Calibration-Free Robotic Eye–Hand Coordination Based on an Auto Disturbance-Rejection Controller

Jianbo Su, Wenbin Qiu, Hongyu Ma, and Peng-Yung Woo

Abstract—This paper addresses the calibration-free robotic eye-hand coordination in a way other than the conventional image Jacobian matrix approach that has been studied extensively in literature. A nonlinear mapping rather than the linear mapping used in the image Jacobian matrix between the image space and the robotic control space is proposed. This mapping is regarded as the system's unmodeled dynamics expressed in system state equations. An extended state observer is designed first to estimate the unmodeled dynamics as well as the external disturbance of the system. With the estimation results as the compensation, a system controller is designed based on the nonlinear state-error feedback control strategy. Convergence of the extended state observer as well as the overall controller for a typical eye-hand coordination system is proved. Compared with the conventional calibration-free robotic eye-hand coordination with a Jacobian matrix, the proposed controller is independent of specific tasks and system configurations. Thus, a general design procedure is proposed for the calibration-free robotic eye-hand coordination. Simulation and experiment results demonstrate the satisfactory performance and effectiveness of the proposed approach.

Index Terms—Auto disturbance-rejection controller (ADRC), calibration-free, eye–hand coordination, robotic vision.

I. INTRODUCTION

The essence of the calibration-free robotic eye–hand coordination is to use the visual information fed from the camera to plan the robotic control for a specific task without any knowledge or with little knowledge of the eye–hand relationship and the camera model [1]. Due to the great potential of the calibration-free robotic eye–hand coordination technology, much research has been done in this area. The basic and the most investigated are the image Jacobian matrix-based approaches [2]. Since a linear Jacobian matrix is used to describe approximately the nonlinear mapping between the position error in the image space and the robotic control, the Jacobian matrix is changing and needs an online real-time estimation in the whole workspace during the whole task [3], [4]. Although in some cases the changes of the image Jacobian matrix do not affect the task process much so that a constant image Jacobian matrix is acceptable [5], [6], an accurate estimation of the Jacobian matrix online, evidently, is the key to the effectiveness of this kind of approach [7].

The estimation of the image Jacobian matrix, in some applications, is based on an online calibration of the eye–hand relationship [8], [9], but it, by far, is based on an observation of the correspondent feature changes of the target in the image plane and robotic workspace. The hand is driven to move in the neighborhood of the robot's present work position with sufficient time to obtain a proper estimation of the image Jacobian matrix at this work position [10]. These exploration movements must be mutually irrelevant to each other to facilitate the estimation algorithm. Since the image Jacobian matrix is closely related to the robot's work positions, this procedure is repeated redundantly during the whole task. Obviously, this strategy is decreasing the system efficacy and inherently prohibitive to dynamic environments.

To improve the system efficacy and extend the application to dynamic environments and tasks, Sutanto *et al.* [4] suggest that the robot movements in fulfilling the task be utilized to estimate the image Jacobian matrix online so that the redundant movements in the estimation process is eliminated. However, exploration movements of the robot at the initial moment are still unavoidable to get an initial estimation of the image Jacobian matrix for numerical iterations. Meanwhile, no mechanism can be invoked to ensure that the sequential robot movements are linearly independent of each other. Hosada *et al.* [11] and Jaegersand *et al.* [12] propose to estimate the increments of the image Jacobian matrix at each instant, which decreases relevance requirements for the sequential robot movements. Qian *et al.* [13] propose to take advantage of the well-known Kalman filter to estimate the image Jacobian matrix by constructing an instrumental system with the elements of the image Jacobian matrix to be the system states. These methods may only overcome the estimation singularities caused by relevant movements of the robot in sequential moments. Since robot movements are basically designed to fulfill specific tasks under specific system configurations, the estimation of the image Jacobian matrix is also related to specific system configurations and tasks. The design of the robot movement at each control moment that satisfies requirements for both task fulfillment and image Jacobian matrix estimation is extremely difficult and sometimes paradoxical.

To reduce the computational complexity of the online estimation of the image Jacobian matrix, the artificial neural network is adopted to map the linear image Jacobian matrix by offline training [14] and use it in online control. Since the image Jacobian matrix describes a temporary linear relationship, the capacity of the neural network to approach a nonlinear function is not sufficiently exploited. In [15], a nonlinear visual mapping model is invented to describe the dynamic relationship between the system error observed by a camera and the system control. The artificial neural network is used to map the nonlinear model and produce control accordingly, which improves the system performance to a great extent. However, in all artificial neural network-based approaches, the neural network must be trained offline before it is used for online control. Thus, enough training data must be accumulated in advance from the whole robotic workspace, which is usually a very tedious procedure [16]. Moreover, the training procedure is normally very time-consuming, and the convergence of the training algorithm should be clarified [17]. These issues prohibit the feasibility of the neural network-based approaches to the calibration-free robotic eye–hand coordination.

Visual servoing has been studied for many years [18]. For calibration-free robotic eye–hand coordination, controller design is to face the unknown and time-varying and spatially varying image Jacobian matrix. Papanikolopoulos *et al.* [19] give an adaptive controller for robotic visual tracking. Hashimoto *et al.* [6] describe visual servoing as a linear constant multiple-input/multiple-output (MIMO) system by considering the image Jacobian matrix as a constant one, so that they adopt a linear quadratic (LQ)-based optimal controller to deal with it. Nonlinear control [20] and robust control [21] theories can also be applied to visual servoing.

In this paper, we propose a new approach to explore the calibration-free robotic eye–hand coordination. It is realized that the unknown, time varying, and spatially varying image Jacobian matrix actually serves as the system's unmodeled dynamics when the eye–hand coordination system is under control. In the conventional control theory, many strategies have been developed so far to deal with unmodeled dynamics. One approach is that the system's unmodeled

J. Su, W. Qiu, and H. Ma are with the Department of Automation, Shanghai Jiaotong University, Shanghai 200030, China (e-mail: jbsu@sjtu.edu.cn).
P.-Y. Woo is with the Department of Electrical Engineering, Northern Illinois University, Dekalb, IL 60115 USA (e-mail: woo@ceet.niu.edu).

dynamics is estimated by a system observer and then is compensated in the system control. Accordingly, an observer is constructed for the eye–hand coordination system, so that the unknown image Jacobian matrix is estimated online [22]. A nonlinear controller is consequently proposed to converge the system error based on the real-time compensation from the system observer [23]. Above all, the system's input signal is first filtered by a tracking differentiator to further improve the system performance [18]. These three parts composed together is named the auto disturbance-rejection controller (ADRC). Since the design and analysis of the state observer and the controller are well developed, this approach offers a standard design procedure for the calibration-free robotic eye–hand coordination. Furthermore, a task-free estimation of the image Jacobian matrix is achieved.

Section II presents the basics of the ADRC, i.e., the nonlinear tracking differentiator, the extended state observer, and the nonlinear controller adopted in this paper. Section III analyzes the nonlinear mapping between the image plane and the robotic workspace. The nonlinear mapping is then transformed to be an appropriate form for the ADRC. A controller design is presented for a single-eye robotic tracking. The stability and convergence of the state observer and the controller are analyzed in Section IV. Simulation and experiment results in Section V demonstrate the effectiveness of this approach. As will be seen, it suppresses successfully the effects of the model uncertainties and system disturbance and therefore has a strong adaptability and robustness.

II. PRELIMINARIES OF THE ADRC

An ADRC is composed of three parts, namely, the nonlinear tracking differentiator (TD), which is used to arrange the transient process of the system, the extended state observer (ESO), which is used for the estimation of the uncertainty and the external disturbance of the system, and the nonlinear state-error feedback (NLSEF), which is used to obtain the control input of the system.

A. Nonlinear TD

The TD is a dynamic system that gives two output signals $x_1(t)$ and $x_2(t)$ for any input signal $v(t)$, where $x_1(t)$ tracks the input signal $v(t)$ and $x_2(t)$ is the differentiation of $x_1(t)$. We have the following theorem [24].

Theorem 1: If any solution of the system

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f(x_1, x_2) \end{cases} \tag{1}$$

satisfies $x_i(t) \to 0 (i = 1, 2)$ when $t \to \infty$, then for any bounded integrable function $v(t)$ and any constant $T > 0$, the solution for the system

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = r^2 f(x_1 - v(t), x_2/r) \end{cases} \tag{2}$$

satisfies

1) $\lim_{r \to \infty} \int_0^T |x_1(t) - v(t)| dt = 0 \, (\forall T > 0)$;
2) when $r \to \infty$, $x_2(t)$ weakly converges to the generalized derivative of $v(t)$.

Hence, when r is sufficiently large in practice, $x_1(t)$ tracks $v(t)$ with certain accuracy and the larger the r, the higher the accuracy.

In accordance with the above theorem, there are many kinds of TD. A second-order TD most in use is

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -r \left[\text{fal}(x_1 - v, \alpha, \delta) + \beta \frac{\text{fal}(x_2, \alpha, \delta)}{r^\alpha} \right] \end{cases} \tag{3}$$

where $0 < \alpha < 1$ and

$$\text{fal}(\varepsilon, \alpha, \delta) = \begin{cases} |\varepsilon|^\alpha \text{sign}(\varepsilon), & |\varepsilon| > \delta \\ |\varepsilon|/\delta^{1-\alpha}, & |\varepsilon| \le \delta. \end{cases} \tag{4}$$

The TD is used to arrange the practical transient process of the system. It also provides the different orders of derivatives of the tracked signal, which are used in the control strategy, which will be discussed later.

B. Extended State Observer (ESO)

The ESO is a new state observer, which tracks the different orders of the state variables of the system and estimates the unmodeled dynamics and external disturbance of the system [22]. Therefore, it is actually the key to the ADRC in controlling a system with an uncertainty. Assume a nonlinear system with an uncertainty, which suffers from some unknown external disturbance

$$x^{(n)} = f\left(x, \dot{x}, \ldots, x^{(n-1)}, t\right) + w(t) + b_0 u(t) \tag{5}$$

where $f(x, \dot{x}, \ldots, x^{(n-1)}, t)$ 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

$$\begin{cases} x_1(t) = x(t) \\ \vdots \\ x_n(t) = x^{(n-1)}(t) \\ x_{n+1}(t) = f(x, \dot{x}, \ldots, x^{(n-1)}, t) + w(t). \end{cases} \tag{6}$$

Then (5) can be transformed to

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \vdots \\ \dot{x}_n(t) = x_{n+1}(t) + b_0 u(t) \\ \dot{x}_{n+1}(t) = \xi(t) \end{cases} \tag{7}$$

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

$$\begin{cases} \dot{z}_1(t) = z_2(t) - g_1(z_1(t) - x_1(t)) \\ \vdots \\ \dot{z}_n(t) = z_{n+1}(t) - g_n(z_1(t) - x_1(t)) + b_0 u(t) \\ \dot{z}_{n+1}(t) = -g_{n+1}(z_1(t) - x_1(t)) \end{cases} \tag{8}$$

where $g_1(e_1), \ldots, g_{n+1}(e_1)$ are all appropriately constructed nonlinear continuous functions. From (7) and (8), we have

$$\begin{cases} \dot{e}_1(t) = e_2(t) - g_1(e_1(t)) \\ \vdots \\ \dot{e}_n(t) = e_{n+1}(t) - g_n(e_1(t)) \\ \dot{e}_{n+1}(t) = -\xi(t) - g_{n+1}(e_1(t)) \end{cases} \tag{9}$$

where $e_i(t) = z_i(t) - x_i(t), (i = 1, \ldots, n+1)$.

For an arbitrarily changing $\xi(t)$ in a certain range, it is proven [22] that if the nonlinear continuous functions $g_1(e_1), \ldots, g_{n+1}(e_1)$ are chosen to satisfy

$$e_1 g_i(e_1) > 0, \quad \forall e \ne 0, \quad \text{and} \quad g_i(0) = 0 (i = 1, \ldots, n+1) \tag{10}$$

then system (9) is stable with respect to the origin. This means that, with appropriate choices of functions $g_1(e_1), \ldots, g_{n+1}(e_1)$, the states of system (8) can track the corresponding states of system (7), i.e.,

$$z_1(t) \to x_1(t), \ldots, z_n(t) \to x_n(t), z_{n+1}(t) \to x_{n+1}(t). \tag{11}$$
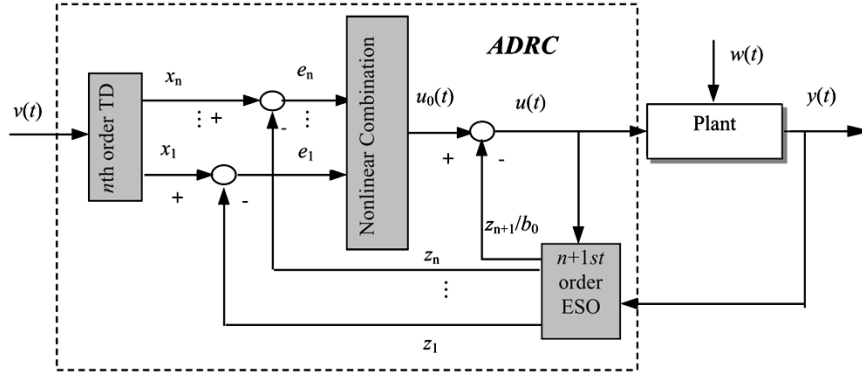
Fig. 1.   Basic structure of the ADRC.

According to the definitions in (6), $x_{n+1}(t)$ is the summation of the unknown function $f(x, \dot{x}, \ldots, x^{(n-1)}, t)$ and the external disturbance $w(t)$ in (5). Thus, the output z_{n+1} in system (8) can have the real-time estimation for $x_{n+1}(t)$ as long as a set of functions $g_1(e_1), \ldots, g_{n+1}(e_1)$ can be found to meet the conditions in (10). Since $x_{n+1}(t)$ is an extended state defined in addition to the necessary states, $x_1(t), \ldots, x_n(t)$, to describe the dynamics in (5), the nonlinear system (8) is hence called the ESO. It is obvious that the ESO is one order higher than that of the system to be observed. The convergence of the ESO is still an open problem for general cases. Analysis of the performance of a second-order ESO can be found in [25].

It is worth mentioning [25] that the system (8) is the classical Luenberger observer when $g_i(e_1) = e_1 (i = 1, \ldots, n+1)$ and is a variable structure observer when $g_i(e_1) = e_1 + k_i \text{sign}(e_1) (i = 1, \ldots, n+1)$.

If the system output $y(t)$ needs to track a given signal $v(t)$, the TD is to arrange the transient process $x_1(t)$ in tracking $v(t)$. The ESO then uses the tracking output $y(t)$ as its input rather than $x_1(t)$ itself. This strategy makes the ESO have a more direct measurement of the system tracking performance.

C. Nonlinear State-Error Feedback (NLSEF)

In this subsection, a controller for system (5) is discussed. Substituting the definition of $x_{n+1}(t)$ into (5), we have

$$x^{(n)}(t) = x_{n+1}(t) + b_0 u(t). \tag{12}$$

If

$$u(t) = u_0(t) - x_{n+1}(t)/b_0 \tag{13}$$

then (12) becomes

$$x^{(n)}(t) = b_0 u_0(t). \tag{14}$$

This means that, if the system output is $y = x^{(n)}$, the open-loop system between $u_0(t)$ and y is a cascaded integrator plant with a gain of b_0. Hereby, we discuss how to design $u_0(t)$ to control the system (14) from the system error feedback. Since $x_1, x_2, \ldots, x_n$ are the different orders of the derivatives of the system input $v(t)$ formed by the TD, and $z_1, z_2, \ldots, z_n$ are the state variables of the system observed by the ESO, the errors between the two groups of variables

$$e_i = x_i - z_i, \quad i = 1, \ldots, n \tag{15}$$

characterize the system's dynamic performance. We take advantage of a nonlinear combination of the errors to realize the control [23]

$$u_0 = b_1 \text{fal}(e_1, \alpha, \delta) + \ldots + b_n \text{fal}(e_n, \alpha, \delta) \tag{16}$$

where the definition of $\text{fal}(e_i, \alpha, \delta)$ is given in (4) and $b_1, \ldots, b_n$ are adjustable parameters. Substituting (16) into (13), the system input $u(t)$ can be obtained, in which b_0 is known, and $x_{n+1}(t)$ is estimated online by the ESO. The nonlinear function $\text{fal}(e_i, \alpha, \delta)$ of the system error e_i is preferred to outline the system control because it provides a small linear region near $e_i = 0$ so that no excessive gain that might lead to high-frequency chattering can occur.

In (16), the controller is constructed by errors of two groups of variables, $x_1, x_2, \ldots, x_n$, and $z_1, z_2, \ldots, z_n$. No derivative of an error is used, in contrast to a proportional integral derivative (PID) controller. Considering that an error signal formed by the system input and output is normally not derivable in practice, the proposed strategy is helpful in improving the system performance. In addition, a nonlinear combination of the errors is adopted in (16), which may lead to a proper tradeoff between the system's response time and overshoots. The overall structure of the ADRC is given in Fig. 1.

From Fig. 1, it is seen that the ADRC is a nonlinear controller that is independent of the system model. Accordingly, if it is used to deal with the calibration-free robotic eye–hand coordination, the structure and the design procedure of the coordination controller are expected to be irrelevant to the system configurations. This is exactly what is to be presented below.

III. SYSTEM MODELING AND CONTROLLER DESIGN

In this section, a dynamic system to describe the calibration-free robotic eye–hand coordination is presented. The controller is designed based on the aforementioned principle of the ADRC, which is a standard procedure and has nothing to do with the configurations and the specific task of the eye–hand system.

A. Visual Mapping Model

Without loss of generality, a robotic eye–hand system with a global monocular camera is taken as an example to demonstrate the system modeling and controller design procedure based on the ADRC. Under the single-eye global visual feedback, the task of the robotic eye–hand coordination is to design a robotic control to make the target position and the hand position to coincide with each other in the camera's image plane according to the error between them observed in the image. Suppose that the hand position is $W = (w_x, w_y, w_z)^T$ in the robotic coordinate system and is $P = (p_x, p_y)^T$ in the image observed by the camera. The target position in the image is $P^* = (p_x^*, p_y^*)^T$, which is also the desired hand position in the image, so that it is the signal for the closed-loop system to track. The relationship between the hand position in the image and that in the robotic coordinate system can be expressed as

$$P = g(W) \tag{17}$$

where $g(\cdot)$ is a function representing all of the effects caused by the eye–hand relationship model, the robot model, and the camera model. Differentiation of the both sides of (17) leads to

$$\begin{cases} \dot{W} = U \\ \dot{P} = J(W) \cdot U \end{cases} \tag{18}$$

where U is the velocity vector in the robotic coordinate system, which is the system control. $J(W)$ is the Jacobian matrix of $g(W)$. (18) describes the differential change of the hand position at a certain instant in the image caused by the differential hand movements in the robotic coordinate system. The essence of calibration-free coordination is to estimate at any instant the current Jacobian matrix. Then, based on its inverse matrix and the specific hand movements in the image, the calculation of the necessary hand movements in the robotic coordinate system is done.

Without a loss of generality, only a three-dimensional (3-D) translational movement of the hand is considered, i.e., $U = (u_x, u_y, u_z)^T$. Since U and P are 3-D and two-dimensional (2-D), respectively, the Jacobian matrix defined in (18) can be expressed as

$$J(W) = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \end{bmatrix}. \tag{19}$$

Then

$$\dot{P} = \begin{bmatrix} \dot{p}_x \\ \dot{p}_y \end{bmatrix} = J(W)U = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \end{bmatrix} \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}. \tag{20}$$

That is

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

B. Controller Design

A calibration-free robotic eye–hand coordination controller based on the principle of the aforementioned ADRC is then proposed. Let us now rewrite the visual mapping model to be in a form suitable to the ADRC. Suppose, in the robotic workspace, a reasonable guess of $J(w)$ is

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

This guess could practically be, for example, the average of all possible values of $J(w)$ in the robotic workspace. Of course, it can only be obtained empirically. System disturbances $w_1(t)$ and $w_2(t)$ are introduced due to the system model inaccuracy, the errors in image detection, and the external disturbance. The system model (21) is rewritten as

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

If we define

$$\begin{cases} a_x(t) = (J_{11} - \hat{J}_{11}) \cdot u_x + J_{12} \cdot u_y + J_{13} \cdot u_z + w_1(t) \\ a_y(t) = J_{21} \cdot u_x + (J_{22} - \hat{J}_{22}) \cdot u_y + J_{23} \cdot u_z + w_2(t) \end{cases} \tag{24}$$

then

$$\begin{cases} \dot{p}_x = a_x(t) + \hat{J}_{11} \cdot u_x \\ \dot{p}_y = a_y(t) + \hat{J}_{22} \cdot u_y. \end{cases} \tag{25}$$

Thus, the original system (21) is transformed to be two decoupled first-order subsystems, each of which has a form similar to (1). The two subsystems in (25) describe the dynamics of the eye–hand coordination in the x and y directions, respectively. $a_x(t)$ and $a_y(t)$ indicate the total disturbances (including the system's unmodeled dynamics and the external disturbance) in the x and y subsystems, respectively.

Consequently, two ADRCs are designed to control the two subsystems in (25). Hereafter, the tracking control in the x direction is taken as the example to demonstrate the design of the ADRC. From (25), we obtain the system equation

$$\begin{cases} \dot{w}_x = u_x \\ \dot{p}_x = a_x(t) + \hat{J}_{11} u_x \\ y_1 = p_x \end{cases} \tag{26}$$

where p_x is the system state variable, y_1 is the system output, and u_x is the system control. It is seen that the system defined by (26) is a first-order system. When $p_x^*(t)$ is the system input, we design the first-order TD as follows:

$$\dot{x}_{x1} = -r_x \text{fal}(x_{x1} - p_x^*, \alpha_{x0}, \delta_{x0}). \tag{27}$$

A second-order ESO to estimate the system uncertainty and external disturbance takes the following form:

$$\begin{cases} \dot{z}_{x1} = z_{x2} - b_{x1}\text{fal}(z_{x1} - y_1, \alpha_{x1}, \delta_{x1}) + \hat{J}_{11} u_x \\ \dot{z}_{x2} = -b_{x2}\text{fal}(z_{x1} - y_1, \alpha_{x2}, \delta_{x2}) \end{cases} \tag{28}$$

where $b_{xi} > 0, 0 < \alpha_{xi} < 1, (i = 1, 2)$. Define the state error of the system tracking as

$$e_{x1} = x_{x1} - z_{x1}. \tag{29}$$

The system control input can be obtained by the following NLSEF control law:

$$\begin{cases} u_{x0} = k_{x0}\text{fal}(e_{x1}, \alpha_x, \delta_x) \\ u_x = (u_{x0} - z_{x2})/\hat{J}_{11} \end{cases} \tag{30}$$

where $\text{fal}(\varepsilon, \alpha, \delta)$ is defined in (4). Thus, the ADRC for target tracking in the x direction is given by (27)–(30). Similarly, the controller in the y direction can be designed.

IV. CONVERGENCE ANALYSIS OF THE COORDINATION CONTROLLER

The coordination controller is composed of three parts: a first-order TD shown in (27), a second-order ESO shown in (28), and a nonlinear controller shown in (30). In this section, the convergence of the ESO is discussed first. Then an explanation for the convergence of the coordination controller is presented.

According to (9), if the ESO [see (28)] is applied to the system (26), the observation error can be expressed as

$$\begin{cases} \dot{e}_{x1} = e_{x2} - b_{x1}\text{fal}(e_{x1}, \alpha_{x1}, \delta_{x1}) \\ \dot{e}_{x2} = \omega_x - b_{x2}\text{fal}(e_{x1}, \alpha_{x2}, \delta_{x2}) \end{cases} \tag{31}$$

where $e_{x1} = z_{x1} - y_1, e_{x2} = z_{x2} - a_x$, and $\omega_x = -\dot{a}_x$. The convergence of (31) is analyzed via the self-stable region (SSR) approach [25].

Definition: Let us assume that G is a region in the state space, which contains the origin. If it satisfies the condition that any system's trajectory, which remains in it after certain time, will eventually converge to the origin, then G is called the self-stable region (SSR) of the system.

The SSR of a system defines the convergence of the system. The convergence of (31) is proved by the following theorem.

Theorem 2: Define

$$h_x(e_{x1}, e_{x2}) = e_{x2} - b_{x1}\text{fal}(e_{x1}, \alpha_{x1}, \delta_{x1}) + kq_x(e_{x1})\text{sign}(e_{x1})$$

where $q_x(e_{x1})$ is a continuous positive definite function, $q_x(0) = 0, k$ is a constant, and $k > 1$. Then region

$$G_x = \{(e_{x1}, e_{x2}) : |h_x(e_{x1}, e_{x2})| \leq q_x(e_{x1})\} \tag{32}$$
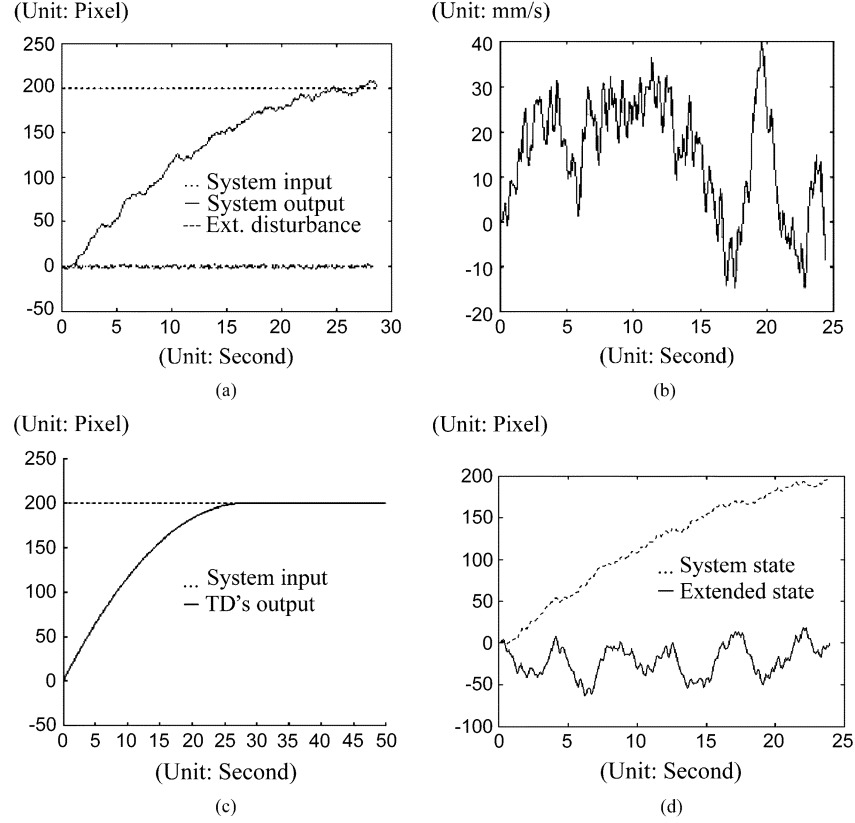
Fig. 2. System responses for the control in the x direction. (a) System input, output, and disturbance. (b) System control. (c) TD's output tracks system input. (d) ESO estimates system state and external disturbance.

is an SSR of system (31).

Proof: Suppose there exists $(e_{x1}(t), e_{x2}(t)) \in G_x, \forall t > T$. From the structure of G_x, we have

$$-kq_x(e_{x1})\text{sign}(e_{x1}) - q_x(e_{x1}) \leq e_{x2} - b_{x1}\text{fal}(e_{x1}, \alpha_{x1}, \delta_{x1})$$
$$\leq -kq_x(e_{x1})\text{sign}(e_{x1}) + q_x(e_{x1}). \tag{33}$$

Select a Lyapunov function of the system (31) as

$$V_x = \frac{1}{2}e_{x1}^2 \tag{34}$$

then

$$\dot{V}_x = e_{x1}\dot{e}_{x1} = e_{x1}(e_{x2} - b_{x1}\text{fal}(e_{x1}, \alpha_{x1}, \delta_{x1}))$$
$$\leq -kq_x(e_{x1})e_{x1}\text{sign}(e_{x1}) + e_{x1}q_x(e_{x1}). \tag{35}$$

Since $q_x(e_{x1})$ is a continuous positive definite function, i.e., $e_{x1}q_x(e_{x1}) \leq |e_{x1}|q_x(e_{x1})$, and $k > 1$, we have

$$\dot{V}_x \leq -(k-1)q_x(e_{x1})|e_{x1}| < 0, \quad \forall e_1(t) \neq 0, \quad t > T. \tag{36}$$

This means

$$e_{x1}(t) \to 0, \quad t \to \infty. \tag{37}$$

Similarly, according to the structure of G_x and the characteristics of the functions $\text{fal}(e_{x1}, \alpha_{x1}, \delta_{x1})$ and $q_x(e_{x1})$, we also have

$$e_{x2}(t) \to 0, \quad t \to \infty. \tag{38}$$

From the definition of the SSR, G_x is then an SSR of (31).

Theorem 2 defines the stable region for the ESO described by (28). It also indicates that the existence of the SSR and the geometric structure determined by (32) are independent of the unknown function ω_x, which

TABLE I
ADRC PARAMETERS

TD			ESO						NLSEF			Est. of b_0	
α_0	δ_0	r	α_1	δ_1	b_1	α_2	δ_2	b_2	α	δ	k_0	$\hat{J}_{11}$	$\hat{J}_{22}$
0.5	0.1	10	0.1	1	10	0.1	2	60	0.1	10	30	0.9	0.9

is determined by the system's unmodeled dynamics and the external disturbances.

When the nonlinear controller (30) is applied to the system (26), a closed-loop system is obtained as follows:

$$\dot{p}_x = a_x + k_{x0}\text{fal}(e_{x1}, \alpha_{x1}, \delta_{x1}) - z_{x2}. \tag{39}$$

If the TD and the ESO are all convergent, we have

$$x_{x1} \to p_x^*, \quad z_{x1} \to y_1 = p_x, \quad z_{x2} \to a_x. \tag{40}$$

Thus, when $t > T$, (39) becomes

$$\dot{p}_x \cong k_{x0}\text{fal}(p_x^* - p_x, \alpha_{x1}, \delta_{x1}). \tag{41}$$

According to the definition of the function $\text{fal}(e_x, \alpha_x, \delta_x)$, it is easy to know that the system (41) is convergent when the parameters are selected appropriately [25]. Thus, the closed-loop system is convergent when the TD, ESO, and nonlinear controller are all applied to the control system (26).

V. SIMULATIONS AND EXPERIMENTS

Simulations and experiments are conducted to demonstrate the feasibility and the performance of the ADRC discussed above for the calibration-free robotic eye–hand coordination.
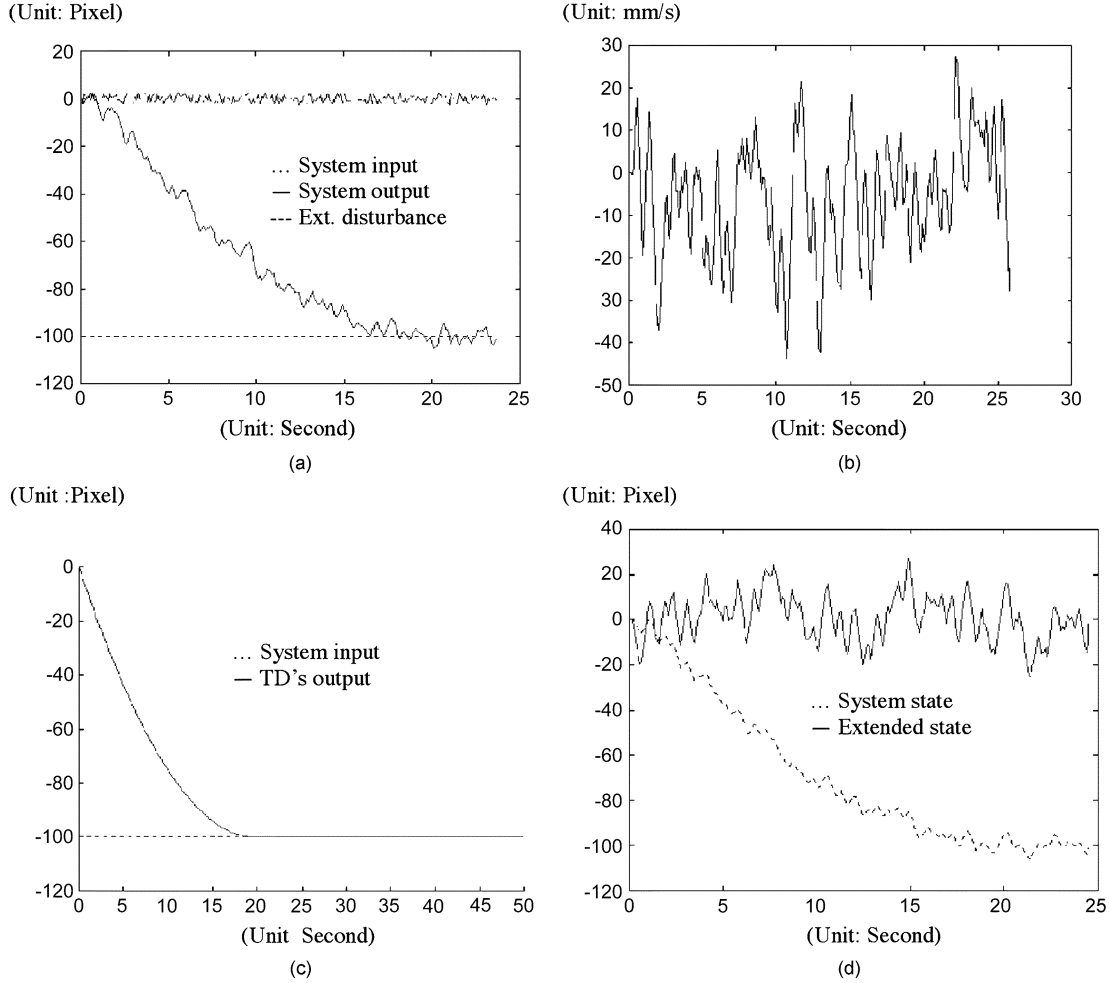
(Unit: Pixel)

(Unit: mm/s)

(a)

(b)

(Unit :Pixel)

(Unit: Pixel)

(c)

(d)

Fig. 3.   System responses for the control in the y direction. (a) System input, output, and disturbance. (b) System control. (c) TD's output tracks system input. (d) ESO estimates system state and external disturbance.

A. Simulations

In the simulation of approaching a static target, the hand and the target are initially located at $(0, 0)$ and $(200, -100)$ in the image plane, respectively. An external disturbance, which is a normal distributed random noise with a maximum magnitude of ± 5 pixels and a zero average, is added to the system in the x and y directions, respectively. The same controllers are used for both the x and y directions. The parameters of the ADRC are chosen as shown in Table I. Since the two directions have the same controller parameters, the subscripts x in (27)–(30) are omitted. Simulation results are depicted in Figs. 2–4.

Fig. 2 shows the system control in the x direction. Fig. 2(a) illustrates the system's input p^*, the system's tracking output p, and the external disturbance w, whereas Fig. 2(b) illustrates the system control u to converge the system error in the x direction. During the control process, the TD is used to arrange a practical input for the system to track, which is shown in Fig. 2(c) with a comparison of the system's true input. The ESO estimates the system state and the summation of the system's unmodeled dynamics and external disturbance, as shown in Fig. 2(d). These estimations are used to compensate for the system control.

In the simulations, the coordination control described by (27)–(30) is to let the hand approach a static target and track a moving target. Without a loss of generality, an eye–hand system with a single global visual feedback is employed. Specifically, the task is to control the hand to approach and track the target so that the distance between the hand
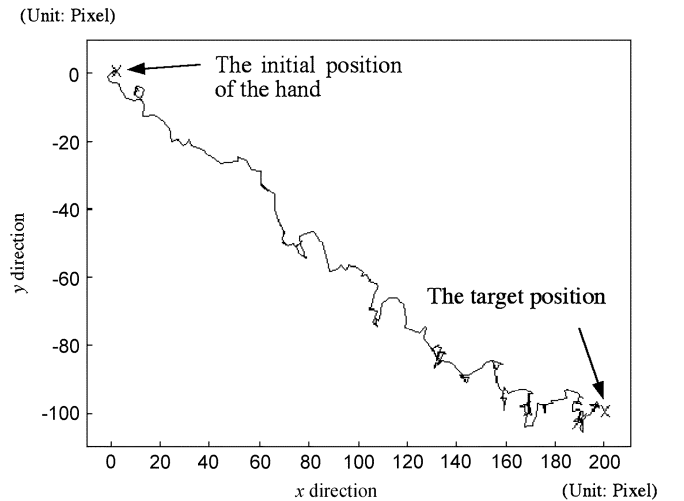
(Unit: Pixel)

Fig. 4.   Approaching trajectory of the hand to the static target.

and the target in the image plane converges to a prescribed threshold. The global visual feedback is conducted with a camera focal length $f = 6$ mm. The camera image-plane quantization resolutions are $N_x = 4.9/582$ mm/pixel and $N_y = 3.7/512$ mm/pixel, respectively. Define
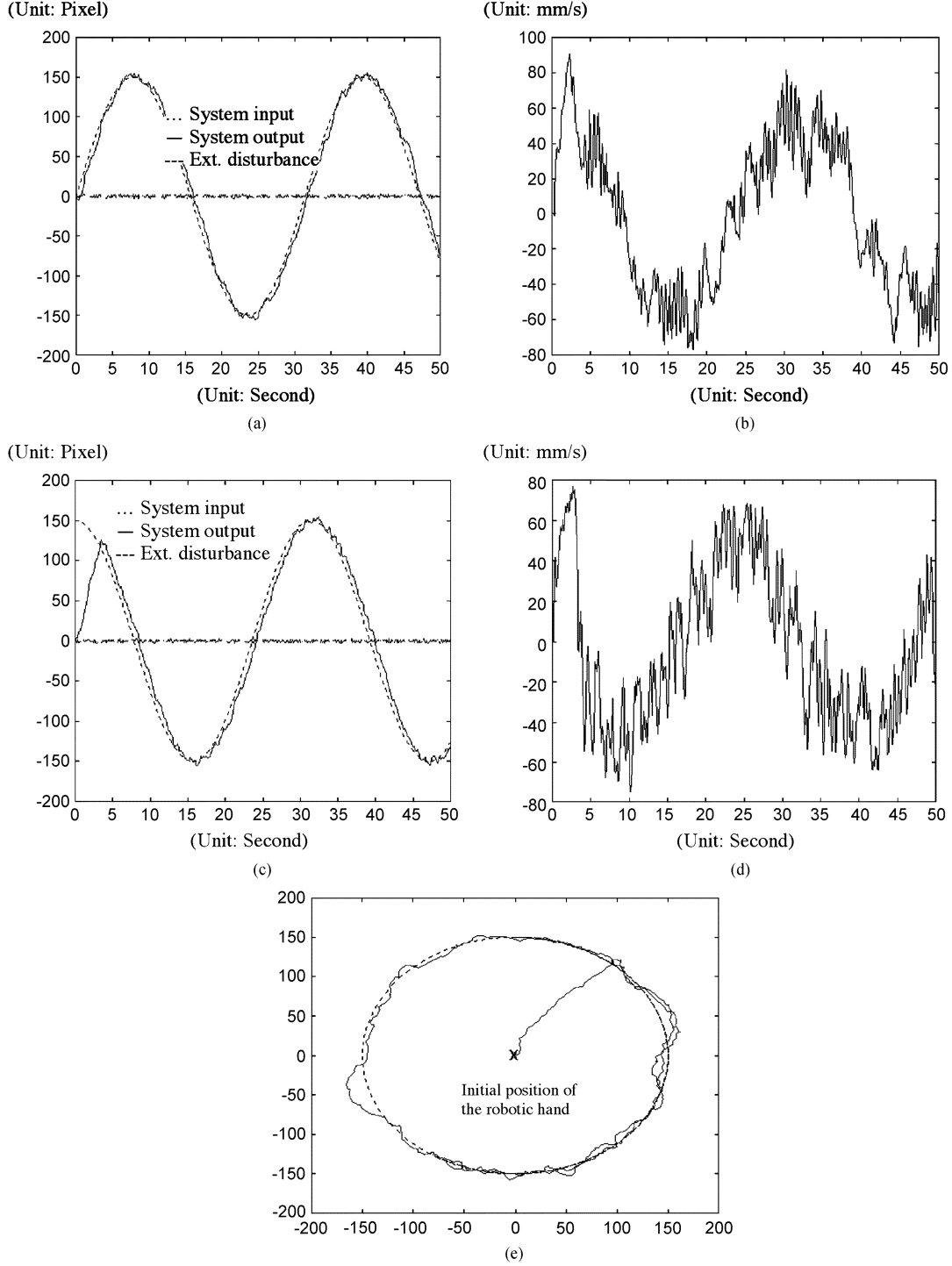
Fig. 5. System response of the ADRC to track a planar moving target. (a) System input, output, and external disturbance in the x direction. (b) System control in the x direction. (c) System input, output, and external disturbance in the y direction. (d) System control in the y direction. (e) The tracking trajectory of the hand in the image plane.

the center of the image plane as the origin of the coordinate system. The pose of the camera with respect to the robotic base coordinates are defined in Euler angles: $\psi = 80°, \theta = -70°$, and $\varphi = -160°$, and the translational movement vector is $T = [40, 10, 1500]^T$ mm.

Simulation results for the control in the y direction are accordingly depicted in Fig. 3. The trajectory of the hand to approach the static target is shown in Fig. 4.

In the simulation of tracking a moving target, the initial positions of the hand and the target are at (0, 0) and (0, 150) in the image plane,

respectively. The target is making a circular movement, which is unknown to the robotic controller.

$$\begin{cases} p_x^*(t) = 150 \sin(t/5) \\ p_y^*(t) = 150 \cos(t/5). \end{cases} \tag{42}$$

An external disturbance in both the x and y directions are normal distributed random noise with a maximum magnitude of ± 5 pixels and a zero average. The same ADRCs with the same control parameters shown in Table I are used in both the x and y directions. The system response is given in Fig. 5.
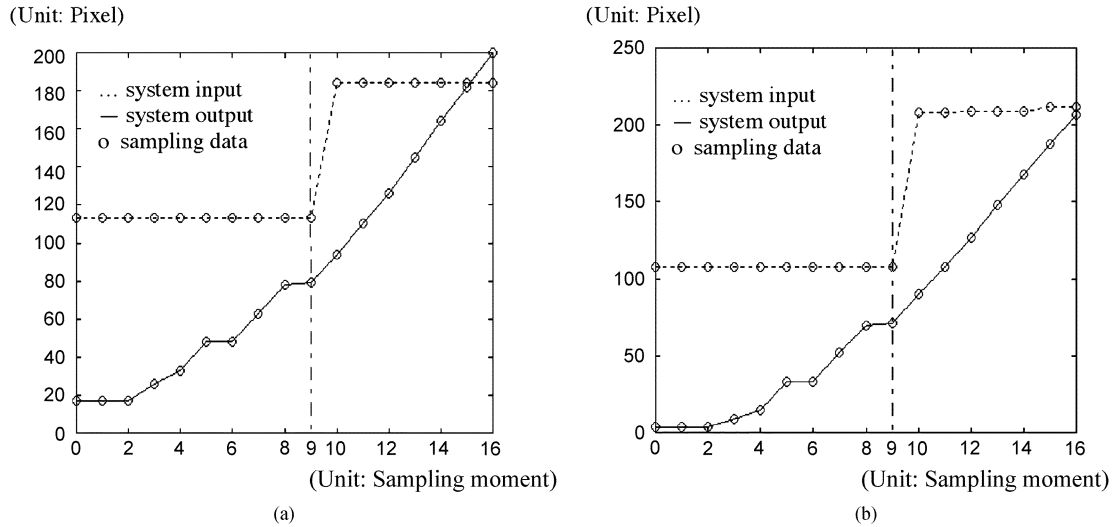
(Unit: Pixel)

(Unit: Pixel)

Fig. 6. System input and output in the (a) x direction and the (b) y direction.
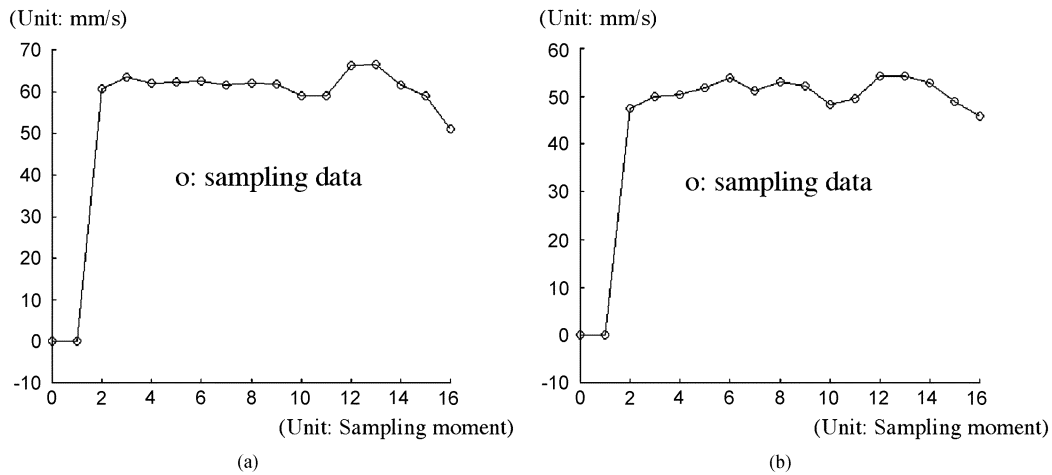
(Unit: mm/s)

(Unit: mm/s)

Fig. 7. Control input in the (a) x direction and the (b) y direction.

In Fig. 5(a) and (c), the transient process of the system input is given by the TD. The system tracks well in the transient process arranged by the TD. The control inputs in the x and y directions are shown in Fig. 5(b) and (d), respectively. Due to the system uncertainty and the external disturbance, there are fluctuations in the control inputs. However, it is seen from Fig. 5(e) that the hand can catch up and track the target quickly. Fig. 5(e) also shows that the tracking in the x direction presents a rather large error systematically, which occurs when the direction of the movement of the x component is changing. This implies that the parameters selected in the x direction for the ADRC are not very reasonable. It is noticed that, sometimes, different parameters are selected in each of the two directions x and y to guarantee a better performance of the system. This will be seen in the experiments below.

B. Experiments

An Adept 604S manipulator with a camera fixed above the robotic workspace is used in an experimental system. The camera's internal parameters and its pose relative to the robotic coordinate system are the same as those in simulations, but are not known to the controller. In our experiments, the robot hand is to reach the target in the workspace, which is now a plane. When the error between the position of the hand image and that of the target image is within a certain range, the task is done. The same order and structure of ADRCs are used in both the

TABLE II
ADRC PARAMETERS IN THE x DIRECTION

TD			ESO						NLSEF			
α_0	δ_0	r	α_1	δ_1	b_1	α_2	δ_2	b_2	α	δ	k_0	$\hat{J}_{11}$
0.5	15	130	0.1	8	80	0.1	8	100	0.1	10	400	0.4

x and y directions. In order to overcome the larger tracking error in the x direction, the parameters for each of the directions in the ADRC are selected differently and empirically, as shown in Tables II and III, respectively. Figs. 6 and 7 show the system response and the control input, respectively. Fig. 8 demonstrates the motion trajectory of the hand in the image.

In Fig. 8, we can see that, during the 0–9 sampling periods, the target is always still at (113, 108). The initial position of the hand is at (17, 4). At the ninth sampling instant, the hand is at (71, 79), which is close to the target, and the target is moved to a new position (184, 212) (in this neighborhood, there is another target position seen, which is the instant target position taken by the camera during the motion). Then the hand keeps on moving toward the new target position. At the 16th sampling instant, the hand moves to (200, 207). The position difference

TABLE III
ADRC PARAMETERS IN THE y DIRECTION

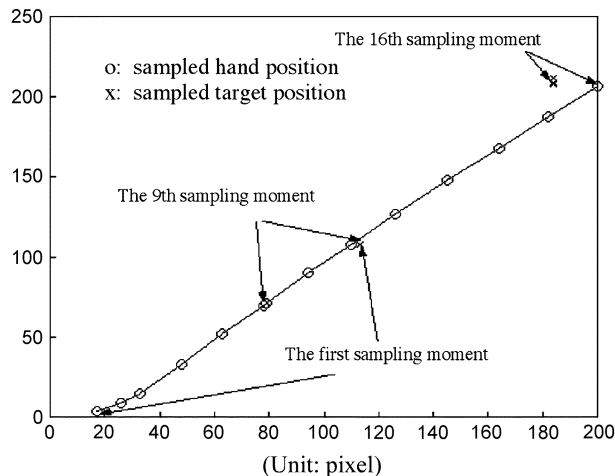| TD			ESO						NLSEF			
α_0	δ_0	r	α_1	δ_1	b_1	α_2	δ_2	b_2	α	δ	k_0	$\hat{J}_{22}$
0.5	10	100	0.1	8	150	0.1	8	200	0.1	8	320	0.6

Fig. 8. Tracking trajectory of the hand in the image plane.

between the hand and the target is now smaller than the prescribed threshold. The system then believes that the hand has reached the target position and the motion stops. In the whole process, both the control inputs and the hand tracking movements are smooth, which confirms the effectiveness of the parameter selections.

VI. CONCLUSION

A new approach to the calibration-free robotic eye–hand coordination is proposed. This approach integrates an ESO to estimate online the system's unmodeled dynamics and the external disturbances. A nonlinear controller is accordingly designed based on the observer's compensation. Convergence of the extended state observer and the overall controller is analyzed and proved so that the performance of the approach is ensured. This approach is advantageous over the conventional approaches to the calibration-free robotic eye–hand coordination in the sense that the proposed solution to the unknown eye–hand relationship is independent of either specific tasks or particular system configurations and thus has a general meaning. The standard design procedure of the ADRC for the calibration-free robotic eye–hand coordination is presented. Simulations and experiments demonstrate that this approach suppresses the effects of the external disturbance and therefore has a strong adaptability and robustness.

Though the ADRC presents a kind of control strategy, which is independent of system model and external disturbance, its superiority over the conventional approaches based on the Jacobian matrix has not been well explored due to its own unsolved problems such as the parameter and nonlinear function selections. However, this paper demonstrates that the ADRC can be successfully applied and offers a new way of thinking for a task-free design in the calibration-free robotic eye–hand coordination. It is believed that along with the development of the ADRC theory itself, its application in the calibration-free robotic eye–hand coordination will surely be further acknowledged.

REFERENCES

[1] L. Hsu and P. L. S. Aquino, "Adaptive visual tracking with uncertain manipulator dynamics and uncalibrated camera," in *Proc. Conf. Decision and Control*, 1999, pp. 1248–1253.

[2] G. D. Hager, W. C. Chang, and A. S. Morse, "Robotic feedback control based on stereo vision: Toward calibration-free hand/eye coordination," in *Proc. IEEE Int. Conf. Robotics and Automation*, 1994, pp. 2850–2856.

[3] C. Schering and B. Kersting, "Uncalibrated hand-eye coordination with a redundant camera system," in *Proc. IEEE Int. Conf. Robotics and Automation*, Leuven, Belgium, 1998, pp. 1366–1372.

[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] A. Kawabata and M. Fujita, "Design of an H_∞ filter-based robust visual servoing system," *Contr. Eng. Practice*, vol. 6, pp. 219–225, 1998.

[6] K. Hashimoto, T. Ebine, and H. Kimura, "Visual servoing with hand-eye manipulator-optimal control approach," *IEEE Trans. Robot. Automat.*, vol. 12, pp. 766–774, Oct. 1996.

[7] 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. In Chinese.

[8] R. Horaud *et al.*, "Visually guided object grasping," *IEEE Trans. Robot. Automat.*, vol. 14, pp. 525–532, Aug. 1998.

[9] 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.

[10] B. H. Yoshimi and P. K. Allen, "Alignment using an uncalibrated camera system," *IEEE Trans. Robot. Automat.*, vol. 11, pp. 516–521, Aug. 1995.

[11] K. Hosada and M. Asada, "Versatile visual servoing without knowledge of true Jacobian," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 1994, pp. 186–191.

[12] M. Jaegersand, O. Fuentes, and R. Nelson, "Experimental evaluation of uncalibrated visual servoing for precision manipulation," in *Proc. IEEE Int. Conf. Robotics and Automation*, Albuquerque, NM, 1997, pp. 2874–2880.

[13] J. Qian and J.-B. Su, "On-line estimation of image Jacobian matrix based on Kalman filter," in *Proc. IEEE Int. Conf. Robotics and Automation*, Washington, DC, 2002, pp. 562–567.

[14] 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. 1992.

[15] J.-B. Su, Y.-G. 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.

[16] C. C. Chan and E. W. Lo, "Visual servo control with artificial neural network," in *Proc. IEEE Int. Conf. Industrial Technology*, 1994, pp. 283–287.

[17] 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.

[18] A. Cretual and F. Chaumette, "Visual servoing based on image motion," *Int. J. Robot. Res.*, vol. 20, pp. 857–877, 2001.

[19] N. Papanikolopoulos and P. Khosla, "Adaptive robotic visual tracking: Theory and experiments," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 429–445, Mar. 1993.

[20] F. Conticelli *et al.*, "Hybrid visual servoing: A combination of nonlinear control and linear vision," *Robot. Auton. Syst.*, vol. 29, no. 4, pp. 243–256, 1999.

[21] E. Grosso, G. Metta, A. Oderra, and G. Sandini, "Robust visual servoing in 3-D reaching tasks," *IEEE Trans. Robot. Automat.*, vol. 12, pp. 732–742, Oct. 1996.

[22] J.-Q. Han, "The extended state observer of a class of uncertain systems," *Control and Decision*, vol. 10, no. 1, pp. 85–88, 1995. In Chinese.

[23] ——, "Nonlinear state error feedback control law—NLSEF," *Control and Decision*, vol. 10, no. 3, pp. 221–225, 1995. In Chinese.

[24] J.-Q. Han and W. Wang, "Nonlinear tracking differentiator," *J. Syst. Sci. Math. Sci.*, vol. 14, no. 2, pp. 177–183, 1994.

[25] Y. Huang and J. Han, "Analysis and design for the second order nonlinear continuous extended states observer," *Chinese Sci. Bull.*, vol. 45, no. 21, pp. 1938–1944, 2000.