J Electr Eng Technol.2018; 13(4): 1740-1751
http://doi.org/10.5370/JEET.2018.13.4.1740

ISSN(Print) 1975-0102
ISSN(Online) 2093-7423

Policy Iteration Algorithm Based Fault Tolerant Tracking Control: An Implementation on Reconfigurable Manipulators

Yuanchun Li*, Hongbing Xia*, and Bo Zhao†**

Abstract – This paper proposes a novel fault tolerant tracking control (FTTC) scheme for a class of nonlinear systems with actuator failures based on the policy iteration (PI) algorithm and the adaptive fault observer. The estimated actuator failure from an adaptive fault observer is utilized to construct an improved performance index function that reflects the failure, regulation and control simultaneously. With the help of the proper performance index function, the FTTC problem can be transformed into an optimal control problem. The fault tolerant tracking controller is composed of the desired controller and the approximated optimal feedback one. The desired controller is developed to maintain the desired tracking performance at the steady-state, and the approximated optimal feedback controller is designed to stabilize the tracking error dynamics in an optimal manner. By establishing a critic neural network, the PI algorithm is utilized to solve the Hamilton-Jacobi-Bellman equation, and then the approximated optimal feedback controller can be derived. Based on Lyapunov technique, the uniform ultimate boundedness of the closed-loop system is proven. The proposed FTTC scheme is applied to reconfigurable manipulators with two degree of freedoms in order to test the effectiveness via numerical simulation.

Keywords: Adaptive dynamic programming, Policy iteration, Fault tolerant tracking control, Reconfigurable manipulators, Neural network.

1. Introduction

As the rapid development of modern technologies, industrial systems are becoming increasingly complex and large-scale, the unavoidable system failures can affect the product quality, damage equipments, or even harm to human beings [1]. As we know, various components such as actuators, sensors and processors may undergo abrupt failures individually or simultaneously during operation. Among all kinds of possible failures, actuator failures are considered as one of the most critical challenges, mainly for the reason that the control performance can be deteriorated by unexpected and unknown actuator actions. Hence, it is urgent to develop fault tolerant control (FTC) methods to deal with such kind of problems.

To achieve higher reliability and better control performance, many research efforts on FTC systems have been made during the past three decades to ensure systems stable and maintain acceptable control [2]. FTC architectures can be classified into passive FTC (PFTC) and active FTC (AFTC). Studies on the significance of FTC are found in [3] and [4], whilst the book [5] provided a useful theoretical framework. For PFTC, a fixed control law is designed to achieve stability and acceptable performance in both normal and fault situations [6, 7]. However, the drawback of PFTC is that it is reliable only for systems with known faults. While AFTC can reconfigure or reconstruct the controller online to recover the system stability and control performance via fault detection and identification (FDI) [8, 9]. It is worthy to mention that the observer technique, which is active in reconfiguring control, plays an important role in achieving active fault tolerance [10-12].

The trajectory tracking control is always required in mobile robots, helicopters, and so on [13, 14], and its goal is to make the system outputs track specified desired trajectories. In spite of the tracking control is widely used, but only a few advanced techniques based FTC are applied for tracking control to ensure system tracking performance when components are faulty. For examples, Liao et al. [15] used a linear matrix inequality (LMI) method to propose a reliable robust tracking control scheme for uncertain discrete-time systems. And inspired by it, Yao et al. [16] applied this control scheme to continuous-time systems. Besides, Jiang et al. [17] proposed an adaptive FTTC scheme based on online estimation for near-space-vehicle (NSV) attitude dynamics systems with actuator failures.

As an effective tool to solve optimal control problems in nonlinear systems, adaptive dynamic programming (ADP) is a useful approximation method proposed by Werbos [18] to obtain the solution to the HJB equation. By using ADP method, the designed controller structure is simple and computational burden is reduced [19]. In recent years,

† Corresponding Author: The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. (zhaobo@ia.ac.cn)
* Department of Control Science and Engineering, Changchun University of Technology, Changchun 130012, China.
** Hytera Communications Co., Ltd, , Nanjing 210000, China.
Received: November 1, 2016; Accepted: April 20, 2018

many scholars have studied ADP approaches in both theory and applications [20-23]. The optimal control methods based on iterative ADP have been developed for both discrete-time [21, 24] and continuous-time systems [20, 25, 27, 29]. Meanwhile, the policy iteration (PI) algorithm is an iterative approach to solve HJB equation by constructing a sequence of stable control policies that converge to optimal control solutions [26]. Wang et al. [27] studied a PI algorithm to handle the robust control problem for nonlinear systems with uncertainties. Based on PI algorithm, Song et al. [28] developed an off-policy integral reinforcement learning algorithm to compensate the disturbances for unknown systems. Besides, Zhang *et al*. [29] proposed an ADP based data-driven robust approximated optimal control algorithm for unknown nonlinear systems. It was the first solution to the tracking control problem of unknown general nonlinear systems based on ADP.

For FTC, only a few results were presented based on ADP or reinforcement learning methods. Chang *et al.* [30] presented an online fault tolerant actor-critic neuro-control scheme for continuous-time nonlinear systems with abrupt nonlinear faults. In [31], in the light of time-varying actuator gain and bias faults, an ADP based FTC scheme was proposed. Wang et al. [32] investigated a RL-based FTC scheme for a class of multiple-input-multiple-output (MIMO) nonlinear discrete-time systems. However, we can see from these researches, most of them were aimed at discrete-time systems or stabilization control, only a few related results for continuous-time systems.

In actual industrial processes, many systems are modeled as nonlinear systems such as robot manipulators [33, 34], missile systems [35], biochemical processes [36] and so on. As a kind of nonlinear systems, reconfigurable manipulators have great application potential in space explorations, smart manufacturing, high risk operations, battle fields and so forth. Compared with traditional manipulators with fixed configuration, reconfigurable manipulators can adapt to diverse task requirements due to their low cost, convenient modification, easy maintenance, portability and durability [37, 38]. Some FTC schemes have been carried out for reconfigurable manipulators. Ahmad et al. [39] investigated a distributed fault detection scheme, which does not require motion states of any other modules. Zhao et al. [40] studied an active FTC scheme for reconfigurable manipulators based on local joint information. While regarding FTC, we should not only focus on guaranteeing the tolerance of fault dynamics, but also the requirements of optimal control performance. Recently, although some researches [20-25] have been developed to study the optimal regulations or control problems for nonlinear systems based ADP, only a few FTTC schemes have been reported, this motivates our research.

To solve this difficult problem, motivated by the aforementioned analysis and observer techniques, in this paper, a fault tolerant tracking (FTT) controller based on ADP with an improved performance index function is investigated for a class of nonlinear systems with actuator failures. The estimated fault from an adaptive fault observer is used to construct the improved performance index function, which reflects the actuator failure, regulation and control. Hence, the FTTC problem is transformed into an optimal control problem. Then, the FTTC scheme based on PI algorithm is designed to deal with the optimal control problem. The improved HJB equation is solved by a critic neural network, whose weights can be derived by gradient descent method. The stability of closed-loop system with actuator failures is analyzed using Lyapunov theorem. Finally, in order to show that the designed controller is effective, two 2-DOF (Degree of Freedom) reconfigurable manipulators with different configurations are used to simulation.

The main contributions of the presented scheme can be summarized as follows.

1) By designing an adaptive fault observer, the estimated unknown actuator failure can be employed to construct the improved performance index function, which reflects the actuator failure, regulation and control. Thus, the FTTC problem can be transformed into an optimal control problem.

2) The control policy can be derived depending only on the critic neural network, and the training of the action neural network is no longer required. Thus, the designed controller structure is simple and the computational burden is reduced. To achieve FTTC, the steady-state control can be achieved within a small set with an approximated optimal control.

The remainder of this paper is organized as follows. Section 2 describes the problem, optimal control and HJB equations. The proposed FTTC scheme is presented in Section 3 and simulations are shown in Section 4. Finally, conclusions are summarized in Section 5.

2. Problem Formulation

Consider the following continuous-time nonlinear system with actuator failures:

$$\dot{x} = f(x) + g(x)(u + f_a) \tag{1}$$

where $x \in R^n$ is the system state vector, $u \in R^m$ is the control input vector, $f_a \in R^m$ is an unknown additive actuator failure. $f(\cdot)$ and $g(\cdot)$ locally Lipchitz and differentiable in their arguments with $f(0) = 0$, and let $x(0) = x_0$ be the initial state.

In [29], the authors mainly studied the trajectory tracking control, but this paper will address a novel FTTC scheme so as to guarantee the closed-loop system (1) to be stable in case of the system suffers from unknown actuator failures. In other words, the main objective is to find an optimal controller u, which enables the system state vector x tracks the desired trajectory x_d even though actuator failures occur.

In order to achieve objective above, inspired by [29], we design the improved infinite horizon performance index function as

$$J(e(t)) = \int_0^\infty \Big(\rho \hat{f}_a^T(\tau)\hat{f}_a(\tau) + \gamma\big(\nabla J^*(e(\tau))\big)^2 + r\big(e(\tau), u_e(e(\tau))\big)\Big) d\tau \tag{2}$$

where $e = x - x_d$ is defined as the tracking error, $r(e,u) = e^T Q e + u_e^T R u_e$ is the utility function, $r(e,u_e) \geq 0$ for all e and u_e with $r(0,0) = 0$, $Q \in R^{n\times n}$, $R \in R^{m\times m}$ are positive definite matrices, $\hat{f}_a(t)$ is the estimation of the actuator failure $f_a(t)$, $\rho > 0$ and $\gamma > 0$ are positive constants.

As in many literature, the following assumptions are given for the system (1).

Assumption 1: The desired trajectory x_d, and its gradient $\dot{x}_d$ are continuous and bounded.

Assumption 2 [10]: The system actuator failure f_a is bounded as $\|f_a\| \leq \varepsilon_1 < \infty$, where ε_1 is a positive constant.

The desired trajectory can be expressed as

$$\dot{x}_d = f(x_d) + g(x_d)u_d \tag{3}$$

Since the desired controller u_d is utilized to ensure the tracking error to a steady state, it can be obtained by (3) as

$$u_d = g^+(x_d)\big(\dot{x}_d - f(x_d)\big) \tag{4}$$

where $g^+(\cdot)$ denotes the generalized inverse of $g(\cdot)$, and the control coefficient matrix function $g(x_d)$ satisfies $g^+(x_d)\times g(x_d)=I$ and $I \in R^{n\times n}$ is the identity matrix.

The trajectory tracking error can be defined as

$$e = x - x_d \tag{5}$$

The time derivative of (5) is

$$\begin{aligned} \dot{e} &= \dot{x} - \dot{x}_d \\ &= f(x) - f(x_d) + g(x)(u - f_a) - g(x_d)u_d \end{aligned} \tag{6}$$

By defining $f_e = f(x) - f(x_d)$, it is worth noticing that the FTT controller u is composed of the desired (steady-state) controller u_d and the approximated optimal feedback one u_e, i.e., $u = u_d + u_e$. Thus, (6) can be rewritten as

$$\begin{aligned} \dot{e} &= f(x) - f(x_d) + g(x)(u - f_a) - g(x_d)u_d \\ &= f_e + [g(x) - g(x_d)]u + g(x_d)u_e - g(x)f_a \end{aligned} \tag{7}$$

The feedback controller u_e is used to make the tracking error dynamics converge to a stable state at transient stage in an optimal manner [29]. Then, the improved infinite horizon performance index function (2) is expressed as

$$\underset{u_e \in \Psi(\Omega)}{J(e(\tau))} = \int_0^\infty \Big(\rho \hat{f}_a^T(\tau)\hat{f}_a(\tau) + \gamma\big(\nabla J^*(e(\tau))\big)^2 + r\big(e(\tau), u_e(e(\tau))\big)\Big) d\tau \tag{8}$$

Definition 1 [29, 41]: For the system (1) with $f_a = 0$, a control policy μ_e is defined to be admissible with respect to (8) on a compact set Ω, if μ_e is continuous on a set Ω with $\mu_e(0) = 0$, μ_e stabilizes (7) on Ω, and $J(e)$ is finite for $\forall e \in \Omega$.

The remainder problem can be presented as: for the error system (7) with a series of admissible control set $\mu_e \in \Psi(\Omega)$ and the improved infinite horizon performance index function (8), find an admissible control policy u_e such that performance index function (8) associated with system (7) is minimized.

If the improved infinite horizon performance index function

$$V(e(\tau)) = \int_0^\infty \Big(\rho \hat{f}_a^T(\tau)\hat{f}_a(\tau) + \gamma\big(\nabla V^*(e(\tau))\big)^2 + r\big(e(\tau), u_e(e(\tau))\big)\Big) d\tau \tag{9}$$

is continuously differentiable, then the infinitesimal version of (9) is the so-called nonlinear Lyapunov equation [25] as

$$\begin{aligned} 0 = {} & \rho \hat{f}_a^T \hat{f}_a + \gamma\big(\nabla V^*(e)\big)^2 + r(e,u_e) + \big(\nabla V(e)\big)^T \\ & \times \big(f_e + [g(x) - g(x_d)]u + g(x_d)u_e\big) \end{aligned} \tag{10}$$

where $V(0) = 0$, and the term $\nabla V(e)$ denotes the partial derivative of $V(e)$ with respect to e, i.e. $\nabla V^T(e) = \dfrac{\partial V(e)}{\partial e}$.

Define the Hamiltonian function and the optimal performance index function as

$$\begin{aligned} H\big(e, u_e, \nabla V(e)\big) = {} & \rho \hat{f}_a^T \hat{f}_a + \gamma\big(\nabla V^*(e)\big)^2 + r(e,u_e) \\ & + \big(\nabla V(e)\big)^T \big(f_e + [g(x) - g(x_d)]u \\ & + g(x_d)u_e\big) \end{aligned} \tag{11}$$

and

$$J^*(e(\tau)) = \min_{u_e} \int_0^\infty \Big(\rho \hat{f}_a^T(\tau)\hat{f}_a(\tau) + \gamma\big(\nabla J^*(e(\tau))\big)^2 + r\big(e(\tau), u_e(e(\tau))\big)\Big) d\tau \tag{12}$$

and let $J^*(e)$ be the optimal performance index function, then

$$0=\min_{u_e} H\left(e,u_e,\nabla J^*(e)\right) \tag{13}$$

where $\nabla J^{*T}(e)=\dfrac{\partial J^*(e)}{\partial e}$, if the optimal solution $J^*(e)$ exists and is continuously differentiable, the PI based optimal control policy can be expressed as

$$u_e^*=-\frac{1}{2}R^{-1}g^T(x)\nabla J^*(e) \tag{14}$$

Combining (10) with (13), we have

$$\begin{aligned}&\left(\nabla J(e)\right)^T\left(f_e+[g(x)-g(x_d)]u+g(x_d)u_e\right)\\&=-e^TQe-u_e^TRu_e-\rho\hat{f}_a^T\hat{f}_a-\gamma\left(\nabla J^*(e)\right)^2\end{aligned} \tag{15}$$

Remark 1: In order to obtain the optimal control policy (14), we first need to solve the HJB (13), then substituting the solution $\nabla J^*(e)$ into (14) to obtain the optimal control policy. For linear systems, the performance index function function is a quadratic form of state and control input, the optimal solution can be obtained by solving a standard Riccati equation. While for nonlinear systems, the difficult HJB equation is solved instead of a Riccati equation. So we use neural network to approximate the optimal performance index function [25, 27-30, 42].

The design process of FTTC can be divided into following two steps:

1) By designing an adaptive fault observer, the estimated unknown actuator failure is employed to construct the improved performance index function, which reflects the actuator failure, regulation and control. Hereafter, the optimal feedback controller is derived by employing a critic neural network.

2) By combining the desired controller with the optimal feedback controller, an optimal FTTC scheme based on PI algorithm is designed.

3. Fault Tolerant Tracking Controller Design and Stability Analysis

3.1 Adaptive fault observer design

Assumption 3: The actuator fault estimation error $e_a=f_a-\hat{f}_a$ is norm-bounded as $\|e_a\|\le\varepsilon_2$, where ε_2 is a positive constant.

For the system with actuator failures (1), we can develop an adaptive fault observer as

$$\dot{\hat{x}}=f(\hat{x})+g(\hat{x})(u-\hat{f}_a)+\alpha_1(x-\hat{x}) \tag{16}$$

where $\hat{x}$ is the observation of the system state x, α_1 is the positive definite observer gain matrix, $\hat{f}_a$ is the estimated actuator failure which can be updated by

$$\dot{\hat{f}}_a=-\alpha_2 g^T(\hat{x})e_F \tag{17}$$

where α_2 is a positive definite matrix, and $e_F=x-\hat{x}$ is the state observation error.

Combining (1) with (16), we have

$$\dot{e}_F=e_f+e_g(u-f_a)-g(\hat{x})(f_a-\hat{f}_a)-\alpha_1 e_F \tag{18}$$

where $e_f=f(x)-f(\hat{x})$ and $e_g=g(x)-g(\hat{x})$ are the observation errors of $f(x)$ and $g(x)$, respectively.

Assumption 4: Defining $\xi=e_f+e_g(u-f_a)$, it is norm-bounded as $\|\xi\|\le\varepsilon_3$, where ε_3 is a positive constant.

Remark 2: In real applications, all signals should be bounded, and the functions of the controllable nonlinear systems are bounded, so it is reasonable to assume that the random unknown actuator failure f_a and ξ are bounded.

Theorem 1. For the system (1) with actuator failures with Assumptions 3 and 4, the fault observation error is guaranteed to be UUB by the developed fault observer (16) with the adaptive law (17).

Proof: Select the Lyapunov function candidate as

$$V_1(t)=\frac{1}{2}e_F^Te_F+\frac{1}{2}e_a^T\alpha_2^{-1}e_a \tag{19}$$

Substituting (18) into the time derivative of (19), we have

$$\begin{aligned}\dot{V}_1(t)&=e_F^T\left(e_f+e_g(u-f_a)-g(\hat{x})(f_a-\hat{f}_a)-\alpha_1e_F\right)-\dot{\hat{f}}_a^T\alpha_2^{-1}e_a\\&\le\varepsilon_3\|e_F\|-e_F^Tg(\hat{x})e_a-\lambda_{\min}(\alpha_1)\|e_F\|^2-\dot{\hat{f}}_a^T\alpha_2^{-1}e_a\\&=-\left(\lambda_{\min}(\alpha_1)\|e_F\|-\varepsilon_3\right)\|e_F\|-\left(e_F^Tg(\hat{x})+\dot{\hat{f}}_a^T\alpha_2^{-1}\right)e_a\end{aligned} \tag{20}$$

where $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue of the matrix, and substituting the adaptive law (17) into (20), it becomes

$$\dot{V}_1(t)\le-\left(\lambda_{\min}(\alpha_1)\|e_F\|-\varepsilon_3\right)\|e_F\| \tag{21}$$

We can observe that $\dot{V}_1(t)\le 0$ when e_F lies outside the compact set $\Omega_1=\left\{e_F:\|e_F\|\le\dfrac{\varepsilon_3}{\lambda_{\min}(\alpha_1)}\right\}$. Therefore, the fault observation error is UUB. This completes the proof.

3.2 Online policy iteration algorithm

Motivated by [25], the PI algorithm consists of policy evaluation based on (10) and policy improvement based on (14). Specifically, its iteration procedure can be presented as Algorithm 1.

Algorithm 1 Online policy iteration

Step 1: Let $i=0$, begin with an initial admissible control policy $u_e^{(i)}$, and select a small positive constant δ.
Step 2: Let $i>0$, based on the control policy $u_e^{(i)}$, solve $V^{(i)}$ from

$$0=\rho\hat{f}_a^T\hat{f}_a+\gamma\left(\nabla V^{(i+1)}(e)\right)^2+r(e,u_e^{(i)})+\left(\nabla V^{(i+1)}(e)\right)^T\left(f_e+g(x)u_e^{(i)}\right),$$

with $V^{(i+1)}=0$.
Step 3: Update the control policy $u_e^{(i)}$ via

$$u_e^{(i+1)}=-\frac{1}{2}R^{-1}g^T(x)\nabla V^{(i+1)}(e).$$

Step 4: If $\left\|V^{(i+1)}(e)-V^{(i)}(e)\right\|\le\delta$, stop and obtain the approximate optimal control; else, let $i=i+1$ and return to Step 2.

This algorithm will converge to the optimal performance index function and optimal control policy, i.e., $V^{(i)}(e)\to J^*(e)$ and $u_e^{(i)}\to u_e^*$ as $i\to\infty$ [42].

3.3 NN approximation for critic neural network

As it is well known, neural networks are powerful for approximating nonlinear functions. Since the performance index function is usually highly nonlinear and nonanalytic, a critic neural network can be utilized to approximate the performance index function $V(e)$. The ideal critic neural network is expressed as

$$V(e)=W_c^T\sigma_c(e)+\varepsilon_c \tag{22}$$

where $W_c\in R^N$ is the ideal weight vector, N is the number of neurons in the hidden-layer, $\sigma_c(e)$ is the activation function, and ε_c is the critic neural network approximation error. Thus, the partial gradient of $V(e)$ respect to *e* is

$$\nabla V(e)=\left(\nabla\sigma_c(e)\right)^T W_c+\nabla\varepsilon_c \tag{23}$$

where $\nabla\sigma_c(e)=\dfrac{\partial\sigma_c(e)}{\partial e}\in R^{N\times n}$ and $\nabla\varepsilon_c$ are the partial gradients of the activation function and the approximation error, respectively.

For the system (1) with actuator failures, (i.e., $f_a\neq 0$), combining (10) with (23), we have

$$0=\rho\hat{f}_a^T\hat{f}_a+\gamma\left(\left(\nabla\sigma_c(e)\right)^T W_c+\nabla\varepsilon_c\right)^T\times\left(\left(\nabla\sigma_c(e)\right)^T W_c+\nabla\varepsilon_c\right)+r(e,u_e)+\left(\left(\nabla\sigma_c(e)\right)^T W_c+\nabla\varepsilon_c\right)^T\dot{e} \tag{24}$$

Thus, the Hamiltonian function can be expressed as

$$\begin{aligned}H&\left(e,u_e,W_c\right)\\&=\rho\hat{f}_a^T\hat{f}_a+\gamma\left(\left(\nabla\sigma_c(e)\right)^T W_c+\nabla\varepsilon_c\right)^T\\&\quad\times\left(\left(\nabla\sigma_c(e)\right)^T W_c+\nabla\varepsilon_c\right)+r(e,u_e)+\left(W_c^T\nabla\sigma_c(e)\right)\dot{e}\\&=-\nabla\varepsilon_c^T\dot{e}=e_{cH}\end{aligned} \tag{25}$$

where e_{cH} is the residual error due to the neural network approximation.

Since the ideal weight vector W_c is unknown, the critic neural network can be approximated by

$$\hat{V}(e)=\hat{W}_c^T\sigma_c(e) \tag{26}$$

Then, the partial gradient of $\hat{V}(e)$ can be expressed as

$$\nabla\hat{V}(e)=\left(\nabla\sigma_c(e)\right)^T\hat{W}_c \tag{27}$$

Thus, the approximate Hamiltonian function can be obtained as

$$\begin{aligned}H\left(e,u_e,\hat{W}_c\right)&=\rho\hat{f}_a^T\hat{f}_a+\gamma\left(\hat{W}_c^T\sigma_c(e)\right)^T\times\left(\hat{W}_c^T\sigma_c(e)\right)\\&\quad+r(e,u_e)+\left(\hat{W}_c^T\nabla\sigma_c(e)\right)\dot{e}=e_c\end{aligned} \tag{28}$$

Denoting $\vartheta=\nabla\sigma_c(e)\dot{e}$, and we assume $\|\vartheta\|\le\vartheta_M$, where $\vartheta_M>0$. We employ the objective function $E_c=\dfrac{1}{2}e_c^Te_c$ to be minimized by the gradient descent algorithm in order to tune the critic neural network weight vector $\hat{W}_c$, which should be updated by

$$\dot{\hat{W}}_c=-\alpha_3e_c\vartheta \tag{29}$$

where $\alpha_3>0$ is the learning rate of the critic neural network.

Define the weight approximation error as

$$\tilde{W}_c=W_c-\hat{W}_c \tag{30}$$

From (25), (28) and (30), one has

$$e_c=e_{cH}-\tilde{W}_c^T\vartheta \tag{31}$$

The weight approximation error can be updated by

$$\dot{\tilde{W}}_c=-\dot{\hat{W}}_c=\alpha_c(e_{cH}-\tilde{W}_c^T\vartheta)\vartheta \tag{32}$$

Hence, according to (14) and (23), the ideal control policy can be described as

$$u_e = -\frac{1}{2}R^{-1}g^T(x)\left(\left(\nabla\sigma_c(e)\right)^T W_c + \nabla\varepsilon_c\right) \tag{33}$$

And it can be approximated as

$$\hat{u}_e = -\frac{1}{2}R^{-1}g^T(x)\left(\nabla\sigma_c(e)\right)^T \hat{W}_c \tag{34}$$

Remark 3: From (34), we can observe that the control policy can be derived only depending on the critic neural network, and the training of the action neural network is no longer required.

Theorem 2. Consider the system (1) without actuator failures, the weights of the critic neural network are updated by (29), the weight approximation error is UUB.

Proof: Select the Lyapunov function candidate as

$$V_2 = \frac{1}{2\alpha_3}\tilde{W}_c^T\tilde{W}_c \tag{35}$$

Its time derivative is

$$\begin{aligned}\dot{V}_2 &= \frac{1}{\alpha_3}\tilde{W}_c^T\dot{\tilde{W}}_c \\ &= \tilde{W}_c^T(e_{cH} - \tilde{W}_c^T\vartheta)\vartheta \\ &= \tilde{W}_c^T e_{cH}\vartheta - \left\|\tilde{W}_c^T\vartheta\right\|^2 \le \frac{1}{2}e_{cH}^2 - \frac{1}{2}\left\|\tilde{W}_c^T\vartheta\right\|^2\end{aligned} \tag{36}$$

Hence, we can observe that $\dot{V}_2 \le 0$ when $\tilde{W}_c$ lies outside the compact set $\Omega_2 = \left\{\tilde{W}_c : \left\|\tilde{W}_c\right\| \le \frac{e_{cH}}{\vartheta_M}\right\}$. Therefore, the weight approximation error is UUB. This completes the proof.

3.4 Stability analysis

Motivated by the aforementioned analysis, the FTT controller consists of the desired controller u_d and the optimal feedback controller u_e. So the control input is written as

$$u = u_d + \hat{u}_e \tag{37}$$

where u_d is obtained directly by (4).

Theorem 3. Assuming that the neural network based HJB solution of the optimal control problem exists, for the considered system (1) and the desired trajectory given by (3), the approximated FTTC policy (37) can guarantee the closed-loop system to be UUB with the improved performance index function (2).

Proof: Select the Lyapunov function candidate as

$$V_3 = \frac{1}{2}e^T e + J^*(e) \tag{38}$$

Its time derivative is

$$\begin{aligned}\dot{V}_3 &= e^T\dot{e} + \left(\nabla J^*(e)\right)^T\dot{e} \\ &= e^T\left[f_e + [g(x) - g(x_d)]u + g(x_d)u_e\right. \\ &\quad \left. - \left(\nabla J^*(e)\right)^T g(x)f_a + \left(\nabla J^*(e)\right)^T\right. \\ &\quad \left. \times\left[f_e + [g(x) - g(x_d)]u + g(x_d)u_e\right]\right.\end{aligned} \tag{39}$$

According to (15), (39) becomes

$$\begin{aligned}\dot{V}_3 &= e^T\dot{e} + \left(\nabla J^*(e)\right)^T\dot{e} \\ &= -e^T g(x)f_a + e^T\left[f_e + [g(x) - g(x_d)]u\right. \\ &\quad \left. + g(x_d)u_e\right] - \left(\nabla J^*(e)\right)^T g(x)f_a - e^T Qe \\ &\quad - u_e^T Ru_e - \rho\hat{f}_a^T\hat{f}_a - \gamma\left(\nabla J^*(e)\right)^2\end{aligned} \tag{40}$$

As $f(x)$ is locally Lipchitz, there exists positive constant L_f, s.t. $\|f_e\| \le L_f\|e\|$. Assuming $\|g(x)\| \le K_g$, $\|g(x_d)\| \le K_{gd}$, then we have $\|g(x) - g(x_d)\| \le \Delta K$, by Young's inequality, we can obtain

$$\begin{aligned}\dot{V}_3 &\le \frac{1}{2}\|e\|^2 + \frac{1}{2}K_g^2\|f_a\|^2 + L_f\|e\|^2 + \Delta K\|u_d + u_e\|\|e\| \\ &\quad + K_{gd}\|u_e\|\|e\| + \frac{1}{2}\left(\nabla J^*(e)\right)^2 + \frac{1}{2}K_g^2\|f_a\|^2 \\ &\quad - e^T Qe - u_e^T Ru_e - \rho\hat{f}_a^T\hat{f}_a - \gamma\left(\nabla J^*(e)\right)^2 \\ &\le -\left[\left(\lambda_{\min}(Q) - L_f - \frac{3}{2}\right)\|e\| - \Delta K\|u_d\|\right]\|e\| \\ &\quad - \left(\lambda_{\min}(R) - \frac{1}{2}K_{gd}^2 - \frac{1}{2}\Delta K^2\right)\|u_e\|^2 \\ &\quad - (\rho - K_g^2)\hat{f}_a^T\hat{f}_a + K_g^2(\|f_a\|^2 - \hat{f}_a^T\hat{f}_a) \\ &\quad - \left(\gamma - \frac{1}{2}\right)\left(\nabla J^*(e)\right)^2\end{aligned} \tag{41}$$

By Assumptions 2 and 3, and assuming $\|u_d\| \le \eta$, where η is a positive constant, one can obtain

$$\begin{aligned}\dot{V}_3 &\le -\left(\psi_1\|e\| - \Delta K\eta - \frac{\varepsilon_l}{\|e\|}\right)\|e\| - \psi_2\|u_e\|^2 \\ &\quad - \left(\rho - K_g^2\right)\hat{f}_a^T\hat{f}_a - \left(\gamma - \frac{1}{2}\right)\left(\nabla J^*(e)\right)^2\end{aligned} \tag{42}$$

where $\varepsilon_l = K_g^2\left(2\varepsilon_1 + \varepsilon_2\right)\varepsilon_2$, $\psi_1 = \lambda_{\min}(Q) - L_f - \frac{3}{2}$ and $\psi_2 = \lambda_{\min}(R) - \frac{1}{2}K_{gd}^2 - \frac{1}{2}\Delta K^2$. Hence, we can obtain that $\dot{V}_3(t) \le 0$ when e lies outside the compact set $\Omega_3 =$

$\left\{ e : \|e\| \le \sqrt{\frac{2\Delta K^2 \eta^2 + 4\psi_1 \varepsilon_l}{4\psi_1^2}} \right\}$ if the following conditions hold:

$$\begin{cases} \lambda_{\min}(Q) \ge L_f + \frac{3}{2} \\ \lambda_{\min}(R) \ge \frac{1}{2} K_{gd}^2 + \frac{1}{2} \Delta K^2 \\ \rho \ge K_g^2 \\ \gamma \ge \frac{1}{2} \end{cases} \tag{43}$$

Therefore, the trajectory tracking error of the closed-loop system under FTTC scheme are UUB. This completes the proof.

Remark 4: In [43], an observer-based FTC scheme via PI algorithm was proposed for stabilization problems of nonlinear systems. By contrast, this paper extends ADP method to solve FTTC problems. By constructing a different performance index function, which reflects the failure, regulation and trajectory tracking control simultaneously, the critic neural network is established to solve the HJB equation to derive the approximated feedback control. And combining with the desired control, the FTTC is realized for reconfigurable manipulators suffer from actuator failures.

4. Simulation Study

Reconfigurable manipulator systems which are regarded as a kind of nonlinear systems are employed to verify the effectiveness of the proposed FTC scheme. The dynamic model of *n*-DOF reconfigurable manipulators with actuator failures in joint space coordination [40] is described as

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) = u - f_a \tag{44}$$

where $q \in R^n$ is the vector of joint positions, $M(q) \in R^{n\times n}$ is the inertia matrix, $C(q,\dot{q})\dot{q} \in R^n$ is the Coriolis and centripetal force, $G(q) \in R^n$ is the gravity term, and $u \in R^m$ is the applied joint torque, f_a is the additive actuator failure.

Now, two 2-DOF reconfigurable manipulators with different configurations shown in Fig. 1 for simulation.

The dynamic model of Configuration a is described as

$$M(q) = \begin{bmatrix} 0.36\cos(q_2) + 0.6066 & 0.18\cos(q_2) + 0.1233 \\ 0.18\cos(q_2) + 0.1233 & 0.1233 \end{bmatrix}$$

$$C(q,\dot{q}) = \begin{bmatrix} -0.36\sin(q_2)q_2 & -0.18\sin(q_2)\dot{q}_2 \\ 0.18\sin(q_2)(\dot{q}_1 - \dot{q}_2) & 0.18\sin(q_2)\dot{q}_1 \end{bmatrix}$$

$$G(q) = \begin{bmatrix} -5.88\sin(q_1 + q_2) - 17.64\sin(q_1) \\ -5.88\sin(q_1 + q_2) \end{bmatrix}$$

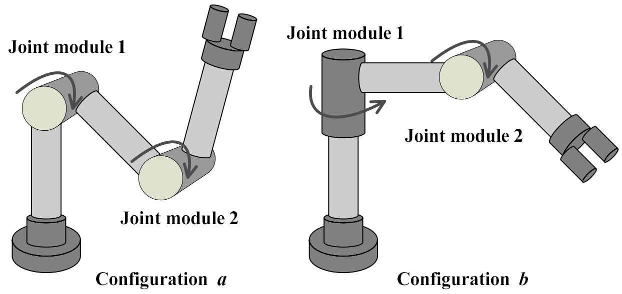

Fig. 1. Configurations of 2-DOF reconfigurable manipulators

Let $x = [x_1, x_2, x_3, x_4]^T = [q_1, q_2, \dot{q}_1, \dot{q}_2]^T$, by simple transformation, the system dynamic model can be expressed as (1) with $f(x) = [0;0;\phi_1(x)]^T$, $g(x) = [0,0;0,0;\phi_2(x)]^T$ where $\phi_1(x) = M^{-1}(q)\left[-C(q,\dot{q})\dot{q} - G(q)\right]$ and $\phi_2(x) = M^{-1}(q)$, $u = [u_1, u_2]^T \in R^2$ is the control input vector, the term $f_a = [f_{a1}, f_{a2}]^T \in R^2$ reflects an unknown additive actuator failure. We choose

$$f_a(t) = \begin{cases} [0;0], & 0 \le t \le 30s \\ [5;0], & 30s < t \le 60s \end{cases} \tag{45}$$

Define the desired trajectory as

$$\begin{aligned} x_{1d} &= 0.4\sin(0.3t) - 0.1\cos(0.5t) \\ x_{2d} &= 0.6\sin(0.2t) + 0.3\cos(0.6t) \end{aligned} \tag{46}$$

Let the initial state be $x = [1;1;0;0]$, the initial observed state be $x = [2;2;0;0]$. We employ a critic neural network to approximate the performance index function, and its weight vector is denoted as $\hat{W}_c = \left[\hat{W}_{c1}, \hat{W}_{c2}, \ldots, \hat{W}_{c10}\right]^T$, whose initial value is $\hat{W}_{c0} = \left[20,30,40,20,30,40,50,40,50,55\right]^T$, the activation function of the critic neural network is chosen as $\sigma_c = \left[e_1^2, e_1e_2, e_1e_3, e_1e_4, e_2^2, e_2e_3, e_2e_4, e_3^2, e_3e_4, e_4^2\right]^T$. The initial value of the actuator failure is chosen as $f_{a0} = [0;0]$. Some related control parameters are listed in Table 1.

Table 1. Parameter list

Variable	Value
Q	$5\times I_4$
R	$0.1\times I_2$
α_1	$20\times I_4$
α_2	$50\times I_2$
α_3	0.001
ρ	2
γ	1

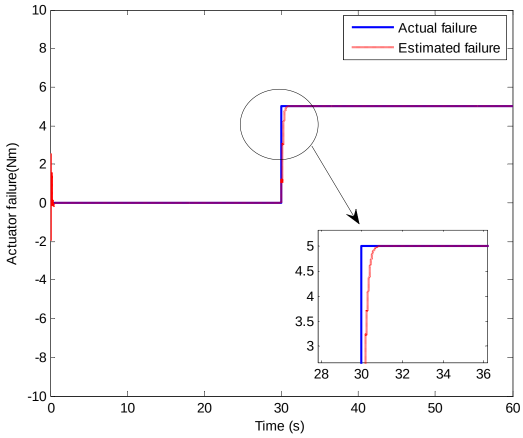

Fig. 2. The estimation of the actuator failure

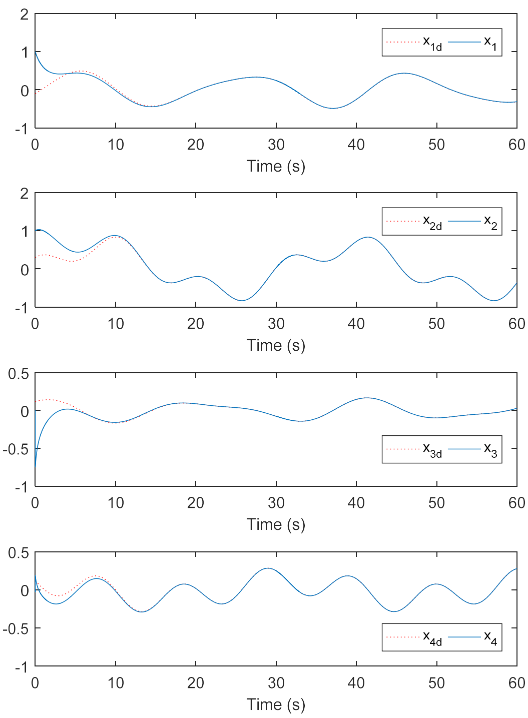

Fig. 3. Tracking curves for Configuration *a*

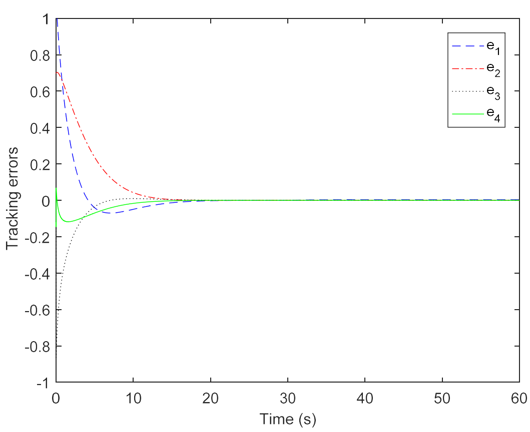

Fig. 4. Tracking error curves for Configuration *a*

The simulation results are shown in Figs. 2-5. The estimated value of actuator failure $\hat{f}_a$ is shown as Fig. 2, we can observe that after the failure occurs at $t = 30s$, owing to lack of the knowledge about the fault function, and then, the estimated failure and the actual failure are almost overlap less than one second. Fig. 3 and Fig. 4 illustrate the trajectory tracking curves and the tracking error curves by the proposed FTTC scheme, and we can observe that the actual trajectories follow the desired ones successfully even after the fault occurs.

By using the proposed scheme, we can see from Fig. 5, the weights of the critic neural network converge to $[25.9151, 34.0755, 21.7413, 1.1632, 32.1819, 24.5659, 33.7582, 45.7669, 49.1382, 46.1396]^T$. Hence, the proposed FTTC scheme is effective to control reconfigurable manipulator systems with actuator failures.

To further test the effectiveness of the proposed scheme based on FTTC, the same scheme is applied to the Configuration *b*, whose dynamics are described as

$$M(q) = \begin{bmatrix} -0.1166\cos^2(q_2) + 0.17 & -0.06\cos(q_2) \\ 0.06\cos(q_2) & 0.1233 \end{bmatrix}$$

$$C(q,\dot{q}) = \begin{bmatrix} 0.1166\sin(2q_2)\dot{q}_2 & 0.06\sin(q_2)\dot{q}_2 \\ 0.06\sin(q_2)\dot{q}_2 - 0.0583\sin(2q_2)\dot{q}_1 & 0.06\sin(q_2)\dot{q}_1 \end{bmatrix}$$

$$G(q) = \begin{bmatrix} 0 \\ -5.88\cos(q_2) \end{bmatrix}$$

The additive actuator failure f_a is injected into joint 2, here we choose it as

$$f_a(t) = \begin{cases} [0;0], & 0 \le t \le 30s \\ [0; 3\sin(0.2t) + \cos t], & 30s < t \le 60s \end{cases} \tag{47}$$

$$\begin{aligned} x_{1d} &= 0.2\cos(0.5t) + 0.2\sin(0.4t) \\ x_{2d} &= 0.3\cos(0.6t) - 0.4\sin(0.6t) \end{aligned} \tag{48}$$

Let $R = 0.01 \times I_2$, and other control parameters are selected as same as those of Configuration *a*.

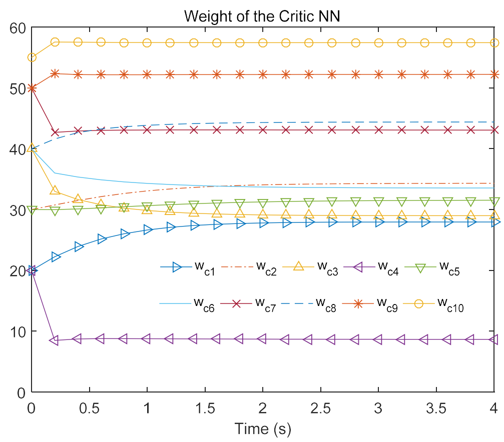

Fig. 5. The weights of the critic neural network for Configuration *a*

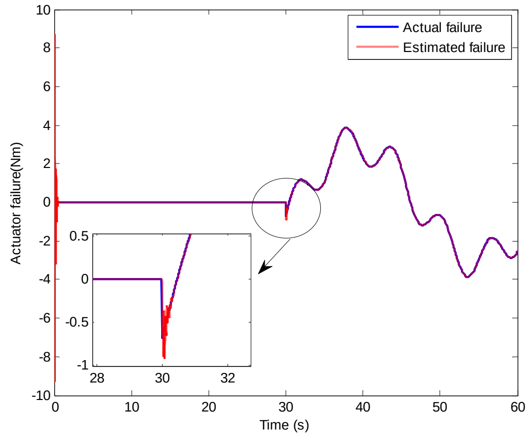

Fig. 6. The estimation of the actuator failure *b* Define the desired trajectory as

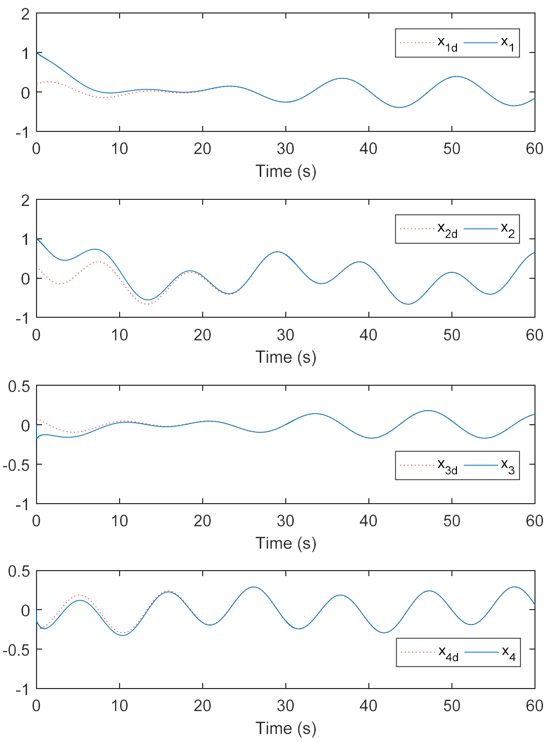

Fig. 7. Tracking curves for configuration *b*

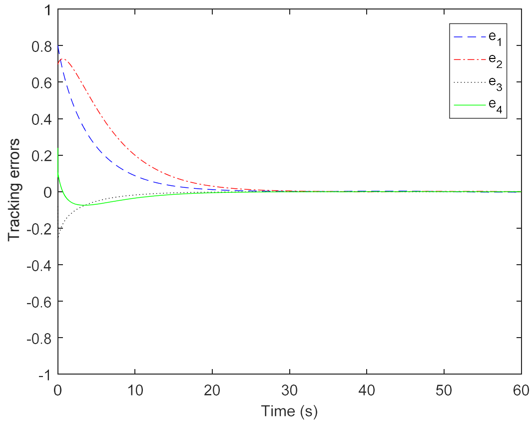

Fig. 8. Tracking error curves for Configuration *b*

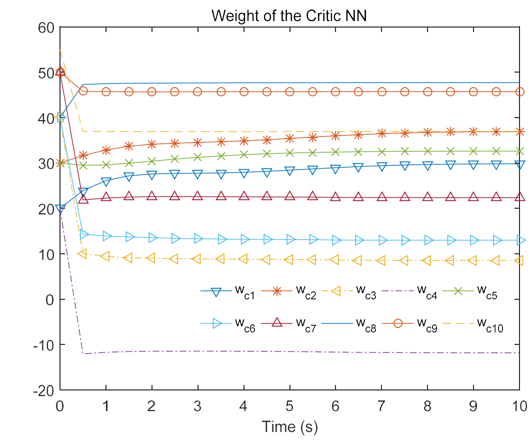

Fig. 9. The weights of the critic neural network for Configuration *b*

The simulation results are shown as Figs. 6-9, and we employ the similar critic neural network to approximate the performance index function, and its initial weight vector is chosen as same as Configuration *a*. By using the proposed algorithm, the weights of the critic neural network converge to $[29.7209, 36.8018, 8.5313, -11.8093, 32.6469, 12.9903, 22.3803, 47.7018, 45.7275, 36.9493]^T$ These two parts are not in same level. From these figures, we can conclude the similar results as Configuration *a*. Therefore, we can declare that the proposed FTTC scheme can be effectively applied to different configurations of reconfigurable manipulators.

Remark 5: Many efforts have been made for FTC methods, such as adaptive dynamic programming [44], hybrid intelligent approach [45] and internal model based approach [46]. Specially, [40] proposed an unknown input state observer based fault identification scheme. And a compensation term was employed to realize FTC of reconfigurable manipulators. Different from them, the PI algorithm based FTTC in this paper does not only emphasize a satisfactory FTC performance, but also guarantee the closed-loop system in an optimal manner. Thus, it is a significant improvement for FTTC of reconfigurable manipulators.

5. Conclusion

This paper develops a novel FTTC scheme for a class of nonlinear systems with actuator failures based on ADP with an improved performance index function. With the help of the estimated failure from the adaptive fault observer, a novel performance index function is constructed to account for system failures, thus the FTTC problem can be transformed into an optimal control problem. Then the PI algorithm is used to solve the optimal control problem. A critic neural network is established to solve the improved HJB equation online, and the approximated optimal feedback controller can be derived directly. It is proven that the closed-loop system is UUB based on the Lyapunov stability theorem. Thus, the FTTC is obtained by combining the optimal feedback controller and the desired controller of fault free system. Finally, simulation results verify the effectiveness of the proposed FTTC scheme. It is shown that the proposed controller is capable of controlling the reconfigurable manipulators which are regarded as highly nonlinear dynamic systems successfully.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 61773075, 61603 387, 61374051), in part by the Scientific and Technological Development Plan Project in Jilin Province of China (Grant Nos. 20150520112JH, 20160414033GH and in part by Early Career Development Award of SKLMCCS (Grant No. 20160520013 JH), and in part by the Opening Funding of State Key Laboratory of Management and Control for Complex Systems (Grant No. 20150102).

References

[1] S. Tong, B. Huo and Y. Li, "Observer-based adaptive decentralized fuzzy fault-tolerant control of nonlinear large-scale systems with actuator failures," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 1, pp. 1-15, Feb 2014.

[2] B. Zhao, Y. Li, and D. Liu, "Self-tuned local feedback gain based decentralized fault tolerant control for a class of large-scale nonlinear Systems," *Neurocomputing*, vol. 235, pp. 147-156, Feb 2017.

[3] R. J. Patton, "Fault tolerant control: The 1997 situation," Proceedings of the 3rd IFAC Symposium on Fault Detection Supervision and Safety for Technical Processes, Hull, United Kingdom pp. 759-762, 1997.

[4] Y. Zhang and J. Jiang, "Bibliographical review on reconfigurable fault-tolerant control systems," *Annual Reviews in Control*, vol. 32, no. 2, pp. 229-252, Dec 2008.

[5] M. Blanke, M. Kinnaert, J. Lunze and M. Staroswiecki, "Diagnosis and fault-tolerant control," *Springer-Verlag*, vol. 49, no. 6, pp. 873-893, 2006.

[6] M. Benosman and K. Y. Lum, "Passive actuators' fault-tolerant control for affine nonlinear systems," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 1, pp. 152-163, Jan 2010.

[7] W. Wang and C. Wen, "Adaptive compensation for infinite number of actuator failures or faults," *Automatica*, vol. 47, no. 10, pp. 2197-2210, Oct 2011.

[8] S. Chatterjee, S. Sadhu and T. K. Ghoshal, "Fault detection and identification of non-linear hybrid system using self-switched sigma point filter bank." *IET Control Theory and Applications*, vol. 9, no. 7, pp. 1093-1102, May 2015.

[9] W. H. Lee, K. H. Kim, G. P. Chan and J. G. Lee, "Two-faults detection and isolation using extended parity space approach," *Journal of Electrical Engineering and Technology*, vol. 7, no. 3, pp. 411-419, May 2012.

[10] K. Zhang, B. Jiang and V. Cocquempot, "Adaptive observer-based fast fault estimation," *International Journal of Control Automation and Systems*, vol. 6, no. 3, pp. 320-326, June 2008.

[11] Y. Xu, S. Tong and Y. Li, "Observer-based fuzzy adaptive control of nonlinear systems with actuator faults and unmodeled dynamics" *Neural Computing and Applications*, vol. 23, no. s1, pp. 391-405, Oct 2013.

[12] Q. Jia, W. Chen, Y. Zhang and H. Li, "Fault reconstruction and fault-tolerant control via learning observers in Takagi-Sugeno fuzzy descriptor systems with time delays" *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3885-3895, June 2015.

[13] D. Q. Khanh and Y. S. Suh, "Mobile robot destination generation by tracking a remote controller using a vision-aided inertial navigation algorithm," *Journal of Electrical Engineering and Technology*, vol. 8, no. 3, pp. 616-620, May 2013.

[14] I. A. Raptis, K. P. Valavanis and G. J. Vachtsevanos, "Linear tracking control for Small-Scale unmanned helicopters," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 4, pp. 995-1010, July 2012.

[15] F. Liao, J. Wang and G. Yang "Reliable robust flight tracking control: an LMI approach," *IEEE Transactions on Control Systems Technology*, vol. 10, no. 1, pp. 76-89, Jan 2002.

[16] B. Yao, F. Wang and Q. Zhang, "LMI-based design of reliable tracking controller," *Acta Automatica Sinica*, vol. 30, no. 6, pp. 863-871, June 2004.

[17] B. Jiang, Z. Gao, P. Shi and Y. Xu, "Adaptive fault-tolerant tracking control of near-space vehicle using Takagi-Sugeno fuzzy models," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 5, pp. 1000-1007, Oct 2010.

[18] P. J. Werbos, "Approximate dynamic programming for real-time control and neural modeling," *Handbook of Intelligent Control Neural Fuzzy and Adaptive Approaches*, ch 13, 1992.

[19] D. Vrabie and F. L. Lewis, "Adaptive dynamic programming for online solution of a zero-sum differential game," *Journal of Control Theory and Applications*, vol. 9, no. 3, pp. 353-360, 2011.

[20] D. Liu, D. Wang, F. Wang and X. Yang, "Neural-network-based online HJB solution for optimal

robust guaranteed cost control of continuous-time uncertain nonlinear systems," *IEEE Transactions on Cybernetics*, vol. 44, no. 12, pp. 2834-2847, Dec 2014.

[21] Q. Wei and D. Liu, "Stable iterative adaptive dynamic programming algorithm with approximation errors for discrete-time nonlinear systems," *Neural Computing and Applications*, vol. 24, no. 6, pp. 1355-1367, May 2014.

[22] L. Yang, J. Si, K. S. Tsakalis, and A.A. Rodriguez, "Direct heuristic dynamic programming for nonlinear tracking conrol with filtered tracking error," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 39, no. 6, pp. 617-1622, Dec 2009.

[23] B. Zhao, D. Wang, G. Shi, D. Liu and Y. Li, "Decentralized control for large-scale nonlinear systems with unknown mismatched interconnections via policy iteration," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, DOI: 10.1109/ TSMC.2017.2690665, Apr 2017.

[24] D. Liu and Q. Wei, "Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 3, pp. 621-634, March 2014.

[25] D. Liu, D. Wang and H. Li, "Decentralized stabilization for a class of continuous-time nonlinear interconnected systems using online learning optimal control approach," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 2, pp. 411-428, Feb 2014.

[26] J. Wang, X. Xu, D. Liu, Z. Sun and Q. Chen , "Self-learning cruise control using kernel-based least squares policy iteration," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 3, pp. 1078-1087, May 2014.

[27] D. Wang, D. Liu and H. Li, "Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 2, pp. 627-632, Apr 2014.

[28] R. Song, F. L. Lewis, Q. Wei and H. Zhang, "Off-policy actor-critic structure for optimal control of unknown systems with disturbances," *IEEE Transactions on Cybernetics*, vol. 46, no. 5, pp. 1041-1050, May 2016.

[29] H. Zhang , L. Cui, X. Zhang and Y. Luo, "Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method," *IEEE Transactions on Neural Networks*, vol. 22, no. 12, pp. 2226-2236, Dec 2011.

[30] S. Chang, J. Y. Lee, J. B. Park and Y. H. Choi, "An online fault tolerant actor-critic neuro-control for a class of nonlinear systems using neural network HJB approach," *International Journal of Control Automation and Systems*, vol. 13, no. 2, pp. 311-318, April 2015.

[31] Q. Fan and G. Yang, "Adaptive fault-tolerant control for affine non-linear systems based on approximate dynamic programming," *IET Control Theory and Applications*, vol. 10, no. 6, pp. 655-663, March 2016.

[32] Z. Wang, L. Liu, H. Zhang and G. Xiao, "Fault-tolerant controller design for a class of nonlinear mimo discrete-time systems via online reinforcement learning algorithm," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 5, pp. 611-622, May 2016.

[33] P. A. Ioannou and J. Sun, "Robust adaptive control," Adaptive and Learning Systems, 1996.

[34] L. Tang, Y. Liu and S. Tong, "Adaptive neural control using reinforcement learning for a class of robot manipulator," *Neural Computing and Applications*, vol. 25, pp. 135-141, July 2014.

[35] H. J. Uang and B. Chen, "Robust adaptive optimal tracking design for uncertain missile systems: a fuzzy approach," *Fuzzy Sets and Systems*, vol. 126, no. 1, pp. 63-87, Feb 2002.

[36] M. Krstic, P. V. Kokotovic and I. Kanellakopoulos, "Nonlinear and adaptive control design," *New York: Wiley*, 1995.

[37] M. Yim, W. M. Shen, B. Salemi, D. Rus and M. Moll , "Modular self-reconfigurable robot systems," *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, pp. 43-52, March 2007.

[38] B. Zhao, C. Li, T. Ma and Y. Li, "Multiple faults detection and isolation via decentralized sliding mode observer for reconfigurable manipulator," *Journal of Electrical Engineering and Technology*, vol. 10, no. 6, pp. 2393-2405, Nov 2015.

[39] S. Ahmad, H. Zhang and G. Liu, "Distributed fault detection for modular and reconigurable robots with joint torque sensing: a prediction error based approach," *Mechatronics*, vol. 23, no. 6, pp. 607-616, 2013.

[40] B. Zhao, and Y. Li, "Local joint information based active fault tolerant control for reconfigurable manipulator," *Nonlinear Dynamics*, vol. 77, no. 3, pp. 859-876, Aug 2014.

[41] M. Abu-Khalaf and F. L. Lewis, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach," *Automatica*, vol. 41, no. 5, pp. 779-791, May 2005.

[42] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878-888, May 2010.

[43] B. Zhao, D. Liu, and Y. Li, "Online fault compensation control based on policy iteration algorithm for a class of affine nonlinear systems with actuator failures," *IET Control Theory & Applications*, vol. 10, no. 15, pp. 1816-1823, October 2016.

[44] B. Zhao, D. Liu, and Y. Li, "Observer based adaptive dynamic programming for fault tolerant control of a class of nonlinear systems," *Information Sciences*, vol. 384, pp. 21-33, 2017.
[45] E. Khalastchi, M. Kalech and L. Rokach, "A hybrid approach for improving unsupervised fault detection for robotic systems," *Expert Systems with Applications*, vol. 81, pp. 372-383, Sep 2017.
[46] D. Kim, D. Lee and K. C. Veluvolu, "Accommodation of actuator fault using local diagnosis and IMC-PID," *International Journal of Control Automation and Systems*, Vol. 12, no. 6, pp. 1139-1149, Dec 2017.

Yuanchun Li He received his Ph.D. degree in the Department of General Mechanics from Harbin Institute of Technology, China in 1990. Now, he is a professor in the Department of Control Science and Engineering, Changchun University of Technology. His research interest covers complex system modeling and robot control.

Hongbing Xia He received his B.S. degree in the Department of mathematics and physics from Bengbu University, China in 2014, and the M.E. degree in the Department of Control Science and Engineering from Changchun University of Technology, China in 2017. Now, he is an algorithm engineer at Hytera Communications Co., Ltd., Nanjing, China. His research interests covers fault diagnosis and fault tolerant control, robot control and adaptive dynamic programming.

Bo Zhao He received his B. E. degree in Automation, and Ph.D. degree in Control Science and Engineering, all from Jilin University, Changchun, China, in 2009 and 2014, respectively. He was a Post-Doctoral Fellow with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China, from November 2014 to June 2017. He is currently an Assistant Professor with the same affiliation. His research interests include adaptive dynamic programming, fault diagnosis and tolerant control, neural-network based control, and robot control. He has published over 50 journal and conference papers.