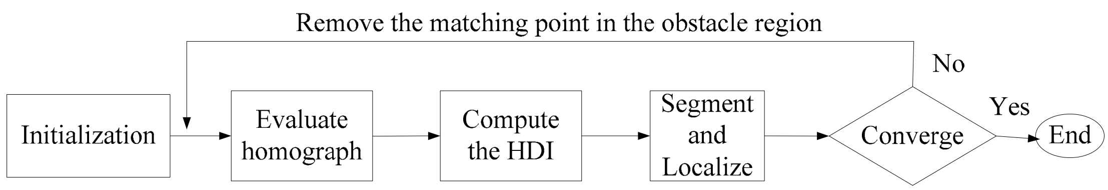

Fig. 2. The block diagram of the algorithm

Primarily, we use the feature points out of the obstacle regions to evaluate the homograph, as for initiation, we set the whole image as the ground if we do not have any prior. Then, the new homograph is used to calculate HDI. Finally, the HDI is segmented to localize the candidate obstacle regions. This process iterates until the obstacle regions is convergence when there is no matching points for the homograph estimation in the obstacles region. Typically, the algorithm will converge after running two to four times.

The algorithm is summarized as in Table 1.

Table 1. The summarization of the algorithm

Algorithm:
Input: A sequence of images.
Output: The location of the obstacle region.
1. Extract the feature points**:** Compute SIFT feature points in each of the given images.
2. **Repeat**
3. Compute the homograph: Use VFC algorithm to estimate the homograph H
4. Compute the homograph disparity image(HDI):
5. (a) match the original image pair by H;
6. (b) get the difference image, which is the absolute value of subtraction of the two matched images.
7. Segment the HDI and Localization the obstacles: Use Felzenszwalb's algorithm and get the obstacles' location.
8. Remove matching points in obstacle regions which is obtained in step 7 and update the feature points set.
9. **Until the result converges when there is no matching points for the homograph estimation in the obstacles region.**

3. EXPERIMENTS AND DISCUSSION

Our experimental results are shown in Figure 3. An image pair is represented in Figure 3(a) (b). Its SIFT matches are depicted in Figure 3(c). Figure 3(d) shows the result after the outliers have been removed by VFC. The initial HDI (Figure 3(e)) and segmentation result is in Figure 3(f). The texture of the ground in initial HDI is obvious, since the homograph is rough and the ground in the image pair hasn't been matched exactly. After using the iterative algorithm, feature points in obstacle region have been removed as Figure 3(g) shows, then the HDI (Figure 3(h)) become better, and the segmentation of ground and obstacles (Figure 3(i)) is much more ideal.

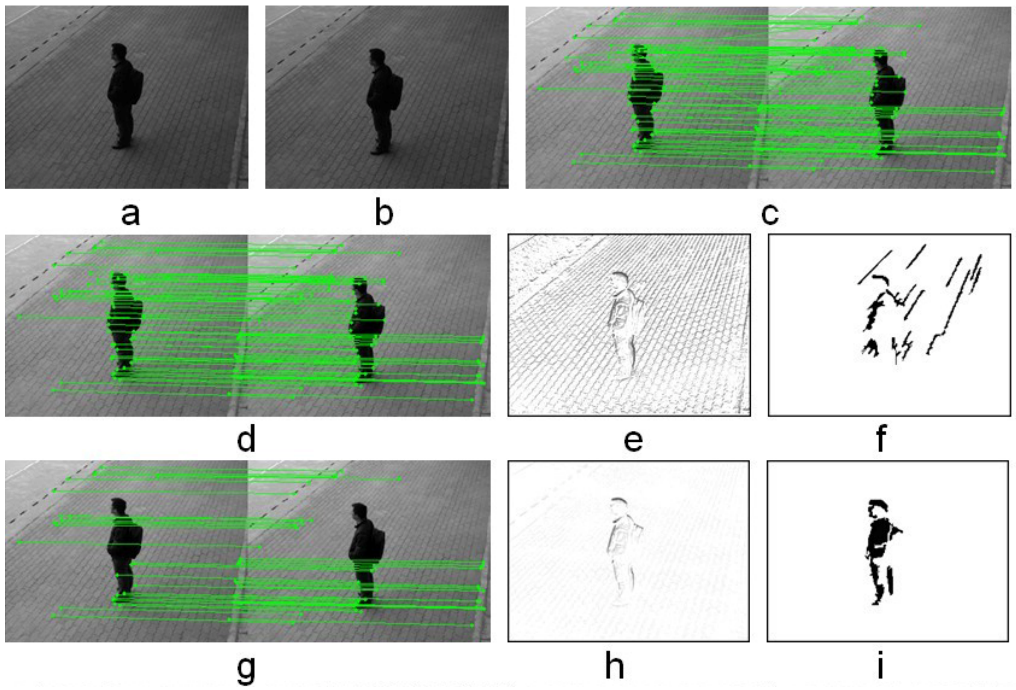

Fig. 3. The experiment process of the algorithm

The final obstacle localization result of Figure 3 and some other obstacle localization results are shown in Figure 4.

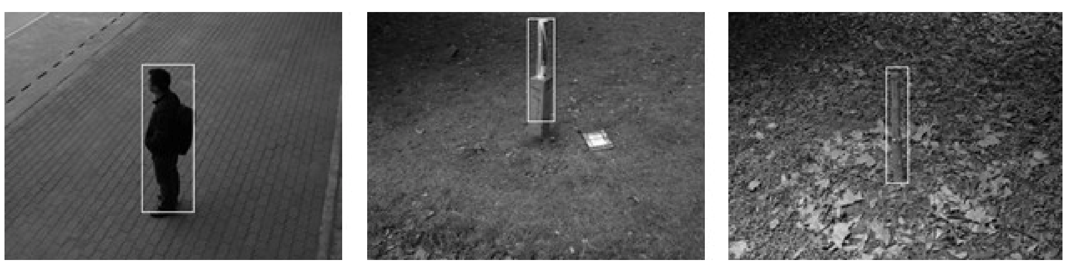

Fig. 4. The results of the algorithm

For our algorithm, it is aware that when the background is approximately located on a plane, it could work quite well. Here we give a quantitative analysis.

As shown in Figure 5(a), f is the focal length, p denotes the image plane, O is the camera optical center, Π denotes the ground plane. Suppose the height of the aircraft relative to the ground is H, the obstacle height is h, the horizon distance of the obstacle to the aircraft in two image is d_1 and d_2, the pitch angle of the camera is θ, pixel resolution is k (m / pixel), then we deduce the disparity(Δ) of the obstacle in two image, which takes the form as (4):

$$
\begin{aligned}
\Delta &= \Delta(x_1 - x_2) \\
&= \Delta\{\frac{f}{k}[\tan(\frac{\pi}{2} - \theta - \theta^{'}) - \tan(\frac{\pi}{2} - \theta - \theta^{"})]\} \\
&= \frac{f}{k}\{[\tan(\frac{\pi}{2} - \theta - \arctan\frac{d_1}{H}) - \tan(\frac{\pi}{2} - \theta - \arctan\frac{d_1}{H-h})] \\
&\quad -[\tan(\frac{\pi}{2} - \theta - \arctan\frac{d_2}{H}) - \tan(\frac{\pi}{2} - \theta - \arctan\frac{d_2}{H-h})]\}
\end{aligned} \tag{4}
$$

Now, we analyze the influence of the ground undulation. In our first experiment, as shown in Figure 3(a)(b), $f = 24mm$, $k = 10^{-5}m/\text{pixel}$, $\theta = 30^{o}$, $d_1 = 4m$, $d_2 = 3.8m$, $H = 4m$ the relationship between Δ an h is shown in Figure 5(b). From that, we can see if the undulation of the background is less than 0.1 m, the disparity will be less than half a pixel. Therefore we could ignore the influence of the undulating ground. In other words, we can treat the ground as a plane approximately in the algorithm.

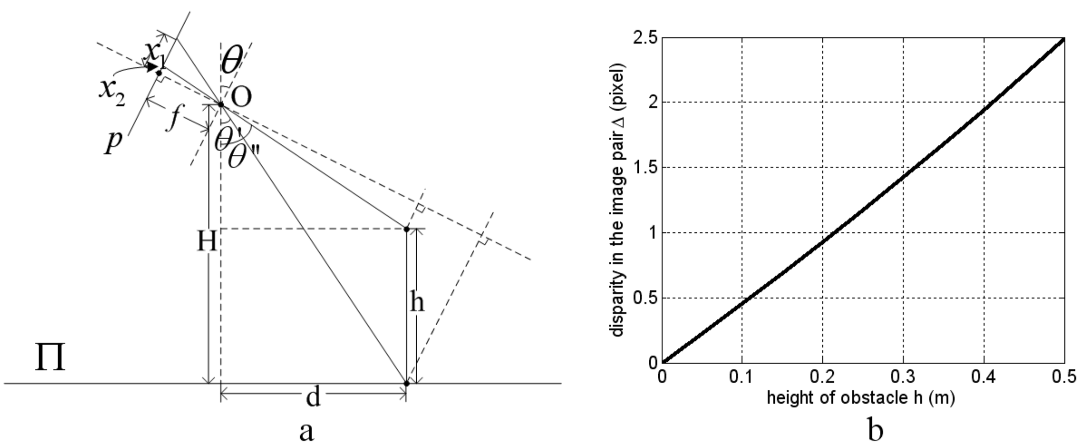

Fig. 5. The schematic diagram and relationship between disparity and height. Fig.6(a): Schematic diagram of imaging model. Fig.6(b) The relationship between disparity Δ and obstacle height h.

4. CONCLUSION

We have shown an iterative scheme for integrating the relationship between obstacles and the background for obstacle detection. The background in an image pair could be filtered by a homograph which could be obtained by our proposed iteration algorithm. We also analyzed the feasibility of our algorithm, and have given a reasonable explanation for why our algorithm could work well. This work can be applied for obstacle avoidance of low flying unmanned aerial vehicles and mobile robots.

ACKNOWLEDGEMENT

This work was funded by National Natural Science Foundation of China (Grant No.50909079).

REFERENCES

1. G. Toulminet, M. Bertozzi, S. Mousset, A. Bensrhair, A. Broggi, “Vehicle Detection by Means of StereoVision-Based Obstacles Features Extraction and Monocular Pattern Analysis,” *IEEE trans on Image Processing*, 15(8), 2364-2375, 2006
2. Q. Li, N. Zheng, and H. Cheng, “Springrobot: A prototype autonomous vehicle and its algorithms for lane detection,” *IEEE Trans. Intell. Transport,* vol. 5, no. 4, pp. 300–308, Dec. 2004.
3. Y. U. Yim and S. Y. Oh, “Three-feature based automatic lane detection algorithm (TFALDA) for autonomous driving,” *IEEE Trans. Intell. Transport. Syst.,* vol. 4, no. 4, pp. 219–225, Dec. 2003.

4. DARPA Grand Challenge: http://www.grandchallenge.org
5. R. Labayrade, and J. Tarel, "Real Time Obstacle Detection on Non-Flat Road Geometry through 'V-Disparity' Representation", *In IEEE Intelligent Vehicle Symposium*, 2002, 2, pp. 646–651
6. A. Broggi , C. Caraffi, P. Porta, and P. Zani, "The Single Frame Stereo Vision System for Reliable Obstacle Detection used during the 2005 DARPA Grand Challenge on TerraMax", *In IEEE Intelligent Transportation System Conference*, 2006, pp. 745–752.
7. R. Hartley, and A. Zisserman, *Multiple View Geometry in Computer Vision.* Cambridge University Press, Cambridge, UK, 2004
8. D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, November 2004.
9. J. Zhao, J. Ma, J. Tian, J. Ma, and D. Zhang, "A Robust Method for Vector Field Learning with Application to Mismatch Removing". *In IEEE Conference on Computer Vision and Pattern Recognition*, 2011.
10. P. Felzenszwalb and D. Huttenlocher, "Efficient Graph-Based Image Segmentation," *International Journal of Computer Vision*, vol. 59, 2004.
11. C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.