Planar Region Extraction for Visual Navigation using Stereo Cameras

Se-Na Lee*,**, Bum-Jae You* and Sung-Jea Ko**

* Intelligent Robotics Research Center, Korea Institute of Science and Technology, Seoul, Korea
(Tel : +82-2-958-5708; E-mail: sena@kist.ac.kr)
(Tel: +82-2-958-5760; E-mail: ybj@kist.re.kr)
** Department of Electronics Engineering, Korea University, Seoul, Korea
(Tel : +82-2-3290-3228; E-mail: sjko@dali.korea.ac.kr)

Abstract: In this paper, we propose an algorithm to extract valid planar regions from stereo images for visual navigation of mobile robots. The algorithm is based on the difference image between the stereo images obtained by applying Homography matrix between stereo cameras. Illegal planar regions are filtered out by the use of labeling of the difference images and filtering of invalid blobs using the size of each blob. Also, illegal large planar regions such as walls are removed by adopting a weighted low-pass filtering of the difference image using the past difference images. The algorithms are experimented successfully by the use of stereo camera system built in a mobile robot and a PC-based real-time vision system.

Keywords: visual navigation, mobile robots, planar region extraction, Homography matrix, sequential difference images

1. INTRODUCTION

Mobile robots are used widely for service robots in indoor environments such as houses, buildings, and offices, hospitals, and etc. Those mobile robots have to be equipped with a capability to navigate dynamic environments to execute a given task while avoiding obstacles. A number of sensors – such as a laser finder, ultrasonic sensors, a stereo-camera-based range sensor, and other range sensors – are used widely in order to detect obstacles in natural environments. Most of sensors are, however, expensive until now and it is not easy to apply for low-cost home service robots such as vacuum cleaning robots or other home service robots. So, visual navigation takes much attention when web cameras were introduced a few years ago since its cost is attractive comparing with previous other sensors.

For visual navigation, there have been proposed a number of approaches based on single camera [1][2], single camera combined with ultrasonic sensors [3] or a laser range finder [4], omnidirectional cameras, stereo cameras, and three-dimensional cameras capturing range data and image data simultaneously. In [1] and [2], visual navigation approaches using single camera were proposed. Ishikawa, Kuwamoto and Ozawa[1] presented a method for the navigation of a mobile robot using white line recognition. An input image is scanned horizontally to detect edges and to extract scan lines. Next, it scans the extracted scan line to recognize whole image pattern. Kutami, Maruya and Okuno[2] proposed an algorithm based on visual tracking of fixed targets as line markers to guide a vehicle along the line markers. These methods, however, are sensitive to intensity changes and do not work in complex environments.

A stereo camera system is adopted for visual navigation based on the extraction of a depth map for environments in [5] and [6]. Tistarelli, Grosso and Sandini[5] developed a cooperation algorithm in which binocular disparity, computed on merged several stereo images over time, is combined with optical flows from monocular image sequences acquired during ego-motion to compute the time to be elapsed before a collision. Burt, Wixson and Salgian[6] proposed that the subspace defined by the disparity gradients of visible surfaces can be used to find these critical plane. Disparity gradients of surfaces at an infinite distance from the camera are zero. They are able to shape the observation window so that it has zero disparity and is made to lie on the ground plane. Camera calibration of stereo cameras was an important procedure to precisely determine camera parameters and the geometrical relationship between stereo cameras. This process needs complex calculations while the robustness and real-time performance of depth estimation were not enough for real-time visual navigation.

By adopting Fundamental matrix and Homography matrix, uncalibrated video camera systems can be used for visual navigation and stereo image processing. Fundamental matrix [7] models the epipolar geometry between two views taken by uncalibrated cameras. The importance of fundamental matrix has been neglected in the sense that almost all of works on motion has been done under the assumption that intrinsic parameters are known [8]. The fundamental matrix is a 3x3 singular matrix whose rank is two while the left and right null-space of the fundamental matrix are generated by the vectors representing two epipoles [9]. Hiroshi Hattori [10] proposed a pseudo-projective camera model based on several assumptions on road scenes and on the linear epipolar geometry for a pair of cameras. It uses traffic lines in order to deal with camera vibrations and inclination to overcome the noise and to estimate the danger of obstacles.

Homography matrix relates the coordinates of points in two images in a planar region. Due to sensitivity to noises in the estimation of fundamental matrix particularly in scenes with a single plane like a corridor, we aim to use primarily homography matrix to detect the planar region [7]. Okutomi, Nakano, Maruyama and Hara have proposed a method to estimates continuously the homography matrix of a pair of images and to find planar regions for the navigation of an autonomous vehicle. They adopted Kalman filter-based filtering technique to remove illegal planes and illegal feature points.

In this paper, a simple and fast algorithm is proposed to extract planar regions such as the floor using a stereo camera system. A difference image between its left image and the transformed image of its right image by homography matrix is used as a basic image. The difference image is converted into a binary image to distinguish desired planar regions from other objects. However, it is observed that undesired regions - large regions without textures - are appeared to be valid. In order to overcome the difficulty, we adopt labeling and filtering based on the size of each label. Also, a low-pass filtering of the difference image using past sequential difference images is applied to distinguish the desired planar regions (floors) and undesired planar regions such as walls since edge information between the floors and the walls is removed while obtaining the difference image and two planar regions are combined together finally. The proposed algorithms are experimented successfully on a real-time PC-based vision system equipped with a Matrox GENESIS-LC frame grabber.

This paper is organized as follow. In section 2, we review the concept of homography matrix and introduce the way of extraction of planar regions using the labeling and filtering based on the size of each blob. Section 3 explains the use of past sequential difference images to differentiate the floor from the walls by weight invariance for accuracy. In section 4, there is proposed an algorithm to compute free spaces from planar regions in the difference image for avoiding obstacles. Section 5 shows some experimental results in indoor environments while the paper is concluded in section 6.

2. PLANAR REGION DETECTION

2.1 Homography matrix

When a stereo camera system is used for a robot, a point on a plane in real world is projected into each image plane of the stereo cameras. Homography is a projective transformation between the two projected points of a point in different two image planes represented as a nonsingular 3×3 matrix as shown in Fig. 1. The origin of the image coordinate is to the center of the image. If P_L and P_R are projected points of a same point on a plane in two image planes, it is represented by the following the homography matrix.

$$P_R = HP_L \tag{1}$$

In the equation (1), the matrix H is called Homography matrix

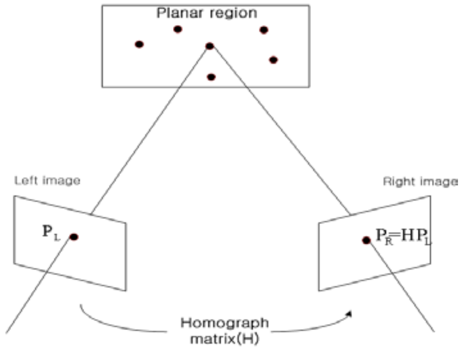


Fig. 1. Homography Matrix

whose dimension is 3 x 3. For each pair of corresponding points in stereo images, equation (1) can be rewritten as the following linear equation.

$$P_R \times HP_L = 0 \tag{2}$$

Using (2), we can make the matrix A as:

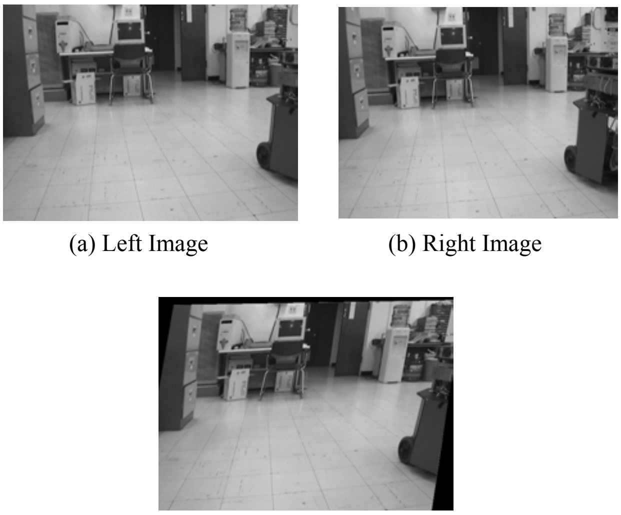
$$A = \begin{bmatrix} x_1' & y_1' & 1 & 0 & 0 & 0 & -x_1'x_1 & -y_1'x_1 & -x_1 \\ 0 & 0 & 0 & x_1' & y_1' & 1 & -x_1'y_1 & -y_1'y_1 & -y_1 \\ x_2' & y_2' & 1 & 0 & 0 & 0 & -x_2'x_2 & -y_2'x_2 & -x_2 \\ 0 & 0 & 0 & x_2' & y_2' & 1 & -x_2'y_2 & -y_2'y_2 & -y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n' & y_n' & 1 & 0 & 0 & 0 & -x_n'x_n & -y_n'x_n & -x_n \\ 0 & 0 & 0 & x_n' & y_n' & 1 & -x_n'y_n & -y_n'y_n & -x_n \end{bmatrix}$$

$(x_1, y_1, 1)$ and $(x_1', y_1', 1)$ are the coordinates of corresponding points P_R and P_L of the same point in real space. n is the number of the points. Given point matches P_L and P_R for stereo images, the homography matrix between two image planes becomes the least eigenvector of $A^T A$. In order to obtain a more stable homography matrix, the coordinates of the corresponding points are normalized by the way as introduced in [9]. Homography matrix has only eight degrees of freedom requiring that we have four corresponding coplanar features in each image [7].

Fig. 2 shows the result of the extraction of the homography matrix based on corresponding points in the floor. Fig. 2-(a) and Fig. 2-(b) are a pair of images taken from the left video camera and the right video camera of a stereo camera, respectively. Fig. 2-(c) is the transformed image of the left image, Fig. 2-(a), by the homography matrix.

2.2 Planar Regions Extraction

If we determine the homography matrix between stereo images of points on a floor and extract the difference image



(a) Left Image (b) Right Image

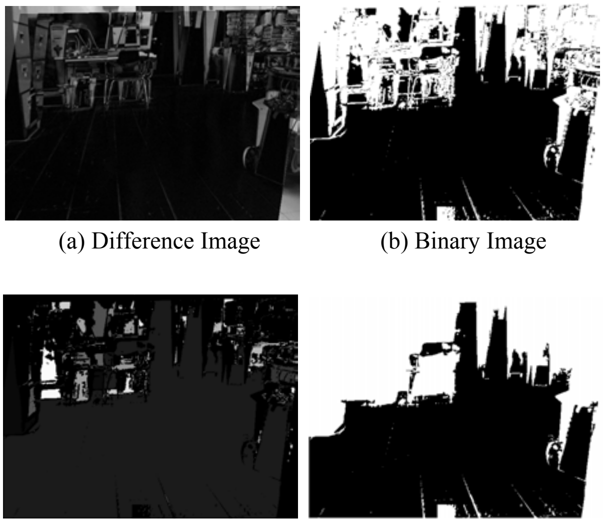
(c) Transformed Image

Fig. 2. Transform by Homography Matrix

between the right image and the transformed image of the left image, the floor region has dark gray levels while other regions have bright gray levels as shown in Fig. 3-(a). So, the difference image can be converted to a binary image, as shown in Fig. 3-(b), to distinguish the floor from non-floor regions by applying thresholding techniques. The existence and location of an obstacle can be determined from the binary image since the black region in the binary image is the floor and white regions shows obstacles. Since most mobile robots are located on a floor, the area of the floor is larger than other regions. So, it is possible to find the floor by analyzing the size of each labeled blob in the black area after labeling. The labeled image is shown in Fig. 3-(c). After labeling, small black blobs are converted into white blobs because those blobs are not located on the floor. It is observed, however, that some large regions without textures such as walls appeared to be black even though those are not located on the floor. Disconnected black blobs are converted into white blobs since most separated black blobs from the largest black blob represent wall regions or vertical planar regions of some objects. After this filtering process, Fig. 3-(d) is obtained.

3. SEQUENTIAL DIFFERENCE IMAGES

By adopting the labeling and the filtering of blobs based on the size of each blob, it is possible to remove a number of non-planar regions. However, some undesired large planar regions such as walls are appeared as black blobs because a portion of the line separating desired planar regions from undesired planar regions disappears during obtaining the difference image. This makes the desired planar regions are merged with the undesired planar regions. To solve this problem, we adopt the weighted sum of past difference images. The procedure is shown in Fig. 4.



(a) Difference Image (b) Binary Image

(c) Labeled Image (d) Size-filtered Image

Fig. 3. Planar Region Extraction

Step 1)

Compute the difference of initial labeled image (t_0) and first frame's labeled image (t_1). The difference is used as the threshold to compute weight of the next frame. At first, the weight of initial image and first image is set 0.5.

Step 2)

Obtain the result image using the weight invariance.

$$result\ image = w_0 * transformed\ image(t_0) - w_1 * transformed\ image(t_1)$$

Step 3)

The planar region is almost same at each frame, so the difference is very small. Furthermore, we decide the weight invariant using the difference of black regions. The difference of the size-filtered image of initial frame(L_0) and the size-filtered image of next frame(L_1) is computed. If there is some change in the next image, then the difference will be greater than the threshold. Hence, for the next image we use a larger weight than the result image in step 2.

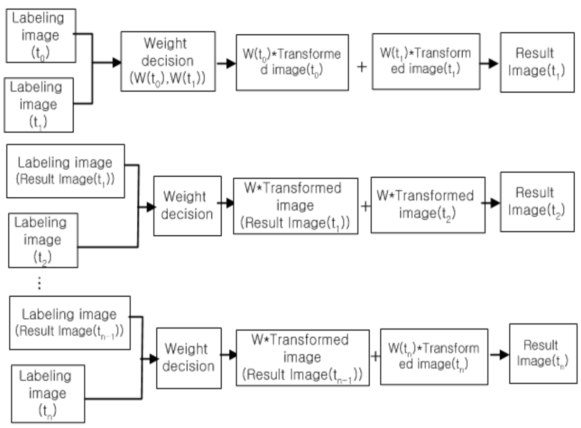


Fig. 4 Proposed approach using sequential images

4. COMPUTATION OF FREE SPACE

Since most of mobile robots move on a flat planes such as roads or corridors, the detection of obstacles on the ground plane is a key issue for many applications.

To decide a direction of a motion for a mobile robot, we use the histogram of white pixels at the same distance from the center of the mobile robot. We suppose the center of the bottom line of the image is the shifted center of the mobile robot. If a camera faces a plane perpendicularly, we can get a half circle by connecting the points with a same distance from the mobile robot. However, the camera is rotated with respect to x-axis, so the half-circle is skewed. Fig. 5 is the process to decide the direction to move. Fig. 5-(a) is the original image. Fig. 5-(b) is the projected half-circle on the labeled image. The black regions are the free space in the floor while the white regions are areas with obstacles. Using the histogram between the lines of the skewed half-circles, we can decide the path for the mobile robot to move. Fig. 5-(c), Fig. 5-(d), Fig. 5-(e), and Fig. 5-(f) are is the histograms for No. 1 circle, No. 2 circle, No. 3 circle, and No. 4 circle, respectively.

The x-axis in the histogram is the width (horizontal line) in the image, and the y-axis is the counter of the white pixels. The block in the histogram is the obstacles in planar region. If the central area of x-axis in the histogram is empty, the mobile robot can go straight. But, if there exist some blocks, the mobile robot have to avoid some obstacles. The direction of the motion for obstacle avoidance is decided by comparing the free space between the right and left area of the center of x-axis in the histogram. If the right region of free space is bigger than the left region the mobile robot goes to the right region of the obstacles. Further, if the widths of two regions are same, the mobile robot can go by either direction.

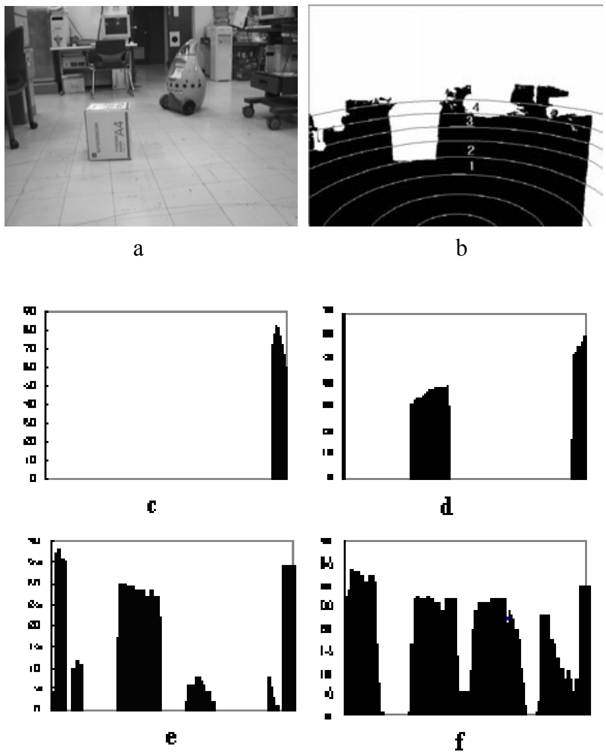


Fig. 5. Histogram of the obstacle between the same distance

Fig. 6. Developed Mobile Robot

5. EXPERIMENTAL RESULTS

We implemented the proposed algorithm on a 2.4GHz PC including a Matrox Genesis-LC frame grabber, and we use stereo cameras shown in Fig. 6 whose focal length is 6mm. Image resolution is 320 × 240 (pixels by pixels) for real-time image processing and a Matrox Genesis-LC frame grabber is used for the interface of stereo black-and-white video cameras. The mobile robot has differential type wheels and the mobile robot is equipped with a laser range finder made by SICK and a web camera. For our research, only the stereo cameras are used.

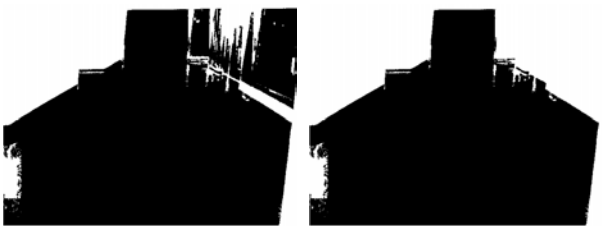
The proposed algorithm is tested in a laboratory and corridors in a building. Fig. 7 shows the results in the corridor and Fig. 8 are experimental results after filtering of labeled difference images. Finally, the robustness with respect to illumination variations is investigated as shown in Fig. 9 by using images under different illumination conditions.

The Fig. 7-(a) and Fig. 7-(b) are a pair of images for a corridor. Fig. 7-(c) is the difference image between transformed image of the left image by the homography matrix and the right image. Fig. 7-(d) is the difference image after labeling and filtering based on the size of each blob. But walls are not disappeared because some parts of boundary lines between floors and walls are disappeared while obtaining the difference image. So, we go to the next step.

(a) (b)

(c) (d)

Fig. 7. Experimental Results in a Corridor



(a) (b)

(c)

Fig. 8. Experimental Results after applying low-pass filtering of the difference image by previous sequential difference images with weight invariance

Fig. 8-(a) is the same image with Fig. 7-(d) including the walls as planar regions that a mobile robot can navigate while Fig. 8-(b) is the result after applying low-pass filtering based on past sequential difference image with weight invariance. Fig. 8-(c) is the original right image superimposed by the detected planar region in Fig. 8-(b).

Fig. 9 shows experimental results under various illumination conditions since the proposed algorithm should be robust in order to apply for real mobile robots. As shown in Fig. 9, the proposed algorithm detects floors in a robust manner. One of difficulties is to remove the effect of the reflection of lights by floors. We can find the effect of the reflection in the bottom line of Fig. 9-(b).

(a) (b)

(c) (d)

(e) (f)

Fig. 9. Experimental results of Floor Detection under Various Illumination Conditions

6. CONCLUSION

In this paper, we presented an algorithm that estimates valid planar regions for a mobile robot to pass using a difference image of uncalibrated stereo images and homography matrix. The algorithm can eliminate illegal planar areas in the input image by using the labeling and filtering based on the size of each blob. However, the resultant image has illegal large black regions that are not floors since some parts of the boundary line are disappeared which made illegal planar regions to have the same label as valid planar regions while getting the difference image. We proposed a low-pass technique using past sequential difference images with weight invariance. It had the advantage in identifying the wall from the floor. Finally, the developed algorithm is experimented successfully in a stereo camera system equipped on a mobile robot by using a PC-based real-time vision system.

In the future study, more robust techniques to remove the effect of light reflection in the floor will be investigated.

REFERENCES

[1] S. Ishikawa, H. Kuwamoto and S. Ozawa, “Visual navigation of and autonomous vehicle using white line

recognition," *IEEE Trans. on PAMI*, vol.10, pp.743-749, Sept. 1988

[2] A. Kutami, Y. Maruya, H. Takahashi and A. Okuno, "Visual navigation of autonomous on-road vehicle," *IEEE Proc. on IROS*, pp.175-180, vol.1, July 1990.

[3] A. Ohya, A. Kosaka and A. Kak, "Vision-based navigation of mobile robot with obstacle avoidance by single camera vision and ultrasonic sensing," *IEEE Proc. on IROS*, pp.704-711, vol.2, Sept. 1997.

[4] N. Bourbakis and R. Andel, "Fusing laser and image data for 3D perceived space representations," *IEEE Proc. on Tools with Artificial Intelligence*, pp.50-58, Nov. 1997.

[5] M. Tistarelli, E. Grosso and G. Sandini, "Dynamic stereo in visual navigation," *IEEE Proc. on Computer vision and Pattern Recognition*, pp.186-193, June 1991.

[6] P. Burt, L. Wixson and S. Salgian, "Electronically directed "focal" stereo," *IEEE Proc. on Computer Vision*, pp.94-101, June 1995.

[7] N. Pears and Bojian Liang, "Ground plane segmentation for mobile robot visual navigation," *IEEE/RSJ Proc. on IROS*, pp.1513-1518, vol.3, Nov. 2001.

[8] Q.-T. Luong and O.D, "Determining the fundamental matrix with planes: instability and new algorithms," *IEEE Proc. on CVPR*, pp.489-494, June 1993.

[9] R.I. Hartley, "In defense of the eight-point algorithm," *IEEE Tran. on PAMI*, pp.580-593, vol.19, June 1997

[10] H. Hattori, "Stereo for 2D visual navigation," *IEEE Proc. on Intelligent Vehicles Symposium*, pp.31-38, Oct. 2000.