OUKA

The University of Osaka
Institutional Knowledge Archive

Title	Environmental map generation and egomotion estimation in a dynamic environment for an omnidirectional image sensor
Author(s)	Yagi, Y.; Shouya, K.; Yachida, M.
Citation	Proceedings - IEEE International Conference on Robotics and Automation. 2000, 4, p. 3493-3498
Version Type	VoR
URL	https://hdl.handle.net/11094/14074
rights	c2000 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE..
Note	

Environmental Map Generation and Egomotion Estimation in a Dynamic Environment for an Omnidirectional Image Sensor

Y. Yagi K. Shouya M. Yachida
Department of Systems and Human Science
Graduate School of Engineering Science, Osaka University
Toyonaka, Osaka 560-8531, Japan
y-yagi@sys.es.osaka-u.ac.jp

Abstract

Generation of a stationary environmental map is one of the important tasks for vision based robot navigation. Under the assumption of known motion of a robot, environmental maps of a real scene can be successfully generated by monitoring azimuth changes in an image. Several researchers have used this property for robot navigation However, it is difficult to observe the exact motion parameters of the robot because of encoder measurement error of the robot. Therefore, observational errors in the generated environmental map accumulate in long movements of the robot. To generate a large environmental map, it is desirable not to assume known robot motion. In this paper, under the assumption of unknown translational motions of the robot, we propose a method to generate a stationary environmental map and estimate the egomotion of a robot in a dynamic environment, by using an omnidirectional image sensor. Since both robot and objects move in the environment, the stationary map generation and the robot egomotion estimation by using a single camera are difficult because of correspondence ambiguity caused by occlusion. The proposed method can detect a moving object and find occlusion and mismatching by evaluating the estimation error of each object location.

1 Introduction

Generation of a stationary environmental map is one of the important tasks for vision based robot navigation [1][2]. For this purpose, a detailed analysis is not necessary but high speed and rough understanding of the environment around the robot is required. If considered from the standpoint of machine perception, autonomous navigation needs the field of view as wide as possible. Thus, a real-time omnidirectional camera, which can acquire an omnidirectional (360 degrees) field of view at video rate, is suitable for autonomous navigation. There have been several attempts to acquire omnidirectional images using a rotating camera, a fish-eye lens, a conic mirror and a spherical mirror. Over the past 15 years, researchers in computer vision, applied optics and robotics have investigated a number of papers related to omnidirectional cameras and their applications [3] [4] [5] [6] [7] [8].

Under the assumption of the known motion of the robot, environmental maps of the real scene are successfully generated by monitoring azimuth changes in the image. Yagi used this property for robot navigation with an omnidirectional image sensor [9]. Delahoche et al have proposed the incremental map building method based on the exploitation of the azimuths data given by omnidirectional vision and by an odometer [10]. The robot position estimation and map updating are based on the use of an Extended Kalman Filter. However, it is difficult to observe the exact motion parameters of the robot because of encoder measurement error of the robot. Observational errors in the generated environmental map accumulate in long movements of the robot. To generate a large environmental map, it is desirable not to assume known robot motion. Furthermore, since both robot and objects move in the environment, the stationary map generation and the robot egomotion estimation by using a single camera are difficult because of correspondence ambiguity caused by occlusion.

In this paper, under the assumption of unknown translational motions of the robot, we propose a method to generate a stationary environmental map and estimate the translational egomotion of a robot in a dynamic environment, by using an omnidirectional image sensor. The method can detect a moving object and find occlusion and mismatching by evaluating estimation error of each object location.

2 Robot System with Omnidirectional Image Sensor

Figure 1 shows the robot system with an omnidirectional image sensor HyperOmni Vision using a TV camera with its optical axis aligned with the hyperboloidal mirror's one [11]. Mounting a HyperOmniVision on a robot so that optical axis is vertical, we can acquire a 360-degree view around the robot. A hyperboloidal mirror yields the image of a point in space on a vertical plane through the point P and its axis. Thus, the point P at (X, Y, Z) is projected onto the image point p at (x, y) such that
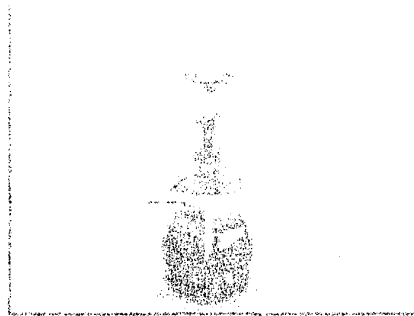
$$\tan \theta = \frac{Y}{X} \qquad (1)$$

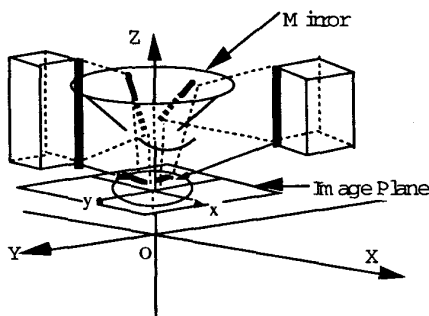Figure 1: Robot System with Omnidirectional Image Sensor
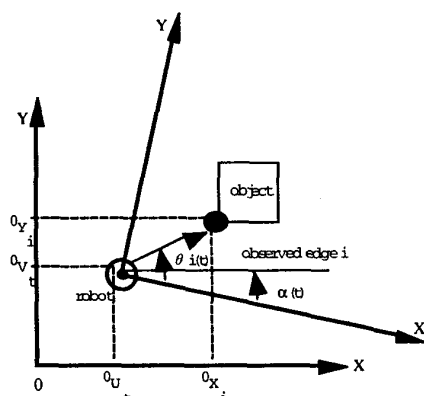

Figure 2: Hyperboloidal Projection

Figure 3: Coordinate System

This means that the angle in the image, which can be easily calculated as y/x shows the azimuth angle θ of the point P in space. Also, it can be easily understood that all points with the same azimuth in space appear on a radial line through the image center as shown in Figure 2. Therefore, with a hyperboloidal projection, the vertical edges in the environment appear radially in the image and the azimuth angles are invariant to changes in distance and height. In this paper, we use the locus of azimuth angle of vertical edges while the robot is moving.

3 Assumption

The following properties of the environment and the mobile robot are assumed for image analysis.

The floor is almost flat and horizontal while walls and stationary objects such as desks or shelves have vertical planes. The robot moves in a man-made environment such as a corridor in a building or a road in down town. Motion parameters are two translational components (U,V) and one rotational component α. Translational components are unknown. The rotational component is given by an internal sensor such as a gyroscope or compass.

4 Principle of Map Generation and Egomotion Estimation

Let us denote the robot location and orientation at time t by $(^0U_t, ^0V_t)$ and $\alpha(t)$, respectively. As shown in Figure 3, defining the position of object i at time $t = 0$ by $(^0X_i, ^0Y_i)$, relation between observed azimuth angle $\theta_i(t)$ of object i at time t and object location relative to the robot is obtained as follows,

$$\tan(\theta_i(t) - \alpha(t)) = \frac{^0Y_i - ^0V_t}{^0X_i - ^0U_t} \qquad (2)$$

Here, unknown parameters are robot location $(^0U_t, ^0V_t)$ and object locations $(^0X_i, ^0Y_i)$. We assume that the robot orientation $\alpha(t)$ is given by an internal sensor, and azimuth angle $\theta_i(t)$ of object i at time t can be obtained by the omnidirectional image sensor. Therefore, the total number of unknown parameters and total number of observational equations are $(2i + 2(t - 1) - 1)$ and $i \times t$, respectively. If the following relation is satisfied, the robot and object locations can be estimated at the same time.

$$i \times t \ >= (2i + 2(t - 1) - 1) \qquad (3)$$

Equivalently, location estimation can be done by observing three object points from three different robot positions.

5 Algorithm for Map Generation

5.1 Estimation of loci of azimuths of vertical edges

As shown in Figure 5, an input omnidirectional image (See Figure 4) is transformed into 2D polar coorinates. Next, as shown in Figure 6, we apply 3x3 Sobl operator to 2D polar image and project onto the horizontal axis to get 1-D projection (7). To estimate the loci of azimuths of vertical edges, the correspondence

of edge length between edges in the 1-D projection of consecutive images is established by using a correlation method in the restricted search field. As consecutive images are sampled densely, one can consider that the azimuth angle of the vertical edge in next frame is in the neighborhood of the azimuth angle in the current frame. Therefore, a certain margin of search field in the next frame is set around the current azimuth angle of the obtained vertical edge. After matching a few frames, the search region can be limited to a narrowed one by calculating the locus of each edge from equation 2. We then evaluate the conformation of neighboring relations.

5.2 Local map generation

The environmental map and robot location have estimated error due to observation error of azimuth angle of vertical edges. Therefore, we estimate the more precise location using consecutive measurements by the least squares method. Mathematically, if we define the squared error f, as shown in 5.2, of the least squares method using 2, values of location $({}^0X_i, {}^0Y_i)$ can be found by solving the following partial differential equations.

$$f = \sum_i \sum_t \{({}^0X_i - {}^0U_t)sin(\theta_i(t) - \alpha(t))$$

$$-({}^0Y_i - {}^0V_t)cos(\theta_i(t) - \alpha(t))\}^2 (4)$$

$$\begin{bmatrix} \frac{\partial f}{\partial {}^0X_k} \\ \frac{\partial f}{\partial {}^0Y_k} \\ \frac{\partial f}{\partial {}^0U_j} \\ \frac{\partial f}{\partial {}^0V_j} \end{bmatrix} = \begin{bmatrix} \sum_i \sum_t A(i,t)\sin(\theta_i(t) - \alpha(t)) \\ -\sum_i \sum_t A(i,t)\cos(\theta_i(t) - \alpha(t)) \\ -\sum_i \sum_t A(i,t)\sin(\theta_i(t) - \alpha(t)) \\ \sum_i \sum_t A(i,t)\cos(\theta_i(t) - \alpha(t)) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$A(i,t) = \left\{ \begin{array}{l} ({}^0X_i - {}^0U_t)\sin(\theta_i(t) - \alpha(t)) \\ -({}^0Y_i - {}^0V_t)\cos(\theta_i(t) - \alpha(t)) \end{array} \right\} \quad (5)$$

$$(k = 1, \cdots i, j = 1, \cdots t)$$

5.3 Global map building by combining local maps

A global map is generated by combining local maps at each position of the robot. Generally, the error of the measurement by triangulation is in inverse proportion to the trigonometric parallax and the distance between view positions. A large estimated error of local map occurs when the azimuth angle of the vertical edges does not change significantly. Therefore, combination of local maps is done by selecting the object location estimated from locus of azimuth with large azimuth change and wide standard deviation of azimuth. Because a least square method give the precise measurement whenobservational azimuth angles of vertical edges distribute randomly. Actually, we use the following evaluation function V_i.
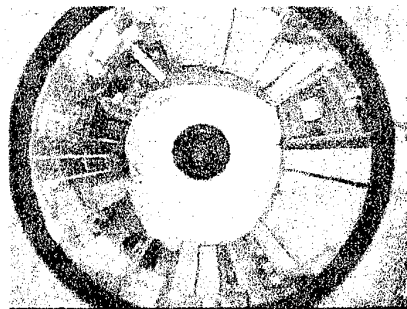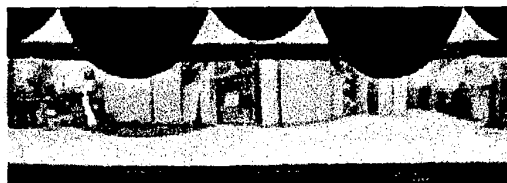

Figure 4: Input Image

Figure 5: Polar Coordinate Transformation

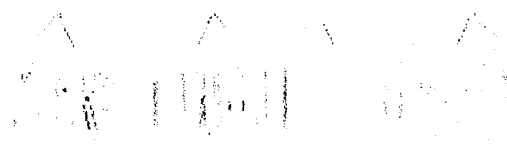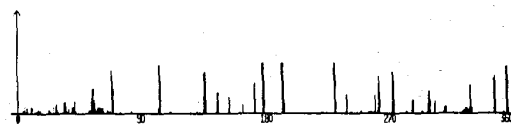
Figure 6: Edge Image

Figure 7: 1-D Projection Data

$$V_i = |\theta_i(t_2) - \theta_i(t_1)| + \sqrt{\frac{1}{t_2 - t_1} \sum_{k=t_1}^{t_2} \left(\theta_i(k) - \frac{\sum \theta_i(k)}{t_2 - t_1}\right)^2}$$

(6)

The first term means the magnitude of azimuth change between time t_1 and time t_2, and the second term means distribution of azimuth within interval between time t_1 and time t_2.

6 Discrimination of Moving Objects and Detection of Mismatching

The principle of map generation and location estimation of the robot above, assumes that the robot moves in a stationary environment. Since both the robot and other objects move in the environment, the stationary map generation and the robot egomotion estimation by using a single camera are difficult because of correspondence ambiguity caused by occlusion.

To use the method in a dynamic environment, we propose the method for discriminating moving objects and finding mismatching. The method consists of two steps. In the first step, the method detects occurrence of mismatching or appearance of moving object. Second, loci of azimuths of vertical edges where mismatching occurs or which correspond to a moving object, are identified by evaluating estimation error. Now, we define the evaluation value as follows

$$D_i = \frac{1}{t_2 - t_1} \sum_{t=t_1}^{t_2} \left(\begin{array}{c} \left(^0X_i - {}^0U_t\right) \sin\left(\theta_i(t) - \alpha(t)\right) \\ - \left(^0Y_i - {}^0V_t\right) \cos\left(\theta_i(t) - \alpha(t)\right) \end{array} \right)^2$$

(7)

$$f_1 = \frac{1}{n} \sum_{i=1}^{n} D_i$$

(8)

$$f_2 = \frac{1}{m} \sum_{i=1}^{m} D_i$$

(9)

D_i is the average square error of each object (vertical edge) i observed between time t_1 and t_2. f_1 is the total amount of average square error of every object used for map generation. If the evaluation value f_1 is bigger than a certain threshold δ_1, the method decides that mismatching occurs or a moving object is included in observational data. Then, every vertical edges whose evaluation value exceed a threshold δ_2 are registered as candidates of mismatching edges or moving objects.

Finally, from (9), if the total amount of average square error of every object with the exception of these candidates is lower than the threshold δ_1, the candidates are determined as mismatching edges or moving objects.
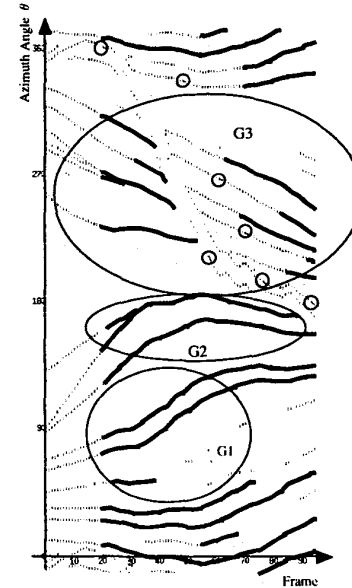

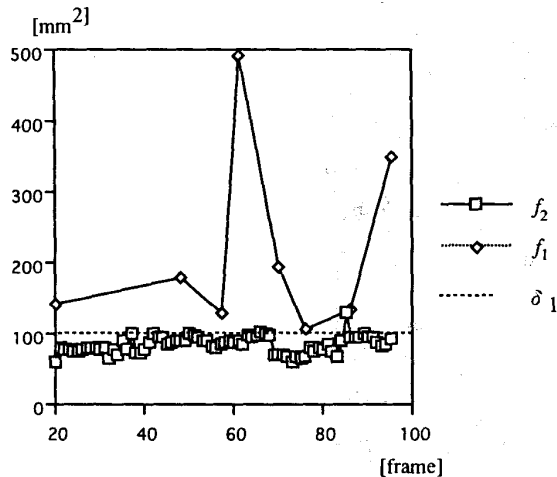Figure 8: Experimental Scene

Figure 9: Locus Map of Vertical Edges

Figure 10: Evaluation Value
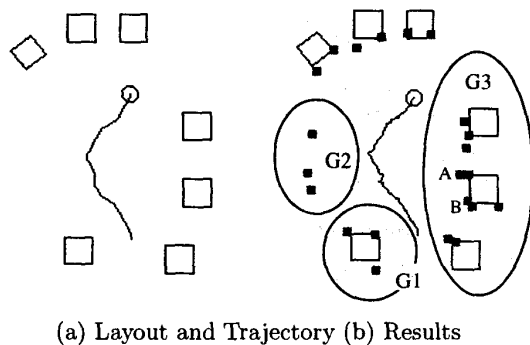

(a) Layout and Trajectory (b) Results

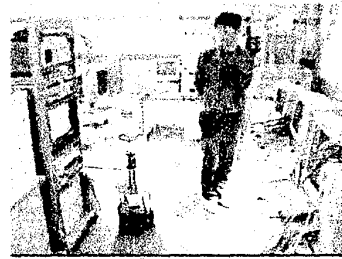Figure 11: Experimental Results of Map Generation and Location Estimation of the Robot

Figure 12: Experimental Scene

7 Experimental Results

Two experiments were conducted for evaluating accuracy of measurements and effectiveness in our computer room (about 6m x 4m).

As shown in Figure 8, the robot moves in an arc toward the left side and avoids colliding with another robot that approaches from the front side of the robot. Figure 9 shows the locus map of azimuth angle of vertical edges. The vertical edges drawn by thick lines were used for map generation and location estimation of the robot. Small black circles are the positions where the robot detects occurrence of mismatching or appearance of moving objects. Figure 10 shows the temporal changes of evaluation value f_1 and f_2, during the robot movement. Figure 11 are results of map generation and location estimation of the robot. Figure 11 (a) shows the correct layout of the stationary environment and the trajectory of the robot calculated by the internal sensor. As shown in Figure 11 (b), a black cluttered line shows the estimated trajectory of the robot and the black squares show the estimated map (location of the vertical edges). Vertical edges within three circles G1, G2 and G3 in Figure 11, are drawn by three elliptical regions (G1, G2 and G3) roughly in Figure 9. An average error of the location measurement of the robot and stationary environmental map were approximately 4 cm and 18cm, respectively.

A second experiment was done in a more real situation. As shown in 12, a person passed by the side of the robot. As shown in Figure 13, the robot could detect occlusion by evaluating estimation error f_1 and edges from walking person. Small black circles are the positions where the robot detect occlusion. As shown in Figure 14, the robot generated the stationary environmental map. In this case, the height of the object (desk) is higher than that of the robot. Therefore, the robot observed the lower part the desk.

8 Conclusions

In this paper, under the assumption of unknown translational motions of the robot, we proposed a method to generate a stationary environmental map and estimate the egomotion of a robot in a dynamic environment, by using an omnidirectional image sensor. Since both robot and objects move in the environment, the method can detect a moving object and
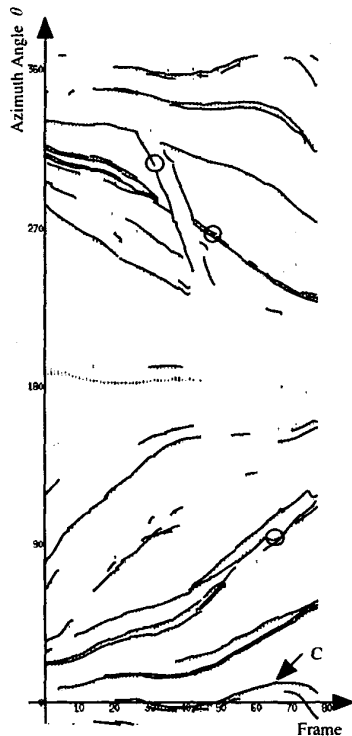
3497

Figure 13: Locus Map of Vertical Edges

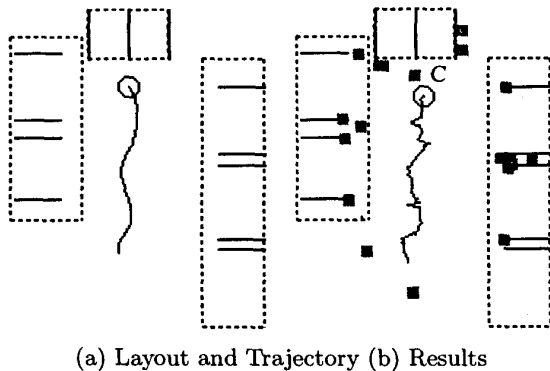(a) Layout and Trajectory (b) Results

Figure 14: Experimental Results of Map Generation and Location Estimation of the Robot

find occlusion and mismatching by evaluating estimation error of each object location. Map generation and egomotion estimation under completely free motions of the robot is on-going study.

Acknowledgments

References

[1] Ayache N., Faugeras O.D., "Building Registration and Fusing Noisy Visual Map," *Proc. Int. Conf. Computer Vision*, pp.73-82, 1987.

[2] Ishiguro, H., Yamamoto H. and Tsuji S., "Omni-directional Stereo," *IEEE Trans. P.A.M.I*, vol.14, no.2, pp.257-262, 1992.

[3] Saraclik, K. B., "Characterizing an Indoor Environment with a Mobile Robot and Uncalibrated Stereo," *Proc. of IEEE Int. Conf. Robotics and Automation*, pp.984-989,1989.

[4] Zheng, J. Y. and Tsuji, S., "Panoramic Representation of Scenes for Route Understanding," *Proc. of Int. Conf. Pattern Recognition*, pp.161-167, 1990.

[5] Kawanishi T., Yamazawa K., Iwasa H., Takemura H. and Yokoya N., "Generation of High-resolution Stereo Panoramic Images by Omnidirectional Imaging Sensor Using Hexagonal Pyramidal Mirrors," *Proc. of Int. Conf. Pattern Recognition*, pp.485-489, 1998.

[6] Peri, V. and Nayar S. K., "Omnidirectional Video System," *Proc. of U.S-Japan Graduate Student Forum in Robotics*, pp.28-31, 1996.

[7] Murray, D. W., "Recovering Range Using Virtual Multicamera Stereo," *Computer Vision and Image Understanding*, vol.61, no.2, pp.285-291, 1995.

[8] Kang, S. B. and Szeliski, R., "3-D Scene Data Recovery Using Omnidirectional Multibaseline Stereo," *Proc. of Computer Vision and Pattern Recognition*, pp.364-370, 1996.

[9] Yagi, Y., Nishizawa, Y. and Yachida, M., "Map-based Navigation for a Mobile Robot with Omnidirectional Image Sensor COPIS," *IEEE Trans. Robotics and Automation*, vol.11, no.5, pp.634-648, 1995.

[10] Delahoche, L., Pegard, C., Mouaddib, E. M. and Vasseur, P., "Incremental Map Building for Mobile Robot Navigation in an Indoor Environment," *Proc. of IEEE Int. Conf. Robotics and Automation*, pp.2560-2565, 1998.

[11] Yamazawa, K., Yagi, Y. and Yachida, M., "Omnidirectional Imaging with Hyperboloidal Projection," *Proc. of IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, no.2, pp.1029-1034, 1993.

3498