영상 피드백에 의한 4축 스카라 로봇의 실시간 궤적추적제어에 관한 연구

A Study on Real-Time Trajectory Tracking Control of SCARA Robot with Four Joints Based on Visual Feedback

정양근[1*], 심현석[2], 이우송[3], 한성현[4]

Yang-Guen Jung, Hyun-Seok Shim, Woo-Song Lee, Sung-Hyun Han

⟨Abstract⟩

This paper proposes a new approach to the designed of visual feedback control system based on visual servoing method. The main focus of this paper is presents how it is effective to use many features for improving the accuracy of the visual feedback control of industrial articulated robot for assembling and inspection of parts. Some rank conditions, which relate the image Jacobian, and the control performance are derived. It is also proven that the accuracy is improved by increasing the number of features. The effectiveness of redundant features is verified by the real time experiments on a SCARA type robot(FARA) made in samsung electronics company.

Keywords : *Visual Feedback Control, Trajectory Tracing Control, Articulated Robot Arm, Real-Time Control*

[1*]정회원, 교신저자, 경남대학교 첨단공학과,
  E-mail:shintech@hanmail.net
[2]정회원, ㈜동산테크 대표이사
[3]정회원, 성산암데코 연구소장, 工博
[4]정회원, 경남대학교 기계공학부 교수, 工博

[1*]Corresponding Author, Dept. of Advanced Engineering, Kyungnam University.
[2]CEO, DongSan Tech Co., Ltd.
[3]Director, R&D Center, SungSanamdeco Co. Ltd., Ph. D.
[4]Prof., School of Mechanical Engineering, Kyungnam University, Ph. D.

1. Introduction

There are mainly two ways to put the visual feedback into practice. One is called look-and-move and the other is visual servoing. Visual feedback is the fusion of results from many elemental areas including high-speed image processing, kinematics, dynamics, control theory, and real-time computing.

In the conventional works, some researchers have presented methods to control the manipulator position with respect to the object or to track the feature points on an object using a hand eye system as the application of visual servoing[1]. These methods maintain or accomplish a desired relative position between the camera and the object by monitoring feature points on the object from the camera.

However, these have been all done by the hand eye system with monocular visions and it is necessary to compensate for the loss of information because the original three-dimensional information of the scene is reduced to two-dimension information on the image. For instance, we must add information of the three-dimension distance between the feature point and the camera in advance or use a model of object stored in the memory [2~3].

This paper presents a method to solve this problem by using a vision. The use of binocular vision can lead to an exact image Jacobian not only at around a desired location but also at the other locations. The suggested technique places a robot manipulator to the desired location without giving such prior knowledge as the relative distance to the desired location or the model of an object even if the initial positioning error is large. This paper deals with modeling of high performance vision and how to generate the visual feedback commands. The performance of the proposed visual feedback system was evaluated by the simulations and experiments and obtained results were compared with the conventional case for a SCARA type robot.

2. Visual Feedback Control

Visual servo systems typically use one of two camera configurations: end-effector mounted, or fixed in the workspace.

The first, often called an eye-in-hand configuration, has the camera mounted on the robot's end-effector. Here, there exists a known, often constant, relationship between the pose of the camera(s) and the pose of the end-effector.

We define the frame of a hand-eye system with the stereo vision and use a standard model of the stereo camera whose optical axes are set parallel each other and perpendicular to the baseline. The focal points of two cameras are apart at distance d on the baseline and the origin of the camera frame $\sum_c$ is located at the center of these cameras[4].

An image plane is orthogonal to the optical axis and apart at distance f from the focal point of a camera and the origins of

frame of the left and right images, $\sum l$ and $\sum r$, are located at the intersecting point of the two optical axes and the image planes. The origin of the world frame $\sum w$ is located at a certain point in the world[5,6].

Now let $^l p = (^l x, ^l y)$ and $^r p = (^r x, ^r y)$ be the projections onto the left and right images of a point p in the environment, which is expressed as $^c p = (^c x\ ^c y\ ^c z)^T$ in the camera frame. Then the following equation is obtained(see Fig. 1).

$$^l x\ ^c z = f(^c x + 0.5d) \tag{1-a}$$

$$^r x\ ^c z = f(^c x - 0.5d) \tag{1-b}$$

$$^l y\ ^c z = f\ ^c y \tag{1-c}$$

$$^r y\ ^c z = f\ ^c y \tag{1-d}$$

Suppose that the stereo correspondence of feature points between the left and right images is found. In the visual servoing, we need to know the precise relation between the moving velocity of camera and the velocity of feature points in the image, because we generate a feedback command of the manipulator based on the velocity of feature points in the image[7].

This relation can be expressed in a matrix form which is called the image Jacobian. Let us consider n feature points $p_k (k = 1, \cdots, n)$ on the object and the coordinates in the left and right images are $^l p_k = (^l x_k, ^l y_k)$ and $^r p_k = (^r x_k, ^r y_k)$, respectively. Also define the current location of the feature points in the

image $^I p$ as

$$^I p = (^l x_1\ ^r x_1\ ^l y_1\ ^r y_1 \cdots\ ^l x_n\ ^r x_n\ ^l y_n\ ^r y_n)^T \tag{2}$$

where each element is expressed with respect to the virtual image frame $\sum p$.

First, to make it simple, let us consider a case when the number of the feature points is one. The relation between the velocity of feature point in image $^I p$ and the velocity of camera frame $^c \dot{p}$ is given as

$$^I \dot{p} = {}^I J_c\ {}^c \dot{p} \tag{3}$$

where $^I J_c$ is the Jacobian matrix which relates the two frames. Now let the translational velocity components of camera be σ_x, σ_y, σ_z and the rotational velocity components be w_x, w_y, w_z then we can express the camera velocity V as

$$V = [\sigma_x\ \sigma_y\ \sigma_z\ w_x\ w_y\ w_z]^T = [^c v\ ^c w_c]^T \tag{4}$$

Then the velocity of the feature point seen from the camera frame $^c \dot{p}$ can be written

$$^c \dot{p} = \frac{d\ ^c p}{dt} = \frac{d}{dt}\ ^c R_w\ (^w p - ^w p_c)$$
$$= \frac{d\ ^c p}{dt} = \frac{d}{dt}\ ^c R_w\ (^w p - ^w p_c) \tag{5}$$

where is the rotation matrix from the camera frame to the world frame and is the location of the origin of the camera frame written in the world frame. As the object is

assumed to be fixed into the world frame, The relation between and is

$$
\begin{aligned}
{}^c\dot{p} &= {}^cR_w\left\{-{}^ww_c\times\left({}^wp-{}^wp_c\right)\right\}-{}^cR_w{}^w\dot{p}_c \\
&= -{}^cw_c\times{}^cp-{}^c\dot{p}_c \\
&= \begin{bmatrix} -w_y{}^cz+w_z{}^cy-v_x \\ -w_z{}^cx+w_x{}^cz-v_y \\ -w_x{}^cy+w_y{}^cx-v_z \end{bmatrix}
\end{aligned} \tag{6}
$$

Therefore, substituting Eq. (6) into Eq. (3), we have the following equation.

$$
{}^I\dot{p} = {}^IJ_c{}^c\dot{p} = JV \tag{7}
$$

In Eq. (7) matrix which expresses the relation between velocity of the feature point in the image and moving velocity of the camera is called the image jacobian.

From the model of the stereo vision Eq. (1), the following equation can be obtained.

$$
2^cx\left({}^lx-{}^rx\right)= d\left({}^lx+{}^rx\right) \tag{8}
$$

$$
{}^cy\left({}^lx-{}^rx\right) = {}^lyd = {}^ryd \tag{9}
$$

$$
{}^cz\left({}^lx-{}^rx\right)= fd \tag{10}
$$

Above discussion is based on the case of one feature point. In practical situation, however, the visual feedback is realized by using plural feature points. When we use feature points, image Jacobian are given from the coordinates of feature points in the image. By combining them, we express the image Jacobian as

$$
J_{im} = \left[J_1, \cdots, J_n \right]^T \tag{11}
$$

Then, it is possible to express the relation of the moving velocity of the camera and the velocity of the feature points even in the case of plural feature points, that is,

$$
{}^I\dot{p} = J_{im}V \tag{12}
$$

where we suppose that the stereo and temporal correspondence of the feature points are found.

In the case of the monocular, the image Jacobian has the following form.

$$
J = f\begin{bmatrix} -\dfrac{1}{{}^cx} & 0 & \dfrac{{}^cx}{{}^cx^2} & \dfrac{{}^cx\,{}^cy}{{}^cz} & -\left(1+\dfrac{{}^cx^2}{{}^cx^2}\right) & \dfrac{{}^cy}{{}^cz} \\ 0 & -\dfrac{1}{{}^cx} & \dfrac{{}^cy}{{}^cx^2} & 1+\dfrac{{}^cy^2}{{}^cz_2} & -\dfrac{{}^cx\,{}^cy}{{}^cx^2} & -\dfrac{{}^cx}{{}^cz} \end{bmatrix} \tag{13}
$$

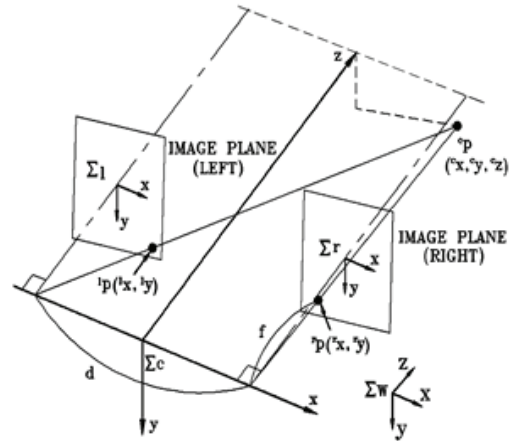The, and axes of the coordinate frames are shown in Fig. 1.

Fig. 1 The coordinates system of vision model

We now introduce the positional vector of the feature point in the image of monocular vision using the symbol. This is the projection

of the point expressed as in the camera frame into the image frame of the monocular vision, and has the following relation.

$$^m x = f^c x^c z^{-1} \qquad (14\text{-}a)$$

$$^m y = f^c y^c z^{-1} \qquad (14\text{-}b)$$

Substituting Eq.s (14-a) and (14-b) into Eq. (13) yields another expression of the image Jacobian for the monocular vision.

$$J = f \begin{bmatrix} -\dfrac{1}{^c z} & 0 & \dfrac{^m x}{^c x} & \dfrac{^m x\, ^m y}{f} & -\dfrac{^m x^2 + f^2}{f^2} & ^m y \\ 0 & -\dfrac{f}{^c x} & \dfrac{^m y}{^c x} & 1 + \dfrac{^m y^2 + f^2}{f^2} & -\dfrac{^m x\, ^m y}{f} & -^m x \end{bmatrix} \quad (15)$$

A disparity which corresponds to the depth of the feature point, is included in J in the case of the stereo vision, but s-term expressed in the camera frame $^c z$ is included in J in the case of the monocular vision.

In the visual servoing, the manipulator is controlled so that the feature points in the image reach their respective desired locations.

We define an error function between the current location of the feature points in image $^I p$ and the desired location $^I p_d$ as

$$E = Q\left(^I p - {}^I p_d\right) \qquad (16)$$

where Q is a matrix which stabilizes the system. Then the feedback law is defined as following equation

$$V = -GE \qquad (17)$$

where G corresponds to a feedback gain.

To realize the visual servoing, we must choose Q so that convergence is satisfied with the error system can be satisfied with

$$\begin{aligned} \dot{E} &= \frac{\delta E}{\delta t} & (18)\\ &= Q\frac{\delta ^I p}{\delta t} \\ &= Q\,{}^I\dot{p} \\ &= QJ_{im}V \\ &= -GQJ_{im}E \end{aligned}$$

We use pseudo-inverse matrix of the image Jacobian J_{im} for Q to make QJ_{im} positive and not to make an input extremely large, that is,

$$Q = J_{im}^+ = \left(J_{im}^T J_{im}\right)^{-1} J_{im}^T \qquad (19)$$
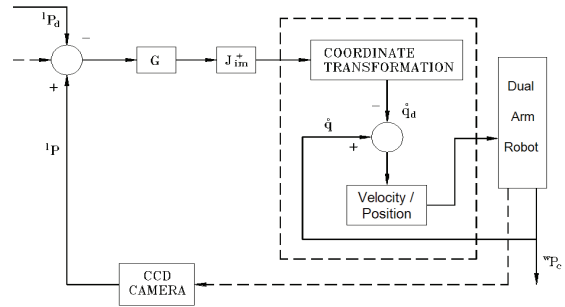

Fig. 2 Block diagram of visual feedback system

Therefore, the feedback command is given as

$$V = -GJ_{im}^+ \left(^I p - {}^I p_d\right) \qquad (20)$$

Fig. 2 shows a block diagram of the control scheme described by Eq. (20). Note that the feedback command is sent to the robot controller and both the transformation of u to

the desired velocity of each joint angle $\dot{q}_d$ and its velocity servo are accomplished in the robot controller as show in Fig. 2.

Furthermore, as J_{im} is a $4n \times 6$ matrix and pseudo-inverse matrix J_{im}^{+} is a $6 \times 4n$ matrix, a feedback command Eq. (20) of 6 degrees of freedom is obtained.

3. Experiment and Results

We have compared the visual feedback using the monocular vision with that using the stereo vision by the experiment. Fig. 3 represents the experimental equipment set-up. In Fig. 3 two DSP vision boards were used, which had been made Samsung Electronics Company in Korea based-on the TMS320C31chips.

In the experiment, feature points of an object are the four corners of a square whose side dimension is $300mm$. In the same condition, we used four feature point seven in the stereo vision. Parameters used the focal length, $f = 20mm$, baseline $d = 150mm$,
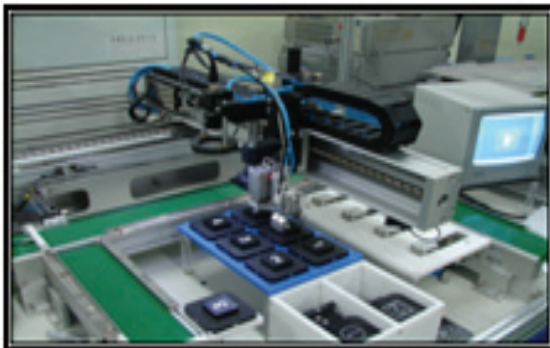

Fig. 3 The experimental equipment set-up

sampling time of $60ms$, gain $\lambda = 1.5$, desired location $^cP_d = (50\,50\,200)^T$, desired orientation In

Euler angle $(\phi, \theta, \psi) = (0, 0, 0)rad$, initial error $(-26, +26, -26)^T mm$ in the translation and $(\phi, \theta, \psi) = (22, 22, 22)rad$ in the orientation.

We select the four corners of a rectangle whose size is $200 \times 200mm$ as the feature points and set the translational error as $(-120\,-120\,-300)mm$ and the other values are the same as before.

The error between the desired location and the current location of the feature points in cases of the monocular and stereo visions are shown in Fig. 4.

Next, we will show the results for the change of the way to choose the feature points and set the initial error image.

Two images were taken and transformed to the binary images in the real time and in parallel by two image input devices and the coordinate of the gravitational center of each feature point was calculated in parallel by two transporters.

We gave the stereo correspondence of the feature point in the first sampling. However, the stereo and temporal correspondence of the feature points in the succeeding sampling was found automatically by searching a nearby area where there were the feature points in the previous sampling frame. The coordinates of the feature points were sent to a transporter for motion control and it calculated a feedback command for the robot.
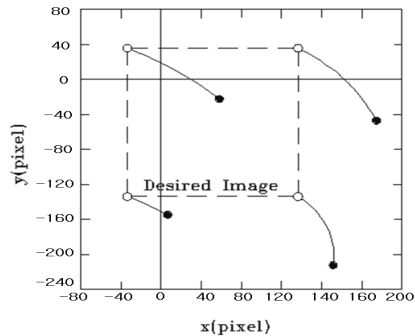
The result was sent to the robot controller by using RS-232C, and the robot was controlled by a velocity servo system in the controller.

Fig. 4 The parts of diverse shapes used on experiment of visual feedback control
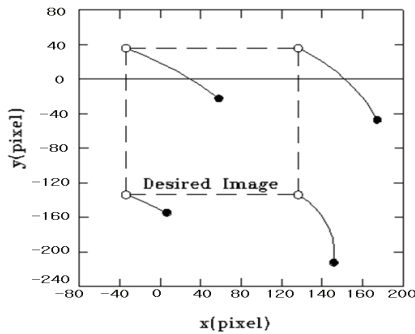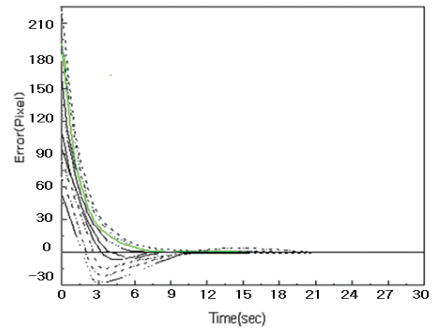

● : Start Point, ○ : Goal Point
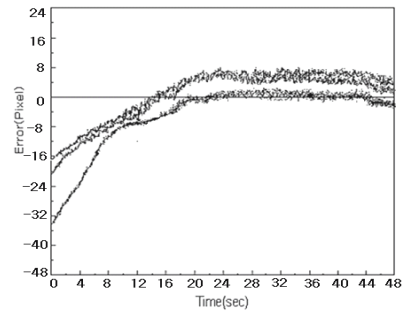(a) Left image

● : Start Point, ○ : Goal Point
(b) Right image

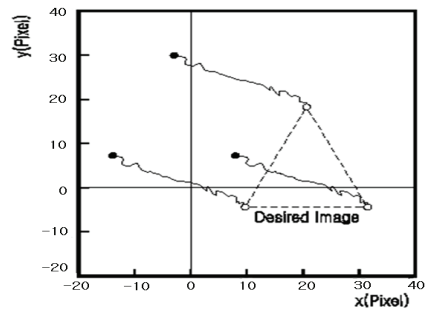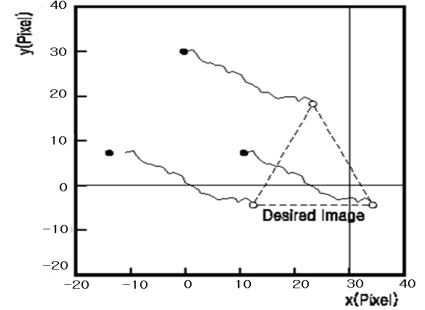Fig. 5 Trajectories of the feature point on the image(Stereo)

(a) Monocular

(b) Stereo

Fig. 6 Positional error in x and y axes(simulation)

● : Start Point, ○ : Goal Point
(a) Left image

● : Start Point, ○ : Goal Point
(b) Right image

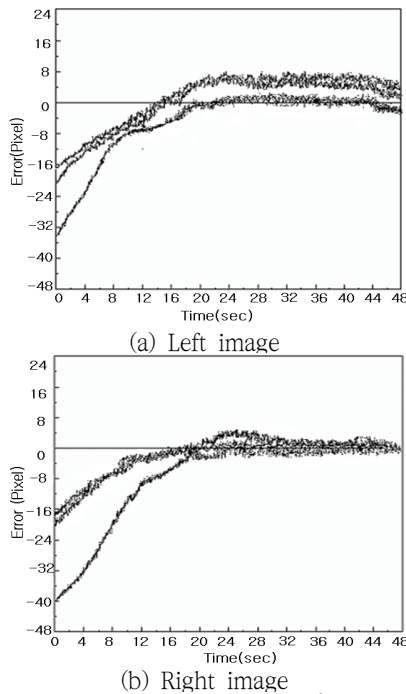Fig. 7 Trajectories of the feature point on the image

(a) Left image

(b) Right image

Fig. 8 Position error in x and y axes

The sampling period of visual feedback was about $50ms$. Details were $16ms$ for taking a stereo image, about $1ms$ for calculating the coordinates of the feature points, $3ms$ for calculating feedback command, about $16ms$ for communicating with the robot controller. If we send a feedback input to the robot controller without using RS-232C, the faster visual feedback can be realized. The desired location was $(0, 0, 550)^T mm$ and the desired orientation in Euler angle, $(\phi, \theta, \psi) = (0, 0, 0)$ degree and the initial error was $(50, 50, 50)^T mm$ for translation. The other parameters were the same as in the simulation. Fig. 4 Shows the parts of diverse shapes used on experiment of visual feedback control. And Fig. 5 Shows trajectories of the feature point on the image(Stereo). Fig. 5(a) shows the results about right image, Fig. 5(b) shows the results about left image. Fig. 6 Shows the position error in x and y axes by simulation test. Fig. 6(a) is the results of monocular vision, and Fig. 6(b) is the results of stereo vision. Fig. 7 shows the result of trajectory tracking of the feature point on the image. Fig. 7(a) is the experiment results about left image, Fig. 7(b) is the experiment results about right image. The error of current and desired location of the feature points are shown in Fig. 8 From these experimental results, we can see that the manipulator converges toward a desired location even if the calibration is not precise.

4. Conclusion

In this papers it has been proposed a new technique of visual feedback control with the vision to control the position and orientation of an SCARA type robot(FARA) with respect to diverse shaped an object. The suggested technique places a robot manipulator to the desired location without giving such prior knowledge as the relative distance to the desired location or the model of an object even if the initial positioning error is large. By using the proposed vision method, the image Jacobian can be calculated at any position. So neither shape information nor desired distance of the target object is required. Also we showed the stability of visual feedback was illustrated even when the initial error is large.

References

1) Weiss, L. E., Sanderson, A. C., and Neuman, C. P., "Dynamic sensor-based control of robots with visual feedback," IEEE Journal of Robotics and Automation, pp. 404-417, 1987.

2) Hashimoto, K., Kimoto, T. Edbine, T., and Kimura, H., "Manipulator control with image-based visual servo," In Proceedings of IEEE International Conference on Robotics and Automation, pp. 2267-2272, 1991.

3) Hashimoto, K., Edbine, T., and Kimura, H., "Dynamic visual feedback control for a hand-eye manipulator," In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1863-1868, 1992.

4) G. Hager, S.Hutchinson, and P. I. Corke."A tutorial on visual servo control, "IEEE Trans. Robotics & Automation, Vol. 12, No. 5, pp. 651-670, 1996.

5) Allen, P. K., Toshimi, B., Timcenko, A., "Real Time Visual Servoing," In Proceeding of the IEEE International Conference on Robotics and Automation, pp. 851-856, 1991.

6) Chaumette, F., Rives, P., and Espiau, B., "Positioning of a robot with respoect to an object, tracking it and estimating its velocity by visual servoing," In Proceedings of the IEEE International Conference on Robotics and Automation, pp. 2248-2253, 1991.

7) Bernard, E., Francois, C., and Patrick, R., "A new approach to visual servoing in robotics," IEEE Transactions on Robotics and Automation, pp. 313-326, 1992.