Visual Servoing of a Mobile Manipulator Based on Stereo Vision

H. J. Lee*, M. G. Park*, M. C. Lee**

*Graduate School of Mechanical and Intelligent Systems Engineering, Pusan National University, Pusan 609-735, Korea
(Tel: 82-51-510-3081; Fax: 82-51-512-9835; E-mail: lhjeong@pusan.ac.kr)
*Graduate School of Mechanical and Intelligent Systems Engineering, Pusan National University, Pusan 609-735, Korea
(Tel: 82-51-510-3081; Fax: 82-51-512-9835; E-mail: pmg@pusan.ac.kr)
**School of Mechanical Engineering, Pusan National University, Pusan 609-735, Korea
(Tel: 82-51-510-2439; Fax: 82-51-512-9835; E-mail: mclee@pusan.ac.kr)

Abstract: In this study, stereo vision system is applied to a mobile manipulator for effective tasks. The robot can recognize a target and compute the position of the target using a stereo vision system. While a monocular vision system needs properties such as geometric shape of a target, a stereo vision system enables the robot to find the position of a target without additional information. Many algorithms have been studied and developed for an object recognition. However, most of these approaches have a disadvantage of the complexity of computations and they are inadequate for real-time visual servoing. However, color information is useful for simple recognition in real-time visual servoing. In this paper, we refer to about object recognition using colors, stereo matching method, recovery of 3D space and the visual servoing.

Keywords: Stereo Vision, Mobile Manipulator, Color, Computer Vision, Visual Servoing

1. INTRODUCTION

In these days, robots have been applied to various areas such as industry, service and entertainment[1-3]. For efficient tasks, they have to obtain much information about environments. The vision system is more effective than ultrasonic sensors or laser sensors for recognition of an object and estimation of the object's position. Visual servoing is to control the pose of the robot's end-effector using visual information. Visual control of the manipulator has substantial advantages for working with targets whose position is unknown[4].

A robot finds a target with a constructed database. This iterative process takes up much computation time due to the complexity of algorithms[5]. In order to simplify this process, informations about color of a target are used. The robot detects pixels of color being similar to a target or a mark. And the robot recognizes a target without respect to a relationship between the pose of the camera and the pose of the target.

The robot must know position of a target to control the manipulator after searching images for a target. Some information such as the geometric relationship between several feature points on a target is needed to estimate position of a target when monocular vision is used. In this paper, stereo vision is used to determine the 3D coordinate corresponding to an image plane point. The use of a stereo system requires less strict camera calibration while monocular vision is concerned with a number of assumptions such as geometric properties[6]. The robot can measure the position of an object by Stereo matching[7]. Stereo matching methods divided into two methods: area-based and feature-based. Stereo matching using area-based method has the advantage of reducing computation quantity but is sensitive to noises because it depends on the intensity. Feature-based method includes complex processes such as interpolation but it is effective to reduce noises. In indoor environment with several sources of light, feature-based method is suitable to applications. The position of the target is computed by cameras' geometry and visual informations.

In section 2, the system is described. Stereo matching and the coordinate recover in 3D space are stated in section 3. Section 4 expresses about the inverse kinematics. Results of experiments with the manipulator are shown in section 5. Finally, conclusion is given in section 6.

2. SYSTEM CONFIGURATIONS

Fig. 1 shows configurations of the robot. The system of robot is separated form the server system and the upper plate as shown in Fig. 2 and Fig. 3. The server system deals with the control of the mobile robot and monitoring as shown in Fig. 2. The upper plate is composed of the CCD camera and manipulator as shown in Fig. 2. The control part consists of 88C166 process of 20MHz, 32k Flash ROM and RAM. The A/D port has two RS232C buses and several DIO ports. The driving part has two servo motors, 12V driving voltage and the maximum permitted voltage is 2.5A. The sensor

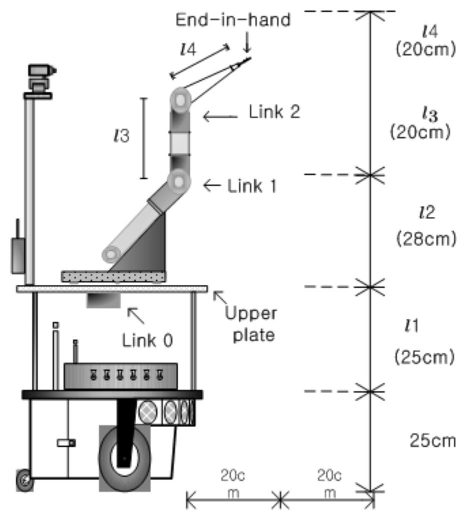


Fig. 1 Robot's configurations

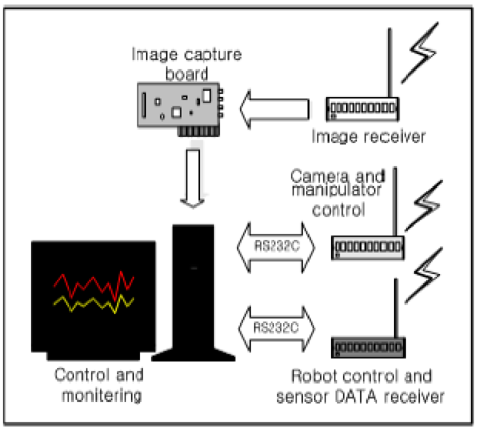


Fig. 2 Control and monitoring system

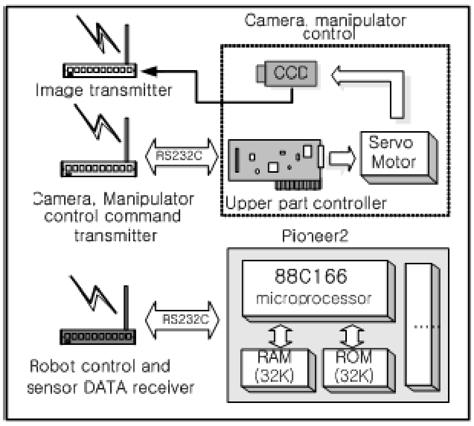


Fig. 3 Mobile robot frame

department consists of sixteen ultrasonic sensors and two encoders. It can accurately measure position and speed because the encoders supply pulses of 9,850 per 1 second. The ultrasonic sensors are 25Hz, it can detect the range of 10cm~5m. The structure of the ultrasonic sensor part is as shown in Fig. 4 and the upper plate is divided into two CCD cameras and manipulator; CCD cameras which stand in a row on 45cm from the upper plate, supports resolutions of 640×480. The distance between two cameras' center is 6.5cm and both of two cameras' lens are set for 6mm focal length. Cameras' specification is as following Table. 1. The manipulator has 4 degrees of freedom that are controlled by a RC servomotor in each articulation. Fig. 5 shows the process of a object recognition and stereo matching in visual servoing.

3. OBJECT RECOGNITION AND RECOVERY OF A 3D POSE

3.1 Object recognition

In this paper, we applied color to extract a target from an image plane. A mark is attached to the target to classify explicitly. The robot analyzes R, G, B values of each pixel and detects pixels corresponding to the mark. In preprocessing, we apply Histogram equalization to input images to reduce effects of darkness[5]. Histogram equalization which is used to redivide the distribution of intensity is useful to an image with a poor intensity's distribution. This process improves the visual appearance of an image by widening peaks, compressing valleys.

In Fig. 6(a), description of end-in-hand and the lever which was used in experiments is shown. Size of the lever is 6×2×0.3cm, the mark is 3×2cm. Fig. 6 shows an original image and the image which the mark was extracted from images by color informations.

Table. 1 Camera specification

Camera Specification	
Image Sensor	1/3" Color CCD SONY
Effective Pixel	510(H) × 492(W)
Cell Size	9.6μm(H) × 7.5μm(W)
TV Type	NTSC
Sync. Type	Internal
Lens (Auto IRIS)	6-12mm Vari focal, F1.4

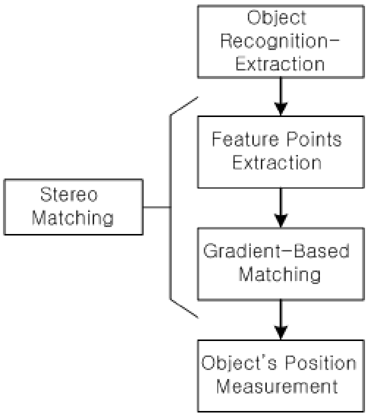
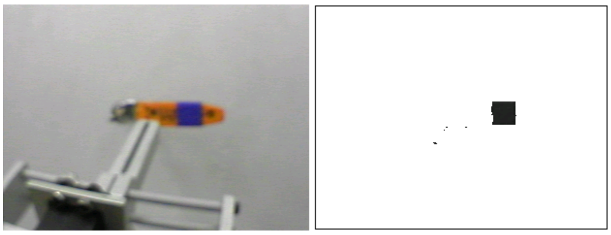


Fig. 5 Block diagram of processes in Section 3



(a) (b)

Fig. 6 Mark detection (a)Original image (b) Image of detected mark

3.2 Feature points extraction

After finding the mark attached to the target, the robot extracts the feature points from the mark image. To find the feature points, 'cornerness' is computed in gray-level[8]. Cornerness is defined as the product of gradient magnitude and the rate of change of gradient direction with gradient magnitude. In order to measure cornerness, interesting operators are used. We use Forstner operator such as following procedures[9].

- Step1: find g_r, g_c for each pixel using first differential operator where g_r and g_c are differential value in row and column.

- Step2: calculate normalized values $< g_r^2 >$, $< g_r g_c >$, $< g_c^2 >$ of g_r^2, g_c^2, $g_r g_c$ using a Gaussian smoothing filter, $G[r,c] = \dfrac{e^{\frac{-(r^2+c^2)}{2\sigma^2}}}{2\pi\sigma^2}$. σ is a parameter adjusting the width of the filter.
  And make out following matrix.

$$A = \begin{bmatrix} < g_r^2 > & < g_r g_c > \\ < g_r g_c > & < g_c^2 > \end{bmatrix} \quad (1)$$

- Step3: find candidates of feature points. A weight and a cornerness are calculated at each pixel as follows

$$\text{weight} \quad W = \frac{Det\,(A)}{Trace\,(A)} \quad (2)$$

$$\text{cornerness} \quad C = \frac{4\,Det\,(A)}{Trace^2\,(A)} \quad (3)$$

  If W and C are larger than threshold, we regard that point as a candidate of a feature point. The threshold value is determined experimentally.

- Step4: select for feature points where are local maxima among candidates.

Fig. 7 shows a result of feature points extraction.

3.3 Gradient-based matching

To get feature points, the left image as a reference is used. We find the right matching point corresponding to the left matching point using gradient-based matching method. This method depends on gradient values to estimate resemblance in gray-level. For finding the most similar point, gradient values of each pixel on the right image are compared with the gradient value of the left matching point within searching window. For efficiency of operation, a size of a searching window is determined in proportion as the sum of whole gradient values' magnitude of an image. If an object is far from cameras, the disparity which is the shifted value with respect to each pixel is small enough to search the matching point by a smaller window's size. Contrary, if an object is close to cameras a searching window being a large size is needed. The size of a window is determined by following expression. a is a constant determined experimently.

$$window = a \sum_{row=1}^{height-1} \sum_{column=1}^{width-1} g_R(row, column) \quad (4)$$

Similarity is expressed as follows and the right matching point is (row, column) where S is the smallest.

$$S = -\alpha \left| g_L(left\ matching\ point) - g_R(row, column) \right| \quad (5)$$

where,
g_L = a gradient value of the left image
g_R = a gradient value of the right image
α = a constant

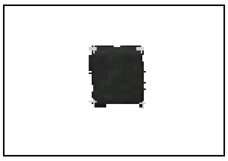


Fig. 7 Feature points extraction

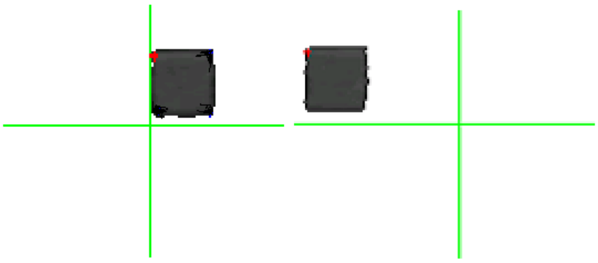


Fig. 8 Matching points on the left and right image

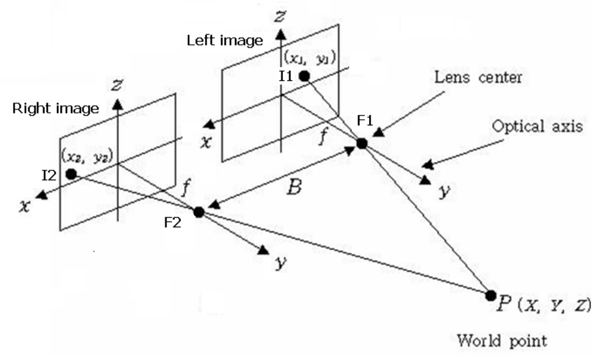


Fig. 9 Stereo vision geometry

Fig. 8 represents the result of finding the matching point.

3.4 Recovery of a 3D pose

We are able to recover the 3D pose parameter from the disparity, the camera's focal length, the distance between two cameras and actual size of CCD cell. These parameters were described in section 2. The position is computed using similarity of $\triangle PI_1I_2$ and $\triangle PF_1F_2$ as follows:

$$f + y : B + x_1 - x_2 = y : B$$
$$y = \frac{f \cdot B}{x_1 - x_2} \quad (6)$$

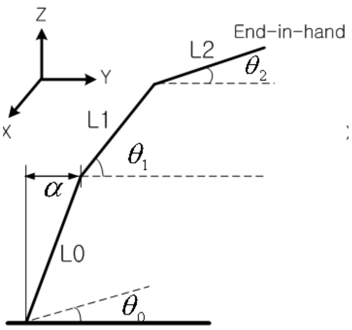


Fig. 10 Structure of manipulator

$$x = \frac{y \cdot x_1}{f} \quad (7)$$

$$z = \frac{y \cdot z_1}{f} \quad (8)$$

The center of the left camera's lens is the origin.

4. INVERSE KINEMATICS

Fig. 10 describes the structure of the manipulator. Lengths of links L0, L1, L2 are represented in section 2. The inverse kinematics is computed as following expressions. Rotation of joint 0 is concerned with yaw of the end effector. Joint 1 and joint 2 have a relation to pitch.

$$\theta_0 = tan^{-1}\left(\frac{x}{y+\alpha}\right) \quad (9)$$

$$r^2 = x^2 + y^2$$

$$r^2 + z^2 = l_2^2\ sin\theta_2^2 + (l_1 + l_2\ cos\theta_2)^2$$

$$cos\theta_2 = \frac{r^2 + z^2 - l_1^2 - l_2^2}{2\,l_1 l_2} = D$$

$$sin\theta_2 = \sqrt{1 - D^2} = C$$

$$\theta_2 = tan^{-1}(C/D) \quad (10)$$

$$\theta_1 = tan^{-1}(z/r) + tan^{-1}\left(\frac{l_2\ sin\theta_2}{l_1 + l_2\ cos\theta_2}\right) \quad (11)$$

5. EXPERIMENTS

The purpose of experiments is to operate a mobile manipulator using the stereo vision system. By recognizing the object and computing the position, the robot may exactly pull down a lever. We established equipments like a following Fig. 11. The distance between the lever and the wall is 4.5cm. The system is described in section 2 and the lever is shown in Fig. 6(a). Distances between center of the joint1 and the wall are represented in y axis. The robot extracts the mark attached to the target and detects feature points. And stereo matching is achieved by gradient-based matching method. The robot measures the target's position based on geometric properties of cameras. We set up so that the manipulator turns the lever to clockwise for 90 degree. The origin of global coordinate is the center of joint1 and each axis' direction is described in Fig. 11. Fig. 12, 13 illustrate errors of measurement in x, y, z-axis where y-axis and z-axis are fixed. This experiment

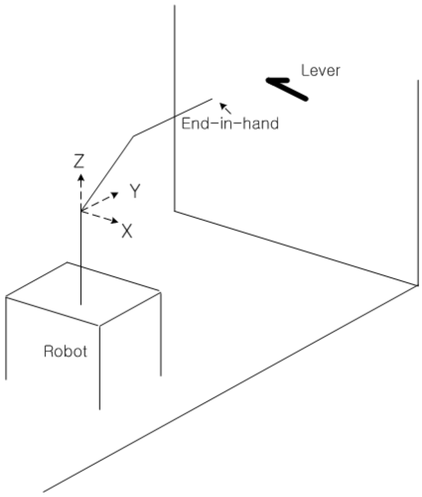


Fig. 11 Experimental set-up

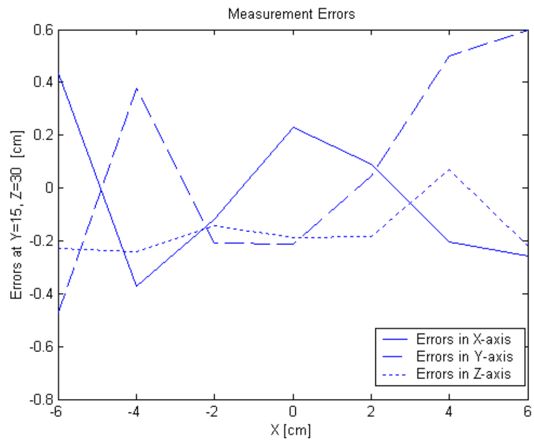


Fig. 12 Errors of measurement at Y=15[cm]

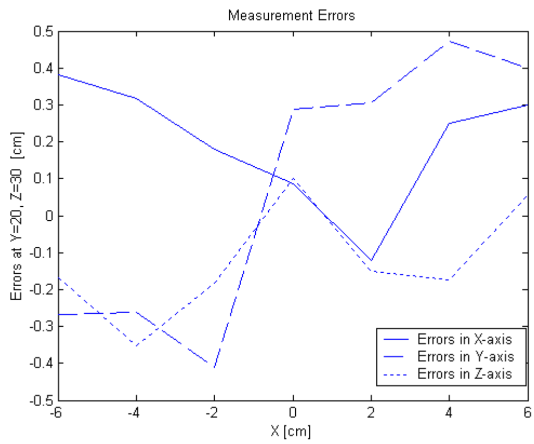


Fig. 13 Errors of measurement at Y=20[cm]

is achieved to estimate accuracy of the visual sensor. We can confirm that errors of measurement are within ±6mm. These errors are small enough to manipulate the lever within the limited end-effector's workspace. The task pulling down the lever was achieved with a probability of near 100% in the limited bound for manipulator independently of depth.

6. CONCLUSIONS

The position of the target which was unknown is computed by using a stereo vision system and the accuracy of measurement is evaluated. We could confirm that the robot with a manipulator exactly turns a lever within the limited end-effector's workspace. However, the accuracy of visual servoing introduced in this paper depends directly on the accuracy of the visual sensor and the manipulator's controller. The study of increasing the accuracy using a visual-feedback control loop is demanded.

REFERENCES

[1] Y.Nakauchi, and R. Simmons, "A Service Robot for People Environments,"Proc. of the IEEE/ASME Int. Conf. on Advanced Intelligent Mechanics, Italy, pp.1093-1098, 2001.

[2] Luo Ren C. Chen Tse Min, and Lin Meng Hsien, "Automatic Guided Intelligent Wheelchair System Using Hierarchical Grey-Fuzzy Motion Decision-Making Algorithms", Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and System, pp.900-905, 1999.

[3] R. C. Luo and T. M. Chen, "Supervisory control of multi sensor integrated intelligent autonomous mobile robot via Internet", Proc. of the 2nd IEEE International Conference on.

[4] Peter I. Corke, "VISUAL CONTROL OF ROBOT MANIPULATORS - A REVIEW," .World Scientific Series in Robotics and Automated Systems Vol.7.

[5] Randy Crane, "A Simplified approach to Image Processing", Prentice-Hall, 1997.

[6] Bart Lamiroy, Bernard Espiau, Nicolas Andreff and Radu Horaud, “Controlling Robots With Two Cameras: How to Do it Properly", Proc. of the 2000 IEEE International Conference on Robotics & Automation San Francisco, pp.2100-2105, 2000.

[7] U. R. Dhonald and J. K. Aggarwal, "Structure from stereo-A review", IEEE Trans. Systems, Man, and Cybermetics, Vol.19, No.6, pp.1489-1510, 1989.

[8] Harris, C G, "Determination of ego-motion from matched points", Proc. Alvey Vision Conf., Cambridge, 1987.

[9] Forstner, W "Quality assessment of object location and point transfer using digital image correction techniques", Proc. 15th ISPRS Congress, Rio De Janeiro, Brazil pp.169-191, 1987.