Research Article

Adaptive Inverse Optimal Control for Rehabilitation Robot Systems Using Actor-Critic Algorithm

Fancheng Meng and Yaping Dai

School of Automation, Beijing Institute of Technology, 5 South Zhongguancun Street, Haidian District, Beijing 100081, China

Correspondence should be addressed to Fancheng Meng; sdmfwzh@163.com

Received 18 April 2014; Revised 15 July 2014; Accepted 21 July 2014; Published 28 August 2014

Academic Editor: Qingsong Xu

Copyright © 2014 F. Meng and Y. Dai. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The higher goal of rehabilitation robot is to aid a person to achieve a desired functional task (e.g., tracking trajectory) based on assisted-as-needed principle. To this goal, a new adaptive inverse optimal hybrid control (AHC) combining inverse optimal control and actor-critic learning is proposed. Specifically, an uncertain nonlinear rehabilitation robot model is firstly developed that includes human motor behavior dynamics. Then, based on this model, an open-loop error system is formed; thereafter, an inverse optimal control input is designed to minimize the cost functional and a NN-based actor-critic feedforward signal is responsible for the nonlinear dynamic part contaminated by uncertainties. Finally, the AHC controller is proven (through a Lyapunov-based stability analysis) to yield a global uniformly ultimately bounded stability result, and the resulting cost functional is meaningful. Simulation and experiment on rehabilitation robot demonstrate the effectiveness of the proposed control scheme.

1. Introduction

Rehabilitation robots are becoming increasingly common in upper extremity rehabilitation [1], because they are not only able to provide intensive rehabilitation consistently for a longer duration, but also able to offer object assessment for patient's motor performance. Relevant studies show they contribute significantly to motor recovery [2–4]. However, most of the existing rehabilitation robots (RRs) so far have still suffered from a number of challenges. Remarkably, an autonomous rehabilitation robot has yet to be developed due to the lack of an advanced and intelligent controller. To achieve this requirement, various control strategies have been proposed, such as impedance control [5], admittance control [6], and PID control [7]. One common aspect of the majority of these RRs control algorithms applied to robot-aided therapies is to allow a variable deviation from a predefined trajectory rather than imposing a rigid training pattern [8]. Relevant studies demonstrate these RRs controllers are effective to some extent [9, 10]. However, their control parameters in essence are predefined by human. From the practical viewpoint, the fixed and predefined control parameters might often lead to either an overdetermined or an underdetermined control performance. Furthermore,

the predefined control pattern does not suit each person. Thus, how to provide controllable, suitable assistance specific to a particular patient is a worth studying problem.

One possible solution to this problem is adaptive functional training. The basic idea of the adaptive functional training is to allow a modification of controller parameters or task difficulty levels according to the patient's motor performance, rather than imposing a rigid predefined training pattern. Typical examples include adaptive admittance control [11], adaptive impedance control [12, 13], adaptive task-oriented therapy [13–15], and adaptive mixed reality rehabilitation training [16, 17], as well as adaptive slide model control [18]. Relevant works show these approaches can adjust their controller parameters or task difficulty levels according to the patients' motor performance or diagnosis and can provide a better control performance. However, their resulting impedance parameters (e.g., mass, damping, and stiffness) are still obtained in a heuristically way and cannot be easily extended to other applications. Recently, the combinations of neural network- (NN-) based intelligent control and adaptive functional training have been proposed to RRs, such as NN-based adaptive sliding mode control [19], neural-fuzzy robust adaptive control [20], and neural-fuzzy control [21]. Relevant studies show these NN-based intelligent RRs controllers can

provide a better performance than the above-mentioned traditional adaptive control. Unfortunately, their results are limited to uniformly ultimately bounded (UUB) stability because of the inevitable NN reconstruction error [22]. Crucially, most of these existing intelligent RRs controllers are in essence adaptive but not optimal for rehabilitation training. However, in practical application, the optimal RRs controller can benefit the patient's rehabilitation training more than the traditional adaptive RRs controllers due to the trade-off between control torques and tracking accuracy [23].

Reinforcement learning (RL) has the potential to overcome the above-mentioned problems [24–28]. Actor-critic (AC-) based controller has, for example, been used for controlling a three-finger gripper [29]; also, this method can be used to find suitable compliance for different contact tasks [30]. Recently, the AC-based controller design problems have been extensively investigated for functional electrical stimulation (FES) [31] and myoelectric devices [32]. Thomas et al. [31] showed the feasibility of using the AC for FES control of upper extremities. Similarly, Pilarski et al. [32] proposed an AC-based reinforcement learning method for optimizing the control of myoelectric devices. Despite such encouraging results, the AC-based control approach so far found relatively limited application in real-time robotic rehabilitation system, because rehabilitation actions are often continuous in robotic rehabilitation process, whereas most of the existing AC-based controllers concern discrete actions. Unfortunately, the discrete AC controllers cannot ensure effectiveness in continuous-time systems. Moreover, during rehabilitation training process, since the exogenous reference inputs for many online applications are event-driven and not known a priori, it is often impossible to monitor online whether a signal is persistence of excitation (PE) and whereas PE condition is vital for AC learning [33]. In this case, the AC-based controller is not able to ensure convergence in a finite time. Recently, a continuous adaptive critic control [34] solves this problem, where a NN-based critic is used to approximate a long-term cost function and a NN-based actor approximates the optimal control. Simulation results demonstrate the feasibility of the proposed AC-based approach. However, a limitation of the method is that it requires the knowledge of the Hamilton-Jacobi-Bellman (HJB) equation. In practice, solving the HJB equation is quite a difficult and time-consuming task in model-free online way [35].

To overcome the above-mentioned difficulties and to develop an AC-based controller for the highly nonlinear RRs, a new adaptive inverse optimal hybrid control (AHC) combining inverse optimal control and AC learning is proposed in this paper, where the inverse optimal controller guarantees the RRs' optimality by minimizing a meaningful cost function; thus, the effort of solving the HJB equation is totally avoided; the AC-based learning method helps the robot select the meaningful cost function despite being uncertain in RRs. In addition, understanding the human motor behavior is essential to assisted-as-needed rehabilitation training. Therefore, an uncertain nonlinear rehabilitation robot model is firstly developed that includes human motor behavior dynamics. Then, based on this model, an open-loop error system is formed; thereafter, an inverse optimal feedback

control input is designed to be responsible for nonlinear static part and minimize the cost functional, and a NN-based AC feedforward control is responsible for the nonlinear dynamic part contaminated by uncertainties. Finally, Lyapunov-based stability analysis proved the stability of the proposed AHC and determined a respective meaningful cost functional. Simulation and experiments on a 2-degree-of-freedom rehabilitation robot indicate the feasibility and effectiveness of the proposed AHC.

Based on the previous discussion, we highlight the contributions of this paper as follows.

- (1) An uncertain nonlinear rehabilitation robot model is developed that includes individual motor behavior dynamics such that the rehabilitation robot can better capture enough information on the real individual (human) motor behavior.
- (2) A data-driven NN is employed to cope with the problem of unknown rehabilitation robotic dynamics such that neither the rehabilitation robot structure nor the physical parameters are required in the control design to be developed.
- (3) AHC is adopted such that the robot can achieve some function task (e.g., tracking trajectory) in an optimal way.

The rest of the paper is organized as follows. In Section 2, a specific rehabilitation robot system under study is described and control objective is discussed. In Section 3, the details of the proposed AHC control are presented, followed by the rigorous analysis. In Section 4, the validity of the proposed method is verified by simulation and experimental studies. Concluding remarks are given in Section 5.

2. Problem Statement

2.1. System Description. In this paper, we investigate a typical robot-assisted system, which includes a human limb and a rehabilitation robot with an end-effector, as shown in Figure 1. Without loss of generality, similar to [36], the following assumption is given.

Assumption 1. The handle (C) of the rehabilitation robot is tightly grasped by the patient and there is no relative motion between the patient's hand and the handle (i.e., end-effector). Based on this assumption, the patient's hand is deemed as "a part" of the rehabilitation robot.

The rehabilitation robot in task space can be modeled as

$$M_R \ddot{x} + C_R \dot{x} + K_R(x) + d = f_H + u, \quad (1)$$

where $x, \dot{x}, \ddot{x} \in \mathbb{R}^n$ are the position, velocity, and acceleration of the end-effector, respectively. $M_R \in \mathbb{R}^{n \times n}$ is the positive inertia matrix of the robot at the end-effector, $C_R \in \mathbb{R}^{n \times n}$ denotes the centripetal and Coriolis force at the end-effector, $K_R \in \mathbb{R}^n$ is stiffness matrices, and $d \in \mathbb{R}^n$ is unknown bounded disturbance. $f_H \in \mathbb{R}^n$ is the user's force. $u \in \mathbb{R}^n$ is the input control vector. In general, in many cases,

FIGURE 1: The system under study.

f_H in (1) is often only considered as a measurable load. However, the patient's motion behavior is typically a time-varying trajectory, which cannot be represented by only force states; thus, this pure load-based control is limited by its poor robustness in practice [37].

To solve this problem, it is necessary to integrate the patient's motor behavior dynamics into system (1). The user's motor dynamics is defined as [38]

$$f_H^* = f_H + M_H \ddot{x} + F_d \dot{x} + K_H x + d_1, \quad (2)$$

where M_H , F_d , and K_H are the mass, damper, and spring matrices of the arm end-effector, respectively, and f_H is the user hand force, f_H^* is the planned force in the patient's CNS, and d_1 is the unknown bounded disturbance (e.g., muscle spasticity). Note that, by using (2), the rehabilitation robot is able to capture the dynamic characteristics of motion behavior.

Substituting (2) into (1), the total rehabilitation robotic dynamics can thus be rewritten as

$$M \ddot{x} + C \dot{x} + G(x) + f + d = u, \quad (3)$$

where M , C , and G denote the combined patient-robot positive inertia matrix, centripetal and Coriolis force, and stiffness matrices at the handle, respectively. And they satisfy the following equation:

$$\begin{aligned} M &= M_R + M_H, & C &= C_R + F_d, \\ G &= K_R + K_H, & f &= f_H^* - d_1. \end{aligned} \quad (4)$$

Remark 2. An uncertain nonlinear rehabilitation robot model (3) is developed that includes human motor behavior dynamics (2). The objective of the models developed is not to replicate human behavior but only to capture enough information to compensate the drawbacks inherent to the most prior rehabilitation robotic dynamics [39, 40]. By

recognizing the patient's motor behavior, RRs can adjust its robotic assistance according to the patient's performance on the task, based on assisted-as-needed principle [39].

Without loss of generality, the following properties are given.

Property 1. The inertia matrix $M(x)$ is symmetric, positive definite and satisfies the following inequality:

$$\underline{m} \|\xi\|^2 \leq \xi^T M(x) \xi \leq \bar{m} \|\xi\|^2, \quad \forall \xi \in \mathbb{R}^n, \quad (5)$$

where $\underline{m}$ is a known positive constant, $\bar{m}(x)$ is a known positive function, and $\|\cdot\|$ denotes the standard Euclidean norm.

Property 2. The following skew-symmetric relationship is satisfied:

$$\xi^T (M(x) - 2C(x, \dot{x})) \xi = 0, \quad \forall \xi \in \mathbb{R}^n. \quad (6)$$

Property 3. The nonlinear disturbance term and its first and second time derivatives are bounded by known constants.

Property 4. The desired trajectory $x_d(t)$ is assumed to be designed such that $\dot{x}_d(t)$, $\ddot{x}_d(t)$ exist and are bounded.

Property 5 (see [41]). A continuous nonlinear function $f(x)$ can be approximated by NN as follows:

$$f(x) = W^T \varphi(\theta) + \varepsilon(\theta), \quad (7)$$

where $W \in \mathbb{R}$ represents weight vector and $\varphi(\cdot) \in \mathbb{R}$ denotes the NN activation function with $\varepsilon(\theta)$ being approximation error satisfying $\|\varepsilon(\theta)\| \leq \varepsilon_m$, where ε_m is the approximation error bound. The NN activation function and the weight vector are upper bounded by a positive constant such that $\|\varphi(\theta)\| \leq \varphi_m$ and $\|W\| \leq W_m$. Note that Property 5 is easily obtained, because any C^1 -function can be expressed as (7).

Remark 3. The above-mentioned properties are the fundamental in this paper, and their validity will be shown later.

2.2. Control Object. The object of rehabilitation robot is to enable a patient to achieve some function task (e.g., trajectory tracking) based on assisted-as-needed principle. Towards this goal, the current control objective is to find a control input which enables a patient to track a desired trajectory, denoted by $x_d(t)$, despite uncertainties in the dynamic model, while also minimizing a given performance index that includes a penalty on the tracking error and the control effort. To quantify the objective, a position tracking error $e_1(t) \in \mathbb{R}$ is defined as

$$e = x_d(t) - x(t). \quad (8)$$

Further, to facilitate the subsequent closed-loop error system development and stability analysis, a filtered tracking error is defined as

$$r = \dot{e} + \lambda e, \quad (9)$$

where λ denotes positive constants, which behaves like the control gain in PID controller.

3. Adaptive Inverse Optimal Hybrid Control Design

In this section, a new AC-based adaptive inverse optimal hybrid control (i.e., AHC) for rehabilitation robot system was given. First, a state-space model is developed based on the tracking errors in (8) and (9). Then, based on this model, direct feedforward terms (i.e., identifier NN and an action NN) with an inverse optimal PD feedback controller are given, and, subsequently, a critic NN is proposed; finally, Lyapunov-based stability and optimality analysis are performed to show that the AHC can yield an asymptotic optimal tracking.

3.1. Actor NN-Based Control. Using the errors in (8) and tracking errors in (9) and the system dynamics (3), an open-loop error for system (1) can be obtained:

$$M\dot{r} = M\ddot{x}_d + C\dot{x} + \lambda M\dot{e} + G + f + d - u. \quad (10)$$

To facilitate the subsequent control design, (10) is rewritten as

$$M\dot{r} = -Cr + f_1 + f_2 + f_3 - u, \quad (11)$$

where some auxiliary terms are defined as follows:

$$f_1 = f, \quad f_2 = M\ddot{x}_d + C\dot{x}_d + G + d, \quad f_3 = \lambda M\dot{e} + \lambda e. \quad (12)$$

Note that f_1 and f_2 depend on the position and desired trajectory and contain the same system parameter (seen from target impedance); instead, f_3 is the error function. Motivation for this segregation is to facilitate the subsequent control design and stability analysis. Nevertheless, since all the terms contain unknown parameters, the controller u given in (10) cannot be implemented. But if the proposed control input can identify and/or cancel these effects, then the optimal control law can be realized. To identify and cancel these effects, two typical NN-based approximators for both f_1 and f_2 are defined:

$$\begin{aligned} f_1 &= W_I \varphi(y) + \varepsilon_I(y), \\ f_2 &= W_a^T \varphi(y) + \varepsilon_a(y), \end{aligned} \quad (13)$$

where $W_a \in \mathbb{R}^{(n+1) \times m}$ and $W_I \in \mathbb{R}^{(n+1) \times m}$ are the ideal weight matrices, $\varphi(\cdot)$ denotes the activation function of NN, $\varepsilon_I(y)$, $\varepsilon_a(y)$ are bounded reconstruction errors, and y is the input vector to NN, which is defined as

$$y = (x, \dot{x}, x_d, \dot{x}_d). \quad (14)$$

The estimates of f_1 and f_2 are designed as

$$\hat{f}_1 = \hat{W}_I^T \varphi(y), \quad \hat{f}_2 = \hat{W}_a^T \varphi(y), \quad (15)$$

where $\hat{W}_a \in \mathbb{R}^{(n+1) \times m}$, $\hat{W}_I \in \mathbb{R}^{(n+1) \times m}$ are the estimates of the ideal weights. Based on the open-loop error system in (10), the control input is defined as

$$u = -u_{01} + \hat{W}_I^T \varphi(y) + \hat{W}_a^T \varphi(y), \quad (16)$$

where u_{01} is the control law to be developed, which is later shown to minimize a meaningful cost, and the estimates for the NN weights in (16) are generated online by the following adaptive update law:

$$\dot{\hat{W}}_I = Y_{Iw} \varphi(y) r + Y_{Iw} \varphi(y) M^{-1} R \hat{W}_c^T \varphi'(z) \hat{V}_c^T, \quad (17)$$

$$\dot{\hat{W}}_a = Y_{aw} \varphi(y) r + Y_{aw} \varphi(y) M^{-1} R \hat{W}_c^T \varphi'(z) \hat{V}_c^T, \quad (18)$$

where $Y_{aw} \in \mathbb{R}^{(n+1) \times (n+1)}$ and $Y_{Iw} \in \mathbb{R}^{(n+1) \times (n+1)}$ are constant, positive definite and symmetric gain matrices, z , $\hat{W}_c^T$, $\hat{V}_c^T$, and R are the input and weight estimates of the critic NN and reinforcement signal, respectively. Note that the weight update law in (17)-(18) consists of a gradient-based term and a reinforcement signal term R ; this makes the subsequent critic NN affect the behavior of the action and/or identifier NN.

As described earlier, AC learning's convergence cannot be ensured in a finite time due to the lack of PE. That is, weight parameter convergence is not guaranteed. Thus, to overcome this difficulty and ensure weight parameter convergence in (17)-(18), a virtual control input with the known noise (i.e., PE) [42, 43] is injected to our control design (16), and the virtual control input is defined:

$$u_v = -k_v r + \omega, \quad (19)$$

where $k_v \in \mathbb{R}$ denotes a positive constant adjustable control gain. ω is any given nonzero measurable signal which is exactly known a priori and bounded by $\|\omega\| \leq \omega_M$.

Thus, the control input with a PE input can be defined as

$$u = -u_{01} + \hat{W}_I^T \varphi(y) + \hat{W}_a^T \varphi(y) - u_v. \quad (20)$$

To facilitate the analysis, an auxiliary function is given:

$$u_o = u_{01} + u_v. \quad (21)$$

Furthermore, substituting (20) and (21) into (11), the closed-loop error for system (1) can hence be obtained:

$$M\dot{r} = -Cr + \hat{W}_a^T \varphi(y) + \hat{W}_I^T \varphi(y) + f_3 + \varepsilon + u_o, \quad (22)$$

where the weight mismatch errors are defined as

$$\widetilde{W}_a^T = W_a^T - \hat{W}_a^T, \quad \widetilde{W}_I^T = W_I^T - \hat{W}_I^T. \quad (23)$$

For simplicity, (22) is expressed as

$$M\dot{r} = S_A + S_I + u_o = S + u_o, \quad (24)$$

where the auxiliary terms S_A and S_I are defined as

$$S_A = \widetilde{W}_a^T \varphi(y) + f_3 + \varepsilon, \quad S_I = \widetilde{W}_I^T \varphi(y) - Cr. \quad (25)$$

According to Properties 1 and 4 and using the mean value theorem, the following upper bound can be developed:

$$\|S_I\| \leq \delta_1 + \delta_2 \|X\|, \quad \|S_A\| \leq \delta_3 + \delta_4 \|X\|, \quad (26)$$

where X is defined as

$$X = [e, r]^T. \quad (27)$$

Since S_A and S_I contain the same parameters, to facilitate the subsequent stability, the term S in (24) can be defined as

$$\|S\| \leq c_1 + c_2 \|X\| \leq \Phi(\|X\|) X, \quad (28)$$

where $c_1 = \delta_1 + \delta_3$ and $c_2 = \delta_2 + \delta_4$ are known constants. Based on (10) and the subsequent stability analysis, an inverse optimal PD feedback control input is designed as

$$u_{01} = -K^{-1}r = -(k_1 + k_2)r. \quad (29)$$

Remark 4. $\widehat{W}_a^T \varphi(y)$ and $\widehat{W}_I^T \varphi(y)$ in (16) belong to the actor NN in practice. The goal of the actor NN is to generate appropriate control signals and approximate the uncertainties in the system; instead, u_{01} in (29) is used to stabilize the system. In contrast to typical controllers, for example, NN + PD control [44] and AC control [45], the proposed AHC can not only ensure system stability but also improve control performance in flexible gains. That is, combining the inverse optimal policy with the AC direct optimal algorithm, the rehabilitation robot can successfully achieve the desired trajectory tracking and minimize a cost functional, despite being uncertain and nonlinear in the dynamic system.

3.2. Critic NN. In this section, a critic NN is employed to approximate the long-term cost function. Similar to [46, 47], the reinforcement signal $R(t) \in \mathbb{R}$ is assumed to be defined as

$$R = \widehat{W}_c^T \varphi(z) + \rho, \quad (30)$$

where $z = \widehat{V}_c^T r$, $\widehat{W}_c \in \mathbb{R}^{(nc+1) \times 1}$, and $\widehat{V}_c \in \mathbb{R}^{m \times nc}$ are the estimates of the ideal weight matrices, $\varphi(z) \in \mathbb{R}^{nc+1}$ denotes the activation function, r is the input to the critic NN, and ρ is an adaptive term (auxiliary critic signal) with known noise (the noise is used to ensure PE condition), to aid the subsequent stability; its time derivative is written as

$$\dot{\rho} = \widehat{W}_c^T \varphi'(z) M^{-1} (K^{-1} + k_v) z, \quad (31)$$

where M and K are as defined before. Based on the subsequent stability analysis, the weight update law for the critic NN can be easily obtained:

$$\begin{aligned} \dot{\widehat{W}}_c &= -Y_{cw} \varphi(z) R - Y_{cw} \widehat{W}_c, \\ \dot{\widehat{V}}_c &= -Y_{cv} r \widehat{W}_c^T \varphi'(z) - Y_{cv} \widehat{V}_c, \end{aligned} \quad (32)$$

where Y_{cw} and $Y_{cv} \in \mathbb{R}$ are positive and known constants. After taking the time derivative of (30)

$$\dot{R} = (\widehat{W}_c^T \varphi'(z))' + \dot{\rho}. \quad (33)$$

To facilitate the subsequent stability analysis, the expression of (33) is rewritten as

$$\begin{aligned} \dot{R} &= \dot{\widehat{W}}_c^T \varphi(z) + \widehat{W}_c^T \varphi'(z) \dot{\widehat{V}}_c^T r + \widehat{W}_c^T \varphi'(z) \widehat{V}_c^T \dot{r} \\ &\quad + \widehat{W}_c^T \varphi'(z) M^{-1} (K^{-1} + k_v) z. \end{aligned} \quad (34)$$

Some auxiliary terms are given:

$$N_{Rc} = \widehat{W}_c^T \varphi'(z) z, \quad N_{Rk} = M^{-1} C R. \quad (35)$$

Using Properties 1–5, and the mean value theorem, the following upper bound can be developed:

$$\|N_{Rc}\| \leq \varsigma_1 r, \quad \|N_{Rk}\| \leq \eta_1 R. \quad (36)$$

Remark 5. The critic network monitors the state of the rehabilitation robot system and affects the behavior of the actor (i.e., action NN and identifier NN). In addition, since the term $K^{-1}r$ in (29) can change the time derivative of R in (34), thus, increasing the term K^{-1} can speed up convergence of the critic. It is concluded that our AHC has faster convergence than the AC in [45]. Furthermore, provided the critic NN converges correctly and r can be adjusted arbitrarily by the actor NN, the proposed AHC will provide global (universal) stable and optimal control [48].

3.3. Stability Analysis. The stability of the proposed AHC given in (20) and (29) can be examined through Theorem 10. First, some assumptions are given.

Assumption 6. The ideal weight values for W_I , W_a , W_c , and V_c are assumed to exist and be bounded by known positive values; that is,

$$\begin{aligned} \|W_I\| &\leq W_{I \max}, & \|W_a\| &\leq W_{a \max}, \\ \|W_c\| &\leq W_{c \max}, & \|V_c\| &\leq V_{c \max}. \end{aligned} \quad (37)$$

Assumption 7. The weight estimate values and mismatch errors for the ideal weights W_I , W_a , W_c , and V_c are assumed to exist and be bounded by known positive values.

Assumption 8. If Property 1 is satisfied, there exists a positive, invertible function Ψ such that

$$\frac{1}{2} |\dot{M}| \leq \Psi(\|X\|). \quad (38)$$

Remark 9. Using Property 1, Assumption 8 is easily obtained. The above-mentioned assumptions are the basis of Theorem 10.

Theorem 10. For the rehabilitation robot system (1)–(3), let the proposed identification scheme in (15) along with the weight update law in (17) be used to identify the human motor behavior dynamics in (2), and let the action-critic controller given in (15) and (30) along with the weight update laws for the action and critic NN given in (18) and (32) and the inverse optimal control u_{01} in (29), respectively, ensure that all system signals are bounded under closed-loop operation and that the position tracking error is regulated in the sense that

$$|e(t)| \leq \zeta_0 \exp(-\zeta_1(t)) + \zeta_2, \quad (39)$$

where $\zeta_0, \zeta_1, \zeta_2 \in \mathbb{R}$ denote positive constants in $\Omega \triangleq \{Y \in \mathbb{R}^2 \mid \|Y\| \leq \Phi^{-1}(2\sqrt{\tau_1 k_1})\}$. Provided the control gains k_1, k_2

introduced in (29) and λ in (9) are selected according to the following sufficient condition:

$$k_1 > c_1, \quad \lambda > \frac{1}{2}, \quad \min\left(k_2 - \frac{1}{2}, \lambda - \frac{1}{2}\right) > c_2 \quad (40)$$

$k_v, \Psi, \eta, \varsigma \in \mathbb{R}$ are chosen to satisfy the following sufficient condition:

$$k_v > \Psi > 0, \quad \varsigma > 0, \quad \eta > 0. \quad (41)$$

Proof. Consider the following positive definite Lyapunov function defined as

$$L(t) = \frac{1}{2}e^2 + \frac{1}{2}Mr^2 + \frac{1}{2}R^2 + P_1 + P_2 + Q, \quad (42)$$

where P_1, P_2 , and Q are defined as, respectively,

$$\begin{aligned} P_1 &= \frac{1}{2} \text{tr} \{ \widetilde{W}_I^T \Upsilon_{Iw}^{-1} \widetilde{W}_I \}, & P_2 &= \frac{1}{2} \text{tr} \{ \widetilde{W}_a^T \Upsilon_{aw}^{-1} \widetilde{W}_a \}, \\ Q &= \frac{1}{2} \text{tr} \{ \widehat{W}_c^T \widehat{W}_c \} + \frac{1}{2} \text{tr} \{ \widehat{V}_c^T \widehat{V}_c \}. \end{aligned} \quad (43)$$

Using Property 1 and Assumptions 6–8, $L(t)$ can be bounded:

$$\gamma_1 \|Y\|^2 \leq L \leq \gamma_2 \|Y\|^2 + \gamma_3, \quad (44)$$

where $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{R}$ are known positive constants and $Y(t) \in \mathbb{R}$ is defined as

$$Y \triangleq [e, r, R]^T. \quad (45)$$

Taking the time derivative of (42) yields

$$\dot{L} = e\dot{e} + \frac{1}{2}\dot{M}r^2 + rM\dot{r} + R\dot{R} + \dot{P}_1 + \dot{P}_2 + \dot{Q}. \quad (46)$$

Based on the previous analysis, using (17)–(18), (32), and (29), the expression of (46) can be bounded:

$$\begin{aligned} \dot{L} &\leq \left(\frac{1}{2} - \lambda\right)e^2 + \left(\frac{1}{2}\dot{M} + \frac{1}{2}\right)r^2 - \Upsilon_{cw} \|\widehat{W}_c\|^2 \\ &\quad - \Upsilon_{cw} |R|^2 \|\varphi(z)\|^2 + 2|R| \Upsilon_{cw} \|\widehat{W}_c \varphi(z)\| \\ &\quad - \Upsilon_{cv} \|\widehat{W}_c^T \varphi'(z)\|^2 \|r\|^2 |R|^2 - r^T C r + r\varepsilon + rd \\ &\quad + r\omega + 2\Upsilon_{cv} \|\widehat{W}_c^T \varphi'(z)\| \|r\| |R| \|\widehat{V}_c\| - \Upsilon_{cv} \|\widehat{V}_c\|^2 \\ &\quad - r^T K^{-1} r - r^T k_v r - R \widehat{W}_c^T \varphi'(z) M^{-1} C z. \end{aligned} \quad (47)$$

Using (24)–(25) and (8)–(9) and canceling similar terms, the expression in (47) can hence be bounded as

$$\dot{L} \leq \left(\frac{1}{2} - \lambda\right)e^2 + \left(\frac{1}{2}\dot{M} + \frac{1}{2}\right)r^2 + rS + ru - N_R, \quad (48)$$

where the auxiliary term N_R is defined:

$$N_R = \widehat{W}_c^T \varphi'(z) M^{-1} C z R. \quad (49)$$

Using inequality (36) and Young's inequality, the auxiliary function N_R can be rewritten as

$$\|N_R\| \leq \varsigma r^2 + \eta R^2, \quad (50)$$

where $\varsigma = \varsigma_1^2, \eta = \eta_1^2$. Furthermore, using (25)–(26), (8)–(9), (19), and (38) and canceling similar terms, the expression in (48) can hence be bounded as

$$\begin{aligned} \dot{L} &\leq -\left(\lambda - \frac{1}{2}\right)e^2 + \left(k_2 - \frac{1}{2}\right)r^2 - \Psi(\|X\|)r^2 - N_R \\ &\quad - (k_1 r^2 - \delta_1 |r| - \delta_3 |r|) + (\delta_2 + \delta_4) \|X\|^2 + k_v r^2. \end{aligned} \quad (51)$$

Based on the above-mentioned analysis and using (26)–(28), the expression in (51) can be rewritten as

$$\begin{aligned} \dot{L} &\leq -\left(\lambda - \frac{1}{2}\right)e^2 + \left(k_2 - \frac{1}{2}\right)r^2 - \left(\sqrt{k_1}r - \frac{c_1}{2\sqrt{k_1}}\right)^2 \\ &\quad + \frac{c_1^2}{4k_1} - \eta R^2 - \varsigma r^2 + c_2 \|X\|^2 - (k_v - \Psi)r^2. \end{aligned} \quad (52)$$

Provided the conditions given in inequality (41) are satisfied, the expression of (52) is hence simplified as

$$\dot{L} \leq -\left[\min\left(\lambda - \frac{1}{2}, k_2 - \frac{1}{2}\right) - c_2\right] \|X\|^2 - \eta R^2 + \frac{c_1^2}{4k_1}. \quad (53)$$

And the following equation is assumed to be satisfied:

$$\tau_1 = \min\left[\left(\min\left(\lambda - \frac{1}{2}, k_2 - \frac{1}{2}\right) - c_2\right), \eta\right]. \quad (54)$$

Thus, the expression of (53) can be rewritten as

$$\begin{aligned} \dot{L} &\leq -\min\left[\left(\min\left(\lambda - \frac{1}{2}, k_2 - \frac{1}{2}\right) - c_2\right), \eta\right] \\ &\quad \cdot \|Y\|^2 + \frac{c_1^2}{4k_1} = -\tau_1 \cdot \|Y\|^2 + \frac{c_1^2}{4k_1}. \end{aligned} \quad (55)$$

To facilitate the subsequent analysis, using inequality (44), the expression of (55) is further simplified as

$$\dot{L} \leq -\frac{\tau_1}{\gamma_2} L + \varepsilon, \quad (56)$$

where $\varepsilon \in \mathbb{R}$ is a positive constant, which is defined as

$$\varepsilon = \frac{c_1^2}{4k_1} + \frac{\tau_1}{\gamma_2} \gamma_3. \quad (57)$$

Furthermore, using inequalities (44) and (56) and assuming that $\tau\tau_1 = \gamma_2$ is satisfied, thus, the above-mentioned inequality (56) can be rewritten as

$$L(t) \leq L(0)e^{-t/\tau} + \varepsilon\tau \left[1 - e^{-t/\tau}\right], \quad \forall \|Y\| \in \Omega, \quad (58)$$

where Ω is defined as before. Provided the sufficient conditions given in (40)–(41) are satisfied, the expressions in (42)

and (58) can be used to prove the control input and all the closed-loop signals are bounded in Ω :

$$\Omega \triangleq \left\{ Y \in \mathbb{R}^2 \mid \|Y\| \leq \Phi^{-1} \left(2\sqrt{\tau_1 k_1} \right) \right\}. \quad (59)$$

According to (56)-(57) and (59), it is observed that larger value of k_1 will expand the size of the domain Ω and reduce the residual error. According to inequality (58), the result in (39) can be easily obtained. $\square$

3.4. Optimality Analysis. Motivated by the works in [49], the ability of the controller u_{01} to minimize a meaningful cost can be examined through the following theorem.

Theorem 11. *The feedback law is given by*

$$u_{01}^* = HK^{-1}r. \quad (60)$$

With the scalar gain constant selected as $H > 2$ and the adaptive weight update law given in (17)-(18), (32), minimizes the meaningful cost functional

$$J = \lim_{t \rightarrow \infty} \left\{ 2HL(t)(t) + \int_0^t \left(\mathfrak{M} + u_{01}^T \Xi u_{01} \right) - 2H \frac{c_1^2}{4k_1} d\sigma \right\}, \quad (61)$$

where a positive function $\mathfrak{M}(Y, t)$ is defined as

$$\mathfrak{M} = -2H \left(e\dot{e} + \frac{1}{2} \dot{M}r^2 + rM\dot{r} + \Theta - \frac{c_1^2}{4k_1} \right) + H^2 r^T K^{-1}r, \quad (62)$$

where $\Theta = \dot{P}_1 + \dot{P}_2 + \dot{Q} + R\dot{R}$.

Proof. The cost function J in (61) is considered to be meaningful if it is a positive control function. That is, J in (61) is a positive function if $\mathfrak{M} > 0$. To examine the sign of $\mathfrak{M}(Y, t)$, utilizing (46) and inequality (55), the following inequality is hence obtained:

$$e\dot{e} + \frac{1}{2} \dot{M}r^2 + rM\dot{r} + \Theta - \frac{c_1^2}{4k_1} \leq -\tau_1 \|Y\|^2. \quad (63)$$

Multiplying both sides by $-2H$ and then adding $H(H-2)r^T K^{-1}r$, inequality (63) can be rewritten as

$$\begin{aligned} & -2H \left(e\dot{e} + \frac{1}{2} \dot{M}r^2 + rM\dot{r} + \Theta - \frac{c_1^2}{4k_1} \right) + H(H-2) \\ & \cdot r^T K^{-1}r \geq 2H\tau_1 \|Y\|^2 + H(H-2)r^T K^{-1}r. \end{aligned} \quad (64)$$

Based on inequality (63), inequality (64) can be simplified as

$$\mathfrak{M} \geq 2H\tau_1 \|Y\|^2 + H(H-2)r^T K^{-1}r. \quad (65)$$

According to inequality (65), if $H \geq 2$, then $\mathfrak{M} \geq 0$ is satisfied. Therefore J is a meaningful cost.

To show that u_{01}^* can minimize $J(\cdot)$, an auxiliary signal μ is defined as

$$\mu = u_{01} + HK^{-1}r. \quad (66)$$

Substituting μ in (66) into (61)

$$\begin{aligned} J &= \lim_{t \rightarrow \infty} \int_0^t \left(\mathfrak{M} + (\mu - HK^{-1}r)^T \Xi (\mu - HK^{-1}r) \right) \\ &\quad - 2H \frac{c_1^2}{4k_1} d\sigma + 2HL(t). \end{aligned} \quad (67)$$

Substituting $\mathfrak{M}$ in (62) into (67) and then utilizing $\dot{L}$ in (46) and $\lim_{t \rightarrow \infty} r = 0$, (67) can be rewritten as

$$\begin{aligned} J &= \lim_{t \rightarrow \infty} \left[2HL(t) - 2H \int_0^t \dot{L}(t) d\sigma + \int_0^t \mu^T \Xi \mu d\sigma \right] \\ &= \lim_{t \rightarrow \infty} \left[2HL(t) - 2HL(t)|_0^t + \int_0^t \mu^T \Xi \mu d\sigma \right] \\ &= \lim_{t \rightarrow \infty} \left[2HL(0) + \int_0^t \mu^T \Xi \mu d\sigma \right]. \end{aligned} \quad (68)$$

According to the result in (68), μ can minimize $J(\cdot)$ and stabilize the system. Therefore μ is optimal control law. Moreover, if the action NN and the identifier NN can precisely approximate the uncertainties and nonlinearities in system (1) or their approximation performance satisfies the proposed condition by [50], then a global optimal control law for rehabilitation robot system is obtained. $\square$

Remark 12. Differing from direct optimal control [33–35], the proposed AHC is hybrid (indirect and direct) optimal control without having to solve an HJB equation but rather minimizing a meaningful cost. Thus computation complex is alleviated.

4. Simulations and Experiments

In this section, the proposed AHC method is examined through simulation and experiments. The objective of the simulation is to examine the performance of the controller in (20); instead, the objective of the experiment is to test the controller's validity. Next we present the simulation and experimental procedure, relevant simulation, and experimental results with some volunteers.

4.1. Simulation. To illustrate the effectiveness and robustness of the proposed control algorithm on robotic rehabilitation system (3), a comparison among the adaptive proportional-derivative (PD) control, the NN-based adaptive control (NN + PD), the AC-based adaptive robust control (NN + AC), and the proposed AC-based inverse optimal control (NN + AHC) is made. Relevant simulation results are shown in Figures 3–6; summary results between NN + AC and NN + AHC are shown in Table 1.

FIGURE 2: (a) The meaningful cost functional. (b) The integral of the meaningful cost functional.

TABLE 1: Summarized simulation results.

Simulation	Position error (degree)		Torque (Nm)	
	AC	AHC	AC	AHC
Maximum (TS)	9.163	7.491	1.163	0.531
Minimum (TS)	6.302	2.872	0.612	0.512
Mean (TS)	7.417	3.431	0.832	0.516
Adjusting time	7–11 s	3–7 s	7–11 s	3–7 s
Maximum (SS)	0.182	0.172	0.523	0.512
Minimum (SS)	0.076	0.057	0.506	0.500

Note: TS represents temporary state; SS denotes steady state.

Presume the desired trajectory is selected as (in rads)

$$x_d = \begin{cases} \text{sgn}(\sin(0.05\pi t - 0.5\pi)); & \Delta > 0, \\ 0; & \Delta \leq 0, \end{cases} \quad (69)$$

where $\Delta = \text{sgn}(\sin(0.05\pi t - 0.5\pi))$, which determines the desired trajectory.

First, the adaptive PD control is applied to the simulation experiment, and the control parameters of the PD are taken as $K_p = 20$, $K_D = 30$, and $\Lambda = 5I$, and its simulation results are shown in Figure 3. From Figure 3, it is obvious that the adaptive PD control cannot track the desired position with the predefined control parameters. Specifically, the maximum position tracking error is about 0.8 rads; the minimum position tracking error is about 0.2 rads, and the steady mean error is about 0.6 rads.

Then, an action NN is added to PD arriving at a NN-based adaptive control (NN + PD), which is applied to the simulation experiment, and its control parameters are taken as $K_p = 10$, $K_D = 20$, with $\lambda = 3$ (learning rate). The simulation results of the NN-based adaptive control are shown in Figure 4. Though the NN-based adaptive control can track the desired trajectory, its control performance still needs to be improved. Specifically, the maximum position tracking error is about 0.2 rads, the minimum position tracking error is less than 0.1 rads, and the steady mean error is about 0.13 rads.

Thereafter, a critic NN is added to NN + PD arriving at an AC control (NN + AC), which is applied to the simulation, and its control parameters are selected as [45]. The simulation results of the AC-based robust adaptive control are shown in Figure 5. Though a favorable tracking performance can be achieved due to the effectiveness of the critic NN,

the unwanted transients (e.g., 10 s, 30 s, ..., 110 s, etc.) can result in unsatisfactory behavior.

Finally, the AC-based inverse optimal control (NN + AHC) is applied to the simulation. At the same time, to furthermore illustrate the robustness of the proposed controller, external disturbance $d = 0.05 \sin(3\pi t) \cos(\pi t)$ is considered. The initial position is (0,0). Both critic NN and actor NN contain 16 nodes in hidden layer. For weight updating, the weights V and W are set to initial random [0, 1]. All the activation functions (sigmoid) for hidden neurons are selected as

$$\varphi(z) = \frac{1}{1 + \exp(-z)}. \quad (70)$$

The gains for inverse optimal controller was selected as

$$\begin{aligned} k_1 &= 2.1, & k_2 &= 1.8, & k_v &= 1.3, \\ H &= 3, & \lambda &= 0.6. \end{aligned} \quad (71)$$

The control gains for the critic NN and actor NN are selected as

$$\Upsilon_{Iw}^{-1} = \Upsilon_{aw}^{-1} = \Upsilon_{Aw}^{-1} = 0.1I, \quad \Upsilon_{Cv}^{-1} = \Upsilon_{Cw}^{-1} = 1I. \quad (72)$$

The choice of these parameters is through some trails and all the gains are chosen in consideration of the requirement of stability. The results of the AC-based inverse optimal control are shown in Figure 6. A favorable high tracking performance can be achieved; the unwanted transients are almost avoided with little control torque. Specifically, the steady position tracking error of the proposed AHC decreases less than 0.001 rads in 43 s. Figure 2 shows that the meaningful cost function is positive, and thus the controller u_{01} in (29) can minimize a meaningful cost, which means the proposed AHC is optimal.

Figures 3–6 show the simulation results of traditional control and our proposed controller. From Figures 3–6, NN + PD, NN + AC, and the proposed controller can make the tracking error decrease during the learning process since these controllers have the learning ability. However, the proposed controller has a faster reduction rate in tracking errors than NN + PD, NN + AC with slightly reduced control torque. The control torque of the proposed controller is smoother and has a smaller vibration than NN + AC to

FIGURE 3: Position tracking, tracking error, control torque, and weights of actor NN with PD.

FIGURE 4: Position tracking, tracking error, control torque, and weights of actor NN with NN + PD.

FIGURE 5: Position tracking, tracking error, control torque, weights of actor NN, and weights of critic NN with NN + AC.

achieve the requested level of performance. It means that NN + AC controller augmented by an inverse optimal can adjust the performance versus control input by minimizing a cost functional. It is concluded that the approximation ability of the proposed NN + AHC is better than the NN + AC control system, and the proposed AHC is optimal.

Table 1 indicates that the position tracking error with NN + AHC is approximately 9% smaller than the one with NN + AC, and resulting force is approximately smaller by 15%, and convergence speed increases by 36%. The control accuracy from the simulation is sufficient for typical functional tasks. Nonetheless, the performance of the proposed method may vary when implemented in the real rehabilitation robot, since

the simulation condition is too ideal compared to the real application.

4.2. Tracking Experiment. Tracking experiments with two controllers (i.e., NN + AC and NN + AHC) were conducted on both arms of five volunteers (two females and three males, ages 22 to 36 yrs.). Designated as A, B, C, D, and E, subject A was used as an example and Figure 7 is the experimental setup.

To achieve better the control objective, each volunteer (e.g., A) first undergoes a standard assessment session before the experiment, in which he/she performs no less than 30 seconds reaching movements in eight specific movement

FIGURE 6: Position tracking, tracking error, control torque, weights of actor NN, and weights of critic NN with the proposed NN + AHC.

directions only by himself/herself [51]. Then, according to the evaluation result, we make a suitable training task, as shown in Figure 8. Thereafter, after having a rest, he/she (e.g., A) is asked to track the desired trajectory (i.e., the green line and the red circle); in the first phase of the experiment (0–50 s), subject A draws circle and in the second phase of the experiment (50–100 s), subject A draws line, according to the instruction on a monitor as accurately as possible. However, since individual motor behavior varies, subject A may not be able to track the desired motion trajectory all along. Once the user's motor behavior varies, the robot will adjust automatically the assistance modes (i.e., assistance or resistance) using the proposed controller. The experiment

results of subject A are shown in Figures 8, 9, 10, and 11, and the whole experimental results of five subjects tested for the desired trajectories are summarized in Table 2.

Figures 8–11 show the experiment results of NN + AC and our proposed controller. From Figures 8–11, NN + AC and the proposed controller can track the desired trajectory during the learning process since both controllers have the learning ability. However, the proposed controller has a faster trajectory tracking velocity than NN + AC, and its control torque is obviously smaller than the NN + AC; this is a favorable property, because high control torque can cause arm to fatigue faster.

FIGURE 7: A subject sitting on a chair with back support, holding the handle of the four-link robot manipulator. The interaction force generated by the subject onto the handle and the robot. The actual and reference handle positions were displayed on a monitor in front of the subject.

FIGURE 8: Trajectory tracking with NN + AC.

FIGURE 9: Trajectory tracking with NN + AHC.

TABLE 2: Summarized experimental results of five volunteers.

Sub.	Maximum				RMS error			
	NN + AC		NN + AHC		NN + AC		NN + AHC	
Units	PEr	F	PEr	F	PEr	F	PEr	F
	cm	N	cm	N	cm	N	cm	N
A (L)	2.12	9.96	2.02	9.33	0.67	0.26	0.66	0.26
A (R)	2.16	10.15	2.10	9.82	0.68	0.27	0.65	0.27
B (L)	1.92	7.08	1.71	6.36	0.75	0.28	0.56	0.19
B (R)	1.96	7.15	1.73	6.53	0.77	0.28	0.48	0.16
C (L)	2.33	9.13	2.23	8.70	0.86	0.34	0.73	0.27
C (R)	2.31	9.03	2.27	8.64	0.78	0.31	0.64	0.25
D (L)	1.78	4.64	1.73	4.53	0.42	0.11	0.56	0.16
D (R)	1.76	4.59	1.72	4.57	0.39	0.10	0.41	0.11
E (L)	2.02	5.83	2.02	5.62	0.62	0.23	0.56	0.17
E (R)	2.03	5.86	1.96	5.65	0.58	0.21	0.58	0.18
Mean	2.04	7.34	1.95	6.98	0.65	0.24	0.58	0.20
S.D.	0.20	2.34	0.16	1.98	0.12	0.07	0.09	0.05

Note: A-B denotes five volunteers, L represents left arm, R represents right arm, PEr represents position tracking error, and F represents the result force; S.D. denotes statistically significant value.

In Table 2, the maximum steady-state position error (PEr) is defined as the maximum absolute value of error that occurs after 2 seconds of the trial. The maximum steady-state position errors range from 1.71 to 2.27 cm with a mean of 1.95 cm and a standard deviation of 0.16 cm. The RMS position errors range from 0.41 to 0.73 cm with a mean RMS error of 0.58 cm and a standard deviation of 0.09 cm. The maximum steady-state force (F) is defined as the maximum absolute value of force that occurs after 2 seconds of the trial. The maximum forces range from 4.53 to 9.82 N with a mean of 6.98 N and a standard deviation of 1.98 N. The RMS forces range from 0.11 to 0.27 N with a mean RMS error of 0.20 N and a standard deviation of 0.05 N. Table 2 indicates that the control accuracy from the experiment is sufficient for typical functional tasks. Nonetheless, these tests were conducted on some healthy individuals; thus, future efforts will focus on the real patients.

Based on these results, it can be concluded that our NN + AHC is able to evaluate the patient's performance during training and automatically change its assistance/resistance based on the volunteer's requirement. However, at initial time, the control performance with the AHC is not better than the one with NN + AC, since RISE feedback structure in AC is more sensible to the disturbance and nonlinear than the inverse optimal controller u_{01} . Nevertheless, with the proceeding of NN + AHC learning, the inverse optimal controller begins to show better performance than the NN + AC's; the control performance of the NN + AC becomes degraded due to its inherent accumulated error.

5. Conclusions

A Lyapunov-based stability analysis indicates that the developed adaptive AC-based inverse optimal hybrid control

FIGURE 10: (a) Resultant force with NN + AC. (b) Velocity of resultant force with NN + AC.

FIGURE 11: (a) Resultant force with NN + AHC. (b) Velocity of resultant force with NN + AHC.

method yields global asymptotic tracking for an unknown nonlinear rehabilitation robot and human motor behavior dynamics, even in the presence of uncertain additive disturbances. Simulation and experimental results clearly illustrate that the proposed AHC controller enables a person to achieve a desired functional task while minimizing a meaningful cost and the proposed AHC performs better in terms of reduced error and convergence speed in comparison to AC controller. However, the performance of the controller on volunteers or patients remains to be seen. Future efforts will focus on improving the practicable intelligent control algorithms for rehabilitation robot and experimental trials on more volunteers, potentially including stroke/hemiplegia patients and so on.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] J. Eschweiler, K. Gerlach-Hahn, A. Jansen-Toy et al., "A survey on robotic devices for upper limb rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 11, no. 1, article 3, 2014.
- [2] J. Zihlerl, D. Novak, A. Olenšek, M. Mihelj, and M. Munih, "Evaluation of upper extremity robot-assistances in subacute and chronic stroke subjects," *Journal of NeuroEngineering and Rehabilitation*, vol. 7, no. 1, article 52, 2010.
- [3] K. Hayward, R. Barker, and S. Brauer, "Interventions to promote upper limb recovery in stroke survivors with severe paresis: a systematic review," *Disability and Rehabilitation*, vol. 32, no. 24, pp. 1973–1986, 2010.
- [4] G. N. Lewis and E. J. Perreault, "An assessment of robot-assisted bimanual movements on upper limb motor coordination following stroke," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 17, no. 6, pp. 595–604, 2009.
- [5] H. I. Krebs, J. J. Palazzolo, L. Dipietro et al., "Rehabilitation robotics: performance-based progressive robot-assisted therapy," *Autonomous Robots*, vol. 15, no. 1, pp. 7–20, 2003.

- [6] R. Loureiro, F. Amirabdollahian, M. Topping, B. Driessen, and W. Harwin, "Upper limb robot mediated stroke therapy—GENTLE/s approach," *Autonomous Robots*, vol. 15, no. 1, pp. 35–51, 2003.
- [7] P. S. Lum, C. G. Burgar, M. Van Der Loos, P. C. Shor, M. Majmundar, and R. Yap, "MIME robotic device for upper-limb neurorehabilitation in subacute stroke subjects: a follow-up study," *Journal of Rehabilitation Research and Development*, vol. 43, no. 5, pp. 631–642, 2006.
- [8] L. E. Kahn, P. S. Lum, W. Z. Rymer, and D. J. Reinkensmeyer, "Robot-assisted movement training for the stroke-impaired arm: does it matter what the robot does?" *Journal of Rehabilitation Research and Development*, vol. 43, no. 5, pp. 619–630, 2006.
- [9] G. Kwakkel, B. J. Kollen, and H. I. Krebs, "Effects of robot-assisted therapy on upper limb recovery after stroke: a systematic review," *Neurorehabilitation and Neural Repair*, vol. 22, no. 2, pp. 111–121, 2008.
- [10] R. Colombo, F. Pisano, S. Micera et al., "Robotic techniques for upper limb evaluation and rehabilitation of stroke patients," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 3, pp. 311–324, 2005.
- [11] T. Tsuji and Y. Tanaka, "Tracking control properties of human-robotic systems based on impedance control," *IEEE Transactions on Systems, Man, and Cybernetics A: Systems and Humans*, vol. 35, no. 4, pp. 523–535, 2005.
- [12] Y. Choi, J. Gordon, D. Kim, and N. Schweighofer, "An adaptive automated robotic task-practice system for rehabilitation of arm functions after stroke," *IEEE Transactions on Robotics*, vol. 25, no. 3, pp. 556–568, 2009.
- [13] F. Meng, Y. Dai, Y. Jin, and Y. Wang, "The design of a new upper limb rehabilitation robot system based on multi-source data fusion," in *Proceedings of the 10th World Congress on Intelligent Control and Automation (WCICA '12)*, pp. 3840–3845, Beijing, China, July 2012.
- [14] Y. Choi, J. Gordon, H. Park, and N. Schweighofer, "Feasibility of the adaptive and automatic presentation of tasks (ADAPT) system for rehabilitation of upper extremity function post-stroke," *Journal of NeuroEngineering and Rehabilitation*, vol. 8, no. 1, article 42, 2011.
- [15] M. Duff, Y. Chen, and S. Attygalle, "An adaptive mixed reality training system for stroke rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 5, pp. 531–541, 2010.
- [16] Y. Chen, M. Duff, N. Lehrer et al., "A novel adaptive mixed reality system for stroke rehabilitation: principles, proof of concept, and preliminary application in 2 patients," *Topics in Stroke Rehabilitation*, vol. 18, no. 3, pp. 212–230, 2011.
- [17] A. Erdogan, A. C. Satici, and V. Patoglu, "Passive velocity field control of a forearm-wrist rehabilitation robot," in *Proceedings of the IEEE International Conference on Rehabilitation Robotics (ICORR '11)*, pp. 1–8, Zurich, Switzerland, July 2011.
- [18] S. Jezernik, R. G. V. Wassink, and T. Keller, "Sliding mode closed-loop control of FES controlling the shank movement," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 2, pp. 263–272, 2004.
- [19] A. Ajoudani and A. Erfanian, "A neuro-sliding-mode control with adaptive modeling of uncertainty for control of movement in paralyzed limbs using functional electrical stimulation," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 7, pp. 1771–1780, 2009.
- [20] C. S. Chen, "Dynamic structure neural-fuzzy networks for robust adaptive control of robot manipulators," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 9, pp. 3402–3414, 2008.
- [21] K. Kiguchi, T. Tanaka, and T. Fukuda, "Neuro-fuzzy control of a robotic exoskeleton with EMG signals," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 4, pp. 481–490, 2004.
- [22] N. Sharma, C. M. Gregory, M. Johnson, and W. E. Dixon, "Closed-loop neural network-based NMES control for human limb tracking," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 3, pp. 712–725, 2012.
- [23] K. N. Arya, S. Pandian, R. Verma, and R. K. Garg, "Movement therapy induced neural reorganization and motor recovery in stroke: a review," *Journal of Bodywork and Movement Therapies*, vol. 15, no. 4, pp. 528–537, 2011.
- [24] R. Davoodi and B. J. Andrews, "Computer simulation of FES standing up in paraplegia: a self-adaptive fuzzy controller with reinforcement learning," *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 2, pp. 151–161, 1998.
- [25] J. Izawa, T. Kondo, and K. Ito, "Biological arm motion through reinforcement learning," *Biological Cybernetics*, vol. 91, no. 1, pp. 10–22, 2004.
- [26] U. Kartoun, H. Stern, and Y. Edan, "A human-robot collaborative reinforcement learning algorithm," *Journal of Intelligent and Robotic Systems*, vol. 60, no. 2, pp. 217–239, 2010.
- [27] S. Wang, W. Chaovaitwongse, and R. Babuška, "Machine learning algorithms in bipedal robot control," *IEEE Transactions on Systems, Man and Cybernetics C: Applications and Reviews*, vol. 42, no. 5, pp. 728–743, 2012.
- [28] T. Tamei and T. Shibata, "Fast reinforcement learning for three-dimensional kinetic human-robot cooperation with an EMG-to-activation model," *Advanced Robotics*, vol. 25, no. 5, pp. 563–580, 2011.
- [29] G. Galan and S. Jagannathan, "Adaptive critic-based neural network object contact controller for a three-finger gripper," in *Proceedings of the IEEE International Symposium on Intelligent Control, (ISIC '01)*, pp. 109–114, September 2001.
- [30] B. Kim, J. Park, S. Park, and S. Kang, "Impedance learning for robotic contact tasks using natural actor-critic algorithm," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 40, no. 2, pp. 433–443, 2010.
- [31] P. S. Thomas, M. Branicky, A. van den Bogert, and K. Jagodnik, "Creating a reinforcement learning controller for functional electrical stimulation of a human arm," in *Proceedings of the Yale Workshop on Adaptive and Learning Systems*, pp. 1–6, 2008.
- [32] P. M. Pilarski, M. R. Dawson, T. Degris, F. Fahimi, J. P. Carey, and R. S. Sutton, "Online human training of a myoelectric prosthesis controller via actor-critic reinforcement learning," in *Proceedings of the IEEE International Conference on Rehabilitation Robotics (ICORR '11)*, pp. 1–7, Zurich, Switzerland, July 2011.
- [33] S. Bhasin, M. Johnson, and W. E. Dixon, "A model-free robust policy iteration algorithm for optimal control of nonlinear systems," in *Proceedings of the 49th IEEE Conference on Decision and Control (CDC '10)*, pp. 3060–3065, IEEE, December 2010.
- [34] 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, 2010.
- [35] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon, "A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems," *Automatica*, vol. 49, no. 1, pp. 82–92, 2013.

- [36] Y. Li and S. S. Ge, "Human—robot collaboration based on motion intention estimation," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 3, pp. 1007–1014, 2013.
- [37] W. S. Newman, "Stability and performance limits of interaction controllers," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 114, no. 4, pp. 563–570, 1992.
- [38] M. Rahman, R. Ikeura, and K. Mizutani, "Investigation of the impedance characteristic of human arm for development of robots to cooperate with humans," *JSME International Journal C: Mechanical Systems, Machine Elements and Manufacturing*, vol. 45, no. 2, pp. 510–518, 2002.
- [39] E. T. Wolbrecht, V. Chan, D. J. Reinkensmeyer, and J. E. Bobrow, "Optimizing compliant, model-based robotic assistance to promote neurorehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 3, pp. 286–297, 2008.
- [40] C. T. Freeman, E. Rogers, A. a. Hughes, and K. L. Meadmore, "Iterative learning control in health care: electrical stimulation and robotic-assisted upper-limb stroke rehabilitation," *IEEE Control Systems Magazine*, vol. 32, no. 1, pp. 18–43, 2012.
- [41] F. L. Lewis, R. Selmic, and J. Campos, *Neuro-Fuzzy Control of Industrial Systems with Actuator Nonlinearities*, Society for Industrial and Applied Mathematics, Philadelphia, Pa, USA, 2002.
- [42] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.
- [43] Y. Jiang and Z. Jiang, "Robust approximate dynamic programming and global stabilization with nonlinear dynamic uncertainties," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC '11)*, pp. 115–120, December 2011.
- [44] S. Puga-Guzmán, J. Moreno-Valenzuela, and V. Santibáñez, "Adaptive neural network motion control of manipulators with experimental evaluations," *The Scientific World Journal*, vol. 2014, Article ID 694706, 13 pages, 2014.
- [45] S. Bhasin, N. Sharma, P. Patre, and W. Dixon, "Asymptotic tracking by a reinforcement learning-based adaptive critic controller," *Journal of Control Theory and Applications*, vol. 9, no. 3, pp. 400–409, 2011.
- [46] J. Campos and F. L. Lewis, "Adaptive critic neural network for feedforward compensation," in *Proceedings of the American Control Conference (ACC '99)*, vol. 4, pp. 2813–2818, June 1999.
- [47] O. Kuljaca and F. L. Lewis, "Adaptive critic design using nonlinear network structures," *International Journal of Adaptive Control and Signal Processing*, vol. 17, no. 6, pp. 431–445, 2003.
- [48] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE Circuits and Systems Magazine*, vol. 9, no. 3, pp. 32–50, 2009.
- [49] M. Krstic and P. Tsiotras, "Inverse optimal stabilization of a rigid spacecraft," *IEEE Transactions on Automatic Control*, vol. 44, no. 5, pp. 1042–1049, 1999.
- [50] J. Hu, D. M. Dawson, and Y. Qian, "Position tracking control for robot manipulators driven by induction motors without flux measurements," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 3, pp. 419–438, 1996.
- [51] P. C. Kung, C. C. K. Lin, and M. S. Ju, "Neuro-rehabilitation robot-assisted assessments of synergy patterns of forearm, elbow and shoulder joints in chronic stroke patients," *Clinical Biomechanics*, vol. 25, no. 7, pp. 647–654, 2010.