

Visual Servoing Utilizing Zoom Mechanism

Koh HOSODA, Hitoshi MORIYAMA and Minoru ASADA

Dept. of Mechanical Engineering for Computer-Controlled Machinery, Faculty of Engineering,
Osaka University, Suita, Osaka 565, Japan

Abstract

A camera, which is used as an artificial vision in visual servoing control, often has a zoom mechanism. A zoom mechanism cannot realize fast motion while an arm mechanism can, and it has only one degree of freedom. On the other hand the arm mechanism cannot cover wide range of change of images while the zoom mechanism can. In this paper, we propose a complementary visual servoing controller of zoom and arm mechanisms. First we discuss on the condition that camera position and zoom setting are considered as redundant. Then a visual servoing controller is proposed making use of complementary characteristics of the both mechanisms. To show the effectiveness of the proposed controller, experimental results are shown.

1 Introduction

Visual information is essential for robot systems to realize given tasks in dynamic environments. There have been many vision researches adopting deliberative approaches in order to describe 3-D scene structure, which are very time consuming, and therefore utilizing these methods to real robot applications seems hard. Recently, there have been many studies on visual servoing control schemes utilizing the visual sensors to increase the capability of the robot systems in the dynamic environments (you can find a summary in [1]). In these days, many studies focus on control schemes which make features on the image planes converge to the desired values [2–9].

On the other hand, a camera, which is used as an artificial vision device, often has a zoom mechanism. A zoom mechanism is a peculiar mechanism of artificial vision that animals do not have. Utilizing a zoom mechanism can change sensitivity of image features w. r. t. manipulator movement, and can compensate the limit of the physical constraint of the manipulator mechanism. It can also change *resolvability* of the cameras [10]. But the zoom mechanism has only one d. o. f. and cannot realize fast motion because of its auto-focusing function, etc. Therefore to utilizing merits of the zoom mechanism with the manipulator mechanism, a certain *combining* controller have to be proposed. In this sense, by utilizing this zoom mechanism, the ability of visual servoing is enhanced. Nelson and Khosla [10] only focused on the resolvability, not on the complementary and/or redundant characteristics of the zoom and arm mechanisms.

In this paper, we propose a complementary visual servoing controller of zoom and arm mechanisms.

First we discuss on the condition that camera position and zoom setting are considered as redundant, which is not strictly true since the former alters the perspective distortion while the latter does not. Then a visual servoing controller is proposed making use of complementary characteristics of the both mechanisms. To show the effectiveness of the proposed controller, experimental results are shown.

2 Equivalence between zoom mechanism and arm mechanism

When the arm has enough degrees of freedom to make the image features converge to the given desired features, and the cameras have *additional* zoom mechanisms, the whole system will be a redundant system to accomplish the given task. But “positioning by the arm mechanism” and “zooming” are not always equivalent (see figure 1). In this section, we will derive the condition to make positioning by the arm mechanism and zooming are *almost* equivalent.

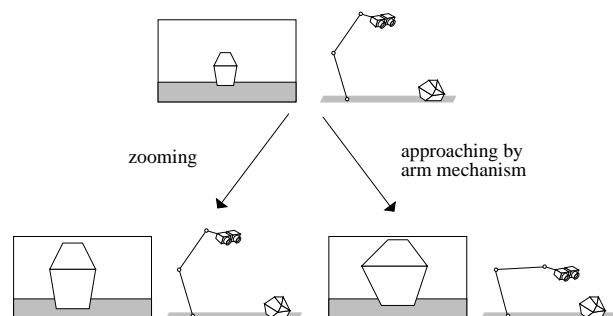


Figure 1: Relation between positioning by arm mechanism and zooming

2.1 Visual servoing system with zoom mechanism

A system consisting of an arm mechanism and cameras is shown in figure 2. Let Σ_w , Σ_c and Σ_I denote the world coordinate frame fixed on the ground, a camera coordinate frame fixed on the camera base, and an image coordinate frame fixed to the image plane, respectively. Let ${}^w\mathbf{p}$ and ${}^w\boldsymbol{\phi}$ denote position and orientation vectors of Σ_c with respect to Σ_w . Let ${}^I\mathbf{x} \in \mathbb{R}^m$ and ${}^I\mathbf{x}_d \in \mathbb{R}^m$ denote an image feature vector consisting

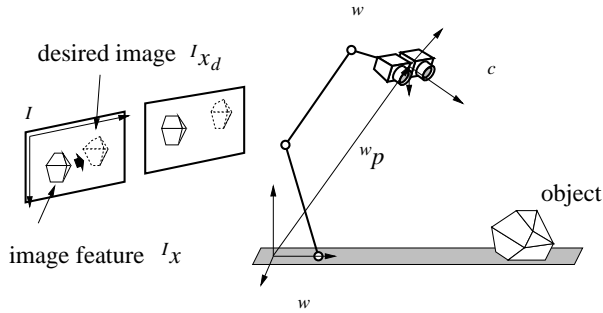


Figure 2: A system consisting of an arm mechanism and camera

of feature points, area of the image, contour etc., and a desired image feature vector, respectively.

Assuming that the object is fixed to the ground. If the system does not have a zoom mechanism, the image feature vector ${}^I\mathbf{x}$ is a function of ${}^w\mathbf{p}$ and ${}^w\phi$,

$${}^I\mathbf{x} = {}^I\mathbf{x}({}^w\mathbf{p}, {}^w\phi). \quad (1)$$

By differentiating this equation, a relation between the time-derivative of image feature vector and velocity vector of cameras,

$${}^I\dot{\mathbf{x}} = \mathbf{J} {}^w\mathbf{v}, \quad (2)$$

can be obtained, where ${}^w\mathbf{v} \in \mathfrak{R}^n$ denotes a velocity vector of Σ_c with respect to Σ_w . The matrix $\mathbf{J} \in \mathfrak{R}^{m \times n}$ is an image jacobian.

If the system has zoom mechanisms, the image feature vector ${}^I\mathbf{x}$ is a function of ${}^w\mathbf{p}$, ${}^w\phi$ and focal length vector $\mathbf{f} \in \mathfrak{R}^l$,

$${}^I\mathbf{x} = {}^I\mathbf{x}({}^w\mathbf{p}, {}^w\phi, \mathbf{f}). \quad (3)$$

Therefore, the relation between the time-derivative of image feature vector and velocity vector of cameras becomes

$${}^I\dot{\mathbf{x}} = \tilde{\mathbf{J}} \tilde{\mathbf{v}}, \quad (4)$$

where $\tilde{\mathbf{v}} = [{}^w\mathbf{v}^T \ \dot{\mathbf{f}}^T]^T \in \mathfrak{R}^{(n+l)}$ is an extended velocity vector, and $\tilde{\mathbf{J}} \in \mathfrak{R}^{m \times (n+l)}$ is an extended image jacobian.

2.2 Equivalence between zoom mechanism and arm mechanism

As we mentioned at the beginning of this section, positioning by the arm mechanism and zooming are not always equivalent. In this subsection, this equivalence will be discussed.

At $t = 0$, initial vector of image features is assumed to coincide with the desired vector. If the system is not equipped with zoom mechanisms, a realizability condition to realize the given desired features, is

$$\text{rank}[\mathbf{J}(t)|{}^I\dot{\mathbf{x}}_d(t)] \leq n. \quad (5)$$

Even if additional desired image features are appended, the realizability condition (5) may be satisfied as far as the desired image features are given dependent to the 3D constraint (for example, as far as the desired image features are given by *teaching by showing*).

If the system is equipped with zoom mechanisms, the condition that zooming and positioning by the arm mechanism are equivalent becomes

$$\text{rank}[\tilde{\mathbf{J}}|{}^I\dot{\mathbf{x}}_d] = \text{rank}[\mathbf{J}|{}^I\dot{\mathbf{x}}_d]. \quad (6)$$

Because zooming is not equivalent to positioning, this equation is not always satisfied. But if cameras are far from the object, and the image features are viewed in small range, as is often the case with visual servoing, zooming and positioning by the arm mechanism are almost equivalent.

Consider the case that the sign of equality holds in eq.(5). By applying singular-value decomposition to the matrix $[\tilde{\mathbf{J}}|{}^I\dot{\mathbf{x}}_d]$, we can calculate singular values $\sigma_1, \dots, \sigma_n, \sigma_{n+1}, \dots$ ($\sigma_1 > \dots > \sigma_n > \sigma_{n+1} > \dots$). The condition that the zooming and positioning by the arm mechanism are almost equivalent becomes

$$\sigma_{n+1} \ll \sigma_1 \cdots \sigma_n. \quad (7)$$

3 Visual servoing controller utilizing zoom mechanism

In the previous section, we have derived the condition that zooming is almost equivalent to positioning by the arm mechanism. In such a case, the whole system is considered to be a redundant system.

The arm mechanism can realize fast motion, but when the dynamic range of the desired image feature becomes large, the system cannot realize visual servoing because of the physical constraint. On the other hand, the zoom mechanism cannot realize fast motion, but by utilizing this, one can compensate the limit of the physical constraint of the arm mechanism.

In this paper, we utilize the redundancy to satisfy the following equations:

$${}^I\dot{\mathbf{x}} + \mathbf{K}_x({}^I\mathbf{x} - {}^I\mathbf{x}_d) = 0, \quad (8)$$

$$\dot{\mathbf{f}} + \mathbf{K}_f(\mathbf{f} - \mathbf{f}_d) = 0, \quad (9)$$

where \mathbf{f}_d is desired focal length vector. The desired focal length is calculated so as to keep the robot arm in an appropriate posture. By tuning the gain matrices \mathbf{K}_x and \mathbf{K}_f , we can make use of the redundancy. For compensating fast and small motion, the arm mechanism is used, and for compensating slow and large motion, the zoom mechanism is used. In this sense, gain \mathbf{K}_x is selected rather large, while \mathbf{K}_f is selected rather small.

Eq.(4) can be rewritten,

$${}^I\dot{\mathbf{x}} = [\mathbf{J}_v \ \mathbf{J}_f] \begin{bmatrix} {}^w\mathbf{v} \\ \mathbf{f} \end{bmatrix}. \quad (10)$$

From eq.(9) we can obtain the input \mathbf{u}_f as

$$\mathbf{u}_f = -\mathbf{K}_f(\mathbf{f} - \mathbf{f}_d). \quad (11)$$

From eqs.(10) and (11), we can calculate the input u_v :

$$\begin{aligned} u_v &= -J_v^+(K_x(Ix - Ix_d) + J_f f) \\ &= -J_v^+\{K_x(Ix - Ix_d) \\ &\quad - J_f K_f(f - f_d)\}, \end{aligned} \quad (12)$$

where J_v^+ denotes a pseudo-inverse matrix of J_v . In figure 3, the structure of the proposed visual servoing scheme is shown.

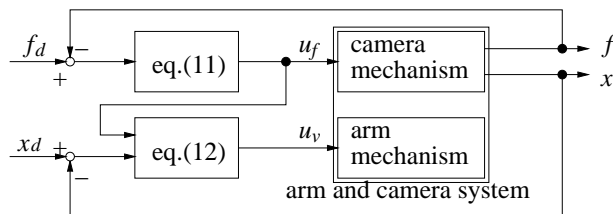


Figure 3: Visual servoing controller utilizing zoom mechanism

4 Experiments

4.1 Experimental system

In this section, some experimental results will be shown to demonstrate effectiveness of the proposed scheme. In figure 4, a camera-manipulator system used for experiments is shown. The robot manipulator is a puma type 6 d. o. f. robot manipulator, Js-5 by Kawasaki Heavy Industry. In this experiment, the manipulator is used as a 3 d. o. f. manipulator. At the tip of the manipulator, a camera module (Sony, EVI-310) is attached.

The structure of the system is shown in figure 5. The video signal from the camera EVI-310 is sent to the tracking module(Fujitsu). The tracking module has a function of block correlation to track some pre-memorized patterns[11]. Then the tracking module feeds coordinates of reference patterns to the controller, MVME167(CPU:68040, 33MHz, motorola) in video plane 512×480 [pixel]. The controller calculates control input for the robot by the proposed scheme and sends it to the robot controller via VME-VME bus adaptor. It also calculates control input for the zoom mechanism and sends via a RS-232C serial port to the camera module EVI-310. A Sun sparc 2 is used to supervise the VME system via ether net. C language is used to write the control program. Sampling time is 49.4 [ms]. In this experiment, the image features are the coordinates of images of two points.

4.2 Experimental results

Preliminary experiments are done to demonstrate responses of zoom and arm mechanisms. Settling time of the zoom mechanism is about 5[s] for compensating 30[pixel] step change of desired image features, while that of the arm mechanism is about 3[s]. They depend on the feedback gain of the controller, which is

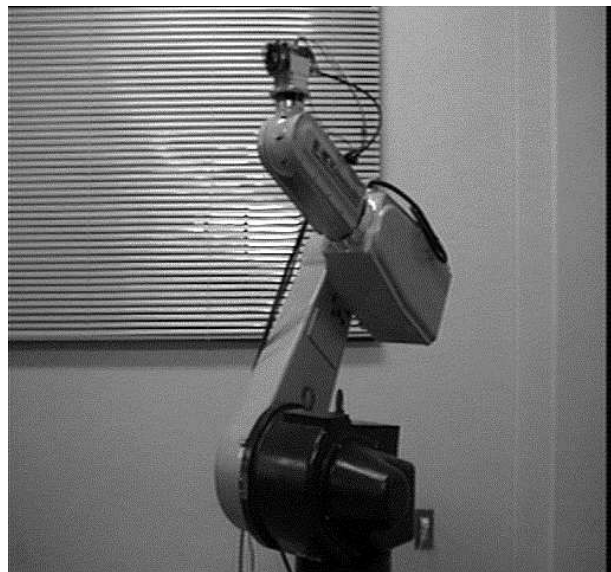


Figure 4: Robot arm with camera, used for experiment

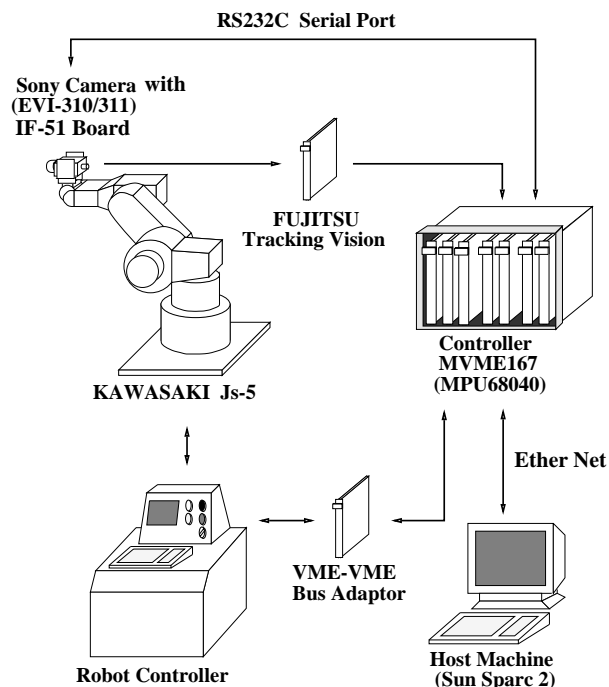


Figure 5: Experimental system

also dependent to the bandwidth of cameras and manipulators. But generally speaking, speed of zoom is rather slow because of the auto-focusing mechanism, which is indispensable for changing focal length.

The initial tip position of the robot is ${}^w p_x, {}^w p_y, {}^w p_z = [9.04 \times 10^{-3}, 1.92 \times 10^{-1}, 7.01 \times 10^{-1}]$ [m]. The initial focal length is 13.75×10^{-3} [m]. Initial image feature vector is

$$\begin{aligned} {}^I \mathbf{x}_0 &= [x_1 \ y_1 \ x_2 \ y_2]^T \\ &= [122 \ 204 \ 350 \ 256]^T, \end{aligned}$$

and at $t = 0$, the desired image feature vector is

$${}^I \mathbf{x}_d = [54 \ 140 \ 306 \ 200]^T,$$

which is given by “teaching by showing”. The extended image jacobian $\tilde{\mathbf{J}}$ at the desired posture is

$$\tilde{\mathbf{J}} = \begin{bmatrix} 2.43 \times 10^3 & 0.00 & 6.93 \times 10^1 & 3.45 \times 10^3 \\ 0.00 & 2.43 \times 10^3 & 1.73 \times 10^2 & 8.61 \times 10^3 \\ 2.40 \times 10^3 & 0.00 & 3.72 \times 10^2 & 1.87 \times 10^4 \\ 0.00 & 2.40 \times 10^3 & 2.40 \times 10^2 & 1.21 \times 10^4 \end{bmatrix}.$$

To examine the equivalence between zooming and positioning, singular values of $[\mathbf{J}|{}^I \mathbf{x}_0 - {}^I \mathbf{x}_d]$ and $[\tilde{\mathbf{J}}|{}^I \mathbf{x}_0 - {}^I \mathbf{x}_d]$ are calculated instead of $[\mathbf{J}|{}^I \dot{\mathbf{x}}_d]$ and $[\tilde{\mathbf{J}}|{}^I \dot{\mathbf{x}}_d]$. Singular values of $[\mathbf{J}|{}^I \mathbf{x}_0 - {}^I \mathbf{x}_d]$ are 3.44×10^3 , 3.42×10^3 , 2.21×10^2 and 1.30. Singular values of $[\tilde{\mathbf{J}}|{}^I \mathbf{x}_0 - {}^I \mathbf{x}_d]$ are 2.43×10^4 , 3.41×10^3 , 1.57×10^3 and 2.00. The fourth singular values are small enough comparing to other three values in both cases, therefore we can apply the proposed controller.

The gain matrix \mathbf{K}_x is selected by trial and error,

$$\mathbf{K}_x = \text{diag}[4.0 \times 10^{-2} \ 4.0 \times 10^{-2} \ 4.0 \times 10^{-2} \ 4.0 \times 10^{-2}].$$

In this paper, results of two cases are shown.

case 1 $K_f = 0$, the case that the zoom mechanism is not utilized.

case 2 $K_f = 1.0 \times 10^{-2}$

In figure 6, error norm of a feature point in the image plane, $\| [x_1 \ y_1]^T - [x_{1d} \ y_{1d}]^T \|$, is shown. In figure 7, change of focal length f is shown. In figure 8, tip position of the robot, ${}^w \mathbf{p}$, is shown. From figure 6, we can show that the capacity of the proposed method to make image features converge to the desired values, is almost the same as that of a method without utilizing zoom mechanisms. Checked in the preliminary experiment, the speed of the zoom mechanism is almost double times comparing to that of the arm mechanism. Therefore the results in figure 6 have demonstrated that the arm can compensate the slow response of the camera from this experiment. From figures 7 and 8, by applying the proposed method, at the beginning of control, the arm is working so as to eliminate error, and after a while the zooming mechanism works

to move the arm to the initial posture. Around 15[s], the posture of the arm is almost the same as initial posture by utilizing the zoom mechanism, while the final posture is different without zoom mechanism. In this sense, the proposed method utilizing zoom mechanisms can make the visual servoing system move in wider range.

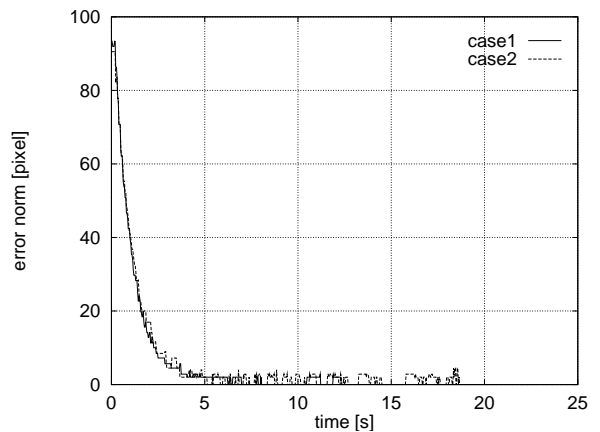


Figure 6: Result 1 (step response, error norm)

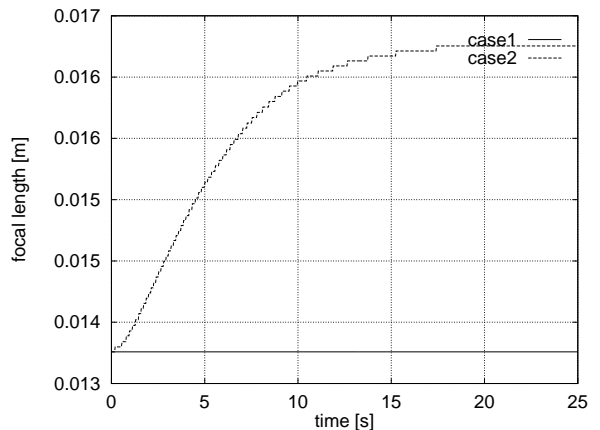
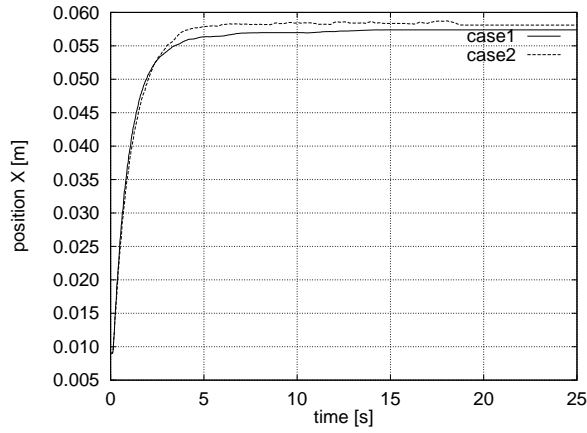
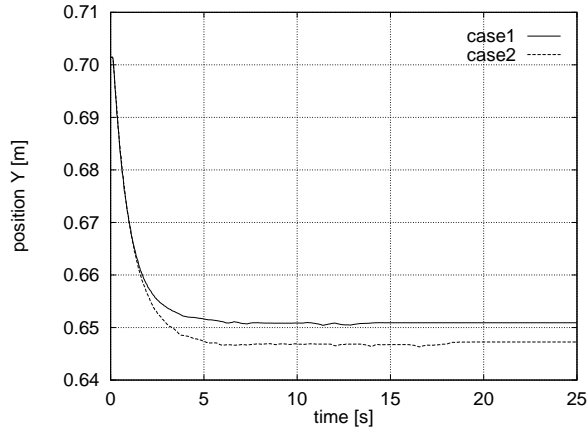


Figure 7: Result 2 (step response, focal length)

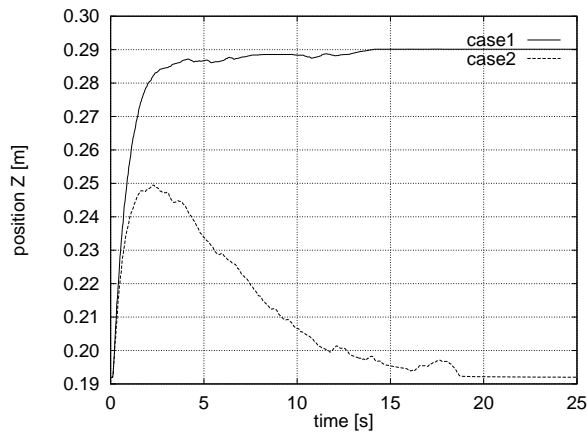
Next, we demonstrate how the zoom mechanism compensate the range of the arm mechanism. We give a desired image features so that the arm mechanism by itself cannot realize. Results are shown in figure 9. With the zoom mechanism the whole system can eliminate the step error, while without the zoom mechanism it stops at around 3[s] because the arm is near a singular posture. From these results, we can show that when change of desired images is large, the zoom mechanism can compensate the physical limitation of the arm mechanism, while the arm by itself cannot



(a) position ${}^w p_x$



(b) position ${}^w p_y$



(c) position ${}^w p_z$

Figure 8: Result 3 (step response, position of tip of robot arm)

eliminate the error.

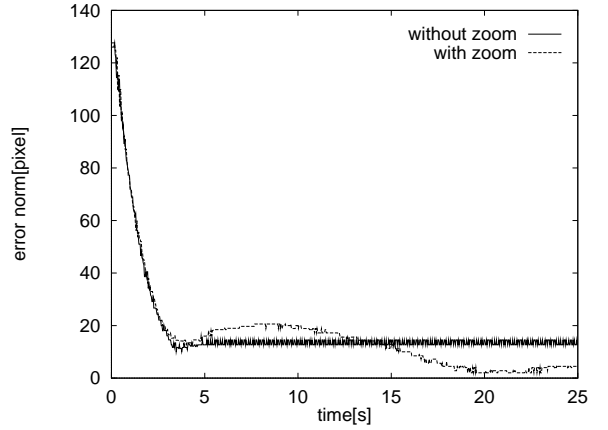


Figure 9: Result 4 (step response, error norm)

5 Discussion and future works

In this paper, a complementary visual servoing controller of zoom and arm mechanisms has been proposed. We have discussed on the condition that camera position and zoom setting are considered as redundant, and on a complementary visual servoing controller.

In this paper we assume that the image jacobian is known. But in the real environment, the image jacobian is generally unknown. To make the system adaptive to environments, we have to think out the estimator of the image jacobian[12].

As mentioned in the introduction, utilizing a zoom mechanism can change sensitivity of image features w. r. t. manipulator movement, resolvability, and can compensate the limit of the physical constraint of the manipulator mechanism. These are partly realized in this paper, at the control level, and can be extended to the phase of trajectory planning and higher layer of the vision-based controller.

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] B. Nelson, N. P. Papanikolopoulos, and P. K. Khosla. Visual servoing for robotic assembly. In *Visual Servoing*, pages 139–164. World Scientific, 1993.
- [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] 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.
- [6] 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.
- [7] 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.
- [8] N. P. Papanikolopoulos and P. K. Khosla. Adaptive robotic visual tracking: Theory and experiments. *IEEE Trans. on Automatic Control*, 38(3):429–445, 1993.
- [9] 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.
- [10] B. Nelson and P. K. Khosla. Integrating sensor placement and visual tracking strategies. In *Proc. of IEEE Int. Conf. on Robotics and Automation*, pages 1351–1356, 1994.
- [11] 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.
- [12] K. Hosoda and M. Asada. Versatile visual servoing without knowledge of true jacobian. In *Proc. of the 1994 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 186–193, 1994.