MVA'98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

13—23

Recognition of 3-D Object Shape and Forward Moving Distance by Monocular Motion Stereo for Mobile Robot

Yasuo SUGA*
Faculty of Science and Technology
Keio University

Nobuhiro OHMORI*
Graduate School of
Keio University

Abstract

As a vision system of autonomous mobile robot which moves along a pipe and carries out pipe works, such as welding, nondestructive inspection and so on, is constructed. The robot can move along a pipe and find obstructions which often exist on the pipe. If it is necessary, the robot can move over them. In order to recognize the position, shape and dimension of the obstruction, a visual sensing system applied monocular motion stereo was constructed as an eye of the autonomous mobile robot. Particularly, a new algorithm, by which 3-D object shape and forward moving distance of robot can be detected simultaneously, was investigated. As the result, it is confirmed that this system is effective as a visual sensor of autonomous mobile robot.

1 Introduction

Authors has been developing an autonomous mobile robot which moves along a pipe and carries out pipe works, such as welding, nondestructive inspection and so on [1]. However, because some obstructions often exist on the pipe, the mobile robot should have ability to find and move over it. In this case, the robot should recognize position, shape and dimension of the obstruction to move over it. Therefore, development of a vision system for the robot was tried. In this study, a visual sensing system applied monocular motion stereo was constructed for a vision of autonomous mobile robot which is movable along a pipe and inspects pipe joints. Particularly, a new algorithm, by which 3-D object shape and forward moving distance of robot can be detected simultaneously, was investigated.

2 Mobile Robot System

Figure 1 shows a photograph of mobile robot which moves along a pipe and moves over obstructions, such as flange or branch pipe. The robot system is constructed by the robot body, a CCD camera, an image digitizer, a personal computer and so on. The CCD camera observes the front area of the robot and can move in lateral direction to realize the monocular motion stereo vision. The image data acquired by the CCD camera is transformed into digital data through the image digitizer (256 × 240 × 6bit), and taken into personal computer (CPU: Pentium 200MHz). The image data are processed in the personal computer, and the shape of the slit light is recognized. Then the motion of robot is controlled.

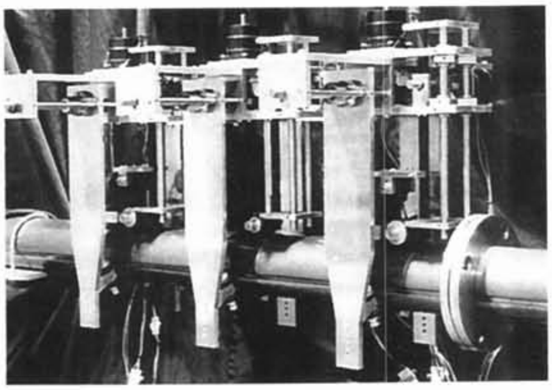

Fig.1 Autonomous mobile robot

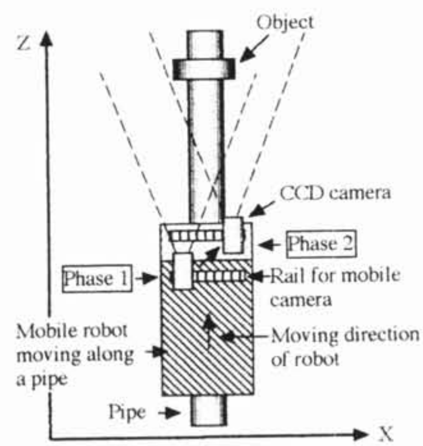

Fig.2 Moving camera installed on a robot

*Address: 3-14-1, Hiyoshi, Kouhoku-ku, Yokohama 223-8522 Japan
Tel. 045-563-1141, Fax. 045-563-5943
e-mail: suga@mech.keio.ac.jp

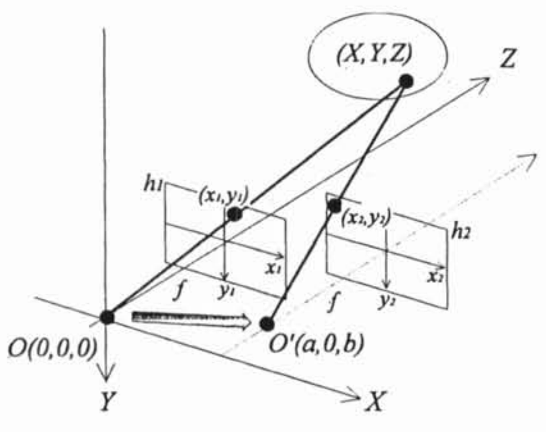

Fig.3 Principle of monocular motion stereo

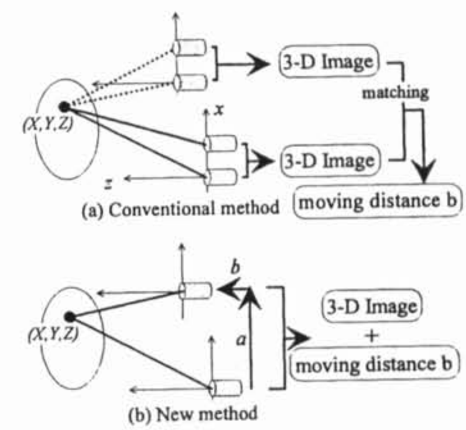

Fig.4 Conventional and new stereo vision

3 Monocular Motion Stereo Vision

Figure 2 shows an schematic illustration of the mobile robot and CCD camera which can move in lateral direction to realize the monocular motion stereo vision. Figure 3 shows principle of stereo vision in this system. At phase 1, the camera exists on the original position O and acquires image 1. Then the camera moves to the right direction (X direction) by a and the robot moves forward (Z direction) by b during movement of the camera. Accordingly the camera moves to a point O' and acquires another image here. In the figure, h_1 and h_2 represent image planes of the camera at O and O' and have coordinate systems of $o\text{-}x_1y_1$ and $o\text{-}x_2y_2$ respectively. Accordingly the position of the remarking point of an object (X,Y,Z) is projected to a point (x_1,y_1) in the image plane h_1 and a point (x_2,y_2) in the image plane h_2 respectively. The transforming equation is shown as follows:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \frac{f}{Z}\begin{pmatrix} X \\ Y \end{pmatrix} \quad (1)$$

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \frac{f}{Z-b}\begin{pmatrix} X-a \\ Y \end{pmatrix} \quad (2)$$

where f is focal distance of the camera lens.

Accordingly, the three dimensional position of object (X,Y,Z) and forward moving distance of the robot b are given by the following equations.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{ay_2}{x_1y_2 - y_1x_2}\begin{pmatrix} x_1 \\ y_1 \\ f \end{pmatrix} \quad (3)$$

$$b = \frac{af}{x_1y_2 - y_1x_2}(y_2 - y_1) \quad (4)$$

Therefore, if the matching between a remarking point in plane h_1 and that in h_2 is realized, the position of the point in a three dimensional space and forward moving distance of the robot b are easily obtained.

Generally, in order to obtain both 3-D shape of an object and moving distance of robot, it is necessary to acquire two images at O and O' respectively in the conventional stereo vision by two cameras as shown in figure 4 (a). In the conventional method, the vision system should process four images and accordingly the processing time increases considerably. On the contrary, using the new method, 3-D shape and moving distance b can be calculated by only two images acquired at O and O'. Therefore, the calculation time can be saved remarkably and it is possible to realize the high speed image processing and real time control of the robot.

4 Detection of Optical Flow

4.1 Template Matching Method

In order to detect the forward moving distance of robot and 3-D shape from two images obtained at position O and O', template matching method was applied. In this method, a template was set in an image h_1 acquired at position O and then most similar area was searched in another image h_2 as shown schematically in figure 5. The similarity between template and the searched image

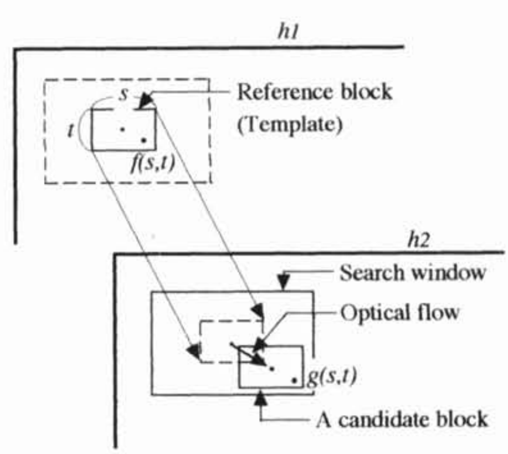

Fig.5 Template matching method

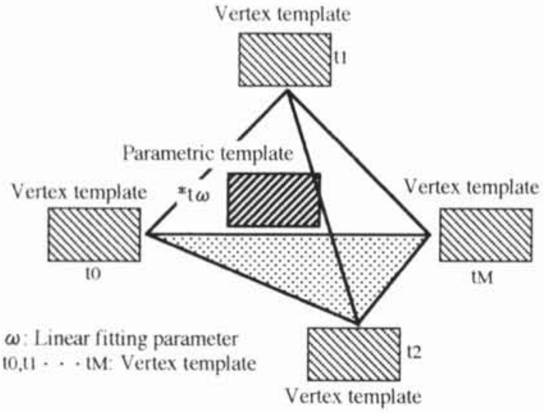

Fig.6 Parametric template space

area was evaluated by normalized correlation. Furthermore, not only gray scale data but also color data expressed by *YUV* color system were applied to calculation of normalized correlation so as to increase the reliability of template matching[2].

Generally, searching resolution of the corresponding points between two images is limited to accuracy of pixel order of the image. However, it is desirable to improve the matching resolution up to sub pixel of the image using some algorithm, such as parabola fitting and parametric template method. In this study, the parametric template method was applied to template matching[3]. In the parametric template matching, plural templates construct a template space. The parametric template space is constructed by plural vertex templates and parametric templates obtained by linear fitting using fitting parameter ω as shown in figure 6. Applying this parametric template matching, high speed and sub pixel matching between reference image and parametric template is realized.

4.2 Template Matching Test

In order to estimate the resolution of the template matching between two images obtained at camera position O and O', template matching test was performed. In this experiments, the camera was moved along only lateral direction to simplify the test condition and estimate accuracy precisely. Template matching test was performed using four conditions mixed with normalized correlations used gray scale data and color scale data, parabola fitting (PF) and parametric template matching (PT). Figure 7 shows the testing results. In the figure, (a) shows histogram of x-direction component of optical flows in an image and (b) that of y-direction. The template matching conditions and standard deviation of x and y component of optical flow are also shown in the figure. From the figure, it is confirmed that the least value of standard deviation is obtained in parametric template using color data.

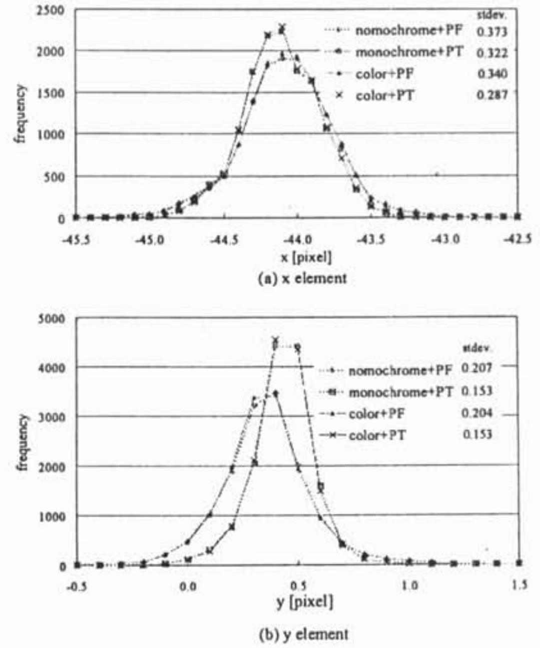

Fig.7 Histogram of optical flow

5 Detection of Forward Moving Distance

Forward moving distance of the robot from position O to O' can be estimated by optical flows obtained by above mentioned method. Accordingly, forward moving distance b was experimentally measured using the method proposed. In this case, a flat object with fine texture surface was put at 500mm from the camera. Generally the value b must have single datum and accordingly the histogram of b should have one peak. However, it have often two peaks practically. Figure 8

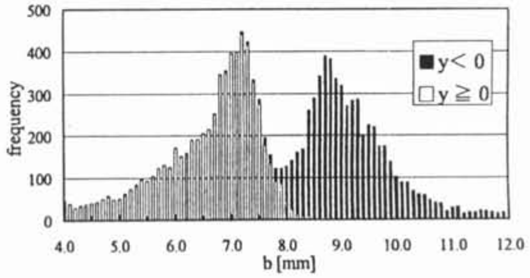

Fig.8 An example of histogram with two peaks

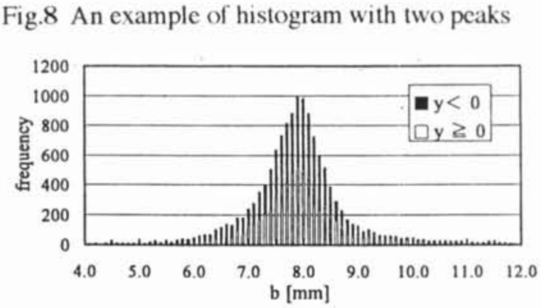

Fig.9 Correction of histogram

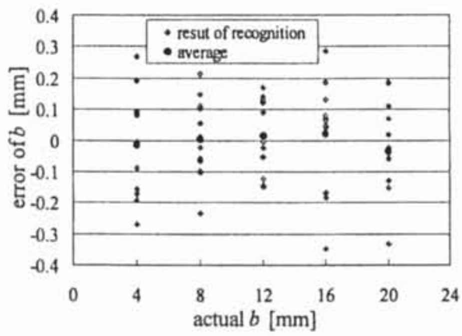

Fig.10 Recognition of forward moving distance

shows an example of histogram of b with two peaks. In this case, the histogram is shown separately in two categories, such as in $y<0$ and $y>0$. The separation occurs by the unexpected movement of the camera (robot) along y-axis. Therefore, b values of these two peaks were calculated and data were corrected. In this system, average value of b at two peaks was adopted as the forward moving distance as shown in figure 9. Figure 10 shows measured results of forward moving distance, where the moving distance was changed from 4mm to 20mm. The detection error ε of b is dispersed within a range of $\pm$ 0.3mm approximately and average error ε was $\pm$0.12mm.

6 Recognition of 3-D Shape of Object

The spatial positions of points in the acquired image can be estimated by optical flows using equation (3). However, practically, some low-reliable data exists in the data obtained by above mentioned method, because of unclear images, matching error and so on. Accordingly, the reliability assessment of data b was performed, and only data with high reliability were used for distance image. Furthermore, in order to eliminate noise in the image, a median filter was applied.

Fig.11 shows recognition result of a part of a steel flange put in front of the robot. In the experiment, the distance from camera to flange was set at 420mm. The lateral moving distance of camera, a, was fixed to 8.0mm and forward moving distance, b, was changed from 8 to 20mm. As the results of preliminary experiments, the template size was selected to 9×9 pixel and matching area 31×31 pixel. As the experimental results, average error of $\pm$ 0.29mm in measurement of distance from camera to flange was obtained, which is calculated by reliable data. From the figure 11, some matching error are recognized due to unclear texture of flange surface. However, approximate shape of the flange was obtained. Therefore, it is recognized that the method is effective to detect forward moving distance and 3-D shape of an object and apply it

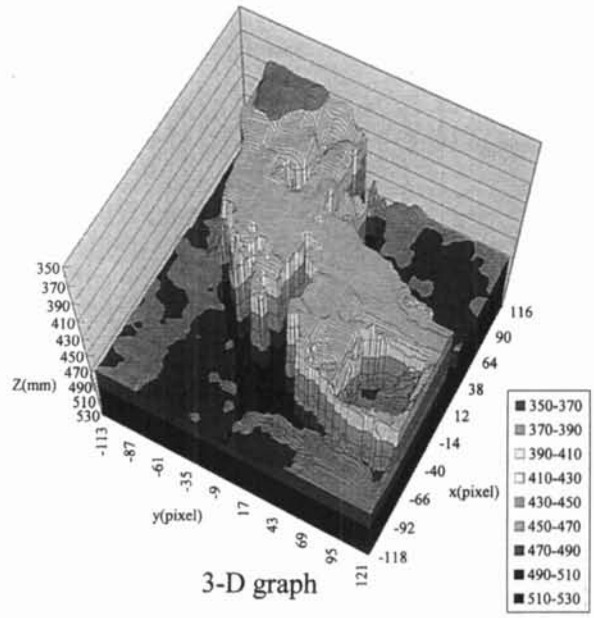

Fig.11 Distance image of a flange

to vision sensor of autonomous mobile robot.

7 Conclusions

Main results obtained are summarized as follows:

1) A visual sensing system applied monocular motion stereo was constructed to detect 3-D object shape and forward moving distance of robot simultaneously.

2) Applying color data and parametric template matching, high reliable template matching with accuracy of sub pixel is realized.

3) In order to compensate measurement error of optical flows due to moving error of camera, a method to correct data by evaluating and correcting forward moving distance obtained was proposed.

4) As the results of measurement experiments of steel flange, the average error of forward moving distance of $\pm$0.29mm and good image of 3-D object shape were obtained.

References

[1] Suga,Y., Sano,T., Proc. 3rd France/Japan Congress on Mechatr. Vol.2 (1996) pp.860

[2] Shimizu, Kuwabara, Gazo-labo, 5-1 (1994) pp.74 (in Japanese)

[3] Sato, and Anan, Trans. Oita Indus. Sci. Tech. Centr. (1994) pp.18 (in Japanese)

[4] Tanaka,K., Sano,M., Ohara,S., Okudaira,M., Proc. 2nd Intelligence & Mechatr. W.S. (1997) pp.74 (in Japanese)