Adaptive tracking control for a class of continuous-time uncertain nonlinear systems using the approximate solution of HJB equation

Chaoxu Mu*a,b,c, Changyin Sunb,a, Ding Wangc, Aiguo Songd

A R T I C L E   I N F O

Article history:
Received 21 June 2016
Revised 15 April 2017
Accepted 29 April 2017
Available online 10 May 2017
Communicated by H. Zhang

Keywords:
Adaptive tracking control
Hamilton-Jacobi-Bellman (HJB) equation
Adaptive dynamic programming (ADP)
Neural networks
Uncertainties

A B S T R A C T

In this paper, an adaptive tracking control scheme is designed for a class of continuous-time uncertain nonlinear systems based on the approximate solution of the Hamilton–Jacobi–Bellman (HJB) equation. Considering matched uncertainties, the tracking control of the continuous-time uncertain nonlinear system can be transformed to the optimal tracking control of the associated nominal system. By building the nominal error system and modifying its cost function, the solution of the relevant HJB equation can be contributed to the adaptive tracking control of the continuous-time uncertain nonlinear system. In view of the complexity on solving the HJB equation, its approximate solution is pursued by the policy iteration algorithm under the adaptive dynamic programming (ADP) framework, where a critic neural network is constructed to approximate the optimal cost function, and an action network is used to directly calculate the approximate optimal control law, which constitutes the tracking control law for the original uncertain system together with the steady control law. The weight convergence of the critic network and the stability of the closed-loop system are provided as the theoretical guarantee based on the Lyapunov theory. Two simulation examples are studied to verify the theoretical results and the effectiveness of the proposed tracking control scheme.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

In the control field, the tracking control problem is considerably important because it is common that the system is required to follow a feasible reference system. It is challenging to generate an effective control law for the tracking problem, but is also significant. A large variety of approaches have been investigated for the tracking control of nonlinear systems. Some classical control strategies, such as variable structure control, model predictive control, and back-stepping control, among others, are all used to deal with the tracking control problem for specific nonlinear systems [1–3]. In these traditional tracking control strategies, it is most implemented by the feedback control technique. Via applying this technique, it requires that the control matrix of system is invertible, which is usually difficult in real conditions. Recently, to some extent, the intelligent control has been utilized to deal with the invertible problem. Neural network control, fuzzy logic control, and so on, have become popular for the tracking problem of nonlinear systems [4–7]. Although intelligent control approaches have been generally applied to the tracking control of nonlinear systems, most of these methods in the past literature do not address optimal properties.

When optimal properties are integrated into nonlinear systems, it leads to solve the nonlinear Hamilton–Jacobi–Bellman (HJB) equation instead of the Riccati equation for linear systems [8,9]. It is difficult to obtain the solution of HJB equation by dealing with partial differential equations. Dynamic programming has been developed for many years to deal with optimal control problems of nonlinear systems, however, it often runs into “curse of dimensionality”. The policy iteration algorithm was proposed by Howard in [10] for Markov decision processes, which obtains the optimal solution via successive approximation. Werbos developed the learning methodology to the policy iteration algorithm in [11,12], which has made great progress to deal with optimal control problems of nonlinear systems. Along this direction, adaptive dynamic programming (ADP) has been widely studied as it can effectively obtain an approximate solution for the optimal control of nonlinear systems [13,14]. By acquiring the approximate optimal control law
within the framework of ADP, the invertible condition of the control matrix is relaxed [15]. Therefore, based on the approximate optimal control law and the relaxed control matrix, the approach of ADP is advantageous to deal with the tracking control of nonlinear systems.

The ADP-based approximate optimal control approaches have been greatly developed to address the stabilization problem of nonlinear systems [14,16–20]. For instance, the heuristic ADP algorithm for discrete-time systems was reported in [14,18]. The algorithm of synchronous policy iteration for continuous-time systems was investigated in [16], and the value iteration technique was studied in [17]. The event-sampled ADP designs were novelly developed for continuous-time affine nonlinear systems and interconnected nonlinear systems in [19,20], respectively. Compared with the stabilization problem, no equilibrium point for the controlled nonlinear system is involved for the tracking control problem, and system states are required to follow prescribed trajectories. Therefore, this class of control problems is more challenging and complex. Some significant results have been reported regarding the ADP-based tracking control of discrete-time nonlinear systems [21–28]. For example, an observer-based output feedback controller was investigated with input constraints in [22]. A heuristic dynamic programming controller with a filter design was considered for the tracking control in [25]. In [26], a value iteration ADP was employed for the finite-horizon tracking control problem of discrete-time affine nonlinear systems. In [27], the online concurrent-learning-based ADP approach was developed for the infinite-horizon approximate optimal tracking control with model-based reinforcement learning. In [28], a fuzzy-based reinforcement learning method was designed for a class of nonlinear discrete-time systems with dead-zone. For continuous-time systems, only a few related research results are reported [29–32]. In [29], optimal tracking controllers were designed for general nonlinear continuous-time systems by considering such a system as the limit of the sequence of linear and time-varying approximations, where the optimal controllers were obtained by solving linear-quadratic regulation problems. In [30], the approximate optimal tracking control for general continuous-time nonlinear systems was firstly addressed using the heuristic ADP approach with unknown system dynamics. In [31], an integral reinforcement learning technique was applied to design the adaptive optimal tracking controller for continuous-time affine systems with input constraints. In [32], a filter-based action network was developed and a goal representation ADP approach was used to cope with the tracking control of double-integral continuous-time nonlinear systems with partially unknown dynamics. The aforementioned literature has provided important insights on the tracking control of continuous-time nonlinear systems. To the best of our knowledge, there is not much research work on the ADP-based tracking control for continuous-time uncertain nonlinear systems, which motivates this research work.

This paper proposes an ADP-based tracking control scheme for continuous-time nonlinear systems with unknown matched uncertainties. The main contributions of this paper are as follows. First, we provide a new formulation for the tracking control problem of continuous-time nonlinear systems with matched uncertainties, where the original problem is transformed to the control of the tracking error dynamics including the transient dynamics and the steady dynamics. This leads to the tracking controller constructed by an approximate optimal controller that stabilizes the transient error dynamics and a traditional steady controller that maintains the tracking performance to the reference system. Second, with unknown and matched uncertainties, the modified cost function as well as the relevant HJB equation is derived in terms of the nominal error system, and then the ADP approach is used to find the approximate solution of the tracking HJB equation by training the weights. Third, the weight convergence of the ADP-based approach and the stability of the closed-loop system are provided based on the Lyapunov theory. It should be noted that the policy iteration algorithm requires an offline weight learning process before applying the derived controller to the plant.

This paper is organized as follows. Section 2 provides the preliminaries and the problem formulation for the tracking control of a class of continuous-time uncertain nonlinear systems. In Section 3, the approximate solution of HJB equation for the tracking control is derived by the policy iteration algorithm, which is implemented by neural networks under the ADP framework, and the convergence analysis is also provided. Two simulation cases are studied in Section 4 and conclusion remarks are given in Section 5.

2. Problem formulation and transformation

In this paper, we investigate a class of continuous-time affine nonlinear systems with matched uncertainties, which has been widely regarded as a class of typical nonlinear systems in the industrial control field. For example, the physical object can be assumed as a multi-area power system, a two-link robot manipulator, and so on. The general mathematical description is formulated as:

\[ \dot{x}(t) = f(x) + g(x)u(t) + \Delta f(x). \] (1)

where \( x(t) \in \Omega \in \mathbb{R}^n \) is the state vector, and \( u(t) \in \mathbb{R}^m \) is the input vector. \( f(x) \in \mathbb{R}^n \) is known and differentiable, \( g(x) \in \mathbb{R}^{n \times m} \) is known and invertible, and its norm is bounded by \( g_M \) and \( \Delta f(x) \) is considered as the uncertainties of system (1).

Each state in system (1) is expected to track its reference value, which is given by the reference system:

\[ \dot{r}(t) = \phi(t). \] (2)

where \( r(t) \in \Omega \in \mathbb{R}^n \) is the reference state vector, and \( \phi(t) \in \mathbb{R}^n \). In this tracking control, it is assumed to be one-to-one mapping between \( x(t) \) and \( r(t) \).

In order to deal with the tracking control problem, we define the tracking error vector as \( \delta(t) = x(t) - r(t) \). The following assumption is given with respect to system uncertainties.

**Assumption 1.** The uncertainties of system (1) are unknown, but matched and bounded, which means

\[ \Delta f(x) = g_2(\tau)\Delta r(t) + \hat{g}_2(\delta)\phi(\tau) \| \phi(\tau) \| \leq D(\delta), \] (3)

where \( \Delta r(t), \phi(\tau) \in \mathbb{R}^n, g_1(\tau) \) and \( g_2(\delta) \) are the steady control matrix and the transient control matrix, respectively. \( \tau(I) \) is known if the reference state vector \( r(t) \) is known, \( \tau(0) = 0 \) and \( D(\delta) = 0 \).

According to (1) and (2), with Assumption 1, the dynamics of the tracking error vector can be described as

\[ \dot{\delta}(t) = \chi(t) - \dot{r}(t) = \hat{f}(\delta) + \hat{g}(\delta)u(t) + \Delta f(x) - \phi(t). \] (4)

By subtracting the tracking error vector \( \delta(t) \) from (4), we can obtain

\[ \dot{\delta}(t) = f_1(\delta) + g_1(\tau)v_1(t) + g_1(\tau)v_1(t) - \phi(t) + f_2(\delta) + g_2(\delta)v_2(t) + g_2(\delta)v_2(t). \] (5)

\( f_2(\delta) \) is the system function with respect to the tracking error state \( \delta(t) \), i.e., \( f_2(\delta) = f(\delta) \), which is associated to the transient control \( v_2(t) \). \( f_1(\tau) \) is the remaining system function with respect to the steady state \( \tau(t) \), i.e., \( f_1(\tau) = f(x) - f(\delta) \), which is associated to the steady control \( v_1(t) \). It is obvious that \( f_1(\tau), f_2(\delta) \in \mathbb{R}^n \). Similarly, \( g_1(\tau) \) is the control dynamics with respect to the tracking error state \( \delta(t) \), i.e., \( g_2(\delta) = \hat{g}(\delta) \), which is associated to the transient control \( v_2(t) \). \( g_1(\tau) \) is the remaining control dynamics with
respect to the steady state $r(t)$, i.e., $g(t) = g(x) - g(\delta(t))$, which is associated to the transient control $v_1(t)$. It can also be obtained that $g_1(t), g_2(\delta) \in \mathbb{R}^{n \times m}$.

Observing system (5), the tracking control law is really divided into the steady control law and the transient control law. The idea has been widely used in the field of the optimal tracking control [21,31]. With this idea, when the controlled system arrives at the steady-state stage, the required control (i.e., the steady control) can be definitely obtained. Therefore, the tracking control aim to design the tracking error control by using near-optimal approaches, which entails the cost function and the subsequent design processes.

When system (1) is steady, it satisfies $f_1(t) + g_1(t) v_1(t) + g_1(\delta(t)) \phi(t) = 0$, which implies that the tracking control is realized. Meanwhile, the transient error satisfies $\delta(t) = \delta(t)$. Considering $f_1(t) + g_1(t) v_1(t) + g_1(\delta(t)) \phi(t) = 0$, therefore $v_1(t)$ can be obtained as

$$v_1(t) = g_1^{-1}(\delta(t)) (\phi(t) - f_1(t)) - t(r).$$

By applying (6) into (5), it can be observed that the steady error is driven to zero. The transient error dynamics can be actually formulated as

$$\dot{\delta}(t) = f_2(\delta) + g_2(\delta) v_2(t) + g_2(\delta) \tau(\delta),$$

which is expected to be stable at zero since $\tau(0) = 0$. Based on this analysis, the control law $u(t)$ can be formulated as $u(t) = u_0(t) + u_1(t)$ with the steady control law $u_0(t) = v_1(t) + t(t)$ and the transient control law $u_1(t) = v_1(t) + \tau(\delta)$. Generally speaking, the control law $u(t)$ is solved to track its reference system (2). However, it is difficult to obtain $\tau(\delta)$ for unknown matched uncertainties. In this paper, we solve this problem by strengthening the control design of its nominal system. The further analysis is given as follows.

The transient error system is presented in (7). Correspondingly, without the uncertainties, the nominal transient error dynamics are

$$\dot{\delta}(t) = f_2(\delta) + g_2(\delta) v_2(t) + g_2(\delta) \tau(\delta),$$

where $f_2(\delta) = f_2(\delta)$, $g_2(\delta) = g_2(\delta)$ and $\tau(\delta) = \tau(\delta)$. Assume that the right side of differentiable Eq. (8) is Lipschitz continuous on the compact set $\Omega \subseteq \mathbb{R}^n$ containing the origin, and system (8) is controllable. In this sense, there exists a continuous control law $\mu(t)$ to asymptotically stabilize the system at the origin.

For system (8), the infinite horizon cost function is given as

$$J(\delta(t)) = \int_0^\infty \left[ R(\delta(\zeta), \mu(\zeta)) + \theta D^2(\delta(\zeta)) \right] d\zeta,$$

where $\theta D^2(\delta)$ satisfies $\theta D^2(\delta) \geq \tau^2(\delta) Q \tau(\delta)$, $\theta > 0$ is a given coefficient and $Q \in \mathbb{R}^{m \times m}$ is a given positive definite matrix, $R(\delta(\zeta), \mu(\zeta))$ is the utility function with $R(0,0) = 0$, which is usually chosen as

$$R(\delta(\zeta), \mu(\zeta)) = \delta^T(\zeta) P \delta(\zeta) + \mu^T(\zeta) Q \mu(\zeta),$$

where $P \in \mathbb{R}^{n \times n}$ is positive. Compared with the common optimal control problem [30,31] the information of system uncertainties has been considered by involving the term $\theta D^2(\delta(\zeta))$.

**Definition 1.** A control law $\mu(t)$ is admissible with regard to (9) on the compact set $\Omega_\mu$, if $\mu(t)$ is continuous on $\Omega_\mu$, $\mu(0) = 0$. $\mu(t)$ can stabilize system (8) on $\Omega_\mu$, and $J(\delta(t))$ is finite for all $\delta(t) \in \Omega$.

Considering a state feedback control law $\mu(t)$ for system (8) based on the cost function (9), $\mu(\delta)$ is required to be admissible. With $\mu(t)$, if the cost function (9) is continuously differentiable, then the nonlinear Lyapunov equation can be obtained as

$$0 = R(\delta(t), \mu(\delta)) + \theta D^2(\delta)$$

and

$$J(0) = 0,$$

where $\gamma \gamma = \delta(t)$ and $\mu(\delta)$. Based on (11), the Hamiltonian function is defined as

$$H(\delta(t), \mu(\delta), \gamma \gamma(\delta(t))) = R(\delta(t), \mu(\delta)) + \theta D^2(\delta) + \gamma \gamma(\delta(t))(f(\delta) + g(\delta) \mu(\delta)).$$

and the optimal cost function is formulated as

$$J^*(\delta(t)) = \min_{\mu(\delta) \in \Omega_\mu} \int_0^T \left[ \theta D^2(\delta) + R(\delta(\zeta), \mu(\zeta)) \right] d\zeta.$$
Since $P$ is a given positive definite matrix, then $J^*(\delta(t)) \leq 0$ holds. Therefore, for system (7), it is asymptotically stable under the condition $\Theta D^2(\delta) \geq \tau^T(\delta)Q^*(\delta)$ by using the optimal control law $\mu^*(\delta)$ of system (8). It means $\lim_{t \to \infty} \delta(t) = 0$ for any $\delta(t) \in \Omega$. Thus, the theorem is completely proved. □

**Remark 1.** Theorem 1 demonstrates the relationship between the robust stabilization of the transient error system (7) and the optimal control of its nominal system (8). The robust control of (7) is equivalent as the optimal control problem of (8) based on a bounded condition for uncertainties and a selected cost function.

In terms of this idea, if the uncertainties satisfy the required condition, then the robust control of system (7) can be replaced by studying the optimal control of system (8). Therefore, in the following part, we carefully investigate the optimal control of system (8) to achieve the asymptotical stability of system (7).

We substitute the optimal control law in (15) into the nonlinear Lyapunov equation (11) to get the modified HJB equation, which is

$$0 = \theta D^2(\delta) + \delta^T(t)P\delta(t) + (\nabla J^*(\delta(t)))^T \tilde{f}(\delta)$$

$$- \frac{1}{2} (\nabla J^*(\delta(t)))^T \tilde{g}(\delta)Q^{-1}\tilde{g}(\delta) \nabla J^*(\delta(t))$$

with the initial condition $J^*(0) = 0$. Observing (15), $\mu^*(\delta)$ can be solved if $\nabla J^*(\delta(t))$ can be obtained from the above HJB equation. However, as the HJB equation is the nonlinear partial differential equation, it is too difficult to directly obtain $J^*(\delta(t))$ by analytically solving the HJB equation. In what follows, an iterative scheme will be introduced to approximately obtain the solution of the HJB equation.

### 3. Adaptive tracking control scheme based on ADP

In this section, a policy iteration algorithm is provided with a neural network implementation to approximately obtain the optimal control law of system (8). The related stability analysis is also presented.

#### 3.1. Derivatives of policy iteration algorithm

The policy iteration algorithm is introduced to iteratively solve the HJB equation, which is made up of the policy evaluation based on (11) and the policy improvement based on (15). Specifically, the procedure of the policy iteration algorithm is described as follows:

1. Set $i = 0$ and $J^{(0)}(\delta(t)) = 0$, and choose a small positive number $\epsilon$ and an initial admissible control policy $\mu^{(0)}(\delta)$, then start the algorithm from $\mu^{(0)}(\delta)$.
2. With $\mu^{(i)}(\delta)$ and $J^{(i+1)}(0) = 0$, solve the nonlinear Lyapunov equation

$$0 = \theta D^2(\delta) + R(\delta(t), \mu^{(i)}(\delta))$$

$$+ (\nabla J^{(i+1)}(\delta(t)))^T \left( \tilde{f}(\delta) + \tilde{g}(\delta)\mu^{(i)}(\delta) \right).$$

3. Update the control law by

$$\mu^{(i+1)}(\delta) = -\frac{1}{2} Q^{-1}\tilde{g}(\delta) \nabla J^{(i+1)}(\delta(t)).$$

4. If $\|J^{(i+1)}(\delta(t)) - J^{(i)}(\delta(t))\| \leq \epsilon$, stop the algorithm and obtain the approximate optimal control policy $\mu^{(i+1)}(\delta)$; else, let $i = i + 1$ and go back to (2).

As the iteration number $i$ goes to infinity, the algorithm can converge to the optimal cost function $J^*(\delta(t))$ and the optimal control policy $\mu^*(\delta)$, i.e., $J^{(i)}(\delta(t)) \to J^*(\delta(t))$ and $\mu^{(i)}(\delta) \to \mu^*(\delta)$ as $i \to \infty$. The convergence of the policy iteration algorithm has been proved in references [33,34].

#### 3.2. Neural-network-based ADP implementation for adaptive tracking control

In this subsection, the above policy iteration algorithm is implemented within the framework of ADP by using the action-critic structure and neural networks.

Based on the universal approximation property of neural networks, a three-layer neural network is used as the critic network to approximate the optimal cost function $J^*(\delta(t))$ for the control of system (8). All states of system (8) are taken as the inputs of the critic network, while $J^*(\delta(t))$ is outputted after the weight regulation. By using $h$ hidden nodes and optimal connection weights, $J^*(\delta(t))$ is formulated as

$$J^*(\delta(t)) = \omega_1^T \sigma_1(\delta) + \zeta_1(\delta),$$

where the optimal hidden-to-output weight vector is denoted as $\omega_1 \in \mathbb{R}^h$. $\sigma_1(\delta) \in \mathbb{R}^h$ is the activation function. $\zeta_1(\delta)$ is the reconstruction error via the neural network approximation.

Since the optimal weights are generally unknown, the estimation weights are suggested to approximate $J^*(\delta(t))$ instead of the optimal weights. Denote $\hat{\omega}_1(t)$ as the estimation weight vector, which is utilized to approximate $J^*(\delta(t))$ as

$$\hat{J}^*(\delta(t)) = \hat{\omega}_1^T(t) \sigma_1(\delta).$$

where $\hat{J}^*(\delta(t))$ denotes the estimation of $J^*(\delta(t))$. Take the derivative of the approximated cost function along $\delta(t)$, which is

$$\nabla \hat{J}^* \delta(t) = \left( \frac{\partial \sigma_1(\delta)}{\partial \delta(t)} \right)^T \hat{\omega}_1(t) = \left( \nabla \sigma_1(\delta) \right)^T \hat{\omega}_1(t),$$

where $\nabla \sigma_1(\delta) \in \mathbb{R}^{h \times n}$, $\nabla \sigma_1(\delta) \in \mathbb{R}^n$. Substituting $\nabla \hat{J}^* \delta(t)$ into (12) to obtain the estimation of Hamiltonian function as

$$\hat{H}(\delta(t), \hat{\mu}^*(\delta), \hat{\omega}_1(t)) = R(\delta(t), \hat{\mu}^*(\delta)) + \theta D^2(\delta) + \hat{\omega}_1^T(t)$$

$$\times \nabla \sigma_1(\delta) \left( \hat{f}(\delta) + \hat{g}(\delta) \hat{\mu}^*(\delta) \right) = \hat{e}_1(t),$$

where $\hat{\mu}^*(\delta)$ is the approximate optimal control law derived from the output of the action network. The estimated Hamiltonian function $\hat{H}(\delta(t), \hat{\mu}(\delta), \hat{\omega}_1(t))$ in (28) is taken as the error function to regulate the critic network. It means that the weights of the critic network are updated by minimizing

$$E_c(t) = \frac{1}{2} \hat{e}_1^T(t) \hat{e}_1(t).$$

An adaptive updating rule is designed to regulate $\hat{\omega}_1(t)$, which is

$$\dot{\hat{\omega}}_1(t) = -\eta_c \frac{\partial E_c(t)}{\partial \hat{\omega}_1(t)} = -\eta_c \hat{e}_1(t) \frac{\partial \hat{e}_1(t)}{\partial \hat{\omega}_1(t)}$$

$$= -\eta_c \hat{e}_1(t) \nabla \sigma_1(\delta) \hat{\delta}(t),$$

where $\eta_c > 0$ is the learning rate of the critic network. Define $\rho(\delta) = \nabla \sigma_1(\delta) \hat{\delta}(t)$, $\rho(\delta) \in \mathbb{R}^h$, then the updating rule in (30) can be abbreviated as

$$\dot{\hat{\omega}}_1(t) = -\eta_c \hat{e}_1(t) \rho(\delta).$$

If the optimal weight vector $\omega_1$ is used, it means that the Hamiltonian function satisfies the nonlinear Lyapunov equation, i.e., the modified HJB equation holds. Using (25), take the derivative of $J^*(\delta(t))$ along $\delta(t)$ to get $\nabla J^*(\delta(t))$ as

$$\nabla J^* \delta(t) = \left( \nabla \sigma_1(\delta) \right)^T \omega_1 + \nabla \zeta_1(\delta).$$

[33,34]
where $\nabla \xi_c(\delta) \in \mathbb{R}^n$. Therefore, we can obtain that
\[
H(\delta(\tau), \mu^*(\delta), \omega_c) = R(\delta(\tau), \mu^*(\delta) + \theta D^2(\delta) + \omega^T c)
\times \nabla \sigma_c(\delta) \delta(\tau) + (\nabla \xi_c(\delta))^T \delta(\tau) = 0.
\] (33)

By letting
\[
e_c(\tau) = R(\delta(\tau), \mu^*(\delta)) + \theta D^2(\delta) + \omega^T c(\delta),
\] (34)

it can be found that $e_c(\tau) = - (\nabla \xi_c(\delta))^T \delta(\tau)$ is only the residual error due to the neural network approximation. Consider both $R(\delta(\tau), \mu^*(\delta))$ and $\theta D^2(\delta)$ are bounded, and assume $r(\delta)$ to be bounded by $\rho_2$, then $e_c(\tau)$ is bounded by $e_2$.

The weight estimation error of the critic network is defined as
\[
\hat{\omega}_c(\tau) = \omega_c - \hat{\omega}_c(\tau).
\] (35)

According to (28) and (34), then we have
\[
e_c(\tau) - \hat{e}_c(\tau) = \hat{\omega}_c^T(\tau) r(\delta).
\] (36)

Therefore, the dynamics of the weight estimation error can be derived as
\[
\dot{\hat{\omega}}_c(\tau) = - \hat{\omega}_c(\tau) \eta c(t) r(\delta) \hat{e}_c(\tau) \rho(\delta)
\] (37)

Since $f(\delta(t))$ has been obtained from the critic network, the action network is used to calculate the optimal control law according to (15), which is
\[
\mu^*(\delta) = - \frac{1}{2} Q^{-1} g^T(\delta) \nabla f(\delta(t))
\] (38)
\[
= - \frac{1}{2} Q^{-1} g^T(\delta) (\nabla \sigma_c(\delta))^T \omega_c + \nabla \xi_c(\delta).
\]

As $\omega_c$ is actually unknown, $\hat{\omega}_c(\tau)$ is used to approximate $f(\delta(t))$. By using $\nabla f(\delta)$ in (27), the approximate optimal control law is formulated as
\[
\hat{\mu}^*(\delta) = - \frac{1}{2} Q^{-1} g^T(\delta) \nabla f(\delta(t))
\] (39)
\[
= - \frac{1}{2} Q^{-1} g^T(\delta) (\nabla \sigma_c(\delta))^T \hat{\omega}_c(\tau).
\]

Eq. (39) implies that the approximate optimal control law can be directly derived by employing the trained weights of the critic network. The action network is actually operated by using the control function (39), while does not involve the function approximation. The ADP-based adaptive tracking control scheme is presented in Fig. 1.

Remark 2. In [30], it did not involve any uncertainties into the general nonlinear continuous-time system $\dot{x}(t) = h(x(t), u(t))$. In technology, it considered the control of the tracking error dynamics, thus the system expression was allowed to be unknown but was identified via recurrent neural network. In [31], it investigated the constrained tracking control of affine continuous-time nonlinear system, described as $\dot{x}(t) = f(x) + g(x)u(t)$. The system function $f(x)$ was not necessary to be known for the integral reinforcement learning method, however, the exact knowledge of the system dynamics was required to find the steady-state control law. Therefore, in [31], the system dynamics were known and it considered the constrained tracking control without involving uncertainties. Mu et al. [32] concerned on the design of filter-based heuristic ADP for the class of double integral chain nonlinear systems $\dot{x}_1(t) = x_2(t)$, $\dot{x}_2(t) = f(x) + u(t) + \Delta f(t)$, where the action network was used to approximate $f(x)$, and $g(x)$ was actually an identity matrix. The filter associated with the reference system was designed to output the tracking control law and reduce the influence of uncertainties. Based on these investigations, this paper contributes on the robust tracking control design of nonlinear continuous-time uncertain systems by using the policy iteration method. By designing the approximate optimal tracking control of nominal systems, this proposed control method is proved to be the robust tracking control strategy of uncertain systems.

3.3. Stability analysis

Assumption 2. The optimal weight vector $\omega_c$, the activation function $\sigma_c(\cdot)$ as well as its derivative $\nabla \sigma_c(\cdot)$, the reconstruction error $\xi_c(t)$ as well as its derivative $\nabla \xi_c(t)$, are all upper bounded i.e., $|\omega_c| \leq \omega_{\mathbf{M}}, |\sigma_c(\cdot)| \leq \sigma_{\mathbf{M}}, |\nabla \sigma_c(\cdot)| \leq \sigma_{\mathbf{D}}, |\xi_c(t)| \leq \xi_{\mathbf{M}}$ and $|\nabla \xi_c(t)| \leq \xi_{\mathbf{D}}$, where $\omega_{\mathbf{M}}, \sigma_{\mathbf{M}}, \sigma_{\mathbf{D}}, \xi_{\mathbf{M}}, \xi_{\mathbf{D}}$ are positive constants.

Theorem 2. For system (8), if the weights of the critic network is updated by (31), then the weight estimation error $\hat{\omega}_c(\tau)$ is uniformly ultimately boundedness.

Proof. Select the Lyapunov function as follows:
\[
V_c(t) = \frac{1}{n_c} \hat{\omega}_c^T(t) \hat{\omega}_c(t).
\] (40)

The time derivative of the Lyapunov function (40) is
\[
\dot{V}_c(t) = \frac{2}{n_c} \hat{\omega}_c^T(t) \dot{\hat{\omega}}_c(t)
\] (41)
\[
= \frac{2}{n_c} \left( \hat{\omega}_c^T(t) \eta c(\tau) - \hat{\omega}_c^T(t) r(\delta) \right) \big( \eta c(\tau) - \hat{\omega}_c^T(t) r(\delta) \big)^T.
\]

By using the Cauchy–Schwarz inequality, we can obtain
\[
\dot{V}_c(t) \leq \frac{1}{n_c} \hat{\omega}_c^T(t) \eta c(\tau) \big( \eta c(\tau) - \hat{\omega}_c^T(t) r(\delta) \big)^T
\]
\[
\leq - (2 - \eta_c) \| \hat{\omega}_c^T(t) r(\delta) \|^2 + \frac{1}{n_c} \eta c^2(t).
\] (42)

Considering inequality (42), it can be concluded that $\dot{V}_c(t) < 0$ as long as $0 < \eta_c < 2$ and
\[
\| \hat{\omega}_c(\tau) r(\delta) \|^2 > \frac{\eta_c^2(t)}{\eta_c(2 - \eta_c)}. \] (43)

By employing the dense property of real numbers, there exists a positive constant $\varphi \in (0, \rho_{\mathbf{M}}]$ to satisfy the inequality
\[
\| \hat{\omega}_c(\tau) r(\delta) \|^2 \geq \varphi^2 \| \hat{\omega}_c(\tau) \|^2 \geq \frac{\eta c^2(t)}{\eta_c(2 - \eta_c)}.
\] (44)

Therefore, when $\hat{\omega}_c(\tau)$ lies outside of the compact set
\[
\Theta_{\delta} = \left\{ \hat{\omega}_c(\tau) : \| \hat{\omega}_c(\tau) \| \leq \frac{\rho_0(t)}{\eta_c(2 - \eta_c)} \right\}
\]
and $0 < \eta_c < 2$, $\dot{V}_c(t) < 0$ holds. According to the Lyapunov theory, it can be obtained that the weight estimation error $\hat{\omega}_c(\tau)$ is uniformly ultimately boundedness. This completes the proof. □
Theorem 3. For system (8), if the approximate optimal control law in (39) is used with the weight updating rule given in (31), then the tracking error \( \delta(t) \) is uniformly ultimately bounded with the boundary M as

\[
M = \frac{\Psi_M}{\sqrt{\alpha^2 \theta + \lambda_{\min}(Q)}}.
\]

where \( \Psi_M \) and \( \alpha \) are positive constants and \( \lambda_{\min}(Q) \) is the minimal eigenvalue of Q.

Proof. Select the positive definite function \( V_c(t) \) as the Lyapunov function of system (8), which is \( V_c(t) = \tilde{f}(\delta(t)). \) Take the derivative of \( V_c(t) \) to obtain that

\[
V_c(t) = (\nabla V_c(t))^T \left( \tilde{f}(\delta) + \tilde{g}(\delta) \mu^*(\delta) \right).
\]

Since \( \nabla V_c(t) \) satisfies (22), therefore we have

\[
0 = \theta D^2(\delta) + \delta^T(t) P \delta(t) + (\nabla V_c(t))^T \tilde{f}(\delta) - \frac{1}{4} (\nabla V_c(t))^T \tilde{g}(\delta) Q^{-1} \tilde{g}(\delta) V(t).
\]

Obviously,

\[
(\nabla V_c(t))^T \tilde{f}(\delta) = -\theta D^2(\delta) - \delta^T(t) P \delta(t) + \frac{1}{4} (\nabla V_c(t))^T \tilde{g}(\delta) Q^{-1} \tilde{g}(\delta) V(t).
\]

By substituting (48) into \( V_c(t) \), it can be derived as

\[
V_c(t) = -\theta D^2(\delta) - \delta^T(t) P \delta(t) + \frac{1}{4} (\nabla V_c(t))^T \tilde{g}(\delta) \mu^*(\delta)
\]

\[
+ \frac{1}{4} \left( (\nabla V_c(t))^T \tilde{g}(\delta) Q^{-1} \tilde{g}(\delta) - (\nabla V_c(t))^T \tilde{g}(\delta) \right) \tilde{g}(\delta)
\]

\[
= -\theta D^2(\delta) - \delta^T(t) P \delta(t) + \frac{1}{4} (\nabla V_c(t))^T \tilde{g}(\delta) Q^{-1} \tilde{g}(\delta) V(t).
\]

Consider \( \mu^*(\delta) \) and \( \mu(\delta) \) are expressed as (38) and (39), then (49) can be further deduced as

\[
V_c(t) = -\theta D^2(\delta) - \delta^T(t) P \delta(t) - \frac{1}{4} (\nabla V_c(t))^T \tilde{g}(\delta) Q^{-1} \tilde{g}(\delta)
\]

\[
\times \tilde{g}(\delta) V(t) + \frac{1}{2} (\nabla V_c(t))^T \tilde{g}(\delta) Q^{-1} \tilde{g}(\delta)
\]

\[
\times \left( \nabla V_c(t) - \nabla \tilde{V}(t) \right)
\]

\[
= -\theta D^2(\delta) - \delta^T(t) P \delta(t) - \frac{1}{4} (\nabla V_c(t))^T \tilde{g}(\delta) Q^{-1} \tilde{g}(\delta)
\]

\[
\times \tilde{g}(\delta) V(t) + \Psi^2(t),
\]

where

\[
\Psi^2(t) = \frac{1}{2} \left( (\nabla V_c(t))^T \tilde{g}(\delta) Q^{-1} \tilde{g}(\delta) \right) \left( \nabla V_c(t) - \nabla \tilde{V}(t) \right)
\]

\[
= \frac{1}{2} \left( \alpha^2 \theta (\nabla \sigma(\delta))^T \tilde{g}(\delta) Q^{-1} \tilde{g}(\delta) \right)
\]

\[
\times \left( (\nabla \sigma(\delta))^T \tilde{\phi}(t) + \nabla \tilde{\sigma}(\delta) \right)
\]

\[
\leq \left\| (\nabla \tilde{\sigma}(\delta))^T G(\delta) \right\|^2 + \frac{1}{2} \left\| (\alpha^2 \theta (\nabla \sigma(\delta))^T G(\delta) \right\|^2
\]

\[
+ \left\| G(\delta) (\nabla \sigma(\delta))^T \tilde{\phi}(t) \right\|^2
\]

\[
\leq \frac{1}{2} \left( \left\| (\nabla \sigma(\delta))^T \right\|^2 + \frac{1}{2} \left\| (\alpha^2 \theta (\nabla \sigma(\delta))^T G(\delta) \right\|^2
\]

\[
+ \left\| G(\delta) (\nabla \sigma(\delta))^T \tilde{\phi}(t) \right\|^2 \right\|^2.
\]

According to Assumption 2 and Theorem 2, \( \omega_e, \nabla \sigma(\delta), \tilde{\sigma}(t) \) and \( \nabla \sigma(\delta) \) are all bounded, therefore \( \Psi^2(t) \) satisfies

\[
\Psi^2(t) \leq G_M^2 \left( \frac{\alpha^2 \theta}{2} + \frac{\varepsilon^2 \theta}{2} \right)
\]

\[
= \Psi^2_M,
\]

where \( G_M = \sup \| q \| \). Then, \( V_c(t) \) can be derived from (50) as

\[
V_c(t) \leq -\theta D^2(\delta) - \delta^T(t) P \delta(t) + \Psi_M^2
\]

\[
\leq -\theta D^2(\delta) - \delta^T(t) P \delta(t) + \Psi_M^2.
\]

In (53), the disturbance boundary \( D(\delta) \) is regard to \( \| \delta(t) \| \), which is assumed that \( D(\delta) = \alpha \| \delta(t) \| \), where \( \alpha \) is a positive constant. Then, (53) becomes

\[
V_c(t) \leq -\left( \alpha^2 \theta + \lambda_{\min}(Q) \right) \| \delta(t) \|^2 + \Psi_M^2.
\]

From (54), it can be observed that \( V_c(t) < 0 \) if \( \delta(t) \) locates outside the compact set

\[
\Theta = \left\{ \delta(t) : \| \delta(t) \| < \frac{\Psi_M}{\sqrt{\alpha^2 \theta + \lambda_{\min}(Q)}} \right\}.
\]

That is to say, using the approximate optimal control policy \( \tilde{\mu}^*(\delta) \), the closed-loop dynamics of the nominal tracking error system (8) is uniformly ultimately bounded with the boundary

\[
M = \frac{\Psi_M}{\sqrt{\alpha^2 \theta + \lambda_{\min}(Q)}}.
\]

Remark 3. According to Theorem 3, the approximate optimal control law \( \tilde{\mu}^*(\delta) \) derived from (39) can realize the stabilization control for system (8) in terms of the modified cost function (9). Based on Theorem 1, system (7) with uncertainties is asymptotically stable. It means that the original system can asymptotically track to its reference system under the designed control law \( u(t) \).

4. Simulation

In this section, two simulation examples are provided to demonstrate the effectiveness of the proposed adaptive tracking control strategy.

Case 1. Consider the following continuous-time nonlinear system with uncertainties:

\[
\dot{x}(t) = f(x) + g(x) \left( u(t) + \tau(t) \right)
\]

where \( x(t) \in \mathbb{R}^3 \) is the state vector, \( u(t) \in \mathbb{R}^2 \) and \( \tau(t) \in \mathbb{R}^2 \) are control vectors corresponding to the system function and the disturbances, respectively. \( r(t) \in \mathbb{R}^2 \) is the reference state vector, and \( u_i(t) \in \mathbb{R}^2 \) is the reference input vector which is given as \( u_i(t) = 0.3 \sin(0.4 t) \). \( 0.1 \) is a constant vector.

Assume that the uncertainties is only regarded to the error state \( \delta(t) \), then \( \tau(t) \) is given as

\[
\tau(t) = \left[ 0.5 \delta_1(t) \sin(\delta_2) 0.5 \delta_2(t) \sin(\delta_1) \right]^T.
\]
with an unknown parameter $p$ belonging to $[-1, 1]$. When the parameter $p$ varies in $[-1, 1]$, $\Delta f(x)$ is unknown and uncertain with the expression $\Delta f(x) = g(x)\tau(\delta)$. According to Theorem 1, $D(\delta)$ is chosen as $D(\delta) = \|\delta(t)\|$ and $\theta = 1$. An adaptive control law is expected to render system (56) track its reference system (57) for all possible uncertainties.

According to the analysis in Section 2, it is obvious that $\iota(t) = 0$ because the uncertainties are only regarded to the transient errors. Thus, we have the nominal error system as

$$\dot{\delta}(t) = \begin{bmatrix} r_2 - r_1 \\ 0.5(r_1 f_2 - r_1) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 + r_1 \end{bmatrix} u(t) - \phi(t) + \begin{bmatrix} 0 \\ 0.5(\delta_1 \delta_2 - \delta_1 + r_2 \delta_1 + r_1 \delta_2) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \mu(\delta). \tag{58}$$

where the steady control law is

$$u(t) = \begin{bmatrix} 0 \\ 1 + r_1 \end{bmatrix}^{-1} \left( \phi(t) - \begin{bmatrix} 3 + r_2 \\ 0 \end{bmatrix} \right).$$

For the nominal error system (58), a feedback control law $\mu(\delta)$ needs to be obtained that minimizes the modified cost function (9) with $P = I_2$ and $Q = I_2$, where $I_2$ represents an identity matrix with two dimensions. The ADP-based iteration algorithm is used to approximately obtain the solution of the optimal control problem, where the critic network is used to estimate the optimal cost function, and the action network outputs the control law according to (39). In the implementation, the critic network is constructed with two input neurons, three hidden neurons and one output neuron. The hidden-to-output weight vector of the critic network is $\hat{\omega} = [\hat{\omega}_{c1} \hat{\omega}_{c2} \hat{\omega}_{c3}]^T$, which is regulated by using (31). The initial weights of the critic network are given between $-1$ and $1$. The learning rate of the critic network is chosen as $\eta_C = 0.05$. The initial state of the control plant is $x(0) = [-0.4 - 1]^T$, and the reference system is started from $r(t) = [0.2 - 0.2]^T$.

During the implementation of the policy iteration algorithm, the offline training process is required with the probing noises satisfying the persistency of excitation condition. As shown in Fig. 2, it can be observed that the convergence of the weights has occurred after 700 $s$, and thus the excitation signal is turned off. The weights of the critic network converge to $\hat{\omega}_c = [0.4203 0.5498 1.9286]^T$ in Fig. 2(a), and the trajectories of system (58) are depicted in Fig. 2(b). It can be seen that the error states converge to zero after turning off the excitation signal.

Next, the approximate optimal control law $\hat{\mu}^*(\delta)$ is obtained from the formula (39). Together with the steady control law $u_0(t)$, we can obtain the designed control law for uncertain system (56), which is $u(t) = u_0(t) + \hat{\mu}^*(\delta)$. In order to further investigate the robust performance of the controller, $\tau(\delta)$ is chosen as

$$\tau(\delta) = [0.5 \delta_1(t) \sin(\delta_2) - 0.5 \delta_2(t) \sin(\delta_1)]^T$$

by separately setting $p = 0.5$ and $p = -0.5$.

The initial state vector of the controlled system (56) is $x(0) = [-1.7 1.8]^T$, and the reference system starts from $r(0) = [-0.9 1]^T$. With this uncertainty, we apply the obtained control law $u_0(t) + \hat{\mu}^*(\delta)$ to system (56) for 30$s$ to test the robust performance of the controller.

Under the control law $\hat{\mu}^*(\delta)$, the error states under the uncertainties $\tau(\delta)$ converge to zero, as presented in Fig. 3(a), and the corresponding control law is shown in Fig. 3(b). The tracking performance of system (56) under the control law $u(t)$ is illustrated in Fig. 4(a) and (b), and the associated control curves are given in Fig. 5. In Fig. 6, it shows the relationship between $\tau(\delta)$ and $D(\delta)$, which illustrates that the bounded condition in Theorem 1 is satisfied. Thus, based on the analysis in Section 3.3, the original uncertain system is asymptotically stable if the approximate optimal control law can guarantee the bounded stability of its nominal system.

Case 2. In this case, a three-order affine nonlinear system is studied with the following formula $\dot{x}(t) = f(x) + g(x)(u(t) + \tau(t))$, where

$$f(x) = \begin{bmatrix} -0.5x_2 - 0.3x_1 \\ -0.8x_2 - x_1 x_3 \\ \sin(x_1) - x_3 \end{bmatrix}, \quad \tau(t) = \begin{bmatrix} px_1 \sin(x_2 x_3) \\ px_3 \\ x_1 \cos(p x_3) \end{bmatrix},$$

$$g(x) = \begin{bmatrix} 0.6 + 0.5 x_3 \\ 0.4 + 0.2 x_1 \\ 2 \end{bmatrix}.$$

The reference system is given as

$$\dot{r}(t) = \phi(t) = \begin{bmatrix} \sin(t) + 0.5 \cos(0.75t) \\ \cos(t) + 0.5 \sin(0.75t) \\ \sin(t) + 0.5 \cos(t) \end{bmatrix}.$$

In order to design the tracking controller, the steady control law $u_0(t)$ can be obtained as $u_0(t) = g^{-1}(r(t)(\phi(t) - f_1(r)), which is responsible for the steady tracking to the reference system. During the transient regulation, the feedback control law $\mu(\delta)$ is
Fig. 3. The control performance for the error system, (a) the convergence of the error system with uncertainties, (b) the transient control law $\hat{\mu}^*(\delta)$.

Fig. 4. The tracking control performance under the uncertainties for Case 1, (a) state $x_1(t)$ tracks to $r_1(t)$, (b) state $x_2(t)$ tracks to $r_2(t)$.

Fig. 5. The tracking control law $u(t)$.

designed to stabilize the transient error system. The sum of the steady control law and the transient control law is taken as the tracking control law for the three-order controlled system to follow its reference system. Therefore, the nominal error system is firstly provided as

$$\dot{\delta}(t) = f_1(r) + g_1(r)u(t) - \phi(t) + f_2(\delta) + g_2(\delta)\mu(\delta)$$

$$= \begin{bmatrix}
-0.5r_2 - 0.3r_1 \\
-0.8r_2 - r_1r_3 \\
r_1 - r_3 - \frac{2}{\delta^2}
\end{bmatrix}
+ \begin{bmatrix}
0.6 + 0.5r_3 & 0 & 1 \\
0 & 0.4 + 0.2r_1 & 0 \\
2 & 1 & 0.1
\end{bmatrix}$$
Theorem tracking control signal critic from consistent solution of the differential equation, \( D(\delta) = \| \delta(t) \| \) and \( \theta = 1 \).

Considering the nominal error system, a modified cost function in [9] with \( P = I_3 \) and \( Q = I_3 \) is used to formulate the optimal control problem, where an approximate optimal control law \( \mu(\delta) \) is obtained based on the ADP method by minimizing the cost function. The critic neural network is built to approximate the optimal control function with the structure of 3-6-1. The action network is used to calculate the feedback control law according to (39). The weights of the critic network, i.e., \( \hat{\omega}_c = \hat{\omega}_{c1} \hat{\omega}_{c2} \hat{\omega}_{c3} \hat{\omega}_{c4} \hat{\omega}_{c5} \hat{\omega}_{c6} \), are initially given in \([-1, 1]\), and then are updated by using (31). The learning rate of the critic network is selected as \( \eta_c = 0.1 \).

We first train the weights of the critic network with the persistent excitation signal. The three-order controlled system starts from \( x(0) = [1.1 \ 0.8 \ -1.1]^T \), and its reference system starts from \( r(0) = [0.3 \ 1 \ -0.3]^T \). Fig. 7 shows the weight updating of the critic network and the convergence of the nominal error system, respectively. It can be seen that the excitation signal lasts 900s and then is turned off. The weights converge to \( \hat{\omega}_c = [1.1849 \ 0.2778 -0.1700 0.6218 -0.1110 0.4958]^T \) under the persistent excitation signal in Fig. 7(a). In Fig. 7(b), the trajectories of the nominal error system under the excitation signal are provided. It can be observed that all error states can converge to zero when the excitation signal is cut off.

Based on the trained weights \( \hat{\omega}_c \), it can obtain the transient control law \( \hat{\mu}^*(\delta) \) according to (39). The steady control law \( u_i(t) \) can be derived from (6) with \( \phi(t), f_1(t) \) and \( g_1(t) \). Therefore, the tracking control law \( u(t) \) can be obtained for the three-order affine nonlinear system with uncertainties.

In order to validate the control performance of the designed controller, the controlled system is assumed to have \( \tau(\delta) = [\delta_1(t) \sin(\delta_2 \delta_3) \ \delta_3(t) \ \delta_1(t) \cos(\delta_3)]^T \). The initial states are set as \( x(0) = [-1.5 \ 0.3 \ -1.6]^T \) and \( r(0) = [-0.5 \ 0.2 \ -1]^T \). The curves of \( r(\delta) \) and \( D(\delta) \) are presented in Fig. 8 to illustrate that Theorem 1 is satisfied. Fig. 9(a) presents the error states converge to zero after 15 s with the control law \( \hat{\mu}^*(\delta) \) even that the controlled system contains the uncertainties, and Fig. 9(b)-(d) show the tracking performance of the controlled system. These results demonstrate the effectiveness of the designed tracking control method.

5. Conclusion

This paper copes with the adaptive tracking control problem of continuous-time nonlinear systems with matched uncertainties. By introducing a modified cost function, the tracking control problem is formulated to an optimal control problem of its associated nominal system. It has been proven that the tracking control can be achieved by applying the solution of the relevant HJB equation, which is approximated by the ADP-based method. The critic neural network is constructed and trained to derive the approximate optimal cost function, and then the action network outputs the approximate optimal control law, which is incorporated to the tracking control law together with the steady control law. The stability of the critic weights as well as the closed-loop system is analyzed in detail. Furthermore, the simulation results are provided to verify the effectiveness of the proposed tracking control strategy. In the future work, it is significant to intensively study the approximate optimal tracking control of nonlinear uncertain systems. There are several interesting problems that can be discussed, such as the approximate optimal tracking control problem of affine nonlinear systems including unmatched uncertainties, unknown
dynamics and control constraints [35,36]. Also, we will investigate some related topics on nonaffine nonlinear systems based on the ADP method.

References


Fig. 9. The tracking control performance under the uncertainties for Case 2, (a) the tracking error curves, (b) state $x_1(t)$ tracks to $r_1(t)$, (c) state $x_2(t)$ tracks to $r_2(t)$, (d) state $x_3(t)$ tracks to $r_3(t)$. 


Chaou Xu received the Ph.D. degree in control science and engineering from Southeast University, Nanjing, China, in 2012. She was a Visiting Ph.D. Student with the Royal Melbourne Institute of Technology University, Melbourne, VIC, Australia, from 2010 to 2011. She has been a Post-Doctoral Fellow with the Department of Electrical, Computer and Biomedical Engineering, The University of Rhode Island, Kingston, RI, USA, from 2014 to 2016. She is currently an Associate Professor with the School of Electrical and Information Engineering, Tianjin University, Tianjin, China. She has authored over 50 journal and conference papers, and co-authored one monograph. Her current research interests include nonlinear system control and optimization, adaptive and learning systems, and machine learning. She is the IEEE member and a member of the Asia-Pacific Neural Network Society.

Changyin Sun received the B.S. degree in 1996 from the Department of Mathematics, Sichuan University, China, the M.S. and Ph.D. degrees in Electrical Engineering from Southeast University, China, respectively in 2001 and 2003. From March 2004 to September 2004, he was a Postdoctoral Fellow in the Department of Computer Science, The Chinese University of Hong Kong. From April 2006 to March 2007 he worked as a Professor in the School of Electrical Engineering, Hohai University, Nanjing, China. From March 2007, he works as a Professor in the School of Automation, Southeast University, China. His research interests include Intelligent Control, Neural Networks, Theory and Design of Intelligent Control Systems, Optimization Algorithms, Pattern Recognition, etc. He is an Associate Editor of IEEE Transactions on Neural Networks (2010-) and Neural Processing Letters (2008-). He is also involved in organizing several international conferences.

Ding Wang received the B.S. degree in mathematics from Zhengzhou University of Light Industry, Zhengzhou, China, the M.S. degree in operations research and cybernetics from Northeastern University, Shenyang, China, and the Ph.D. degree in control theory and control engineering from Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2007, 2009, and 2012, respectively. He was a Visiting Scholar with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI, USA, from 2015 to 2017. He is currently an Associate Professor with The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interests include adaptive and learning systems, intelligent control, and neural networks. He has published over 90 journal and conference papers, and coauthored two monographs. He was a recipient of the Excellent Doctoral Dissertation Award of Chinese Academy of Sciences in 2013, and a nomination of the Excellent Doctoral Dissertation Award of Chinese Association of Automation (CAA) in 2014. He serves as an Associate Editor of IEEE Transactions on Neural Networks and Learning Systems and Neurocomputing. He is a member of Institute of Electrical and Electronics Engineers (IEEE), Asia-Pacific Neural Network Society (APNNS), and CAA.

Aiguo Song received the B.S. degree in Automatic Control in 1990, the M.S. degree in Measurement and Control in 1993 from Nanjing Aeronautics and Astronautics University, Nanjing, China, and the Ph.D. degree in Measurement and Control from Southeast University, Nanjing, China, 1996. He is currently a Professor and Director of Robot Sensor and Control Laboratory, and Dean of School of Instrument Science and Engineering, Southeast University, China. His research expertise and interests are in the areas of telerobot, rehabilitation robot, human computer interface, robot force/tactile sensor, haptic display, and signal processing. He has published over 160 refereed journal and conference papers and held 45 patents. He is a senior member of IEEE Robotics and Automation Society.