

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Adaptive Impedance Control of Robots with Reference Trajectory Learning

QING XU¹ AND XIAOMING SUN²

¹School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China (e-mail: xuq0909@buaa.edu.cn)

²College of Engineering Science and Technology, Shanghai Ocean University Shanghai, China (e-mail: xmsun@shou.edu.cn)

Corresponding author: Qing Xu (e-mail: xuq0909@buaa.edu.cn).

ABSTRACT In this paper, we propose an adaptive impedance control with reference trajectory learning for the robots interacting with unknown environment. A cost function considering its tracking errors and interaction force is introduced and a reference trajectory learning law based on iterative learning is presented to minimize it. Also, an adaptive impedance control is designed to follow the target impedance model with the adaptive reference trajectory to implement the convergence of tracking errors and interaction force. Through simulation and experimental studies, we find that the robot can autonomously adjust its trajectory or interaction mode when the environment types or cost function parameters are tuned, so that a more humanized and intelligent interaction can be realized.

INDEX TERMS Adaptive impedance control, human-robot interaction control, reference trajectory learning.

I. INTRODUCTION

IN the foreseeable future, robots will be increasingly integrated into human society, serving human beings in various fields such as old-age care, education and entertainment, forming a new harmonious social environment with us [1], [2]. In recent years, more and more robots have appeared in human environment of nursing homes, hospitals, shopping malls and even families [3].

When robots are deployed in open human environments, intentional or accidental contact with humans is unavoidable [4]. As we all know, the closer the robot gets to a human, the more likely he or she will be hurt by the robot, which brings a new challenge to the control of robots [5]. Under this circumstance, the objectives of the robot control cannot only be defined as realizing programmable, repeatable and high-precision motions [6]. Thus, in order to ensure the safety of human beings and the surrounding environment, it is very necessary to investigate the robot interaction control, so that the robots can behave compliantly to the surrounding environment [7], [8]. Also, an efficient approach is needed to be developed to seek for a balance between the robot's compliance and its accuracy.

In the literature on robot interaction control, the following two approaches are mainly utilized by researchers, hybrid position/force control [9] and impedance control [10], [11]. And impedance control is more widely used due to its simpleness,

robustness and no need to design a separate force/position loop [12], [13]. Moreover, the robot under the impedance control can adjust its motion according to the external forces, so that a stable, safe and compliant interaction with the environment can be achieved [14].

At early phase, the efforts of impedance control were made on following a fixed impedance model with a given reference trajectory for a robot with unknown model or uncertain parameters, and the environment information, which is usually unknown, was needed in the acquisition of the target model [15], [16]. In order to solve this issue, researchers tried to figure out how humans interact with unknown environment. Thus, neuroscientists have conducted several studies, showing that humans are able to be adapted to the dynamically changing environment by constantly adjusting their body movements and impedance [17], [18]. Inspired by this, researchers have made their recent efforts on achieving an ideal human-robot interaction through two typical strategies, impedance learning and reference trajectory learning, where the expected impedance model is obtained by optimizing its model parameters and reference trajectory respectively [19].

In the enlightenment of research results in [17], [18], impedance learning has been taken into account and extensively studied in the literature. In [20], a reinforcement learning approach called natural actor-critic algorithm was proposed to continuously optimize the impedance param-

ters of the robot when conducting interacting task with the environment. Also, another improved reinforcement learning with integrated path was put forward in [21], where variable parameter impedance control problem was solved through the optimizing of the cost function. In other literatures, such as [22], biomimetic adaptive controller was designed using motion features of humans to achieve impedance adaption, so that human-robot collaboration could be realized.

Additionally, reference trajectory learning has also been investigated to acquire the optimized impedance model. In [23], adaption of joint reference trajectory was proposed to implement robot trajectory tracking. However, the interaction force was regarded as a disturbance and the relationship between tracking errors and interaction force was not taken into account. In [24], reference trajectory along with position control of the robot was optimized by utilizing Kalman filter and the estimation of human's moving direction. Similarly, making use of hidden Markov model, human's moving intentions were estimated and reference trajectory learning was further designed in [25]. Nevertheless, the reference trajectory learning performance relied on the prediction of human's motion, and the interaction force was only aimed to be minimized without considering the tracking errors. In [26], an impedance model with fixed parameters was put forward, and reference trajectory learning was obtained through the optimization of the introduced cost function. However, the expected impedance model could not be acquired unless the environment information is known.

In order to improve the compliance of the robot interacting with unknown environment, and search for a trade-off between interaction force and trajectory tracking errors, an adaptive impedance control is proposed in this paper to follow a target impedance model with adaptive reference trajectory, of which the main contributions can be concluded as the following.

- Firstly, a cost function which can describe the robot interaction behavior is proposed.
- Secondly, the cost function is parameterized and then minimized through iterative learning and parameter optimization.
- Finally, an adaptive impedance controller is designed to follow the proposed target impedance model.

Compared with the work in the literature, the proposed approach is greatly simplified as the human-robot interaction problem is transformed into a parameter optimization problem. Moreover, the introduced learning algorithm requires little environment information, which is more applicable.

The remainder of the paper is organized as below. In Section II, the description of the studied system is proposed, together with the preliminaries of operation-space impedance control. In Section III, the reference trajectory learning is studied. In Section IV, operation-space impedance controller is designed. In Section V and VI, the simulation and experimental studies are respectively conducted to verify the effectiveness of the proposed approach. The conclusion and future work of this paper are given in Section VII.

II. PROBLEM FORMULATION

A. ROBOT KINEMATICS AND DYNAMICS

Using forward kinematics, the relationship between n degrees-of-freedom (DoF) robot manipulator's joint-space and operating-space can be obtained as

$$x(t) = \phi(q(t)) \quad (1)$$

where $x(t) \in \mathbb{R}^n$ denotes position and orientation of the manipulator in operation-space, and $q(t) \in \mathbb{R}^n$ is the joint variable.

Differentiating (1) and further differentiating it with respect to time, we have

$$\begin{aligned} \dot{x}(t) &= J(q(t))\dot{q}(t) \\ \ddot{x}(t) &= \dot{J}(q(t))\dot{q}(t) + J(q(t))\ddot{q}(t) \end{aligned} \quad (2)$$

where $J(q(t)) = \frac{\partial \phi(q(t))}{\partial q(t)} \in \mathbb{R}^{n \times n}$ is the so-called Jacobian matrix.

Through Lagrange modeling, the dynamic model of rigid robot system in joint-space can be expressed as

$$\begin{aligned} M(q(t))\ddot{q}(t) + C(q(t), \dot{q}(t))\dot{q}(t) + G(q(t)) \\ = \tau(t) + \tau_{ext}(t) \end{aligned} \quad (3)$$

where $M(q(t)) \in \mathbb{R}^{n \times n}$ is the inertia matrix; $C(q(t), \dot{q}(t))\dot{q}(t) \in \mathbb{R}^n$ denotes the Coriolis and centrifugal forces; $G(q(t)) \in \mathbb{R}^n$ is the gravitational force; $\tau(t) \in \mathbb{R}^n$ is the control input and $\tau_{ext}(t) = J^T(q(t))f_{ext}(t) \in \mathbb{R}^n$ denotes the external torque exerted by the environment, in which $f_{ext}(t)$ is the external force. In this paper, we suppose that the torque/force sensors are equipped on the body of the manipulator so as to measure the $\tau_{ext}(t)$ directly.

Substituting the kinematic model (1) and (2) into (4), the operation-space dynamics can be described as

$$\begin{aligned} M_x(q(t))\ddot{x}(t) + C_x(q(t), \dot{q}(t))\dot{x}(t) + G_x(q(t)) \\ = u(t) + f_{ext}(t) \end{aligned} \quad (4)$$

Property 1: [27] The matrix $M_x(q(t))$ is symmetric and positive definite.

Property 2: [27] The matrix $C_x(q(t), \dot{q}(t)) - \frac{1}{2}\dot{M}_x(q(t))$ is skew-symmetric, i.e., $\forall r \in \mathbb{R}^n, r^T(C(q(t), \dot{q}(t)) - \frac{1}{2}\dot{M}(q(t)))r = 0$.

Property 3: [28] If the relevant parameters of the robot are appropriately selected, the dynamic model in (5) can be converted to the following linear system, i.e., for any vectors $a, b \in \mathbb{R}^n$, we have

$$\begin{aligned} M_x(q(t))a + C_x(q(t), \dot{q}(t))b + G_x(q(t)) \\ = Y(a, b, q, \dot{q})\Psi \end{aligned} \quad (5)$$

where $\Psi \in \mathbb{R}^{n_\Psi}$ is the vector composed of the relevant parameters of the robot which are unknown, n_Ψ is the number of the parameters, and $Y(\cdot)$ denotes the regression matrix independent of these parameters.

B. OPERATION-SPACE IMPEDANCE CONTROL

In this paper, the target impedance model in operation-space is described as follows

$$M_D \ddot{x}(t) + C_D \dot{x}(t) + G_D(x(t) - x_r(t)) = f_{ext}(t) \quad (6)$$

where $M_D \in \mathbb{R}^{n \times n}$, $C_D \in \mathbb{R}^n$ and $G_D \in \mathbb{R}^n$ represent the expected inertia, damping, and stiffness matrices respectively, which are determined by the specific interaction requirements of the robot and the environment. Once the robot can successfully follow the given target impedance model, the robot is regarded as being compliant to the surrounding environment.

From (6), we find that the trajectory $x(t)$ and interaction force $f_{ext}(t)$ of the robot can be adjusted under the influence of impedance model parameters M_D , C_D , G_D and reference trajectory $x_r(t)$. Thus, the actual trajectory and the interaction force can be optimized through the proper adaption of model parameters and reference trajectory, which are respectively called impedance learning and reference trajectory learning. In this paper, the effort is put on reference trajectory learning.

III. REFERENCE TRAJECTORY LEARNING

The main purpose of reference trajectory learning is to optimize the $x_r(t)$ in (6) without utilizing the environment information, so as to obtain the expected human-robot interaction performance, which can be realized through the following three steps:

- Firstly, a cost function which contains the tracking errors and interaction force of the robot is introduced.
- Secondly, the cost function is parameterized through Bezier curve approach.
- At last, the cost function is optimized by iterative learning with gradient follow principle.

A. COST FUNCTION

We firstly propose a cost function $V(t)$, which includes the tracking errors and interaction force, to effectively quantify and evaluate the human-robot interaction performance

$$\begin{aligned} V(t) &= \int_{t_0}^{t_f} ((x(t) - x_t(t))^T Q (x(t) - x_t(t)) \\ &\quad + f_{ext}^T(t) R f_{ext}(t)) dt \end{aligned} \quad (7)$$

where t_0 and t_f are the start and end time, $x_t \in \mathbb{R}^n$ is the given task trajectory to be tracked, Q and R are both positive constants. Through minimizing the cost function (7), the contradictions between reducing trajectory tracking errors and the interaction force can be effective balanced, so as to achieve a safe and harmonious human-robot interaction. *Remark 1:* From (7), we find that the respective weights of tracking errors and interaction force can be adjusted through the design of parameters Q and R , which is similar to the classical Linear Quadratic Regulator (LQR) problem in optimal control. Also, some trials are required in order to achieve the exactly desired interaction performance.

B. PARAMETERIZATION

The cost function needs to be parameterized so that its optimization can be facilitated. However, from (7) we find that the terms $x(t)$ and $f_{ext}(t)$ in the cost function are difficult to be parameterized.

But according to the impedance model in (6), they and their derivative can both be determined by the reference trajectory once the parameters M_D , C_D , G_D are fixed, which means that once the $x_r(t)$ is parameterized with respect to θ , $x(t)$ and $f_{ext}(t)$ can also be parameterized as $x(\theta)$ and $f_{ext}(\theta)$, then the mapping between $V(t)$ and θ can be established. However, it is noted that the specific analytic expression of $V(\theta)$ is not required, and we only need to guarantee that after $x_r(t)$ being parameterized as $x_r(\theta)$, $V(t)$ can be determined by θ .

In the literature of robot trajectory planning, Bezier curve has been widely employed into trajectory parameterization [29], [30]. The approach is generally realized by sampling the original trajectory at different points, and linking them by smooth segments. Thus, the reference trajectory in this paper can be parameterized as

$$\begin{aligned} x_r(\theta) &= \sum_{i=0}^N P_i \frac{N!}{i!(N-i)!} \rho^i (1-\rho)^{N-i} \\ &= \sum_{i=0}^N \begin{bmatrix} \theta_{ni} \\ \vdots \\ \theta_{ni+n-1} \end{bmatrix} \frac{N!}{i!(N-i)!} \rho^i (1-\rho)^{N-i} \end{aligned} \quad (8)$$

where N is the number of sampling points, $P_i = [\theta_{ni}, \dots, \theta_{ni+n-1}]^T$ is the i -th sampling point, $\theta = [\theta_0, \dots, \theta_{N+n-1}]^T$ is the trajectory parameter, and $\rho \in [0, 1]$ is an arbitrary constant, which is defined as $\rho = \frac{t-t_0}{t_f-t_0}$ in this paper. Then, the parameterized reference trajectory can be described as

$$\begin{aligned} x_r(\theta) &= \sum_{i=0}^N \begin{bmatrix} \theta_{ni} \\ \vdots \\ \theta_{ni+n-1} \end{bmatrix} \frac{N!}{i!(N-i)!} \\ &\quad \times \left(\frac{t-t_0}{t_f-t_0} \right)^i \left(1 - \frac{t-t_0}{t_f-t_0} \right)^{N-i} \end{aligned} \quad (9)$$

After the $x_r(t)$ being parameterized, the $V(t)$ can be also parameterized as $V(\theta)$ according to the above analysis. In the following, the learning algorithm of the parameter θ is going to be designed, so that the minimum optimization of the $V(\theta)$ can be obtained.

C. PARAMETER LEARNING ALGORITHM

In order to minimize the cost function, the following mapping relation is established

$$V(\theta^{i+1}) = \lambda V(\theta^i) \quad (10)$$

where i describes the number of iterations, and λ denotes the rate of convergence. From (10), it is easy to find that once $|\lambda| < 1$, the mapping relation is converged and $V(\theta) \rightarrow V^*(\theta)$ will be achieved as long as $i \rightarrow \infty$.

Through simple transformation, we have

$$\begin{aligned} V(\theta^{i+1}) &= V(\theta^i) + (V(\theta^{i+1}) - V(\theta^i)) \quad (11) \\ &= V(\theta^i) + \frac{\partial V(\theta^i)}{\partial \theta^i} |_{\theta^i=\theta_a^i} (\theta^{i+1} - \theta^i) \\ &= V(\theta^i) + g(\theta_a^i)(\theta^{i+1} - \theta^i) \end{aligned}$$

where $g(\theta^i) = \frac{\partial V(\theta^i)}{\partial \theta^i}$ is the gradient of the cost function, and $\theta_a^i \in [\min\{\theta^i, \theta^{i+1}\}, \max\{\theta^i, \theta^{i+1}\}]$.

Thus, if we design the parameter learning algorithm as

$$\theta^{i+1} = \theta^i - \sigma^i V(\theta^i) \quad (12)$$

where σ^i is the learning rate at the i -th iteration, and we substitute it into (11), it can be obtained that

$$V(\theta^{i+1}) = (1 - g(\theta_a^i)\sigma^i)V(\theta^i) \quad (13)$$

Also, according to the above analysis, the learning rate σ^i needs to be carefully chosen, so that $|1 - g(\theta_a^i)\sigma^i| < 1$ is satisfied, and the convergence of $V(\theta^i)$ can be realized.

Remark 2: As the specific value of $g(\theta_a^i)$ is usually unknown, the following strategy is introduced to choose the learning rate σ^i , where three steps are involved.

- Firstly, the magnitude of $g(\theta_a^i)$ is estimated through the information of the last iteration, i.e., we suppose that

$$\begin{aligned} |g(\theta^i)| &\approx \left| \frac{\partial V(\theta^{i-1})}{\partial \theta^{i-1}} \right| \quad (14) \\ &= \left| \frac{V(\theta^{i-1}) - V(\theta^{i-2})}{\theta^{i-1} - \theta^{i-2}} \right| \end{aligned}$$

- Secondly, using the above estimated magnitude and assuming two different signs of it, compute two appropriate learning rates respectively, which satisfies the inequation $|1 - g(\theta_a^i)\sigma^i| < 1$.
- Finally, substituting the two learning rates into (12), then generate the new reference trajectory and choose one correct learning rate from the last step under the principle of minimizing the cost function.

IV. OPERATION-SPACE ADAPTIVE IMPEDANCE CONTROL

In the last section, the reference trajectory has been parameterized using Bezier curve in (9), and the parameter's learning algorithm has been designed as (12). In this section, based on our previous work and inspired by [31] and [32], we will propose an operation-space adaptive impedance controller for the robot to follow the target impedance model in (6) with the self-learned reference trajectory proposed above. The control diagram is illustrated as Fig. 1.

First, an impedance error $e(t) \in \mathbb{R}^n$ is defined as

$$e(t) = \ddot{x}(t) + C' \dot{x}(t) + G'(x(t) - x_r(t)) - F' f_{ext}(t) \quad (15)$$

where $C' = M_D^{-1}C_D$, $G' = M_D^{-1}G_D$ and $F' = M_D^{-1}$.

Then we introduce the following positive definite matrices $\Xi, \Omega \in \mathbb{R}^{n \times n}$, with satisfying $\Xi + \Omega = C'$ and $\Xi\Omega = G'$.

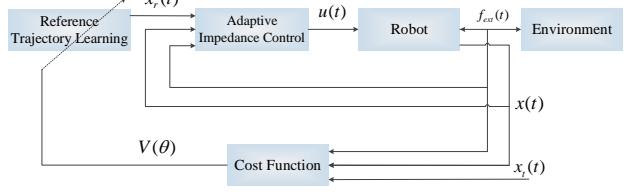


FIGURE 1: Control Diagram.

Also, an auxiliary variable $x_a \in \mathbb{R}^n$ is proposed and defined as

$$G' x_r(t) + F' f_{ext}(t) = \dot{x}_a(t) + \Xi x_a(t) \quad (16)$$

Substituting them into (15), we have

$$\begin{aligned} e(t) &= \ddot{x}(t) + \Xi \dot{x}(t) - \dot{x}_a(t) \quad (17) \\ &+ \Omega(\dot{x}(t) + \Xi x(t) - x_a(t)) \end{aligned}$$

In order to simplify the following formulas, another impedance error $\epsilon(t) \in \mathbb{R}^n$ is proposed as

$$\epsilon(t) = \dot{x}(t) + \Xi x(t) - x_a(t) \quad (18)$$

and the original impedance error in (17) can be denoted as

$$e(t) = \dot{\epsilon}(t) + \Omega \epsilon(t) \quad (19)$$

As Ω is positive definite, it is easy to find that if $\lim_{t \rightarrow \infty} \epsilon(t) = \mathbf{0}$ and $\lim_{t \rightarrow \infty} \dot{\epsilon}(t) = \mathbf{0}$, we will obtain $\lim_{t \rightarrow \infty} e(t) = \mathbf{0}$, where $\mathbf{0}$ is a zero matrix with a proper dimension.

Thus, the objective of the impedance control is changed into realizing

$$\lim_{t \rightarrow \infty} \epsilon(t) = \mathbf{0}, \lim_{t \rightarrow \infty} \dot{\epsilon}(t) = \mathbf{0} \quad (20)$$

In order to achieve the above objective, a Lyapunov function is proposed as

$$L(t) = \frac{1}{2} \epsilon^T(t) M_x(q(t)) \epsilon(t) + \frac{1}{2} \tilde{\Psi}^T \Omega \tilde{\Psi} \quad (21)$$

where $\tilde{\Psi} = \hat{\Psi} - \Psi$, Ψ is the robot's parameter vector in *Property 3*, and $\hat{\Psi}$ is its estimation. And it is obvious to find that $L(t)$ is a positive definite function.

Differentiating (21) with respect to time, we have

$$\begin{aligned} \dot{L}(t) &= \tilde{\Psi}^T \Omega \dot{\tilde{\Psi}} + \epsilon^T(t) M_x(q(t)) \dot{\epsilon}(t) \quad (22) \\ &+ \frac{1}{2} \epsilon^T(t) \dot{M}_x(q(t)) \epsilon(t) \end{aligned}$$

Substituting (18) in to the robot operation-space dynamics in (5), we find that

$$\begin{aligned} M_x(q) \dot{\epsilon}(t) + C_x(q, \dot{q}) \epsilon(t) &= u(t) + f_{ext}(t) \quad (23) \\ -M_x(q) \dot{x}_a(t) - C_x(q, \dot{q}) x_a(t) - G_x(q) \end{aligned}$$

where $x_a(t) = \dot{x}(t) - \epsilon(t) = -\Xi x(t) + x_a(t)$.

Then, substituting (23) into (22) and considering *Property 2*, we acquire that

$$\begin{aligned} \dot{L}(t) &= \tilde{\Psi}^T \Omega \dot{\tilde{\Psi}} + \epsilon^T(t) [u(t) + f_{ext}(t)] \quad (24) \\ &- M_x(q) \dot{x}_a(t) - C_x(q, \dot{q}) x_a(t) - G_x(q) \end{aligned}$$

Utilizing *Property 3*, it can be found that

$$M_x(q)\dot{x}_e(t) + C_x(q, \dot{q})x_e(t) + G_x(q) = Y(\dot{x}_e(t), x_e(t), \dot{q}, q)\Psi \quad (25)$$

Substituting it into (24), we obtain that

$$\begin{aligned} \dot{L}(t) &= \tilde{\Psi}^T \Omega \dot{\tilde{\Psi}} + \epsilon^T(t)[u(t) + f_{ext}(t)] \\ &\quad - Y(\dot{x}_e(t), x_e(t), \dot{q}, q)\Psi \end{aligned} \quad (26)$$

Therefore, the control input can be designed as

$$u(t) = -f_{ext}(t) - K_1\epsilon(t) + Y(\dot{x}_e(t), x_e(t), \dot{q}, q)\hat{\Psi} \quad (27)$$

where K_1 is a positive parameter to be designed.

Substituting (25) and (27) into (23) yields

$$\begin{aligned} M_x(q)\dot{\epsilon}(t) &= -C_x(q(t), \dot{q}(t))\epsilon(t) - K_1\epsilon(t) \\ &\quad + Y(\dot{x}_e(t), x_e(t), \dot{q}, q)\tilde{\Psi} \end{aligned} \quad (28)$$

It indicates that once the convergence to zero of $\epsilon(t)$ and $\tilde{\Psi}$ can be obtained, so will $\dot{\epsilon}(t)$. Thus, in the following, we will design the update law of $\hat{\Psi}$ to achieve it.

Firstly, we convolve the dynamics (5) by $\phi(t)$ to avoid using the acceleration variable in the update law design

$$\begin{aligned} \int_0^t \phi(t-\sigma)(M_x(q)\ddot{x} + C_x(q, \dot{q})\dot{x} + G_x(q))d\sigma &= (29) \\ &= \int_0^t \phi(t-\sigma)(u(\sigma) + f_{ext}(\sigma))d\sigma \triangleq y(t) \end{aligned}$$

where $\phi(t) = \alpha e^{-\alpha t}$, which is a low-pass filter and described as $\Phi(s) = \frac{\alpha}{s+\alpha}$ in Laplace domain.

Find a bounded signal $W(t) \in \mathbb{N}^{n \times n_\Psi}$ which satisfies

$$W(t)\Psi = y(t) \quad (30)$$

Also, for $t \in [T_e - \tau_e, T_e]$, there exists $\lambda > 0$ such that

$$\int_{T_e - \tau_e}^{T_e} W(\sigma)^T W(\sigma)d\sigma \geq \lambda \mathbf{I} \quad (31)$$

where \mathbf{I} is an unit matrix with a proper dimension.

A prediction error e_p is introduced and defined as $e_p = Q_e \tilde{\Psi}$ where

$$Q_e = \begin{cases} 0, & \text{for } t < T_e \\ Q(T_e), & \text{for } t \geq T_e \end{cases} \quad (32)$$

with

$$Q(t) = \int_{t-\tau_d}^t W^T(\sigma)W(\sigma)d\sigma \quad (33)$$

Thus, the update law of Ψ is designed as

$$\dot{\tilde{\Psi}} = -\Omega^{-1}(Y^T(\dot{x}_e(t), x_e(t), \dot{q}, q)\epsilon(t) + K_2 e_p) \quad (34)$$

where K_2 is also a positive parameter to be designed.

Remark 3: In order to derive the $e_p = Q_e \tilde{\Psi}$ in (34), $Q_e \Psi$ is required. We can multiply both sides of (30) by $W^T(t)$, then integrate it over $[t - \tau_d, t]$, we have

$$Q(t)\Psi = \int_{t-\tau_d}^t W^T(\sigma)y(\sigma)d\sigma \quad (35)$$

TABLE 1: Parameters of the robot links

Parameter	Description	Value
m_1	Mass of link 1	8.20 kg
m_2	Mass of link 2	8.50 kg
l_1	Length of link 1	0.50 m
l_2	Length of link 2	0.40 m
I_1	Inertia of link 1	4.33×10^{-2} kgm ²
I_2	Inertia of link 2	2.50×10^{-2} kgm ²

Thus, e_p can be further obtained to derive $\dot{\tilde{\Psi}}$ in (34).

Substituting (27) and (34) into (26), we obtain

$$\dot{L}(t) = -K_1\epsilon^T(t)\epsilon(t) - K_2\tilde{\Psi}^T e_p \quad (36)$$

According to (32), we find $\dot{L}(t) = -K_1\epsilon^T(t)\epsilon(t) \leq 0$ over $t \in [0, T_e]$, indicating that $L(t)$ is bounded.

While for $t > T_e$, from (32), we have

$$\begin{aligned} \dot{L}(t) &= -K_1\epsilon^T(t)\epsilon(t) - K_2\tilde{\Psi}^T Q_e \tilde{\Psi} \\ &\leq -K_1\epsilon^T(t)\epsilon(t) - K_2\lambda\Psi^T \Psi \\ &\leq -\rho L(t) \end{aligned} \quad (37)$$

where

$$\rho = \min\left(\frac{2K_1}{\|M_x\|}, \frac{2K_2}{\|\Omega\|}\right) \quad (38)$$

Hence, *Theorem 1* can be obtained.

Theorem 1: For the robot described as (5), under the control input (27) with the update law (34), all the close-loop signals are bounded, and the control objective described in (20) can be achieved.

Proof 1: Multiplying $e^{\rho t}$ on both sides of (37), we obtain

$$e^{\rho t}\dot{L}(t) \leq -\rho e^{\rho t}L(t) \quad (39)$$

$$e^{\rho t}\dot{L}(t) + \rho e^{\rho t}L(t) = \frac{d}{dt}e^{\rho t}L(t) \leq 0 \quad (40)$$

Integrating (40) over $[0, t]$ on both sides, we can obtain

$$0 \leq L(t) \leq L(0)e^{-\rho t} \quad (41)$$

meaning that $L(t)$, together with $\epsilon(t)$ and $\tilde{\Psi}$ tends to zero when $t \rightarrow \infty$.

Then, it is easy to find that $\lim_{t \rightarrow \infty} \dot{\epsilon}(t) = 0$. Thus, the control objective in (20) is achieved, which concludes the proof.

V. SIMULATION

In order to verify the effectiveness of the approach proposed in this paper, a simulation study with two different kinds of scenarios is going to be carried out in this section, where a robot manipulator with two revolute joints in [28] is going to be studied.

A. SETTINGS

The parameters of the robot links are shown as Table 1. Let m_i and l_i represent the mass and length of link i , and I_i stands for the inertia of link i passing through the center of mass of link i in the direction of coming out of the page.

The initial position of the robot are given as $q(0) = [\frac{\pi}{3}; -\frac{2\pi}{3}]$ rad in the joint-space, and $x(0) = [0 \ 0]^T$ m in the operation-space. The task trajectory is a point-to-point movement, with the start point being $[0 \ 0]^T$ m and the end point being $[0 \ 0.5]^T$ m. The movement starts at $t_0 = 0$ s and lasts for $t_f = 4$ s.

For the reference trajectory, substituting the above start and end points into (9), with the number of sampling points being $N = 2$, we have the parameterized reference trajectory as

$$x_r(t) = (\frac{t}{2} - \frac{t^2}{8})\theta + \frac{t^2}{16} \cdot [0, 0.5]^T \quad (42)$$

where θ is initialized as $[5 \ 3]^T$. Also, by setting $\theta = [0 \ 0]^T$, the task trajectory is defined as

$$x_t(t) = \frac{t^2}{16} \cdot [0, 0.5]^T \quad (43)$$

In terms of the impedance model, the parameters are chosen as

$$\begin{aligned} M_D &= \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, C_D = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}, \\ G_D &= \begin{bmatrix} 0.6 & 0 \\ 0 & 0.6 \end{bmatrix} \end{aligned} \quad (44)$$

And we set the related controller parameters as $\Xi = 3\mathbf{I}_{2 \times 2}$, $\Gamma = 2\mathbf{I}_{2 \times 2}$ and $K_1 = K_2 = 1$.

For the environment, the divergent force field (DF) model in [33] is utilized in this paper, whose center is at $[-0.1 \ 0.25]^T$ m and the radius of the workspace is set as 0.3 m.

$$f_{ext}(t) = \begin{cases} K_s(r_0 - r)\vec{n}, r \leq r_0 \\ 0, r > r_0 \end{cases} \quad (45)$$

where \vec{n} is the unit vector from center to the interaction point, K_s is a constant to describe the stiffness of different environment, and r denotes the distance from the interaction point to the center. Note that the force decreases progressively along with the r increases, and finally to zero when it reaches the boundary of the workspace region, i.e. when $r \geq r_0$.

B. SCENARIO 1: DIFFERENT TYPES OF THE ENVIRONMENT MODEL

In this scenario, we will focus on investigating the influence of different types of the environment through setting different values of the stiffness parameter K_s in (45), and the parameters in the cost function (7) are all set as $Q = R = 1$.

1) $K_s = 100$ N/m

Firstly, we set the environment as a moderate type by choosing $K_s = 100$ N/m. The simulation results are illustrated from Fig. 2 to Fig. 4.

Fig. 2 illustrates the trajectory of the robot under the proposed controller. The task trajectory is shown as a purple dash dot line. The force field around the environment is shown by green lines with the arrows denoting its direction, and the boundary of the force field is described by the blue dashed line. The trajectory of the robot is represented by the

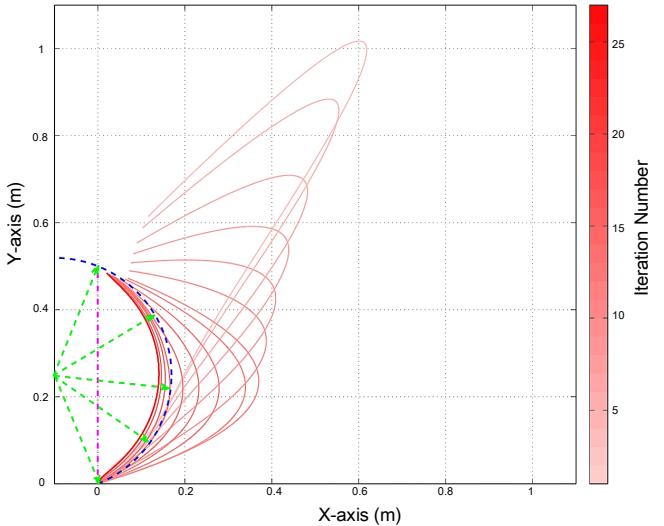


FIGURE 2: Trajectory of the robot in Scenario 1 with $K_s = 100$ N/m.

red curve. The color bar on the right of the figure shows the number of iterations, and the darker the color is, the more iterations have been conducted.

From Fig. 2, we can easily find that the trajectory of the robot gradually converge to the task trajectory iteratively, and is finally equilibrated near the force field boundary, where the contradiction of reducing tracking errors and interaction force can be best balanced.

Also, as the robot moves towards the direction of getting close to the force field, the interaction force increases iteratively, which can be shown in Fig. 3. However, from Fig. 4, we find due to the tracking errors being reduced, the cost function decreases iteratively, meaning the control objective is achieved.

Remark 4: The variation tendencies of interaction force, the cost function and the trajectory parameter in the following scenarios are almost the same as the ones in Figs. 3 and 4, thus they are omitted to reduce the paper length.

2) $K_s = 500$ N/m

Then, we change the environment type into a strong one by setting $K_s = 500$ N/m. The simulation result is illustrated in Fig. 5.

We find that the robot trajectory also converges iteratively, but it is equilibrated at the place which is more far away from the task trajectory compared to that in Fig. 2. It shows that in spite of the existence of tracking errors, the robot refuses to move too close to the strong environment and avoids bringing about a huge interaction force and damage to the environment.

3) $K_s = 10$ N/m

In this case, the environment is changed into a soft one by setting $K_s = 10$ N/m, and the result is shown as Fig. 6.

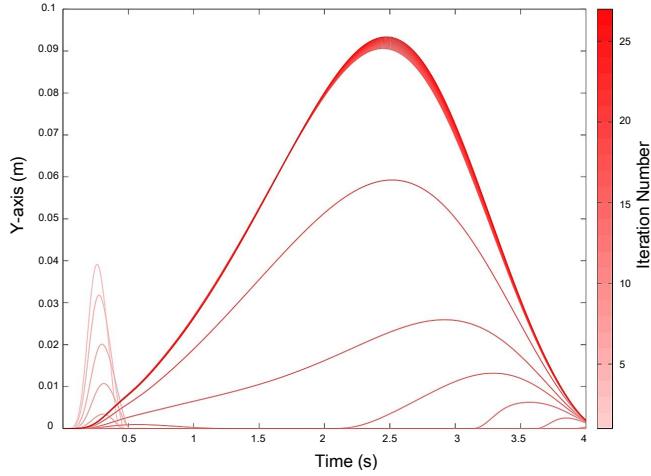


FIGURE 3: Interaction force in Scenario 1 with $K_s = 100$ N/m.

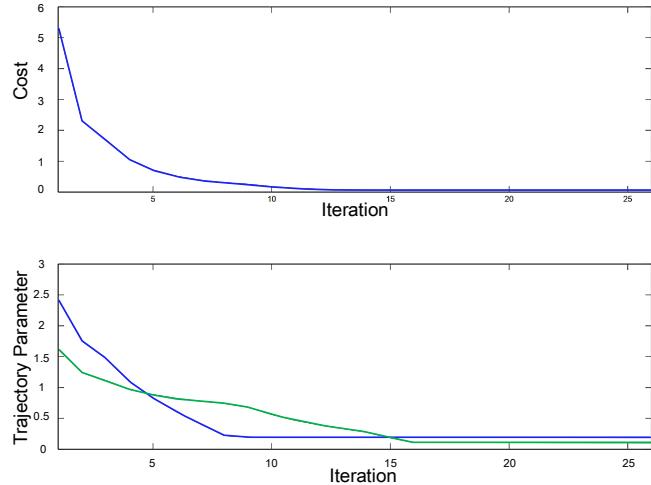


FIGURE 4: Cost Function and Reference Trajectory Parameter in Scenario 1 with $K_s = 100$ N/m.

We find that the robot trajectory gets even closer compared to the one in Fig. 2, in order to further reduce the tracking errors. As the environment is soft, there is no worry of the interaction force being rocket up when the robot approaches the force field center.

In conclusion, the three simulation results are coincident with the relationship between the tracking errors and interaction force described in (7), i.e., when the environment is strong, the interaction force is the key to the cost function and the robot trajectory will finally move along the force field boundary; on the contrary, when the environment is soft, the tracking error plays a more important role in the cost function and the equilibrium trajectory will get closer to the task trajectory to reduce the tracking errors. In other words, the robot under the controller proposed in this paper is able to autonomously adjust its movement and balance the contradiction between trajectory tracking errors and interaction force according to different environment types.

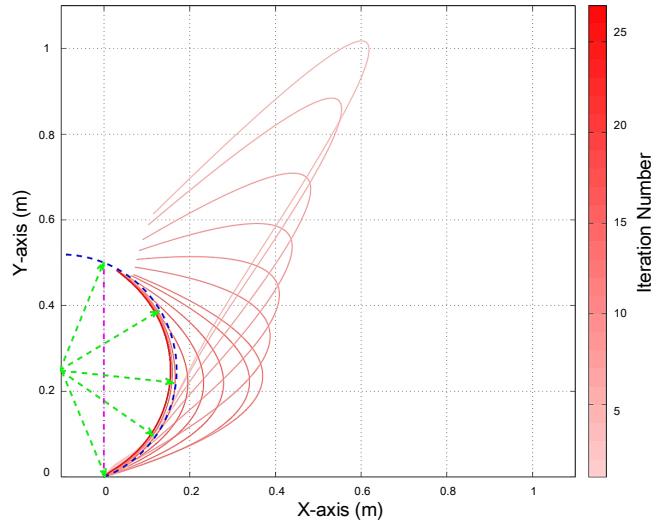


FIGURE 5: Trajectory of the robot in Scenario 1 with $K_s = 500$ N/m.

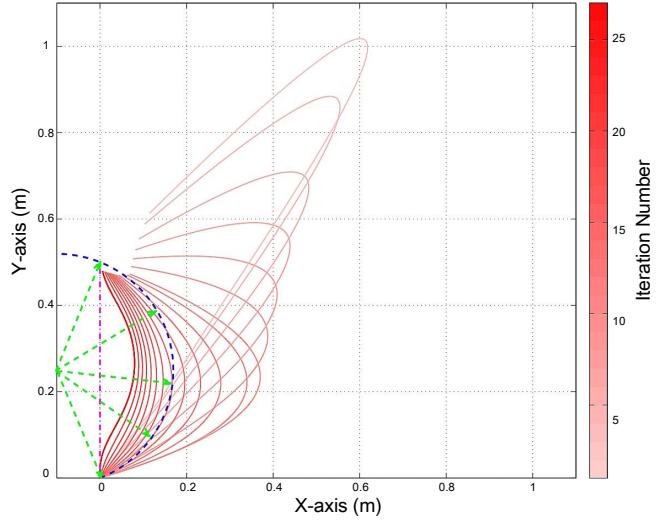


FIGURE 6: Trajectory of the robot in Scenario 1 with $K_s = 10$ N/m.

C. SCENARIO 2: DIFFERENT PARAMETERS OF THE COST FUNCTION

In this scenario, the effort is put on investigating the effect of the cost function by tuning the related parameters in (7), where two sets of specific parameters are considered. And the environment parameters are both set as $K_s = 100$.

1) $Q = 100, R = 1$

First, we let the trajectory tracking account for a larger percentage, by setting Q as a relative large value. The result is illustrated in Fig. 7.

2) $Q = 1, R = 100$

Then, the value of Q and R is exchanged so that the interaction force plays a more important role in the cost function.

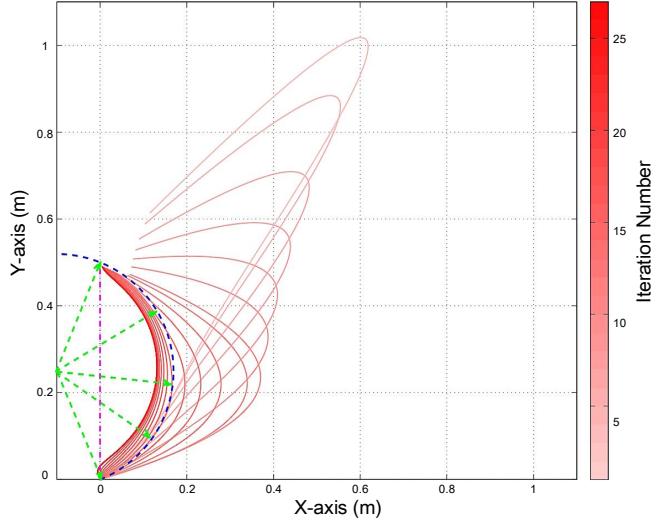


FIGURE 7: Trajectory of the robot in Scenario 2 with $Q = 100, R = 1$.

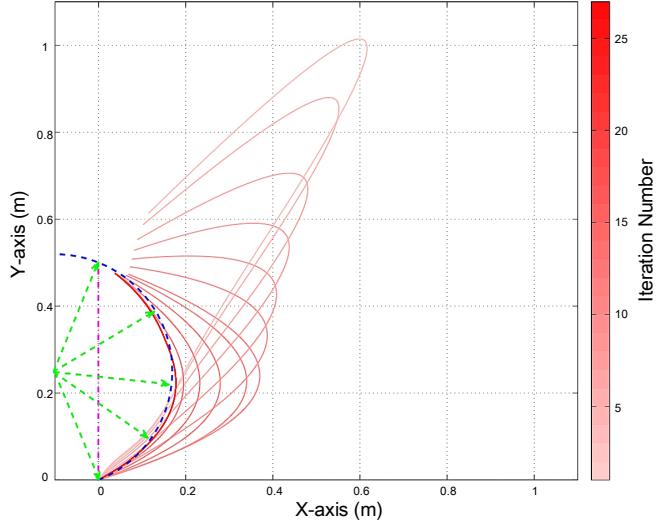


FIGURE 8: Trajectory of the robot in Scenario 2 with $Q = 1, R = 100$.

The result is illustrated in Fig. 8, showing that the trajectory is equilibrated at the place where the interaction force is smaller, compared to the one in Fig. 7.

Facing the same environment, the robot can interact in different manners by tuning the relative values of Q and R in the cost function, so as to pursue different interaction objectives, such as compliant interaction forces or accurate trajectory tracking. It is similar to the result in [34], which indicates that humans from various cultures or at different ages would like to adjust their interaction manners when facing the same type of environment.

VI. EXPERIMENT

In order to further verify the effectiveness of the proposed approach, an experimental study is conducted on a two DoF

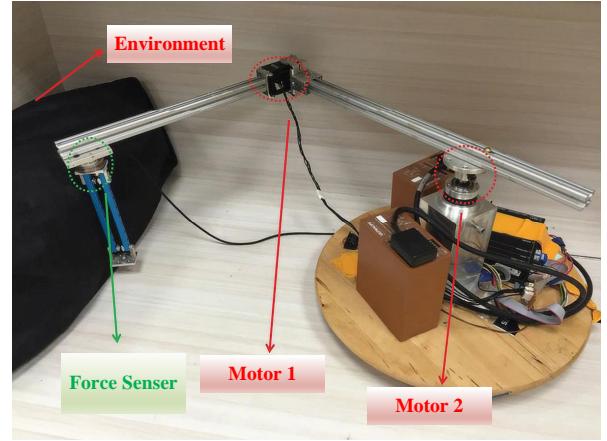


FIGURE 9: Experiment settings.

robot manipulator, shown as Fig. 9.

A. EXPERIMENT SETTINGS

The robot consists with two MAXON DC motors controlled by MAXON EPOS2 70/10. Also, it is equipped with an ATI Mini40-E sensor to measure the interaction force. The lengths of the robot links are $l_1 = 0.35$ m and $l_2 = 0.35$ m, with their mass being $m_1 = 0.32$ kg and $m_2 = 0.44$ kg. A soft and deformable toy is used to represent the external environment. And a Lenovo laptop is served as an external computing unit.

The robot position is initialized at $[0.63, -1.26]^T$ rad in the joint-space, and $[0.55, 0]^T$ m in the operation-space. The task trajectory is a point-to-point uniform linear movement along the X-axis from $[0.55, 0]^T$ m to $[0.60, 0]^T$ m. The movement starts at $t_0 = 0$ s and lasts for $t_f = 5$ s.

The reference trajectory is parameterized as $[0.60, 0]^T + p(t)\theta[0.05, 0]^T$, where $p(t) = 6(\frac{t}{5})^5 - 15(\frac{t}{5})^4 + 10(\frac{t}{5})^3$, with θ being initialized as 1. For the impedance model, the parameters are chosen as

$$\begin{aligned} M_D &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, C_D = \begin{bmatrix} 11 & 0 \\ 0 & 11 \end{bmatrix}, \\ G_D &= \begin{bmatrix} 30 & 0 \\ 0 & 30 \end{bmatrix} \end{aligned} \quad (46)$$

And the related controller parameters are set as $Q = 100$, $R = 0.1$, $\Xi = 5\mathbf{I}_{2 \times 2}$, $\Omega = 6\mathbf{I}_{2 \times 2}$ and $K_1 = K_2 = 1$.

B. EXPERIMENT RESULTS AND DISCUSSIONS

The experiment results are illustrated from Figs. 10 to 12. As the robot movement is along with the X-axis, only the X-axis trajectory is shown.

From Fig. 11, we find that the contact between the robot and the environment starts at $t = 1.5$ s, and the interaction force continuously increases as the contact is gradually getting tight. However, under the proposed approach, the robot is moving far away from its task trajectory iteratively, in order to reduce the interaction force, which can be demonstrated from Figs. 10 and 11.

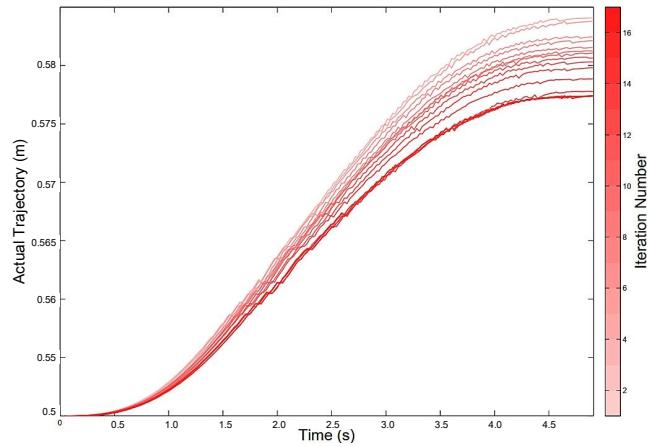


FIGURE 10: Trajectory of the robot in X-axis.

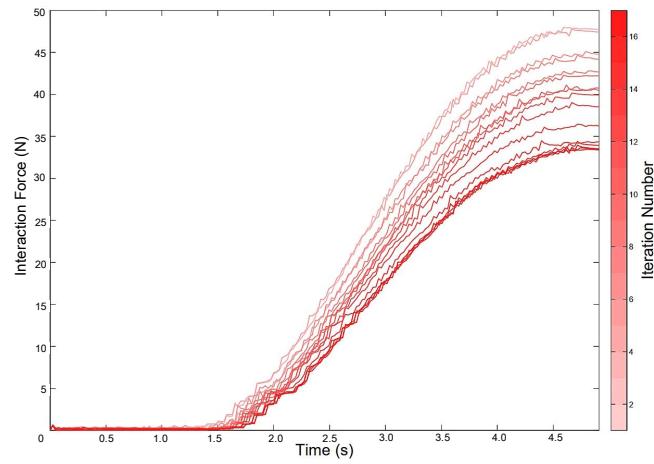


FIGURE 11: Interaction Force.

Moreover, the trajectory and the interaction force are converged after about 12 iterations, and the cost function is successfully minimized at the same time, illustrated from Fig. 12. Thus, it is proved that the robot under the proposed controller in this paper can autonomously adjust its trajectory to improve the interaction performance without using the environment information.

VII. CONCLUSION AND FUTURE WORK

Aiming at the interaction control of robots with unknown environmental information, an adaptive impedance control with reference trajectory learning was proposed in this paper. A cost function, which can describe the interaction performance was introduced and minimized through the parameter optimization, then the autonomous balance of the contradictions between the tracking errors and interaction force can be achieved. Then, the effectiveness of the proposed approach was proved by a simulation study involved two interaction scenarios and an experiment. The approach provides a solution to improving the human-robot interaction performance only by optimizing the reference trajectory, and a possibility of changing the interaction manners only by tuning the cost

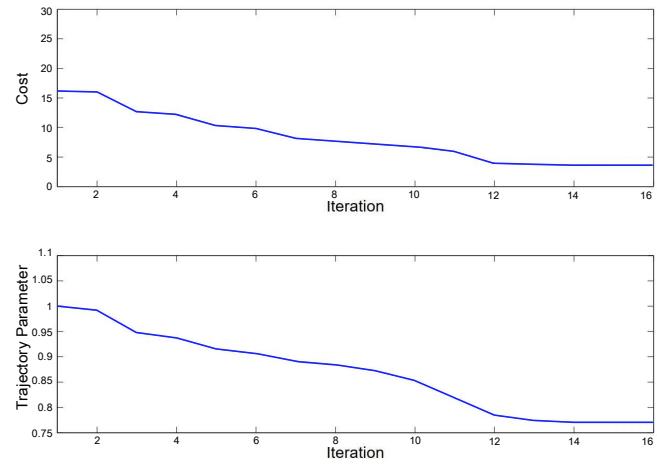


FIGURE 12: Cost Function and Reference Trajectory Parameter.

function parameters.

The performance of the proposed strategy depends on the merits of the cost function to a large extend. Thus, in the future, it is worth thinking about how to take good advantages of the available environment information as a guiding role in the optimization of robot interaction performance. Also, according to [17], humans can adapt their impedance and movement simultaneously when facing unknown environment. How to effectively combine the reference trajectory learning with the impedance learning is another interesting and challenging field to be studied.

REFERENCES

- [1] W. He and Y. Dong, "Adaptive fuzzy neural network control for a constrained robot using impedance learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 1174–1186, 2018.
- [2] Z. Li, B. Huang, A. Ajoudani, C. Yang, C.-Y. Su, and A. Bicchi, "Asymmetric bimanual control of dual-arm exoskeletons for human-cooperative manipulations," *IEEE Transactions on Robotics*, vol. 34, no. 1, pp. 264–271, 2018.
- [3] Q. Xu and S. S. Ge, "Adaptive control of redundant robot manipulators with null-space compliance," *Assembly Automation*, vol. 38, no. 5, p. 615–624, 2018.
- [4] W. He, S. S. Ge, Y. Li, E. Chew, and Y. S. Ng, "Neural network control of a rehabilitation robot by state and output feedback," *Journal of Intelligent & Robotic Systems*, vol. 80, no. 1, pp. 15–31, 2015.
- [5] Y. Li, K. P. Tee, W. L. Chan, R. Yan, Y. Chua, and D. K. Limbu, "Continuous role adaptation for human–robot shared control," *IEEE Transactions on Robotics*, vol. 31, no. 3, pp. 672–681, 2015.
- [6] Z. Li, B. Huang, Z. Ye, M. Deng, and C. Yang, "Physical human–robot interaction of a robotic exoskeleton by admittance control," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 12, pp. 9614–9624, 2018.
- [7] Z. Li, J. Liu, Z. Huang, Y. Peng, H. Pu, and L. Ding, "Adaptive impedance control of human–robot cooperation using reinforcement learning," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 10, pp. 8013–8022, 2017.
- [8] S. S. Ge, Y. Li, and C. Wang, "Impedance adaptation for optimal robot–environment interaction," *International Journal of Control*, vol. 87, no. 2, pp. 249–263, 2014.
- [9] H. Navvabi and A. H. Markazi, "Hybrid position/force control of stewart manipulator using extended adaptive fuzzy sliding mode controller (eafsmc)," *ISA transactions*, vol. 88, pp. 280–295, 2019.
- [10] N. Hogan, "Impedance control: An approach to manipulation: Part ii' implementation," 1985.

- [11] Y. Li and S. S. Ge, "Impedance learning for robots interacting with unknown environments," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 4, pp. 1422–1432, 2013.
- [12] T. Boaventura, J. Buchli, C. Semini, and D. G. Caldwell, "Model-based hydraulic impedance control for dynamic robots," *IEEE Transactions on Robotics*, vol. 31, no. 6, pp. 1324–1336, 2015.
- [13] W. He, Y. Dong, and C. Sun, "Adaptive neural impedance control of a robotic manipulator with input saturation," *IEEE Transactions on Systems Man and Cybernetics Systems*, vol. 46, no. 3, pp. 334–344, 2016.
- [14] Y. Li, S. Sam Ge, and C. Yang, "Learning impedance control for physical robot–environment interaction," *International Journal of Control*, vol. 85, no. 2, pp. 182–193, 2012.
- [15] S. P. Buerger and N. Hogan, "Complementary stability and loop shaping for improved human–robot interaction," *IEEE Transactions on Robotics*, vol. 23, no. 2, pp. 232–244, 2007.
- [16] A. Saccon, J. Hauser, and A. Beghi, "Trajectory exploration of a rigid motorcycle model," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 2, pp. 424–437, 2011.
- [17] D. W. Franklin, G. Liaw, T. E. Milner, R. Osu, E. Burdet, and M. Kawato, "Endpoint stiffness of the arm is directionally tuned to instability in the environment," *Journal of Neuroscience*, vol. 27, no. 29, pp. 7705–7716, 2007.
- [18] E. Burdet, R. Osu, D. W. Franklin, T. E. Milner, and M. Kawato, "The central nervous system stabilizes unstable dynamics by learning optimal impedance," *Nature*, vol. 414, no. 6862, pp. 446–449, 2001.
- [19] H. Vallery, E. H. Van Asseldonk, M. Buss, and H. Van Der Kooij, "Reference trajectory generation for rehabilitation robots: complementary limb motion estimation," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 17, no. 1, pp. 23–30, 2008.
- [20] 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, Part B (Cybernetics)*, vol. 40, no. 2, pp. 433–443, 2009.
- [21] J. Buchli, F. Stulp, E. Theodorou, and S. Schaal, "Learning variable impedance control," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 820–833, 2011.
- [22] C. Yang, G. Ganesh, S. Haddadin, S. Parusel, A. Albu-Schaeffer, and E. Burdet, "Human-like adaptation of force and impedance in stable and unstable interactions," *IEEE transactions on robotics*, vol. 27, no. 5, pp. 918–930, 2011.
- [23] S. Ito, M. Darainy, M. Sasaki, and D. J. Ostry, "Computational model of motor learning and perceptual change," *Biological cybernetics*, vol. 107, no. 6, pp. 653–667, 2013.
- [24] K. Wakita, J. Huang, P. Di, K. Sekiyama, and T. Fukuda, "Human-walking-intention-based motion control of an omnidirectional-type cane robot," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 1, pp. p.285–296.
- [25] 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.
- [26] Y. Li, G. Ganesh, N. Jarrassé, S. Haddadin, A. Albu-Schaeffer, and E. Burdet, "Force, impedance, and trajectory learning for contact tooling and haptic identification," *IEEE Transactions on Robotics*, vol. 34, no. 5, pp. 1170–1182, 2018.
- [27] Q. Xu and X. Sun, "Adaptive operation-space control of redundant manipulators with joint limits avoidance," in *2018 Tenth International Conference on Advanced Computational Intelligence (ICACI)*, pp. 358–363, IEEE, 2018.
- [28] M. Li, Y. Li, S. S. Ge, and T. H. Lee, "Adaptive control of robotic manipulators with unified motion constraints," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 1, pp. 184–194, 2016.
- [29] K. Jolly, R. S. Kumar, and R. Vijayakumar, "A bezier curve based path planning in a multi-agent robot soccer system without violating the acceleration limits," *Robotics and Autonomous Systems*, vol. 57, no. 1, pp. 23–33, 2009.
- [30] S. Liu and D. Sun, "Minimizing energy consumption of wheeled mobile robots via optimal motion planning," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 2, pp. 401–411, 2013.
- [31] Y. Pan and H. Yu, "Composite learning robot control with guaranteed parameter convergence," *Automatica*, vol. 89, pp. 398–406, 2018.
- [32] T. Sun, L. Peng, L. Cheng, Z.-G. Hou, and Y. Pan, "Composite learning enhanced robot impedance control," *IEEE transactions on neural networks and learning systems*, 2019.
- [33] E. Burdet, G. Ganesh, C. Yang, and A. Albu-Schäffer, "Interaction force, impedance and trajectory adaptation: by humans, for robots," in *Experimental Robotics*, pp. 331–345, Springer, 2014.
- [34] V. S. Chib, J. L. Patton, K. M. Lynch, and F. A. Mussa-Ivaldi, "Haptic identification of surfaces as fields of force," *Journal of neurophysiology*, vol. 95, no. 2, pp. 1068–1077, 2006.



QING XU received the B.S. degree in automation from Beijing Forestry University, Beijing, China, in 2014. He is currently pursuing the Ph.D. degree in control science and engineering with Beihang University, Beijing, China. His research interests include human-robot interaction control, social robot and intelligent control.



XIAOMING SUN received her bachelor degree from Beijing Forestry University, Beijing, China, in 2007, and Ph.D degree in the Department of Automation Science and Electrical Engineering, Beihang University (BUAA) in 2016. Since 2016, she has been with the school of Engineering Science and Technology, Shanghai Ocean University as a lecturer. Her research interest covers adaptive control, intelligent control and control applications.

...