

# Correspondence

## Task-Independent Robotic Uncalibrated Hand-Eye Coordination Based on the Extended State Observer

Jianbo Su, Hongyu Ma, Wenbin Qiu, and Yugeng Xi

**Abstract**—This paper proposes a standard method to approach the uncalibrated robotic hand-eye coordination problem that is system configuration- and task-independent. The unknown hand-eye relationship is first modeled as the modeling errors of a dynamic system. An extended state observer is then implemented to estimate summation of the system's modeling error and the system's external disturbances. With the estimation results as the compensation, the system control is accomplished from a nonlinear combination of the system state errors. A universal framework of controller design is provided for decoupled and coupled hand-eye systems of different configurations to execute dynamic tracking task.

**Index Terms**—Calibration-free, dynamic tracking, hand/eye coordination, state observer.

### I. INTRODUCTION

The uncalibrated hand-eye coordination technique is not only essential for a robotic hand-eye system to transplant into applications, but also a basic property worth of being integrated into many robotic systems, e.g., humanoid robot and mobile robot, to make them more adaptive and intelligent. So it attracts more and more attentions in recent years that many promising approaches have been worked out. Most of the methods toward the uncalibrated hand-eye coordination problem can be classified into two categories. One is the image Jacobian matrix-based approach [1], [2], in which the unknown static and nonlinear relations between the robot (hand) and visual sensors (eye) are approximated by the linear image Jacobian matrix that is dynamic and should be estimated iteratively online. The other is the artificial neural network (ANN)-based approach, in which the hand-eye relations are modeled by a neural network [6], and/or the coordination controllers are realized in a neural network way [8].

For the image Jacobian matrix based approaches, it has been shown in the literatures that how to estimate the image Jacobian matrix is the key to the performance of these methods [15]. Currently, the estimation algorithms developed so far rely on system configuration and specific tasks. Data accumulations for estimation are normally redundant procedures to task fulfillment that increases system expenses [3]. The estimation accuracy is related to the position of the robot manipulator in its workspace and the camera's visual field, and thus no consistent system performance can be achieved [4]. Moreover, since the estimation of the image Jacobian matrix is an iterative online procedure, it gives rise to some problems [5], [16] such as time delay, computation singularity and convergence ability, etc. These problems need to be addressed when the estimation algorithms and procedures are designed. And they become especially serious in dynamic circumstances.

In the ANN-based approaches, the hand-eye relationship model is trained off-line, no matter whether the model is explicitly described

[7] or not [8], and used as a part of online control. However, since the off-line training of the ANN needs to gather a large set of training samples, feasibility and applicability of this approach is restricted in practice.

In this paper, we explore a general framework to address the uncalibrated hand-eye coordination problem. In classical control theory, a lot of strategies have been developed to deal with a system with uncertainty. One of the solutions is to estimate the system uncertainty by a state observer [11], [12], and then compensate it in system control. Though the formulation and the structure of the state observer are closely related to the system model, there exist general theories and standard procedures to instruct the design of it [10]. Thus if the unknown hand-eye relationship can be modeled as a system state and then estimated by an online system state observer, it is not related to robotic hand-eye system configuration and specific tasks to be executed. This will result in a system configuration- and task-independent method to deal with the unknown hand-eye relations.

The idea is to be realized by involving the extended state observer (ESO) to estimate the system's unmodeled dynamics and external disturbances. The ESO was first proposed in [10] based on the nonlinear (nonsmooth) continuous structure of the output error. It is a novel observer for a class of uncertain systems, which has found successful applications in attitude control of aircraft, magnetic suspension control [9], and motion/force control [18]. This paper takes advantages of the ESO to deal with the unknown hand-eye relations for the first time in order to develop a universal way to address the uncalibrated hand-eye coordination problem, irrespective of system configurations and specific tasks to be fulfilled.

The strategy proposed in this paper has no restrictions on system modeling so long as a dynamics system can be obtained to describe the hand-eye coordination procedure. Here we adopt the image Jacobian matrix model to discuss the uncalibrated hand-eye coordination since it has been proven to be an efficient tool [17]. The unknown time- and spatially varying image Jacobian matrix is estimated by an ESO together with the system's external disturbances. A nonlinear controller is then designed by using estimation of the state observer as compensation [11], so that the calibration-free robotic hand-eye coordination is achieved. This approach offers a task-free strategy for the estimation of the Jacobian matrix, which effectively solves the above problems.

Section II presents preliminaries of the ESO. Section III analyzes the nonlinear visual mapping models for the uncalibrated hand-eye coordination problem under two kinds of system configurations, i.e. hand-eye coordination under monocular visual feedback and stereo visual feedback. These models are used as the bases in Section IV to design the corresponding ESOs and controllers. Although the resultant ESOs and controllers for the two systems are different in formations from each other due to system orders and complexities, design procedures are similar, thus general enough to be standard ones. Experiments are provided in Section V to demonstrate the effectiveness and performance of this scheme, followed by the Conclusion in Section VI.

### II. PRELIMINARIES

The ESO is a kind of state observer that tracks different orders of the state variables of the system and estimates the unmodeled dynamics and external disturbance of the system [10]. Thus, it is the key to controlling a system with uncertainties. Assume a second-order nonlinear

Manuscript received May 20, 2003; revised November 24, 2003. This work was supported by the National Natural Science Foundation of China under Grant 69875010. This paper was recommended by Associate Editor H. Qiao.

The authors are with the Department of Automation and Research Center of Intelligent Robotics, Shanghai Jiaotong University, Shanghai, 200030, China (e-mail: jbsu@sjtu.edu.cn).

Digital Object Identifier 10.1109/TSMCB.2004.827615

system with an uncertainty, which suffers from some unknown external disturbances

$$\ddot{x} = f(x, \dot{x}, t) + w(t) + b_0 u(t) \quad (1)$$

where  $f(x, \dot{x}, t)$  is an unknown function,  $w(t)$  is the unknown external disturbance,  $u(t)$  is the control input, and  $b_0$  is a known constant. Let

$$\begin{cases} x_1(t) = x(t) \\ x_2(t) = \dot{x}(t) \\ x_3(t) = f(x, \dot{x}, t) + w(t) \end{cases} \quad (2)$$

where an extended state  $x_3(t)$  is formed by the summation of the unknown function  $f(x, \dot{x}, t)$  and the system's unknown external disturbances  $w(t)$ . Then (1) can be transformed to be

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = x_3(t) + b_0 u(t) \\ \dot{x}_3(t) = \xi(t) \end{cases} \quad (3)$$

where  $\xi(t)$  is an unknown function. Construct a nonlinear system

$$\begin{cases} \dot{z}_1(t) = z_2(t) - g_1(e_1(t)) \\ \dot{z}_2(t) = z_3(t) - g_2(e_1(t)) + b_0 u(t) \\ \dot{z}_3(t) = -g_3(e_1(t)) \end{cases} \quad (4)$$

where  $e_1(t) = z_1(t) - x_1(t)$ ,  $g_i(e_1(t))$  ( $i = 1, 2, 3$ ) is a set of suitably constructed nonlinear continuous functions satisfying

$$e_1 g_i(e_1) > 0, \quad \forall e \neq 0, \quad \text{and} \quad g_i(0) = 0, \quad (i = 1, 2, 3). \quad (5)$$

Then from (3) and (4), we have

$$\begin{cases} \dot{e}_1(t) = e_2(t) - g_1(e_1(t)) \\ \dot{e}_2(t) = e_3(t) - g_2(e_1(t)) \\ \dot{e}_3(t) = -\xi(t) - g_3(e_1(t)) \end{cases} \quad (6)$$

where

$$e_i(t) = z_i(t) - x_i(t), \quad (i = 1, 2, 3). \quad (7)$$

It is proven [9] that with appropriate selections of the set of functions  $g_i(e_1)$  ( $i = 1, 2, 3$ ), system (6) converges to the origin as time goes to infinity, i.e.,  $e_1(t) \rightarrow 0$  ( $i = 1, 2, 3$ ) as  $t \rightarrow \infty$ . Therefore, the states of system (4) can track the corresponding states of system (3), i.e.,

$$z_1(t) \rightarrow x_1(t), \quad z_2(t) \rightarrow x_2(t), \quad z_3(t) \rightarrow x_3(t). \quad (8)$$

According to the definition in (2),  $x_3(t)$  is the summation of the unknown function and the external disturbance in system (1). Thus, although  $f(x, \dot{x}, t)$  and  $w(t)$  in (1) are unknown, the extended state  $z_3$  in (4) can still have real-time estimation for their summation  $x_3(t)$ . System (4) is therefore called the ESO of (1). From (4), we know that the ESO is of one order higher than that of the nonlinear system.

### III. VISUAL MAPPING MODEL

In image-based visual servoing, the task of the robotic hand-eye coordination is to design a robot control so that the target and the hand become coincident with each other in the image according to the error between them observed in the image. Suppose that the hand position is  $W$  in the robotic coordinate system and is  $P$  in the image(s) observed by the camera(s). The target position in the image(s) is  $P^*$ , which is also the desired hand position in the image. The relation between the hand position in the image and that in the robotic coordinate system can be expressed as

$$P = g(W) \quad (9)$$

where  $g(\cdot)$  is a function representing all the effects caused by the eye-hand relationship model, the robotic manipulator model and the camera model. Differentiation of both sides of (9) leads to

$$\begin{cases} \dot{W} = U \\ \dot{P} = J(W) \cdot U \end{cases} \quad (10)$$

where  $U$  is the velocity vector of the robot hand in the robotic coordinate system, which is the system control input and  $J(W)$  is the Jacobian matrix of  $g(W)$ . Equation (10) is a general form to describe the differential change of the hand position in the image(s) caused by the differential motion of the hand in the robotic coordinate system. Without loss of generality, only translational motion of the robot hand in the robotic coordinate system is considered for simplicity hereafter. Thus,  $U$  can be stated as  $U = (u_x, u_y, u_z)^T$ .

If we take a monocular global visual feedback as an example to analyze the visual mapping model of the system,  $P$  is a two-dimensional (2-D) vector. In this case, the Jacobian matrix defined in (10) can be expressed as

$$J(W) = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \end{bmatrix}. \quad (11)$$

Let  $P = (p_x, p_y)^T$ , we have

$$\dot{P} = \begin{bmatrix} \dot{p}_x \\ \dot{p}_y \end{bmatrix} = J(W)U = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \end{bmatrix} \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}. \quad (12)$$

That is

$$\begin{cases} \dot{p}_x = J_{11} \cdot u_x + J_{12} \cdot u_y + J_{13} \cdot u_z \\ \dot{p}_y = J_{21} \cdot u_x + J_{22} \cdot u_y + J_{23} \cdot u_z \end{cases} \quad (13)$$

If the robotic hand-eye coordination system has visual feedback from a stereovision system,  $P$  must be a four-dimensional (4-D) vector formed by stacking the hand positions described in two image frames, i.e.,  $P = (p_x^1, p_y^1, p_x^2, p_y^2)^T$ . In this case, the Jacobian matrix defined in (10) has the form

$$J(W) = \begin{bmatrix} J^1 \\ J^2 \end{bmatrix} = \begin{bmatrix} J_{11}^1 & J_{12}^1 & J_{13}^1 \\ J_{21}^1 & J_{22}^1 & J_{23}^1 \\ J_{11}^2 & J_{12}^2 & J_{13}^2 \\ J_{21}^2 & J_{22}^2 & J_{23}^2 \end{bmatrix}. \quad (14)$$

Thus, the visual mapping model can be described as

$$\dot{P} = \begin{bmatrix} \dot{p}_x^1 \\ \dot{p}_y^1 \\ \dot{p}_x^2 \\ \dot{p}_y^2 \end{bmatrix} = J(W)U = \begin{bmatrix} J_{11}^1 & J_{12}^1 & J_{13}^1 \\ J_{21}^1 & J_{22}^1 & J_{23}^1 \\ J_{11}^2 & J_{12}^2 & J_{13}^2 \\ J_{21}^2 & J_{22}^2 & J_{23}^2 \end{bmatrix} \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}. \quad (15)$$

That is

$$\begin{cases} \dot{p}_x^1 = J_{11}^1 u_x + J_{12}^1 u_y + J_{13}^1 u_z \\ \dot{p}_y^1 = J_{21}^1 u_x + J_{22}^1 u_y + J_{23}^1 u_z \\ \dot{p}_x^2 = J_{11}^2 u_x + J_{12}^2 u_y + J_{13}^2 u_z \\ \dot{p}_y^2 = J_{21}^2 u_x + J_{22}^2 u_y + J_{23}^2 u_z \end{cases} \quad (16)$$

Equations (13) and (16) give two examples of the visual mapping model from the general form of (10). It is easy to see that although the formations of (13) and (16) are different due to system configurations, the modeling procedures are similar to each other. Thus, a general modeling procedure is reached for the uncalibrated hand-eye coordination problem. The following section will discuss the system state observer and controller based on system models discussed above.

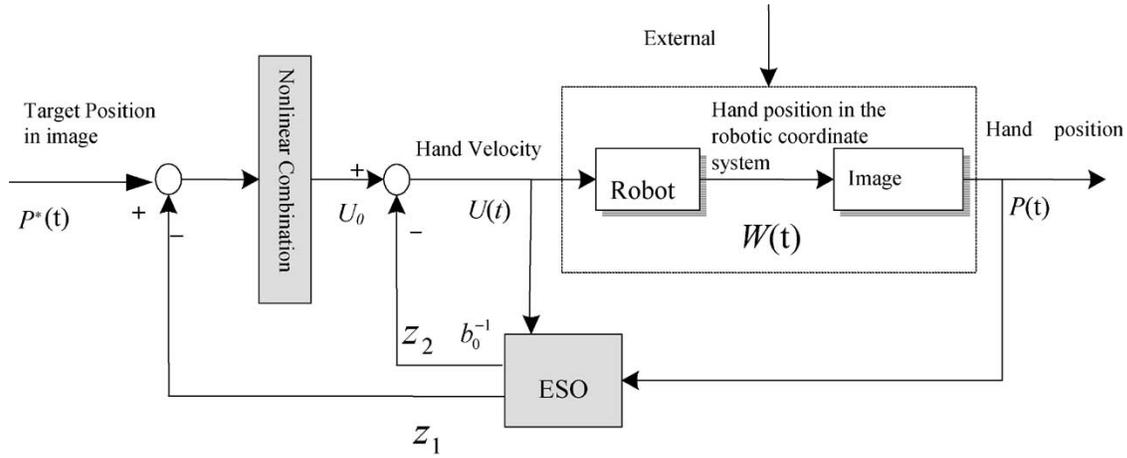


Fig. 1. The ESO-based control structure.

#### IV. THE ESO AND CONTROLLER DESIGN

Since the essence of image Jacobian matrix based approach is to design an online estimation for the image Jacobian matrix, lengthy try-and-error or iterative procedures are used. The former increases the cost of the system and works only for static or slowly varying system, while the latter has time delay, singularity, and convergence problems. Their common drawbacks are that both of them are related to particular tasks and lack general design rules. Here we use the ESO to estimate the system's unmodeled dynamics and the external disturbance. As shown in Section II, design of the ESO is irrelevant to system configuration or a particular task. Thus, a general way to address the unknown hand-eye relations may result.

We first consider the simpler case of system (13) of single-eye visual feedback to show the design procedure of the observer and the controller. We should rewrite the visual mapping model of (13) to be a form suitable for the ESO. Suppose, in the robotic workspace, a rough but a reasonable estimate of  $J(W)$  is

$$\hat{J}(W) = \begin{bmatrix} \hat{J}_{11} & \hat{J}_{12} & \hat{J}_{13} \\ \hat{J}_{21} & \hat{J}_{22} & \hat{J}_{23} \end{bmatrix}. \quad (17)$$

This can be obtained empirically from prior knowledge of the system. Here, we also invoke  $w_1(t)$  and  $w_2(t)$  to describe system disturbances, which include the system modeling inaccuracy, image processing errors and the external disturbances. Then (13) is rewritten as

$$\begin{cases} \dot{p}_x = (J_{11} - \hat{J}_{11}) \cdot u_x + J_{12} \cdot u_y + J_{13} \cdot u_z + w_1(t) + \hat{J}_{11} \cdot u_x \\ \dot{p}_y = J_{21} \cdot u_x + (J_{22} - \hat{J}_{22}) \cdot u_y + J_{23} \cdot u_z + w_2(t) + \hat{J}_{22} \cdot u_y \end{cases}. \quad (18)$$

Supposing

$$\begin{cases} a_x(t) = (J_{11} - \hat{J}_{11}) \cdot u_x + J_{12} \cdot u_y + J_{13} \cdot u_z + w_1(t) \\ a_y(t) = J_{21} \cdot u_x + (J_{22} - \hat{J}_{22}) \cdot u_y + J_{23} \cdot u_z + w_2(t) \end{cases} \quad (19)$$

and substituting (19) into (18), we have

$$\begin{cases} \dot{p}_x = a_x(t) + \hat{J}_{11} \cdot u_x \\ \dot{p}_y = a_y(t) + \hat{J}_{22} \cdot u_y \end{cases}. \quad (20)$$

Thus, the original system is decoupled into two first-order subsystems in  $x$  and  $y$  directions. Note that each of the equations in (20) has a similar form to that in (1), where  $a_x(t)$  and  $a_y(t)$  are respectively the total effects of the system's unmodeled dynamics and the external disturbances in the  $x$  and  $y$  directions thus can be estimated by ESOs.  $\hat{J}_{11}$  and  $\hat{J}_{22}$  are the estimations for  $J_{11}$  and  $J_{22}$  respectively, which are both similar to  $b_0$  in (1).  $u_x$  and  $u_y$  are the system controls for each of the two subsystems, respectively.

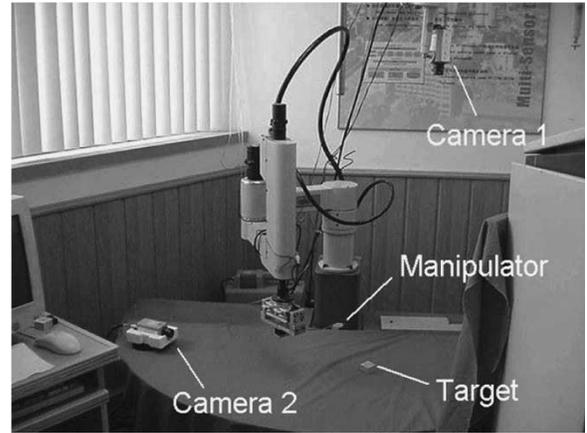


Fig. 2. Experimental test-bed configuration.

TABLE I  
ESO AND CONTROLLER PARAMETERS  
USED IN SIMULATIONS

ESO			Controller			Est. of $b_0$				
$\alpha_1$	$\delta_1$	$b_1$	$\alpha_2$	$\delta_2$	$b_2$	$\alpha$	$\delta$	$k_0$	$\hat{J}_{11}$	$\hat{J}_{22}$
0.5	1.5	10	0.5	0.9	10	0.5	1.5	30	0.7	0.9

For the two decoupled subsystems, it is straightforward to design two controllers for the  $x$  and  $y$  subsystems, respectively. Here, control in the  $x$  direction is taken as the example to demonstrate the controller design procedure. We formulate the  $x$  subsystem in (20) to a standard dynamic system for consequent discussions as

$$\begin{cases} \dot{w}_x = u_x \\ \dot{p}_x = a_x(t) + \hat{J}_{11} u_x \\ y_1 = p_x \end{cases} \quad (21)$$

where  $w_x$  is an instrumental state of the system and  $y_1$  is the system output. It is seen that the system defined by (21) is a first-order system. According to (4) and (5), a second-order ESO should be employed to estimate the uncertainty and external disturbance of the system (21)

$$\begin{cases} \dot{z}_{x1} = z_{x2} - b_{x1} \text{fal}(z_{x1} - y_1, \alpha_{x1}, \delta_{x1}) + \hat{J}_{11} u_x \\ \dot{z}_{x2} = -b_{x2} \text{fal}(z_{x1} - y_1, \alpha_{x2}, \delta_{x2}) \end{cases} \quad (22)$$

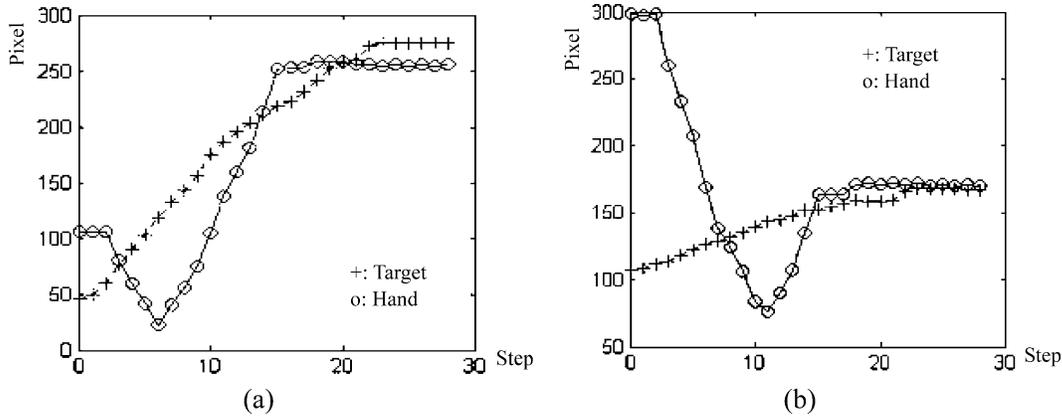


Fig. 3. Tracking process in  $x$  and  $y$  directions. (a) Control of the  $x$  subsystem. (b) Control of the  $y$  subsystem.

where  $b_{x1}$  and  $b_{x2}$  are parameters to be chosen, and

$$fal(\varepsilon, \alpha, \delta) = \begin{cases} |\varepsilon|^\alpha \text{sign}(\varepsilon) & |\varepsilon| > \delta \\ \frac{\varepsilon}{\delta^{1-\alpha}} & |\varepsilon| \leq \delta \end{cases}, \quad 0 < \alpha < 1, \delta > 0. \quad (23)$$

Here we choose the nonlinear function  $fal(\varepsilon, \alpha, \delta)$  as an example to formalize  $g_i(e_1(t))$  ( $i = 1, 2$ ) in (4), since it basically satisfies the necessary conditions for choosing  $g_i(e_1(t))$  shown in (5).

In (22),  $z_{x1}$  gives an estimation for  $y_1$ , thus  $p_x$ , which is the state of system (21), whereas  $z_{x2}$  gives an estimation for  $a_x(t)$  which is the total effect of the system's unmodeled dynamics and external disturbances of (21). These estimations are used as compensations to construct system controller. If  $p_x^*(t)$  is the system input, define the system's tracking error as

$$e_x = p_x^*(t) - z_{x1} \quad (24)$$

then the system control can be obtained by the following nonlinear state error feedback control law:

$$\begin{cases} u_{x0} = k_{x0} fal(e_x, \alpha_x, \delta_x) \\ u_x = u_{x0} - \frac{z_{x2}}{J_{11}} \end{cases} \quad (25)$$

where  $k_{x0}$  is a parameter to control the dynamic performance of the system. Here, we choose the nonlinear function  $fal(e_x, \alpha_x, \delta_x)$  of the system error  $e_x$  to form the system control because it provides a small linear region near  $e_x = 0$  so that no excessive gain that might lead to high-frequency chattering occur [13]. Thus, controller of system (21) designed based on ESO for target tracking in the  $x$  subsystem is given by (22)–(25). Similarly, the controller in the  $y$  subsystem can be obtained. The overall control diagram is given in Fig. 1, where  $z_i = [z_{xi}, z_{yi}]$ , ( $i = 1, 2$ ),  $U_o = [u_{x0}, u_{y0}]$ ,  $U = [u_x, u_y]$ ,  $b_o = \text{diag}[\hat{J}_{11}, \hat{J}_{22}]$ .

A similar procedure can be traced to design the ESO and the controller for system (16). Since (16) is much more complex than (13), coupling between each subsystem would rather not be neglected. If a reasonable estimation for the  $J$  matrix in (15) is  $\hat{J}$ , a vector-form ESO is adopted as follows:

$$\begin{cases} Z_1(t) = Z_2(t) - B_1 \bullet fal(Z_1(t) - P(t), \alpha_1, \delta_1) + \hat{J} \cdot U(t) \\ Z_2(t) = -B_2 \bullet fal(Z_1(t) - P(t), \alpha_2, \delta_2) \end{cases} \quad (26)$$

where “ $\bullet$ ” stands for multiplication operations of the corresponding elements of two vectors. Suppose  $P^*(t)$  is the system input, we have the system controller

$$\begin{cases} E(t) = P^*(t) - Z_1(t) \\ U_o(t) = K \bullet fal(E(t), \alpha_3, \delta_3) \\ U(t) = U_o(t) - (\hat{J}^T \cdot \hat{J})^{-1} \cdot \hat{J}^T \cdot Z_2(t) \end{cases}. \quad (27)$$

The controller defined by (26) and (27) has a similar structure to that shown in Fig. 1, except that all variables are in vector forms.

## V. EXPERIMENTS

An Adept 604 S robotic manipulator is used to construct a robotic hand-eye coordination system, as shown in Fig. 2. Two cameras are fixed to observe the situations of the robot hand and the target in the robot workspace. Dynamic tracking task is conducted in experiments with the uncalibrated hand-eye coordination systems of different configurations: 1) monocular global visual feedback and 2) stereo visual feedback. These two system configurations are modeled by (13) and (16), thus controls given by (22)–(25) and (26), (27) are, respectively, adopted here in experiments. The image size is  $320 \times 320$  pixels. The robot hand and the target are respectively identified by a red and a green color block in order to simplify the image processing and object recognition tasks. In the experiments, the visual sampling period is 0.4 s.

### A. Monocular Global Visual Feedback

The target is moving in a 2-D working plane in a speed of 30 mm/s. Parameters for ESO and the controller are selected as shown in Table I. Since the same parameters in the  $x$  and  $y$  subsystems are used, the subscripts  $x$  in (22), (24), (25) are omitted. Since only a single eye is used, the robot hand is driven to approach the target in a 2-D plane parallel to the working plane.

Fig. 3(a) and (b) shows the system controls for the  $x$  subsystem and the  $y$  subsystem, respectively. Fig. 4 shows the tracking process observed in image plane. To evaluate only the controller's performance, no prediction for the target motions or path planning for hand tracking is included in the tracking procedure. The system response has a 3-step delay at the beginning of the task due to image processing. The task is finished at the 20th step where the tracking errors are about three pixels in  $x$  direction and 8 pixels in  $y$  direction, which are both less than predefined thresholds. From the figures, we can see that system control can successfully drive the hand to the target and finally suppress the system errors to a predefined threshold. But there are also overshoots in both the  $x$  and  $y$  subsystem controls. These may be overcome by better selections of control parameters or a proper planning for the tracking path during control.

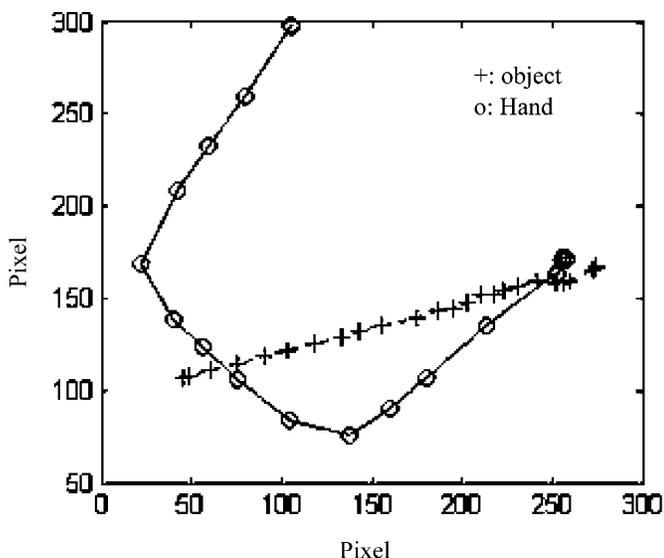


Fig. 4. Tracking process in the image plane.

TABLE II  
SELECTIONS OF ESO AND CONTROLLER PARAMETERS

ESO						Controller		
$\alpha_1$	$\delta_1$	$B_1$	$\alpha_2$	$\delta_2$	$B_2$	$\alpha_3$	$\delta_3$	$K$
0.5	3	[10 10 10] <sup>T</sup>	0.5	5	[20 20 20] <sup>T</sup>	0.5	5	[12 12 12] <sup>T</sup>

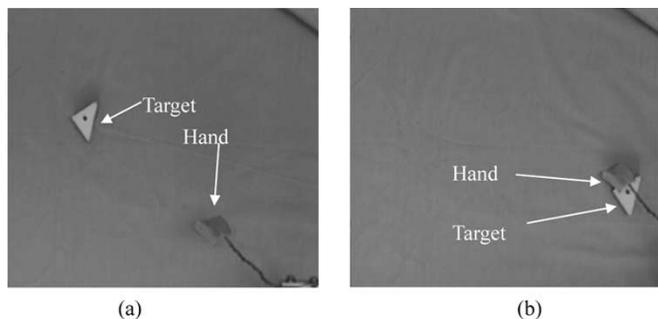


Fig. 5. Visual tracking observed by the first camera. (a) At the beginning of the task. (b) At the end of the task.

**B. Stereo Visual Feedback**

In this experiment, the target moves in a 2-D plane in a speed of about 20 mm/s. All the parameters for the ESOs and the tracking controllers in  $x$  and  $y$  directions are selected to be the same, which are shown in Table II. These parameters are basically selected empirically at present. Relations of the parameters to the system performance are still under investigation [14]. The initial estimation of the image Jacobian matrix is set as

$$\hat{J} = \begin{bmatrix} 0.7 & -0.2 & 0 \\ 0.2 & 0.9 & -0.5 \\ 0.7 & 0.1 & 0 \\ -0.1 & 0.9 & 0.5 \end{bmatrix}$$

Since a stereovision feedback is available, a 3-D translational movement for the robot hand can be achieved to approach the target. Figs. 5 and 6 show the scenes captured by the two cameras at the beginning

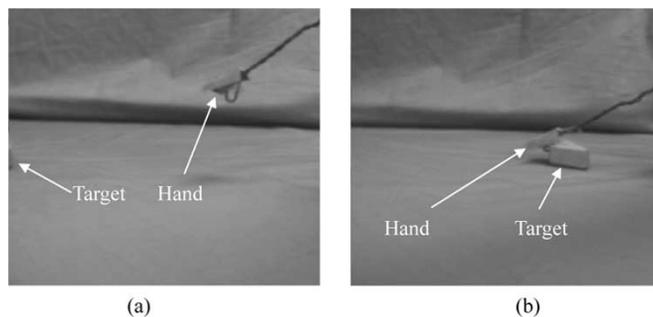


Fig. 6. Visual tracking observed by the second camera. (a) At the beginning of the task. (b) At the end of the task.

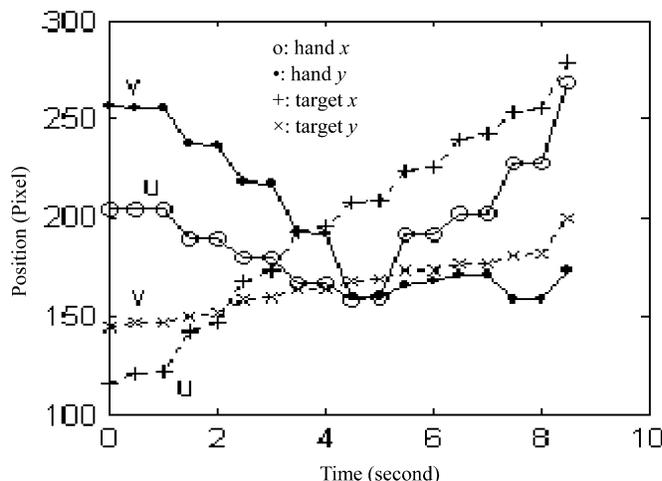


Fig. 7. Hand and target positions captured by camera 1.

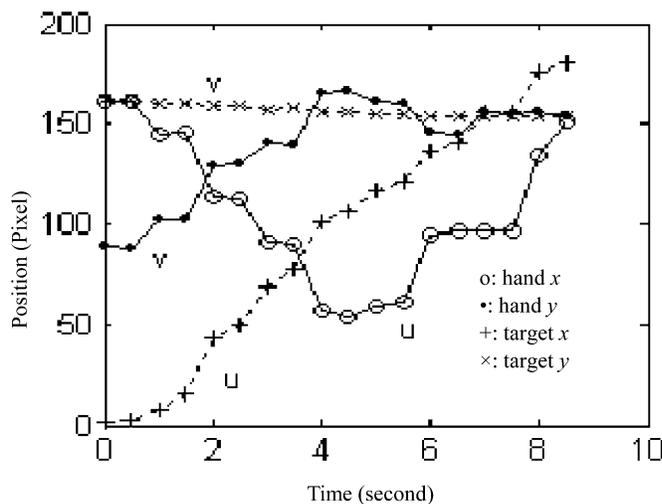


Fig. 8. Hand and target positions captured by camera 2.

and end of the task. From Fig. 5(b) and Fig. 6(b), we can see that the robot hand has already touched the object at the end of the task. Figs. 7 and 8 show the position variations of both the target and hand in the two image planes. The task is done at 8.5 s with the position error between the target and the hand in each image plane of about 28 pixels, which are mostly due to the size of their representative color blocks. Thus, from the experiments, we can confirm the feasibility and effectiveness of the ESO-based approach for uncalibrated robotic hand-eye coordination system.

## VI. CONCLUSIONS

A universal framework is proposed in this paper toward the uncalibrated robotic hand-eye coordination problem. The unknown hand-eye relationship and camera parameters are considered as unmodeled system dynamics. Together with the system's external disturbances, the unmodeled system dynamics are estimated online by an extended state observer, and a nonlinear controller is thus designed based on the observer's compensation. Typical cases are analyzed to show that designs of the observer and the controller are independent of specific tasks and system configurations and thus have general meanings. Simulations and experiments demonstrate that the proposed controller can suppress the effects of the external disturbance, and therefore has a strong adaptability and robustness. Though the image Jacobian matrix model is exemplified to obtain the system models and the control is discussed upon it, it is obvious that the methodology is far beyond it, so long as a dynamic system could be obtained to describe the uncalibrated hand-eye coordination problem.

Although the ESO presents a base for a kind of control design theory, which is independent of system model and external disturbance, its superiority over the conventional approaches has not been well explored due to its own unsolved problems such as the parameter and nonlinear function selections. However, ESO-based control theory has been demonstrated to successfully apply to a task-free design in the uncalibrated robotic hand-eye coordination control and does offer a new way of thinking for this problem. It is believed that along with the development of the ESO and the nonlinear control theory itself, its application in the uncalibrated robotic hand-eye coordination control will surely be further acknowledged.

## REFERENCES

- [1] G. D. Hager, W. C. Chang, and A. S. Morse, "Robot feedback control based on stereo vision: toward calibration-free hand/eye coordination," in *Proc. IEEE Int. Conf. Robotics and Automation*, 1994, pp. 2850–2856.
- [2] B. H. Yoshimi and P. K. Allen, "Alignment using an uncalibrated camera system," *IEEE Trans. Robot. Automat.*, vol. 11, pp. 516–521, Aug. 1995.
- [3] J.-B. Su and Y.-G. Xi, "Image tracking for a 3-D moving object based on uncalibrated global visual feedback," *High Technol. Lett.*, vol. 10, no. 7, pp. 85–87, 2000.
- [4] H. Sutanto, R. Sharma, and V. Varma, "Image based autodocking without calibration," in *Proc. IEEE Int. Conf. Robotics and Automation*, 1997, pp. 974–979.
- [5] L. Hsu and P. L. S. Aquino, "Adaptive visual tracking with uncertain manipulator dynamics and uncalibrated camera," in *Proc. 38th Conf. Decision and Control*, 1999, pp. 1248–1253.
- [6] H. Hashimoto, T. Kubota, M. Sato, and F. Hurashima, "Visual control of robotic manipulator based on neural networks," *IEEE Trans. Ind. Electron.*, vol. 39, pp. 490–496, Nov./Dec. 1992.
- [7] K. Stanley, Q. M. J. Wu, A. Jerbi, and W. A. Gruver, "Neural network-based vision guided robotics," in *Proc. 1999 IEEE Int. Conf. Robotics and Automation*, 1999, pp. 281–286.
- [8] J. Su, Y. Xi, U. Hanebeck, and G. Schmidt, "Nonlinear visual mapping model for 3-D visual tracking with uncalibrated eye-in-hand robotic system," *IEEE Trans. Syst., Man, Cybern. B*, vol. 34, pp. 652–659, Feb. 2004.
- [9] H. Yi and H. Jingqing, "Analysis and design for the second order nonlinear continuous extended states observer," *Chinese Sci. Bull.*, vol. 45, no. 21, pp. 1938–1944, 2000.
- [10] J.-Q. Han, "The extended state observer of a class of uncertain systems," *Control Decis.*, vol. 10, no. 1, pp. 85–88, 1995.
- [11] M. Corless and J. TU, "State/input estimation for a class of uncertain systems," *Automatica*, vol. 34, no. 6, pp. 757–764, 1998.
- [12] M. Darouach, M. Zasadzinski, and S. J. Xu, "Full-order observers for linear systems with uncertain inputs," *IEEE Trans. Automat. Cont.*, vol. 39, pp. 606–609, Mar. 1994.
- [13] Z. Gao *et al.*, "An alternative paradigm for control system design," in *Proc. 40th IEEE Int. Conf. Decision and Control*, vol. 5, 2001, pp. 4578–4585.
- [14] H. Ma and J. SU, "Research about adjusting the parameters of auto disturbance rejection controller," *Elect. Automat.*, vol. 24, pp. 10–13, 2002.
- [15] J. Feddema and C. Lee, "Adaptive image feature prediction and control for visual tracking with a hand-eye coordinated camera," *IEEE Trans. Syst., Man, Cybern.*, vol. 20, pp. 1172–1183, Sept./Oct. 1990.
- [16] T. Drummond and R. Cipolla, "Real-time tracking of complex structures with on-line camera calibration," *Image Vis. Comput.*, vol. 20, pp. 427–433, 2002.
- [17] C. Colombo and B. Allotta, "Image-based robot task planning and control using a compact visual representation," *IEEE Trans. Syst., Man, Cybern. A*, vol. 29, pp. 92–99, Jan. 1999.
- [18] Y. Huang, M. Svinin, Z. Luo, and S. Hosoe, "A design of an extended state observer for the motion/force control of constrained robotic systems," in *Proc. ICASE Workshop on Intelligent Control and Systems*, Oct. 2002, pp. 114–119.