Adaptive Visual Servoing for Various Kinds of Robot Systems

Koh Hosoda and Minoru Asada

Dept. of Adaptive Machine Systems, Osaka University

Yamadaoka 2–1, Suita 565, Japan

hosoda@mech.eng.osaka-u.ac.jp

Abstract: This paper propose an adaptive visual servoing method consisting of an on-line estimator of the robot/image Jacobian matrix and a feedback/feedforward controller for uncalibrated camera-manipulator systems. The estimator does not need *a priori* knowledge on the kinematic structure nor on parameters of the camera-manipulator system. The controller consists of feedforward and feedback terms to make the image features converge to the desired trajectories using the estimated results. Some experimental results are given to show the validity of the proposed method.

1. Introduction

Visual information plays an important role for a robot to accomplish given tasks in an unknown/dynamic environment. Many vision researchers have been adopting deliberative approaches to the problem of reconstructing 3-D scene structure from visual information, which are not only very time consuming but also brittle to noise, therefore it seems hard to apply these methods to real robot tasks. Recently, there have been several studies on visual servoing, using visual information in the dynamic feedback loop to increase robustness of the closed loop system [1]. For vision-based robots, image features on the image planes are primitive descriptions of the environments. In this sense, feature-based visual servoing control is the most fundamental one for the vision-based robots in which image features are controlled to converge to the desired ones, and therefore has been focused by a number of researchers [2–13].

In most of the previous work on visual servoing, they assumed that the system structure and parameters were known [2–7], or that the parameters could be identified in an off-line process [14]. Such a controller, however, is not robust for disturbances and changes of the parameters. To overcome this problem, some on-line parameter identification schemes have been proposed [8–13]. Weiss et al. [8] assumed that the system could be modeled by linear SISO (Single Input Single Output) equations, and applied an MRAC (Model Reference Adaptive Control) controller. In [9], the structure and parameters of a camera-manipulator system were assumed to be known, and an ARMAX (auto-regressive with external inputs) model was used to estimate disturbances and positions of the target points. Papanikolopoulos et al. [10, 11] modeled the system using an ARMAX model and estimated the coefficients of the model.

They also estimated the depth related parameters in [12]. Yoshimi and Allen [13] used a special camera-manipulator setup to realize an uncalibrated camera system. Thus, in these approaches, there were restrictions and assumptions on the system that the camera-manipulator system was described as SISO equations [8], that *a priori* knowledge on the system structure was required [9, 12, 13], or that the depth was assumed to be constant [10, 11].

On the other hand, most of the previous work have paid their attentions only to the feedback servoing. They sensed positions of targets and made feedback inputs by subtracting the sensed positions from the desired ones. Using their controllers, the manipulator does not move until the error is observed, which can be considered to be *reactive.* To increase the ability of trajectory tracking, there have been several researches on *feedforward,* in which the dynamic motion of the target is predicted [15–17], but no one has mentioned to feedforward control to predict the motion of the robot itself to the best of our knowledge. If the estimator can obtain a kinematic model of the robot system appropriately, the servoing controller can feedforward the obtained kinematics to realize smooth trajectory tracking motion along the desired trajectories designed to accomplish a certain task. For example, a trajectory generator for obstacle avoidance is proposed by authors utilizing the adaptive visual servoing method [18]. In this paper, we propose an adaptive visual servoing method consisting of an on-line estimator and a feedback/feedforward controller for uncalibrated camera-manipulator systems. It has the following features:

1. The estimator does not need *a priori* knowledge on the system parameters nor on the kinematic structure of the system. That is, we need not to devote ourselves to tedious calibration processes, or to separate the unknown parameters from the system equations, which depends on the detailed knowledge on the kinematic structure of the system.

2. There is no restriction on a camera-manipulator system: the number of cameras, kinds of images features, structure of the system (camera-in-manipulator or camera-and-manipulator), the numbers of inputs and outputs (SISO or MIMO). The proposed method is applicable to any kinds of systems.

3. The aim of the estimator is not to obtain the true parameters but to ensure asymptotical convergence of the image features to the desired values under the proposed controller. Therefore, the estimated parameters do not necessarily converge to the true values. In [8–12], they tried to estimate the true parameters, and therefore they need their restrictions and assumptions.

4. The proposed controller can realize smooth tracking motions along the desired trajectories because not only the feedback terms but also feedforward terms are utilized based on the estimated results.

This paper is organized as follows. First, we propose an estimator for an image Jacobian that represents the relation between the image features
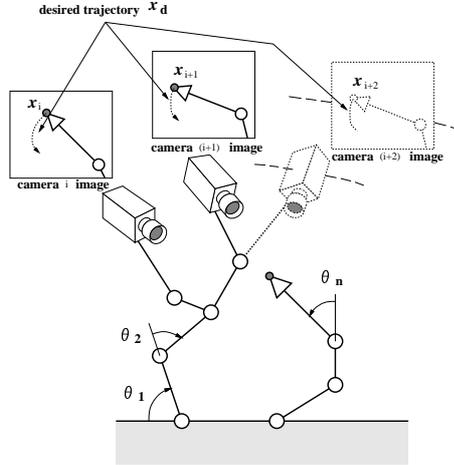
Figure 1. Camera-manipulator system

and the variables that describe the state of the system. Then, a visual feedforward/feedback servoing controller is proposed based on the estimated Jacobian. Finally, experimental results show the validity of the proposed method.

2. Adaptive visual servoing

2.1. Estimation of the relation between image features and system describing variables

A camera-manipulator system consists of manipulators and cameras, as shown in Figure 1. From cameras, one can observe quantities of image features such as position, line length, contour length, and/or area of certain image patterns. The task of the system is to make the quantities of image features converge to the given desired values. The image features are assumed to be on the tip/body of the manipulators or fixed to the ground. That is, a set of measurable variables (which we call system describing variables in the rest of this paper) can describe the state of the camera-manipulator system. Changes of the image features caused by an independently moving object with unknown velocity are not dealt with here.

Let $\boldsymbol{\theta} \in \Re^n$ and $\boldsymbol{x} \in \Re^m$ denote the vectors of the system describing variables and the image features obtained from visual sensors, respectively. A relation between $\boldsymbol{\theta}$ and $\boldsymbol{x}$ is

$$\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{\theta}), \tag{1}$$

because we assume that the system describing variables can describe the state of the system. Differentiating eq.(1), we obtain a velocity relation,

$$\dot{\boldsymbol{x}} = \boldsymbol{J}(\boldsymbol{\theta})\dot{\boldsymbol{\theta}}, \tag{2}$$

where $\boldsymbol{J}(\boldsymbol{\theta}) = \partial\boldsymbol{x}/\partial\boldsymbol{\theta}^T \in \Re^{m \times n}$ is a Jacobian matrix of time-derivatives of the quantities of image features with respect to those of system describing variables.

This Jacobian matrix depends on the kinematic structure of the system, the internal camera parameters such as focal length, aspect ratio, distortion coefficients, and the kinematic parameters such as the length of links and the relative position and orientation of cameras with respect to the tip of the manipulator.

Assuming that movement of the camera-manipulator system is slow enough to consider the Jacobian matrix $\boldsymbol{J}$ to be constant during the sampling time, we obtain

$$\boldsymbol{x}(k+1) = \boldsymbol{x}(k) + \boldsymbol{J}(k)\boldsymbol{u}(k), \tag{3}$$

as a discrete model of the system, where $\boldsymbol{J}(k)$ and $\boldsymbol{u}(k)(=\dot{\boldsymbol{\theta}}\Delta T)$ denote the constant Jacobian matrix and a control input vector in k-th step during sampling rate ΔT, respectively. From eq.(3), i-th row vector of the matrix $\boldsymbol{J}$, $\boldsymbol{j}_i{}^T$, satisfies

$$\{\boldsymbol{j}_i(k+1)^T - \boldsymbol{j}_i(k)^T\}\boldsymbol{u}(k+1) = \{\boldsymbol{x}(k+2) - \boldsymbol{x}(k+1) - \boldsymbol{J}(k)\boldsymbol{u}(k+1)\}_i. \tag{4}$$

Among an infinite number of solutions of eq.(4), we pick up one to make the norm of weighted time-derivatives of $\widehat{\boldsymbol{j}}_i$ as small as possible by the iteration,

$$\widehat{\boldsymbol{j}}_i(k+1) - \widehat{\boldsymbol{j}}_i(k) = \frac{\{\boldsymbol{x}(k+1) - \boldsymbol{x}(k) - \widehat{\boldsymbol{J}}(k)\boldsymbol{u}(k)\}_i}{\boldsymbol{u}(k)^T\boldsymbol{W}_i(k)\boldsymbol{u}(k)}\boldsymbol{W}_i(k)\boldsymbol{u}(k). \tag{5}$$

Theoretically, the right-hand side of eq.(5) does not tend to infinity when $\parallel \boldsymbol{u} \parallel$ tends to 0, because $\mid \{\boldsymbol{x}(k+1) - \boldsymbol{x}(k) - \widehat{\boldsymbol{J}}(k)\boldsymbol{u}(k)\}_i \mid$ also tends to 0 at the same or faster speed. In real situations, however, the right-hand side is prone to be unstable because of disturbances. To increase the stability of estimation (5), particularly when $\parallel \boldsymbol{u} \parallel$ tends to 0, the estimating law is modified as

$$\widehat{\boldsymbol{j}}_i(k+1) - \widehat{\boldsymbol{j}}_i(k) = \frac{\{\boldsymbol{x}(k+1) - \boldsymbol{x}(k) - \widehat{\boldsymbol{J}}(k)\boldsymbol{u}(k)\}_i}{\rho_i + \boldsymbol{u}(k)^T\boldsymbol{W}_i(k)\boldsymbol{u}(k)}\boldsymbol{W}_i(k)\boldsymbol{u}(k), \tag{6}$$

where ρ_i is an appropriate positive constant that makes the iteration (6) stable. When $\parallel \boldsymbol{u} \parallel$ tends to 0, the denominator tends to ρ_i and the stability is ensured even if the numerator does not tend to 0 because of disturbances. The positive constant ρ_i is determined so small that ρ_i can be neglected with respect to $\parallel \boldsymbol{u} \parallel$ when $\parallel \boldsymbol{u} \parallel$ is large. Note that when ρ is in the range $0 < \rho \leq 1$ and the matrix $\boldsymbol{W}_i$ is a covariance matrix, the proposed estimator coincides with the least-mean-square method[19].

The proposed estimator is intended not to obtain the true Jacobian matrix/parameters, but to estimate a matrix that satisfies eq.(3). This is the main difference from [9], [10], and [11], in which they tried to estimate the true parameters. To estimate the true parameters, one have to make restrictions and assumptions on the camera-manipulator system. The proposed estimator, however, is not intended to estimate the true parameters, but to make the closed loop system consisting of this estimator and a controller stable. Therefore, there is neither restrictions nor assumptions on the camera-manipulator system.
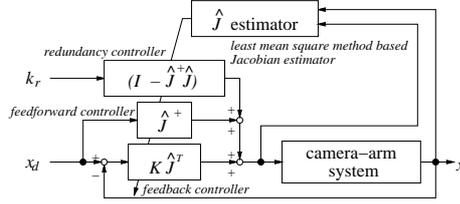
Figure 2. Block diagram of the proposed method

2.2. Feedforward/feedback visual controller

In this section, a feedforward/feedback visual controller is proposed, based on the estimated Jacobian matrix $\widehat{\boldsymbol{J}}$. The aim of the controller is to ensure convergence of the image feature vector $\boldsymbol{x}(k)$ to the desired vector $\boldsymbol{x}_d(k)$.

From eq.(3), we can derive a feedforward/feedback controller,

$$
\begin{aligned}
\boldsymbol{u}(k) \quad = \quad & \widehat{\boldsymbol{J}}(k)^+\{\boldsymbol{x}_d(k+1) - \boldsymbol{x}_d(k)\} \\
& +\{\boldsymbol{I}_n - \widehat{\boldsymbol{J}}(k)^+\widehat{\boldsymbol{J}}(k)\}\boldsymbol{k}_r \\
& +\boldsymbol{K}\widehat{\boldsymbol{J}}(k)^T\{\boldsymbol{x}_d(k+1) - \boldsymbol{x}(k)\},
\end{aligned} \tag{7}
$$

where $\widehat{\boldsymbol{J}}(k)^+$, $\boldsymbol{I}_n$, and $\boldsymbol{K}$ denote a pseudo-inverse matrix of $\widehat{\boldsymbol{J}}(k)$, an $n \times n$ identity matrix, and a positive-definite gain matrix, respectively. Let $\boldsymbol{k}_r$ be an arbitrary vector.

The first and second terms on the right-hand side are feedforward terms. The second term on the right-hand side denotes the redundancy of the camera-manipulator system. The third term on the right-hand side is a feedback term that ensures stability of the closed loop system. Note that one can use $\widehat{\boldsymbol{J}}(k)^+$ instead of $\widehat{\boldsymbol{J}}(k)^T$ to ensure the closed loop stability [3].

We propose an adaptive visual servoing method consisting of the proposed estimator and controller shown in Figure 2.

3. Experiments

To show the validity of the proposed method, some experimental results are given in this section. First step response results of two kinds of camera-manipulator systems are given to show how the estimator and the feedback terms of the controller can realize *reactive* tasks well and how the proposed method can be applied to various kinds of system structures. Then, a result of trajectory tracking is given to show how the feedforward terms work.

3.1. Experimental equipment

In Figure 3, a camera-manipulator system used for the experiments is shown. The video signals from two cameras (UN401, ELMO) are sent to an image processing board MV200 (DataCube, image size: 512[pixel] $\times$ 480[pixel]) and compressed into the half along the horizontal axis (256[pixel] $\times$ 480[pixel]). Two images are pasted onto one image (512[pixel] $\times$ 480[pixel], see Figure 5 for example), which is sent to a tracking module equipped with a high-speed
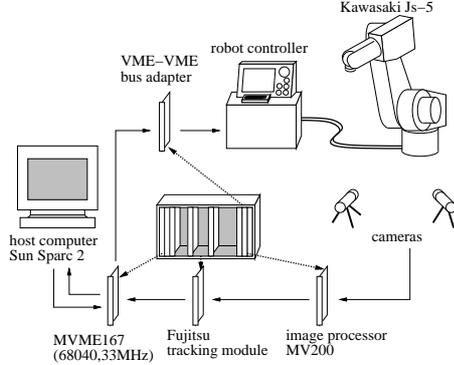
Figure 3. Experimental equipment

correlation processor utilizing a SAD (Sum of Absolute Difference) measure by Fujitsu [20]. Before starting an experiment, we specify target images to be tracked by the module. During the experiment, the module tracks the target images, and it feeds coordinates of the images in the image plane to the main control board MVME167 (CPU:68040, 33MHz, motorola). The board calculates a desired posture of the manipulator by the proposed method and sends it to the manipulator controller through a VME-VME bus adapter. We use a 6 d. o. f. manipulator Js-5 (Kawasaki Heavy Industry Co.) as a 3 d. o. f. manipulator, maintaining fixed desired orientation of the tip of the manipulator. Therefore, the system describing variable vector $\boldsymbol{\theta}$ is a tip position vector of the manipulator in these experiments. Using this experimental equipment and writing programs using C language on VxWorks (Wind River), the sampling ratio is 30[Hz].

3.2. Step responses of two kinds of systems

To show that the estimator and the feedback terms of the controller can realize *reactive* tasks well and that the estimator and controller can be applied to different system structures, step responses of two kinds of systems are given in this subsection. At $t = 0$, the desired image feature coordinates which are *taught by showing* are fed to the controller. In these experiments, we show step responses, and therefore the feedforward terms in eq.(7) equal zeros.

The positive constants $\rho_i = 0.8$ $(i = 1, \cdots, 4)$ are selected as small as possible in the trial and error manner. We set the weighting matrices $\boldsymbol{W}_i(k) = \boldsymbol{I}_3 (i = 1, \cdots, 4)$. The feedback gain matrix $\boldsymbol{K}$[m/pixel] in eq.(7) is also selected in the trial and error manner,

$$\boldsymbol{K} = \mathrm{diag} \left[\begin{array}{cc} 1.5 \times 10^{-4} & 1.5 \times 10^{-4} \\ 1.5 \times 10^{-4} & 1.5 \times 10^{-4} \end{array} \right].$$

An initial Jacobian matrix, which is needed at the beginning of the control phase, can be given arbitrarily as far as its rank is full. In the experiment,

therefore, we give the initial Jacobian matrix as

$$\hat{\boldsymbol{J}}(0) = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \\ 0 & 0 & 0.1 \end{bmatrix}.$$

First, we fix the two cameras on the ground. The reference images are windows of a pattern (a cross) which is fixed at the tip of the manipulator (see Figure 4).

In Figure 5, we can find the initial posture of the manipulator, the initial positions of the reference images, and the desired positions of the reference images. Responses of two cases, (a) applying the proposed controller with the proposed estimator, and (b) applying the proposed controller without on-line estimation, are shown in Figure 6. Because the initial Jacobian matrix is given arbitrary, the controller cannot eliminate error without on-line estimation. On the other hand, using the proposed method, the manipulator can be controlled to make the image features converge to the desired ones.

Second, in order to show that it can be applied to various kinds of system structures without *a priori* knowledge on the systems, we apply the proposed method to a different system in which two cameras are mounted on the tip of the manipulator (see Figure 7). The reference images are the windows of a pattern (a cross) fixed on the ground. The positive constants ρ_i, the gain matrix $\boldsymbol{K}$ and the weighting matrices $\boldsymbol{W}_i$ are the same as the previous case. That is, the estimator and the controller are all the same as in the previous experiment. The reference images, their initial positions, and the desired positions are shown in Figure 8, and the results are shown in Figure 9.

From these experimental results, we can conclude that the proposed on-line estimator and the feedback term of the proposed controller are effective to realize a reactive task, and that the proposed method is applicable to different kinds of systems.

3.3. Trajectory tracking task

In this subsection, a result on trajectory tracking task is given to show the validity of the proposed purposive visual control.

We fix two cameras on the ground. The reference images are windows of a pattern (a cross) fixed at the tip of the manipulator (see Figure 4). The desired trajectories must be realizable. Therefore, the desired trajectories are *taught by showing* to satisfy this constraint (Figure 10). The squares in the figure indicate the location along the desired trajectory every 0.2[s]. The target moves along each trajectory in 12[s] in each image.

The initial Jacobian matrix is the same one given in the previous experiments. The tracked image feature vector is $\boldsymbol{x} \in \Re^4$, and the controlled tip position vector of the manipulator is $\boldsymbol{\theta} \in \Re^3$, therefore the second feedforward term on the right-hand side of eq.(7) equals zero. Because our experimental system has a time-delay problem, we cannot stabilize the closed loop system with 100% feedforward terms in eq.(7). Therefore we apply 30% feedforward terms in the following experiment.

One of experimental results is given in Figure 11, where desired trajectories and realized ones are indicated. From this figure, we can see how the proposed method can track the desired trajectories better than the controller without the feedforward terms.

4. Conclusion and Discussion

In this paper, we have proposed an adaptive visual servoing method consisting of an on-line estimator and a feedback/feedforward controller for uncalibrated camera-manipulator systems. We have proposed an estimator for the Jacobian matrix that describes the relation between the image features and the system describing variables. Then, a feedforward/feedback controller has been proposed making use of the estimated relation. Finally experimental results are given to show that the proposed method is validity to various kinds of robot systems.

We have to mention to the redundancy of the system. If the camera-manipulator system is redundant to accomplish given tasks, the redundancy can be utilized to realize other sub-tasks such as obstacle avoidance. In the proposed controller, the redundancy is denoted as the second term on the right-hand side of eq.(7), but we have not mentioned how to utilize the redundancy in this paper. *A study on redundancy in the uncalibrated system* is one of our future major work.

One alternative to deal with such redundancy is to introduce another servoing method, and to build a hybrid servoing controller. The authors have shown some theoretical and experimental results: (1) hybrid adaptive visual servoing/force servoing control [21], and (2) adaptive visual servoing control for legged robots [22]. When we build a fast/robust robot system, such kinds of hybrid servoing controllers would be powerful and essential.

References

[1] P. I. Corke. Visual control of robot manipulators – a review. In *Visual Servoing*, pages 1–31. World Scientific, 1993.

[2] W. Jang and Z. Bien. Feature-based visual servoing of an eye-in-hand robot with improved tracking performance. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 2254–2260, 1991.

[3] K. Hashimoto, T. Kimoto, T. Ebine, and H. Kimura. Manipulator control with image-based visual servo. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 2267–2272, 1991.

[4] N. Maru, H. Kase, et al. Manipulator control by visual servoing with the stereo vision. In *Proc. of the 1993 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 1865–1870, 1993.

[5] P. Allen, A. Timcenko, B. Yoshimi, and P. Michelman. Automated tracking and grasping of a moving object with a robotic hand-eye system. *IEEE Trans. on Robotics and Automation*, RA–9(2):152–165, 1993.

[6] A. Castano and S. Hutchinson. Visual compliance: Task-derected visual servo control. *IEEE Trans. on Robotics and Automation*, 10(3):334–342, 1994.

[7] G. D. Hager, W.-C. Cang, and A. S. Morse. Robot feedback control based on

stereo vision: Towards calibration-free hand-eye coordination. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 2850–2856, 1994.

[8] L. E. Weiss, A. C. Sanderson, and C. P. Neuman. Dynamic sensor-based control of robots with visual feedback. *IEEE J. of Robotics and Automation*, RA-3(5):404–417, 1987.

[9] J. T. Feddema and C. S. G. Lee. Adaptive image feature prediction and control for visual tracking with a hand-eye coordinated camera. *IEEE Trans. on System, Man, and Cybernetics*, 20(5):1172–1183, 1990.

[10] N. P. Papanikolopoulos and P. K. Khosla. Adaptive robotic visual tracking: Theory and experiments. *IEEE Trans. on Automatic Control*, 38(3):429–445, 1993.

[11] B. Nelson, N. P. Papanikolopoulos, and P. K. Khosla. Visual servoing for robotic assembly. In *Visual Servoing*, pages 139–164. World Scientific, 1993.

[12] N. P. Papanikolopoulos, B. Nelson, and P. K. Khosla. Six degree-of-freedom hand/eye visual tracking with uncertain parameters. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 174–179, 1994.

[13] B. H. Yoshimi and P. K. Allen. Alignment using an uncalibrated camera system. *IEEE Trans. on Robotics and Automation*, 11(4):516–521, 1995.

[14] R. Y. Tsai and R. K. Lenz. A new technique for fully autonomous and efficient 3d robotics hand/eye calibration. *IEEE Trans. on Robotics and Automation*, 5(3):345–358, 1989.

[15] C. Brown. Gaze controls with interactions and delays. *IEEE Trans. on System, Man, and Cybernetics*, 20(1):518–527, 1990.

[16] E. D. Dickmanns, B. Mysliwetz, and T. Christians. An integreted spatiotemporal approach to automatic visual guidance of autonomous vehicles. *IEEE Trans. on System, Man, and Cybernetics*, 20(6):1273–1284, 1990.

[17] W. J. Wilson. Visual servo control of robots using kalman filter estimates of robot pose relative to work-pieces. In *Visual Servoing*, pages 71–104. World Scientific, 1993.

[18] K. Hosoda, K. Sakamoto, and M. Asada. Trajectory generation for obstacle avoidance of uncalibrated stereo visual servoing without 3d reconstruction. In *Proc. of the 1995 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 29–34, 1995.

[19] P. Eykhoff. *System Identification*, chapter 7. John Wiley & Sons Ltd., 1974.

[20] M. Inaba, T. Kamata, and H. Inoue. Rope handling by mobile hand-eye robots. In *Proc. of Int. Conf. on Advanced Robotics*, pages 121–126, 1993.

[21] Koh Hosoda, Katsuji Igarashi, and Minoru Asada. Adaptive hybrid visual servoing/force control in unknown environment. In *Proc. of the 1996 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 1097–1103, 1996.

[22] K. Hosoda, M. Kamado, and M. Asada. Vision-based servoing control for legged robots. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 3154–3159, 1997.
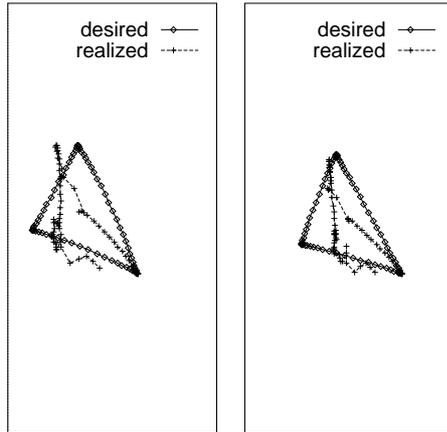
image of camera 1   image of camera 2

(a) Realized trajectories without feedforward
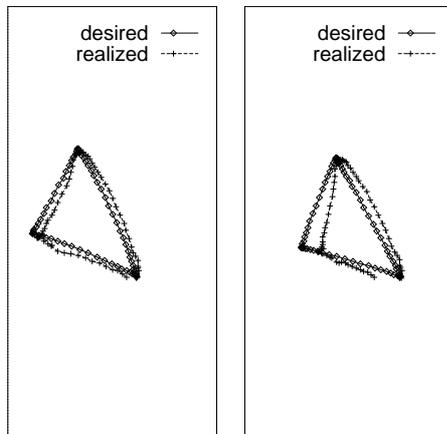
image of camera 1   image of camera 2

(b) Realized trajectories with feedforward

Figure 11. Realized trajectories with and without feedforward terms
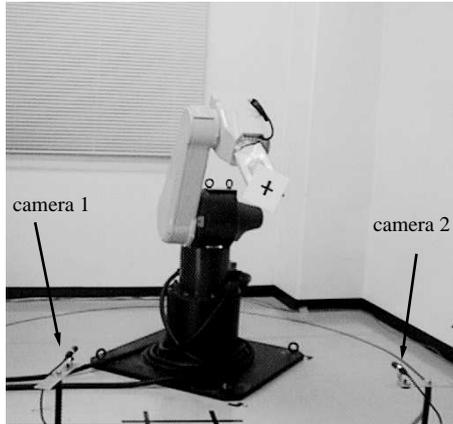
Figure 4. Eye and arm system used for experiments
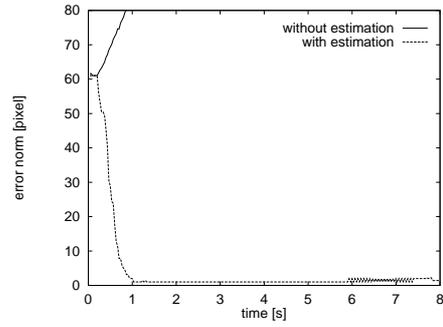
Figure 6. Experimental result 1 (step response, eye and arm system, with and without estimation)
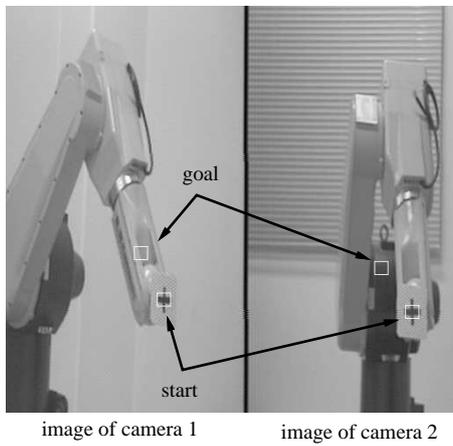
Figure 5. Initial and final position of prememorized image pattern(initial posture, eye and arm system)
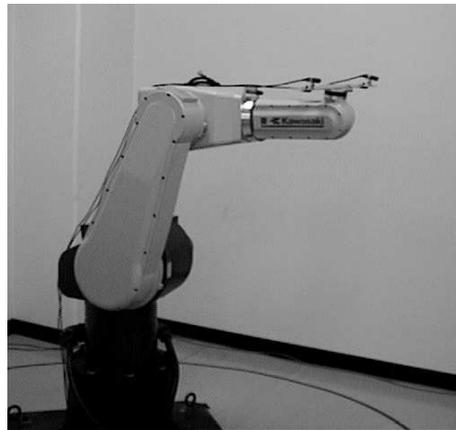
Figure 7. Eye on arm system used for experiments

image of camera 1 image of camera 2
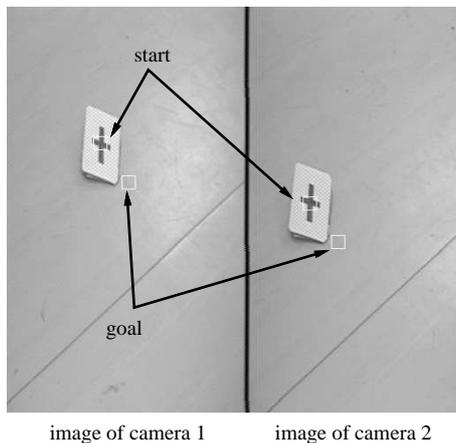
Figure 8. Initial and final position of prememorized image pattern(initial posture, eye on arm system)

start and goal

image of camera 1 image of camera 2
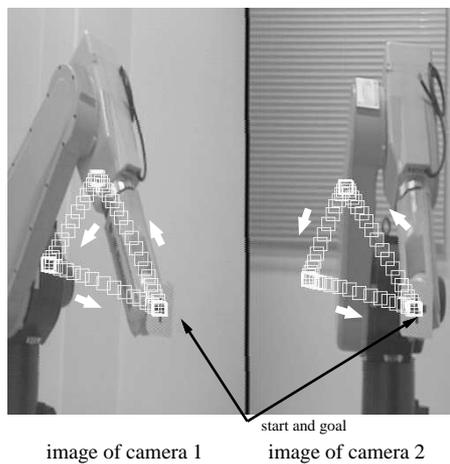
Figure 10. Given desired trajectories for tracking control
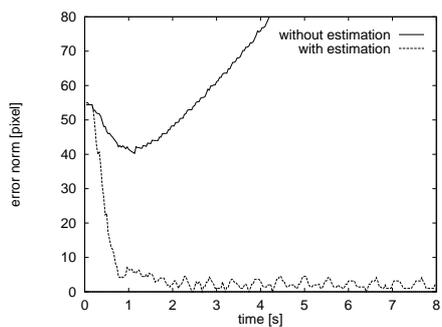
Figure 9. Experimental result 2 (step response, eye on arm system, with and without estimation)