Reference Adaptation for Robots in Physical Interactions With Unknown Environments

Chen Wang, Yanan Li, *Member, IEEE*, Shuzhi Sam Ge, *Fellow, IEEE*, and Tong Heng Lee, *Member, IEEE*

***Abstract*—In this paper, we propose a method of reference adaptation for robots in physical interactions with unknown environments. A cost function is constructed to describe the interaction performance, which combines trajectory tracking error and interaction force between the robot and the environment. It is minimized by the proposed reference adaptation based on trajectory parametrization and iterative learning. An adaptive impedance control is developed to make the robot be governed by the target impedance model. Simulation and experiment studies are conducted to verify the effectiveness of the proposed method.**

***Index Terms*—Impedance control, interaction control, reference adaptation.**

I. Introduction

As ROBOTS make their way into daily applications such as health care, elderly care, education, etc., there is an increasing demand for research in control of robots in physical interactions with surrounding environments [1]. In a typical interaction task such as table cleaning, surface exploration, and environment identification, the robot is supposed to track a predefined task trajectory while at the same time maintaining certain compliance to the environment force.

In the previous research of interaction control, there are two methods that are widely used: 1) hybrid position/force control [2] and 2) impedance control [3]. Compared to hybrid position/force control, impedance control is well recognized due to its robustness and the fact that no direct decomposition for position control and force control is required [4]. Under the framework of impedance control, robots are controlled to modulate their motion according to the force from the environment, and stable interactions between the robots and the environments are achieved [5]. Early research effort has been made on tracking a given impedance model with a fixed reference trajectory in the presence of unknown robot model and uncertainties (see [6]–[8]). In many situations, to impose a fixed impedance model to the robot is too conservative and the knowledge of the environment dynamics is required to specify a target impedance model [9], [10]. In recent studies, how to obtain an impedance model and a reference trajectory that result in desired interaction performance has attracted much attention. To understand the mechanisms that humans use in physical interactions with environments, neuroscientists have investigated human motor control and adaptation using controlled force fields [11]–[14]. It has been shown that the central nervous system of humans has an excellent ability to repetitively adjust and tune the motion and impedance of the limb subject to changing environments and uncertain internal dynamics.

In the enlightenment how humans adapt to physical interactions with environments, impedance adaptation/learning has been investigated in [10] and [15]–[18]. In [15], a natural actor-critic algorithm is adopted to determine the optimal impedance parameters for robotic contact tasks. In [16], a reinforcement learning algorithm called policy improvement with path integrals is developed for variable impedance control which focuses on optimizing a cost function designed for a specific task. In [17], a novel human-like learning controller is proposed for robots interacting with unknown environments which minimizes motion error and effort without requiring force sensing. In [10], impedance adaptation is proposed for robots interacting with unknown time-invariant environments.

Besides impedance adaptation/learning, reference adaptation/learning also has to be taken into account to achieve desirable learning/adaptation performance [19]. Trajectory planning and learning have been studied extensively in autonomous robotics, where physical interactions between environments and robots are not taken into consideration [20]–[23]. In [24], adaptation of desired joint-angular trajectories is proposed to achieve trajectory tracking. However, the major control goal is to guarantee the joint trajectory tracking with the interaction force treated as a disturbance and the tradeoff between trajectory tracking and compliance to the interaction force is not considered. Reference adaptation/learning has also been studied in the field of physical human–robot interaction, where the human motion is modeled and estimated, and the robot's reference trajectory is updated accordingly to synchronize the robot's motion with the human's motion intention. In [25], motion characteristics of humans are considered for reference adaptation of robots in human–robot co-manipulation and the robot estimates the intended human motion and uses this identified motion to move along with the operator besides using

Manuscript received September 22, 2015; revised January 13, 2016 and March 15, 2016; accepted April 23, 2016. This paper was recommended by Associate Editor P. X. Liu.

C. Wang, S. S. Ge, and T. H. Lee are with the Department of Electrical and Computer Engineering, and the Social Robotics Laboratory, Interactive and Digital Media Institute, National University of Singapore, Singapore 117576 (e-mail: wang_chen09@nus.edu.sg; samge@nus.edu.sg; eleleeth@nus.edu.sg).

Y. Li is with the Department of Bioengineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: yli4@imperial.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2016.2562698

admittance control to react to interaction forces generated by its operator. In [26], human's motion intention is estimated using the interaction force and it is used for reference adaptation of the robot. In [27] and [28], a hidden Markov model is implemented to estimate the human intention and the robot's reference trajectory is modified accordingly. In [29], human's moving direction is estimated using the Kalman filter and it is used for the position control of the robot. In [30] and [31], human's motion intention is estimated by minimizing the interaction force and the robot's reference trajectory is adapted accordingly. The above research works [25]–[31] mainly focus on reference adaptation based on human's motion intention, and the control objective is to minimize the interaction force between human and robot and the desired reference trajectory is considered to be the same as human's intention. In [32], reference shaping is developed only to make the robot be governed by a given impedance model while at the same time guaranteeing the constraint satisfaction. In [33], an impedance model with fixed impedance parameters is obtained by minimizing a cost function, and the reference trajectory is adapted to make the robot dynamics follow this given impedance model. This method is only applicable when the environment is known because otherwise the target impedance model cannot be obtained.

Based on above discussions, we propose a method to adapt the reference trajectory subject to unknown environments. This method is based on iterative learning which was first proposed in [34]. The proposed reference adaptation includes three steps. First, a cost function is defined to evaluate the desired interaction performance, which combines the trajectory tracking error and the interaction force. Second, an adaptation law is developed to update the reference trajectory of the robot, such that the defined cost function is minimized in an iterative manner. Unlike [33], the knowledge of the environment is not required in this step. Lastly, an adaptive impedance control in the Cartesian space is developed so that the robot's dynamics are governed by the target impedance model.

The rest of this paper is organized as follows. Section II describes the system to be studied in this paper and the control objective. In Section III, the proposed reference adaptation is introduced. In Section IV, an adaptive impedance control is developed. In Sections V and VI, simulation and experimental results of the proposed method are presented and discussed. Section VII concludes this paper.

II. Preliminaries

A. System Description

In this paper, we consider a general scenario where a rigid robot arm is in physical interaction with an environment. A force sensor is mounted at the end-effector of the robot arm, which is used to measure the interaction force between the robot arm and the environment. The kinematics of the robot arm are given by

$$X(t) = \phi(q(t)) \tag{1}$$

where $X(t) \in \mathbb{R}^{n_C}$ and $q(t) \in \mathbb{R}^n$ are positions of the robot arm in the Cartesian space and coordinates in the joint space, respectively, n_C is the dimension of the Cartesian space, and n is the degree-of-freedom (DOF). Differentiating (1) with respect to time results in

$$\dot{X}(t) = J(q(t))\dot{q}(t) \tag{2}$$

where $J(q(t)) \in \mathbb{R}^{n_C \times n}$ is the Jacobian matrix. Further differentiating (2) results in

$$\ddot{X}(t) = \dot{J}(q(t))\dot{q}(t) + J(q(t))\ddot{q}(t). \tag{3}$$

The dynamics of the robot arm in the joint space are given by

$$M(q(t))\ddot{q}(t) + C(q(t), \dot{q}(t))\dot{q}(t) + G(q(t)) = \tau(t) + J^T(q(t))F(t) \tag{4}$$

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 $F(t) \in \mathbb{R}^{n_C}$ denotes the interaction force exerted by the environment. By substituting the kinematics (1)–(3) into (4), we have the dynamics of the robot arm in the Cartesian space, as follows:

$$M_R(q(t))\ddot{X}(t) + C_R(q(t), \dot{q}(t))\dot{X}(t) + G_R(q(t)) = u(t) + F(t) \tag{5}$$

where

$$\begin{aligned} M_R(q(t)) &= J^{-T}(q(t))M(q(t))J^{-1}(q(t)) \\ C_R(q(t), \dot{q}(t)) &= J^{-T}(q(t))\Big(C(q(t), \dot{q}(t)) \\ &\quad - M(q(t))J^{-1}(q(t))\dot{J}(q(t))\Big)J^{-1}(q(t)) \\ G_R(q(t)) &= J^{-T}(q(t))G(q(t)) \\ u(t) &= J^{-T}(q(t))\tau(t). \end{aligned} \tag{6}$$

Property 1 [35]: Matrix $M_R(q(t))$ is symmetric and positive definite.

Property 2 [35]–[38]: Matrix $2C_R(q(t), \dot{q}(t)) - \dot{M}_R(q(t))$ is a skew-symmetric matrix if $C(q(t), \dot{q}(t))$ is in the Christoffel form, i.e., $\rho^T(2C_R(q(t), \dot{q}(t)) - \dot{M}_R(q(t)))\rho = 0,\ \forall \rho \in \mathbb{R}^{n_C}$.

Property 3 [35], [39], [40]: The dynamics are linear in terms of a suitably selected set of the physical parameters of the robot arm, that is

$$M_R(q(t))a + C_R(q(t), \dot{q}(t))b + G_R(q(t)) = Y(a, b, \dot{q}(t), q(t))\Psi \tag{7}$$

for any $a, b \in \mathbb{R}^{n_C}$, where $\Psi \in \mathbb{R}^{n_\Psi}$ is a vector of the physical parameters of the robot arm; n_Ψ is a positive integer denoting the number of these parameters; and $Y(a, b, \dot{q}(t), q(t)) \in \mathbb{R}^{n_C \times n_\Psi}$ is the regression matrix, which is independent of the physical parameters.

The other part of the system is the environment that the robot is supposed to physically interact with. Without loss of generality, the following environment model is considered [18]:

$$M_E\ddot{X}(t) + C_E\dot{X}(t) + G_E X(t) = -F(t) \tag{8}$$

where M_E, C_E, and G_E are inertia, damping, and stiffness matrices of the environment which are supposed to be unknown in this paper.

B. Impedance Control

Impedance control is usually implemented in the control of robots in physical interactions with the environment. In particular, the dynamics of the robot arm (5) follow a target impedance model, as follows:

$$M_D\ddot{X}(t) + C_D\dot{X}(t) + G_D(X(t) - X_r(t)) = F(t) \tag{9}$$

where M_D, C_D, and G_D are desired inertia, damping, and stiffness matrices, respectively, and $X_r(t)$ is the reference trajectory.

Remark 1: Besides (9), impedance models in other forms are also studied in [3], [4], and [41]. For example, the following are two simplified impedances models:

$$\begin{aligned} &C_D\dot{X}(t) + G_D(X(t) - X_r(t)) = F(t) \\ &G_D(X(t) - X_r(t)) = F(t). \end{aligned} \tag{10}$$

From the given impedance model (9), it can be easily derived that the actual position of the robot arm $X(t)$ will be refined according to the interaction force $F(t)$ and the reference trajectory $X_r(t)$. To modulate the response of the robot's motion [$X(t)$, $\dot{X}(t)$, and $\ddot{X}(t)$] to the interaction force $F(t)$, i.e., the interaction performance, we may design impedance parameters M_D, C_D, and G_D, as well as the reference trajectory $X_r(t)$. As discussed in Section I, we focus on the design of the reference trajectory $X_r(t)$ in this paper to achieve the desired interaction performance.

Remark 2: In impedance control, a fixed impedance model (M_D, C_D, and G_D) and a fixed reference trajectory $X_r(t)$ to the robot are too conservative, and the environment dynamics have to be considered for desired interaction performance. To address this issue, both impedance adaptation/learning and reference adaptation/learning are required. In previous works [10], [15]–[18], impedance adaptation/learning (to optimize the impedance parameters M_D, C_D, and G_D) is investigated, and in this paper, reference adaptation [to optimize the reference trajectory $X_r(t)$] is studied.

C. Control Objective

The following cost function is defined to quantify the interaction performance:

$$V = \int_{t_0}^{t_f} \left((X(t) - X_t(t))^T Q(X(t) - X_t(t)) + F^T(t)RF(t)\right)\mathrm{d}t \tag{11}$$

where t_0 and t_f are the starting and ending times of each iteration, respectively, $X_t(t)$ is a given task trajectory, Q is a positive semi-definite matrix, and R is a positive definite matrix. By minimizing V, a tradeoff between trajectory tracking and minimization of the interaction force can be achieved, and thus the desired interaction performance achieved.

Remark 3: The rational behind introducing a cost function in interaction control is similar to that in the linear-quadratic regulator (LQR) problem where a cost function is often defined to quantify the control performance. For a feedback controller, we can specify the feedback gains which will have a similar impact on the control performance, however, the LQR provides a systematic way to find the feedback gains that guarantee the optimal control performance. Similarly, a cost function is defined in interaction control to quantify the interaction performance and by minimizing it the proposed reference adaptation guarantees the desired interaction performance. It is possible to achieve the same interaction performance by tuning feedback gains but it must rely on trials and errors. The advantage of the cost-function-based method is especially obvious when the environments are changing, since the desired interaction performance is guaranteed with the defined cost function while it is not with the predefined feedback gains.

Remark 4: The experiments in [12] have demonstrated that humans tend to compensate for the original task trajectory when the environment is compliant. On the contrary, when the environment is stiff, humans will adjust the trajectory in an effort to decrease the interaction force and the reference trajectory will gradually deform to the environment surface. The observed phenomenon can be modeled by a maintained balance between the trajectory tracking error and the interaction force, as in (11).

The control objective is described by the cost function with different combinations of Q and R. Based on this cost function, a target impedance model can be obtained through reference adaptation without the knowledge of the environment, which will be detailed in the following section.

III. Reference Adaptation

The aim of reference adaptation is to update the reference trajectory according to the dynamics of the environment, such that the desired interaction performance can be achieved. In the following, the reference trajectory and thus the defined cost function in Section II-C is parameterized and then the parameterized cost function is minimized by developing an adaptation method. These two steps will be introduced in the following two sections, respectively.

A. Parametrization of Cost Function

By considering (8) and (9), we obtain

$$\begin{aligned} (M_E + M_D)\ddot{X}(t) + (C_D + C_E)\dot{X}(t) + (G_D + G_E)X(t) \\ = -G_D X_r(t) \end{aligned} \tag{12}$$

from which we see that the actual trajectory of the robot arm $X(t)$ can be obtained based on $X_r(t)$, and thus can be represented as $X(\theta)$ where θ is the trajectory parameter. From the environment model (8), we see that the interaction force $F(t)$ can be also obtained based on $X(\theta)$, and thus can be represented as $F(\theta)$. Then, it is obvious that the cost function $V(t)$ given in (11) can be determined by the trajectory parameters θ. Therefore, the objective becomes looking for an optimal set of θ such that the corresponding cost function $V(\theta)$ can be minimized, that is

$$\theta^* = \arg\min_{\theta} V(\theta). \tag{13}$$

The key idea of this paper is to first parameterize the reference trajectory X_r using $X_r(\theta)$ and then optimize the parameter θ as to improve the interaction performance which is represented using $V(\theta)$. In this regard, as long as the trajectory

can be parameterized as $X_r(\theta)$, the proposed method could be applied for reference adaptation.

Remark 5: In motion and path planning of autonomous robots, Bezier curves have been widely used in order to interpolate and to parameterize the trajectory [42]–[44]. By using a Bezier curve, the simplest method to approximate a trajectory is to evaluate it at several control points and form an approximated trajectory by connecting a sequence of line segments. Based on the above idea, we approximate the reference trajectory X_r as follows:

$$\begin{aligned} X_r(\theta) &= \sum_{i=0}^{N} P_i \frac{N!}{i!(N-i)} \rho^i (1-\rho)^{N-i} \\ &= \sum_{i=0}^{N} \begin{bmatrix} \theta_{ni} \\ \vdots \\ \theta_{ni+n-1} \end{bmatrix} \frac{N!}{i!(N-i)} \rho^i \times (1-\rho)^{N-i} \end{aligned} \tag{14}$$

where N is the number of control points connected by a sequence of line segments to form the trajectory, $P_i = [\theta_{ni}, \ldots, \theta_{ni+n-1}]^T$ is the ith control point, $\theta = [\theta_0, \ldots, \theta_{m-1}]^T$ is the trajectory parameter where $m = (N+1)n$ is the dimension of θ and $\rho \in [0, 1]$. For example, we can define $\rho = (t - t_0/t_f - t_0)$. Then, the reference trajectory X_r in the time sequence becomes

$$\begin{aligned} X_r(\theta) = \sum_{i=0}^{N} \begin{bmatrix} \theta_{ni} \\ \vdots \\ \theta_{ni+n-1} \end{bmatrix} \frac{N!}{i!(N-i)} \left(\frac{t-t_0}{t_f-t_0} \right)^i \\ \times \left(1 - \frac{t-t_0}{t_f-t_0} \right)^{N-i}. \end{aligned} \tag{15}$$

When the reference trajectory is parameterized using Bezier curves, it will suffer from the computation cost due to the combinatorial explosion. In practice, a tradeoff between the computation cost and the inclusion of various types of reference trajectory should be maintained. If the goal is to have more accurate reference shapes, more control points should be chosen. However, if the goal is to reduce the computation cost, the number of control point should be reduced. Besides Bezier curves, there are other methods for trajectory parametrization, e.g., polynomial parametrization, Fourier approximation, Quintic Bezier splines and dynamic representations such as dynamical movement primitives. In different applications, we can select appropriate parametrization methods based on specific task requirements.

B. Adaptation Law

This section is dedicated to develop an adaptation law to obtain θ^*. The basic idea is to construct a mapping

$$V^* - V\left(\theta^{j+1}\right) = \lambda\left(V^* - V\left(\theta^j\right)\right) \tag{16}$$

where $V^* = V(\theta^*)$ denotes the minimum of $V(\theta)$, j is the iteration index, and λ is the convergence rate. The convergence of the mapping is discussed in the following lemma.

Lemma 1 [45]: If $|\lambda| < 1$, $V \to V^*$ as $j \to \infty$.

To achieve the above mapping, a simple adaptation law can be designed as

$$\theta^{j+1} = \theta^j + \gamma^j\left(V^* - V\left(\theta^j\right)\right) \tag{17}$$

where $\theta^j = [\theta_0^j, \ldots, \theta_{m-1}^j]^T$ and $\gamma^j = [\gamma_0^j, \ldots, \gamma_{m-1}^j]^T$ is the adaptation rate at the jth iteration. By defining the gradient

$$g\left(\theta^j\right) = \left(\frac{\partial V\left(\theta^j\right)}{\partial \theta^j} \right)^T \tag{18}$$

we have

$$\begin{aligned} V^* - V\left(\theta^{j+1}\right) &= V^* - V\left(\theta^j\right) - \left(V\left(\theta^{j+1}\right) - V\left(\theta^j\right)\right) \\ &= V^* - V\left(\theta^j\right) - \left(\frac{\partial V\left(\theta^j\right)}{\partial \theta^j} \right)^T \Big|_{\theta j = \theta_a^j} \\ &\quad \times \left(\theta^{j+1} - \theta^j\right) \\ &= \left(1 - g\left(\theta_a^j\right)\gamma^j\right)\left(V^* - V\left(\theta^j\right)\right) \end{aligned} \tag{19}$$

where $\theta_a^j \in [\min\{\theta^j, \theta^{j+1}\}, \max\{\theta^j, \theta^{j+1}\}]$. According to Lemma 1, as long as $|\lambda| = |1 - g(\theta_a^j)\gamma^j| < 1$, the convergence to the minimized cost function is achieved.

However, V^* is used in the adaptation law, which is unknown. To avoid this limitation, we revise the adaptation law as follows:

$$\theta^{j+1} = \theta^j - \sigma^j V\left(\theta^j\right) \tag{20}$$

where $\sigma^j = [\sigma_0^j, \ldots, \sigma_{m-1}^j]^T$ is the new adaptation rate. Then, the constructed mapping becomes

$$\begin{aligned} V\left(\theta^{j+1}\right) &= V\left(\theta^j\right) + \left(V\left(\theta^{j+1}\right) - V\left(\theta^j\right)\right) \\ &= V\left(\theta^j\right) + g\left(\theta_a^j\right)\left(\theta^{j+1} - \theta^j\right). \end{aligned} \tag{21}$$

By substituting (20) into (21), we have

$$V\left(\theta^{j+1}\right) = \left(1 - g\left(\theta_a^j\right)\sigma^j\right)V\left(\theta^j\right). \tag{22}$$

Similarly as in Lemma 1, the new iteration rate σ^j must also satisfy the relationship $|1 - g(\theta_a^j)\sigma^j| < 1$. In the following section, we discuss the selection of σ^j based on [46].

C. Selection of Adaptation Rate

The selection of σ^j depends on the knowledge of $g(\theta_a^j)$. When $g(\theta_a^j)$ is completely known, σ^j can be selected such that

$$\left|1 - g\left(\theta_a^j\right)\sigma^j\right| = 0 \tag{23}$$

which will lead to the fastest convergence.

When the sign and bounds of $g(\theta_a^j)$ are known, the convergence of the trajectory adaptation can also be assured. For example, if

$$0 < \alpha_k \le g_k\left(\theta_a^j\right) \le \beta_k < \infty \tag{24}$$

where α_k and β_k are the lower and upper bounds of the kth gradient component $g_k(\theta_a^j)$, respectively, then we can select $\sigma_k^j = (1/m\beta_k)$.

When neither the bounds nor the sign of $g(\theta_a^j)$ is known, special treatment for the adaptation rate σ^j is needed. A solution to this problem is to perform extra learning to determine

Algorithm 1 Learning of Gradient $g(\theta)$

1: Choose two initial trajectory parameters θ^0 and θ^1 and perform the robot motion. Compute $V(\theta^0)$ and $V(\theta^1)$ and Let $j = 2$.
2: Estimate the amplitude of each component of the gradient amplitude using

$$\left|g_k\left(\theta_a^j\right)\right| = \left|\frac{V\left(\theta^{j-1}\right) - V\left(\theta^{j-2}\right)}{\theta_k^{j-1} - \theta_k^{j-2}}\right|$$

3: Choose different sign combinations of the gradient $g(\theta_a^j)$ and determine the adaptation rate σ^j by making

$$\left|1 - g\left(\theta_a^j\right)\sigma^j\right| < 1$$

4: Update the trajectory parameter θ^{j+1} with

$$\theta^{j+1} = \theta^j - \sigma^j V\left(\theta^j\right)$$

and generate new trajectory $X_r^{j+1}(t)$.
5: Perform the robot motion and select the sign of $g(\theta_a^j)$ and θ^{j+1} corresponding to the minimum $V(\theta^j)$.
6: Let $j \leftarrow j + 1$ and go to Step 2.

the correct gradient sign. From (21), we know that when the sign of σ^j is selected wrongly, the cost function will increase. Therefore, certain extra learning trials are sufficient to determine the correct sign of σ^j. In general, if $g(\theta_a^k)$ is m-dimensional, there will be 2^m sets of trials that are needed to determine the correct sign of σ^j.

To speed up the adaptation process, the gradient component $g_k(\theta_a^j)$ can be also numerically estimated using the previous cost function and previous parameters, as follows:

$$g_k\left(\theta_a^j\right) = \frac{V\left(\theta^{j-1}\right) - V\left(\theta^{j-2}\right)}{\theta_k^{j-1} - \theta_k^{j-2}}. \tag{25}$$

Then, the adaptation rate σ^j can be adjusted accordingly. The sign of the gradient is critically important in obtaining the desired trajectory parameters, but the estimation of the gradient may not always result in a correct sign. Therefore, the aforementioned extra learning can be combined with the estimation. In particular, learning is used to determine the sign of the gradient, while (25) is used to obtain the magnitude, that is

$$\left|g_k\left(\theta_a^j\right)\right| = \left|\frac{V\left(\theta^{j-1}\right) - V\left(\theta^{j-2}\right)}{\theta_k^{j-1} - \theta_k^{j-2}}\right|. \tag{26}$$

We summarize the above procedures to learn the gradient $g(\theta)$ for the reference trajectory adaptation, such that the desired interaction performance is achieved subject to unknown environments.

IV. Adaptive Impedance Control in Cartesian Space

As the reference trajectory $X_r(t)$ is obtained according to (14) with the adaptation law (20) in the previous section, the effort will then focus on designing adaptive impedance control to make the robot dynamics (5) track the given impedance model (9). The following design is an adaptive counterpart of the learning version in our previous work [47].

Define the impedance error

$$\varepsilon(t) = \ddot{X}(t) + K_C\dot{X}(t) + K_G(X(t) - X_r(t)) - K_F F(t) \tag{27}$$

where $K_C = M_D^{-1}C_D$, $K_G = M_D^{-1}G_D$, and $K_F = M_D^{-1}$. Choose two positive definite matrices which satisfy $\Lambda + \Gamma = K_C$ and $\Lambda\Gamma = K_G$, and define the filtered auxiliary variable $X_l(t)$ as

$$K_G X_r(t) + K_F F(t) = \dot{X}_l(t) + \Lambda X_l(t). \tag{28}$$

Then, (27) can be rewritten as

$$\varepsilon(t) = \ddot{X}(t) + \Gamma\dot{X}(t) - \dot{X}_l(t) + \Lambda\left(\dot{X}(t) + \Gamma X(t) - X_l(t)\right). \tag{29}$$

By defining another impedance error

$$Z(t) = \dot{X}(t) + \Gamma X(t) - X_l(t) \tag{30}$$

the following equation can be obtained:

$$\varepsilon(t) = \dot{Z}(t) + \Lambda Z(t). \tag{31}$$

According to (31), if $\lim_{t\to\infty} Z(t) = 0$ and $\lim_{t\to\infty} \dot{Z}(t)$ exists, then $\lim_{t\to\infty} \varepsilon(t) = 0$, since Λ is positive definite. Therefore, the control objective of the adaptive impedance control is to make

$$\lim_{t\to\infty} Z(t) = 0. \tag{32}$$

By considering (30), we can rewrite (5) as

$$\begin{aligned} M_R(q(t))\dot{Z}(t) + C_R(q(t), \dot{q}(t))Z(t) &= u(t) + F(t) \\ &- M_R(q(t))\dot{X}_v(t) - C_R(q(t), \dot{q}(t))X_v(t) - G_R(q(t)) \end{aligned} \tag{33}$$

where

$$X_v(t) = -\Gamma X(t) + X_l(t). \tag{34}$$

In addition, we have

$$Z(t) = \dot{X}(t) - X_v(t). \tag{35}$$

We propose an adaptive impedance control in the Cartesian space as follows:

$$u(t) = -F(t) - KZ(t) + Y(q(t), \dot{q}(t), \dot{X}_v(t), X_v(t))\hat{\Psi} \tag{36}$$

where $\hat{\Psi}$ is updated as follows:

$$\dot{\hat{\Psi}} = -\Gamma^{-1}Y^T\left(q(t), \dot{q}(t), \dot{X}_v(t), X_v(t)\right)Z(t) \tag{37}$$

and K is a positive definite matrix. In the update law (37), $\hat{\Psi}$ is the estimate of Ψ in (7).

Theorem 1: Considering the robot dynamics (5), the control input (36) with the parameter updating law (37), the following results are guaranteed: 1) $Z(t)$ asymptotically converges to 0 as $t \to \infty$ and 2) all the closed-loop signals are bounded.

Proof: See the Appendix. ■

The proposed control framework is summarized in Fig. 1. In this framework, the first step is to generate the reference trajectory $X_r(t)$ in the Cartesian space based on the evaluation of the interaction performance $V(\theta)$. After that, the developed adaptive impedance control is implemented to make the robot dynamics follow the target impedance model.

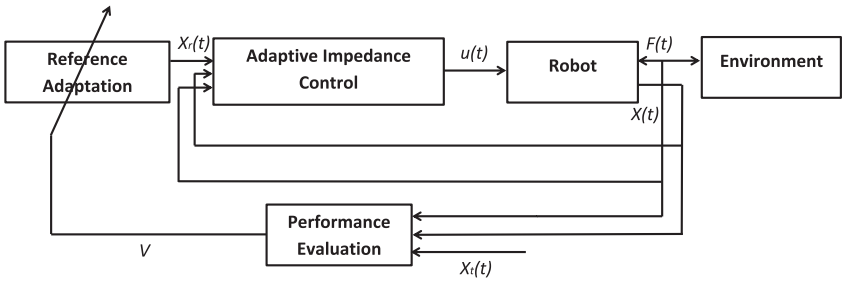


Fig. 1. Control diagram.

TABLE I
PARAMETERS OF THE ROBOT ARM IN SIMULATION

Parameter	Description	Value
m_1	Mass of link 1	2.00kg
m_2	Mass of link 2	0.85kg
l_1	Length of link 1	0.40m
l_2	Length of link 2	0.40m
I_1	Inertia moment of link 1	0.02kgm^2
I_2	Inertia moment of link 2	0.02kgm^2

V. SIMULATION

To verify the proposed method, we consider a robot arm with two revolute joints in physical interaction with an unknown environment. The simulation scenario is inspired by the tasks in which a predefined trajectory is expected to be tracked and at the same time a contact force needs to be maintained between the robot and the environment. As discussed in Section I, these tasks can be found in applications such as table cleaning, surface exploration and environment identification. The simulation is conducted with the Robotics Toolbox [48].

A. Settings

The parameters of the robot arm are given in Table I, where m_j, l_j, and I_j, $j = 1, 2$, represent the mass, the length, and the inertia moment passing through the center of mass, respectively. The initial position of the robot arm in the joint space are given as $q_1(0) = (\pi/3)$ and $q_2(0) = -(2\pi/3)$. The initial position in the Cartesian space is $x(t_0) = [0\ 0]^T$m. The task trajectory is a point to point movement from $x_t(t_0) = [0\ 0]^T$m to $x_t(t_f) = [0\ 0.5]^T$ m. The movement duration is 4 s. The reference trajectory is parameterized using a second-order Bezier curve as discussed in Section III-A. To make the reference trajectory coincide with the task trajectory at the beginning and end points with $X_r(t_0) = X_t(t_0)$ and $X_r(t_f) = X_t(t_f)$, we have the reference trajectory represented as

$$X_r(t) = \left(1 - \frac{t-t_0}{t_f-t_0}\right)^2 X_t(t_0) + 2\left(1 - \frac{t-t_0}{t_f-t_0}\right)\frac{t-t_0}{t_f-t_0}\theta + \left(\frac{t-t_0}{t_f-t_0}\right)^2 X_t(t_f) \quad (38)$$

where θ is the trajectory parameter to be optimized which is initially set as $\theta = [2\ 3]^T$. The predefined task trajectory is given by setting $\theta = [0\ 0]^T$, as follows:

$$X_t(t) = \left(1 - \frac{t-t_0}{t_f-t_0}\right)^2 X_t(t_0) + \left(\frac{t-t_0}{t_f-t_0}\right)^2 X_t(t_f). \quad (39)$$

In the impedance model (9), we set

$$M_D = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, C_D = \begin{bmatrix} 0.7 & 0 \\ 0 & 0.7 \end{bmatrix}, G_D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (40)$$

The parameters Γ, Λ, and K are chosen as

$$\Gamma = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix},\ \Lambda = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},\ K = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}. \quad (41)$$

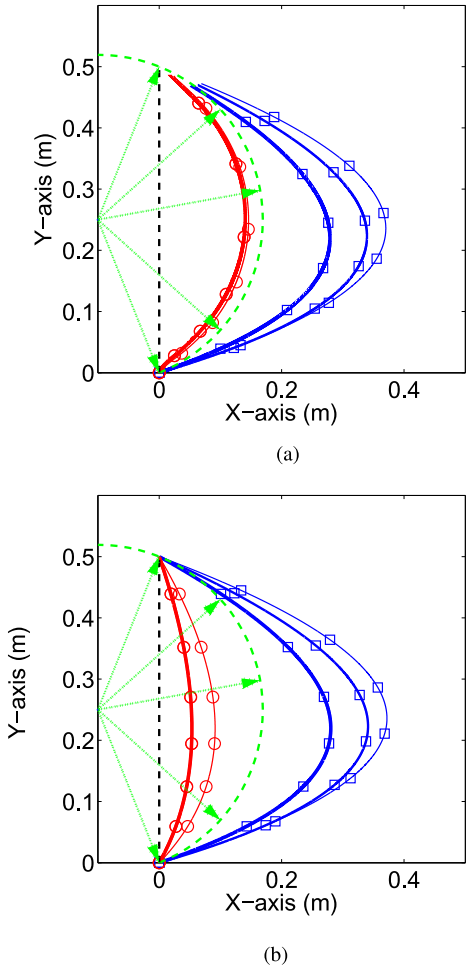
The adaptive impedance control discussed in Section IV is applied to make the dynamics of the robot arm be governed by the target impedance model. Similarly to the experiment in [12], we consider the environment as a radial force field centered at $x_c = [-0.1\ 0.25]^T$m and with a radius of $r_0 = 0.27$ m, that is

$$F = \begin{cases} K_E(r_0 - r)\vec{n}, & r \le r_0 \\ 0, & r > r_0 \end{cases} \quad (42)$$

where $\vec{n}$ is the unit vector pointing from the force field center to the interaction point, K_E is the stiffness constant, and r is the distance between a point and the force field center.

B. Different Environments

In the first case, the environment stiffness is chosen as $K_E = 110$ N/m. The performance parameters in (11) are selected as $Q = 1$ and $R = 1$. The simulation results are shown in Figs. 2–4. From Fig. 2, we can see that trajectories converge iteratively under the proposed reference adaptation. As the initial actual trajectory is far away from the force field, it gradually deforms to the task trajectory in order to reduce the tracking error. This leads to increase of the interaction force as shown in Fig. 3. The equilibrium reference trajectory and equilibrium actual trajectory are obtained after ten iterations where a tradeoff between the interaction force and the tracking error is achieved. By comparing the equilibrium reference trajectory in Fig. 2(b), we also notice that the equilibrium actual trajectory in Fig. 2(a) deviates to the direction in which the interaction force decreases. This can be explained by studying the impedance model in (9) which defines a compliant behavior of the robot arm. From Fig. 4, it is found that the cost function becomes smaller with respect to iterations, which is followed by the convergence of the trajectory parameters.



(a)

(b)

Fig. 2. Actual trajectory and reference trajectory of first and last three iterations with $K_E = 110$ N/m. The first three iterations are denoted using the blue lines with square markers (line width increases as iteration number increases from 1 to 3) and the last three iterations are denoted using the red lines with circle markers (line width increases as iteration number increases from 8 to 10). The predefined task trajectory is denoted by the black dashed line. The radial force field is represented using green dashed arrows pointing to the direction in which the interaction force decreases and encircled by the green dashed line which is the boundary of the force field. (a) Actual trajectory. (b) Reference trajectory.

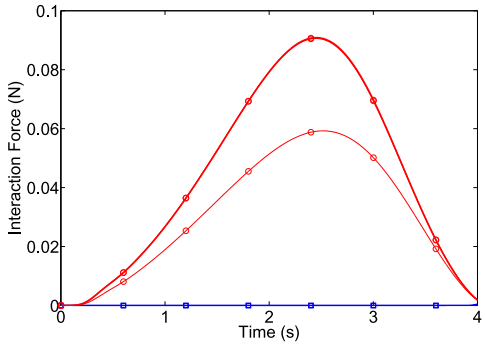


Fig. 3. Interaction force of first and last three iterations with $K_E = 110$ N/m. The first three iterations are denoted using the blue lines with square markers (line width increases as iteration number increases from 1 to 3) and the last three iterations are denoted using the red lines with circle markers (line width increases as iteration number increases from 8 to 10).

Simulation studies are conducted with another two environment stiffness K_E: 1) 300 N/m and 2) 10 N/m, and the results are shown in Figs. 5 and 6, respectively. In the circumstance

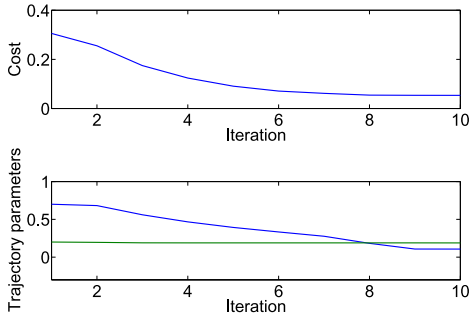
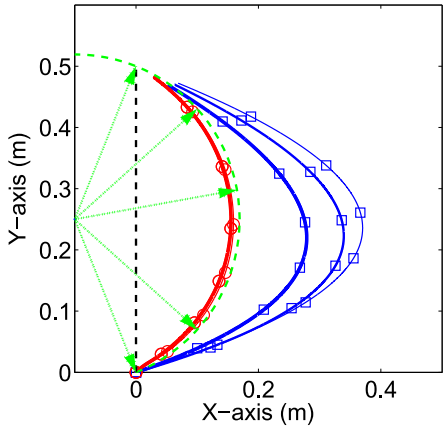
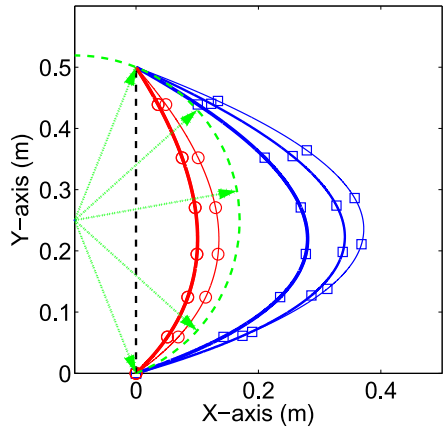


Fig. 4. Cost function and trajectory parameters. The two trajectory parameters are denoted in blue and green lines in the below subfigure.



(a)



(b)

Fig. 5. Actual trajectory and reference trajectory of first and last three iterations with $K_E = 300$ N/m. The first three iterations are denoted using the blue lines with square markers (line width increases as iteration number increases from 1 to 3) and the last three iterations are denoted using the red lines with circle markers (line width increases as iteration number increases from 8 to 10). The predefined task trajectory is denoted by the black dashed line. The radial force field is represented using green dashed arrows pointing to the direction in which the interaction force decreases and encircled by the green dashed line which is the boundary of the force field. (a) Actual trajectory. (b) Reference trajectory.

of a strong force field (Fig. 5), the reference trajectory and the actual trajectory deviate more in the direction in which the interaction force decreases (compared to Fig. 2). When the force field is weak (Fig. 6), the reference trajectory

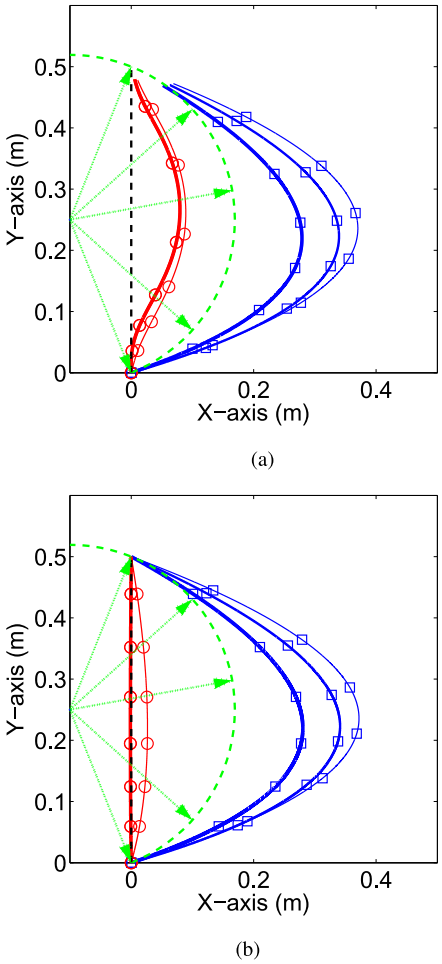


Fig. 6. Actual trajectory and reference trajectory of first and last three iterations with $K_E = 10$ N/m. The first three iterations are denoted using the blue lines with square markers (line width increases as iteration number increases from 1 to 3) and the last three iterations are denoted using the red lines with circle markers (line width increases as iteration number increases from 8 to 10). The predefined task trajectory is denoted by the black dashed line. The radial force field is represented using green dashed arrows pointing to the direction in which the interaction force decreases and encircled by the green dashed line which is the boundary of the force field. (a) Actual trajectory. (b) Reference trajectory.

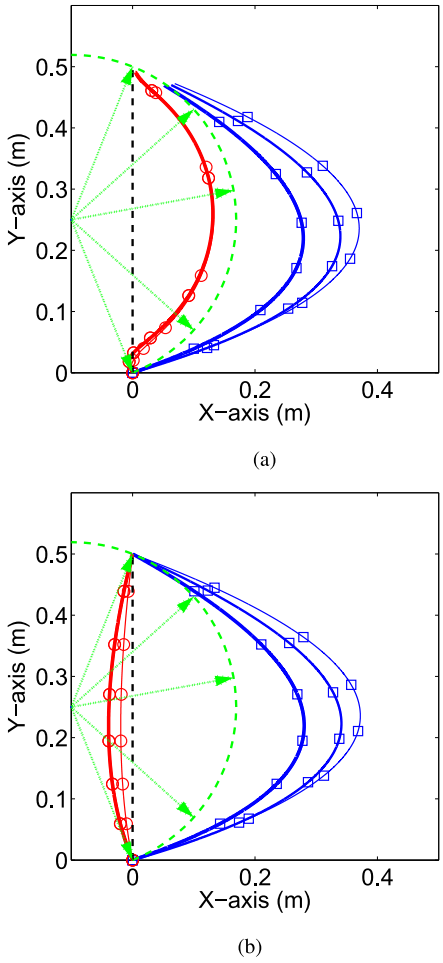
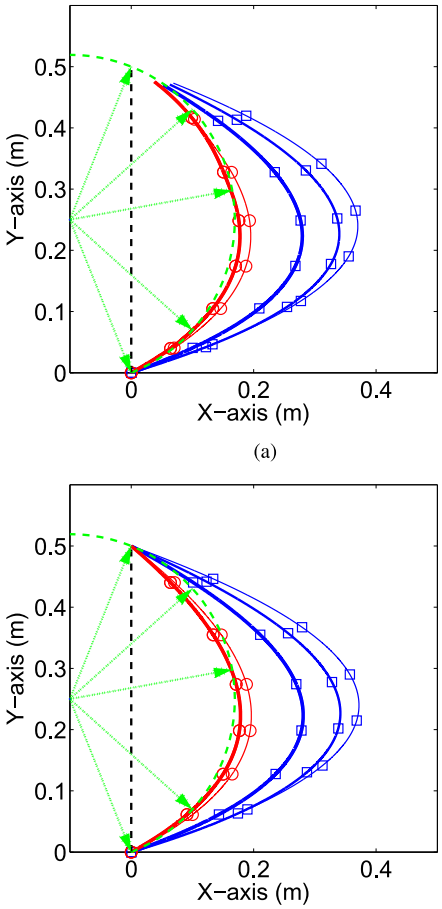


Fig. 7. Actual trajectory and reference trajectory of the first and last three iterations with $Q = [100\ 0; 0\ 100]$ and $R = 1$. The first three iterations are denoted using the blue lines with square markers (line width increases as iteration number increases from 1 to 3) and the last three iterations are denoted using the red lines with circle markers (line width increases as iteration number increases from 8 to 10). The predefined task trajectory is denoted by the black dashed line. The radial force field is represented using green dashed arrows pointing to the direction in which the interaction force decreases and encircled by the green dashed line which is the boundary of the force field. (a) Actual trajectory. (b) Reference trajectory.

and the actual trajectory are closer to the predefined task trajectory and the equilibrium reference trajectory almost coincides with the task trajectory. These results are in line with the performance requirement described by (11). When the force field is strong, the interaction force plays a major role in (11), so the equilibrium reference trajectory will be closer to the force field boundary where the interaction force is minimized; conversely, when the force field is weak, the tracking error plays a major role so the equilibrium reference trajectory will be closer to the task trajectory as to minimize the tracking error. With more simulation studies, it can be further shown that when there is no interaction force, the equilibrium reference trajectory and the equilibrium actual trajectory will be identical to the task trajectory. This is similar to the human experiment results observed in [13], where it shows that humans tend to make compensatory movements with small interaction forces, and seek a tradeoff between tracking errors and interaction forces in force fields with moderate stiffness. Based on the above observations, the proposed method could be used for force boundary (object surface) exploration and environment identification.

C. Different Cost Functions

In this section, we consider different performance requirements defined by different cost functions in (11): $Q = 100$, $R = 1$ and $Q = 1$, $R = 100$. The environment stiffness is chosen as $K_E = 110$ N/m. From simulation results in Figs. 7 and 8, it can be observed that when Q is relatively larger, the equilibrium reference trajectory and the actual trajectory gradually deviate in the direction in which the tracking error decreases. Conversely, when R is relatively larger, the trajectories deviate in the direction in which the interaction force decreases. By recalling the cost function (11) again, we



(a)

(b)

Fig. 8. Actual trajectory and reference trajectory of the first and last three iterations with $Q = [1\ 0; 0\ 1]$ and $R = 100$. The first three iterations are denoted using the blue lines with square markers (line width increases as iteration number increases from 1 to 3) and the last three iterations are denoted using the red lines with circle markers (line width increases as iteration number increases from 8 to 10). The predefined task trajectory is denoted by the black dashed line. The radial force field is represented using green dashed arrows pointing to the direction in which the interaction force decreases and encircled by the green dashed line which is the boundary of the force field. (a) Actual trajectory. (b) Reference trajectory.

know that the tracking error plays a major role when Q is relatively larger, so the equilibrium reference trajectory and the actual trajectory will be closer to the task trajectory. When R is relatively larger, the interaction force plays a major role in (11), so the equilibrium reference trajectory and the actual trajectory will be closer to the force field boundary where the interaction force is minimized. It can be concluded that different Q and R can be chosen to realize different interaction performances, e.g., either "softer" interaction or more accurate trajectory tracking [10]. This is similar to the human experiment results where the interaction performance can be also adjusted by humans [13].

VI. Experiment

A. Settings and Results

In this section, we conduct an experimental study of the proposed method with a 2-DOF robot arm, as shown in Fig. 9. Two dc motors are controlled by an EPOS2 70/10 motor controller. An ATI mini-40 force/torque sensor is mounted at the end-effector of the robot arm. The environment is a stuffed toy with a deformable surface. The parameters of the robot arm are given in Table II.

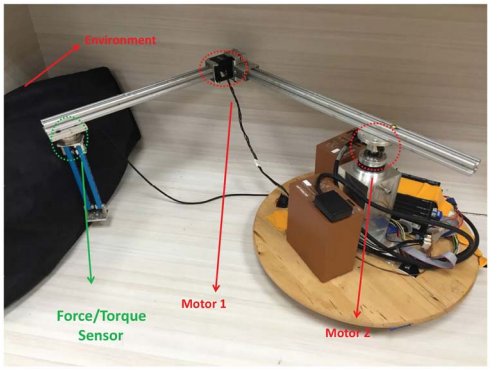


Fig. 9. Experiment setup. Two dc motors are controlled by an EPOS2 70/10 motor controller. An ATI mini-40 force/torque sensor is mounted at the end-effector of the robot arm. The environment is a stuffed toy with a deformable surface.

TABLE II
Parameters of the Robot Arm in Experiment

Parameter	Description	Value
m_1	Mass of link 1	0.32kg
m_2	Mass of link 2	0.44kg
l_1	Length of link 1	0.35m
l_2	Length of link 2	0.35m
I_1	Inertia moment of link 1	0.01kgm^2
I_2	Inertia moment of link 2	0.02kgm^2

The initial joint coordinates of the robot arm are $q_1 = 0.63$ rad and $q_2 = -1.26$ rad. The initial position in the Cartesian space is $[0.55\ 0]^T$m. The robot's task trajectory X_t is from $[0.55\ 0]^T$m to $[0.60\ 0]^T$m in uniform motion. The reference trajectory is parameterized as a minimal jerk trajectory with θ as the trajectory parameter and the movement duration of 50 s, that is

$$X_r(t) = X_t(t_0) + \big(X_t(t_f) - X_t(t_0)\big)p(t)\theta$$
$$p(t) = 10\left(\frac{t-t_0}{t_f-t_0}\right)^3 - 15\left(\frac{t-t_0}{t_f-t_0}\right)^4 + 6\left(\frac{t-t_0}{t_f-t_0}\right)^5. \quad (43)$$

The force exerted by the environment is only along the X-axis so the robot arm along the Y-axis is interaction-free. The weights in (11) are set as $Q = 2000$ and $R = 0.001$. The trajectory parameter θ is initially selected as $\theta = 1$. The impedance model is selected as $3000(X(t) - X_r(t)) = F(t)$.

The experimental results are shown in Figs. 10–12. Fig. 10 shows the iterative adaptation of the reference trajectory and the actual trajectory. In particular, the reference trajectory and the actual trajectory deviate from the task trajectory iteratively. As a result, the interaction force decreases iteratively, as shown in Fig. 11. The robot arm is initially not in contact with the environment. The interaction starts at around $t = 15$ s which can be seen from Fig. 11. After that, the interaction force gradually increases as the robot arm moves against the environment and the movement stops at $t = 50$ s. Both the trajectories and

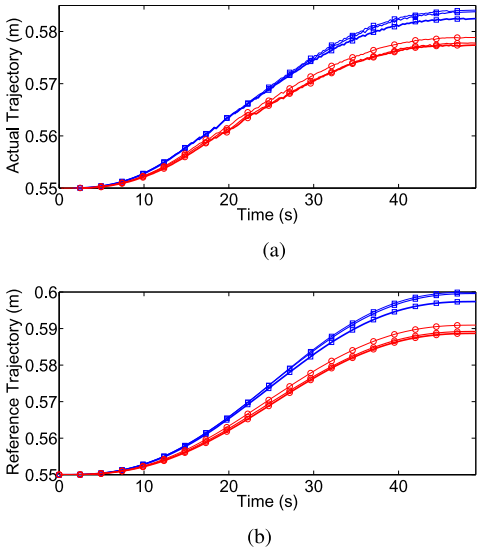


Fig. 10. Actual trajectory and reference trajectory of the first and last three iterations. The first three iterations are denoted using the blue lines with square markers (line width increases as iteration number increases from 1 to 3) and the last three iterations are denoted using the red lines with circle markers (line width increases as iteration number increases from 14 to 16). (a) Actual trajectory. (b) Reference trajectory.

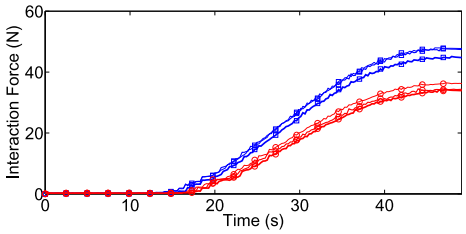


Fig. 11. Interaction force of the first and last three iterations. The first three iterations are denoted using the blue lines with square markers (line width increases as iteration number increases from 1 to 3) and the last three iterations are denoted using the red lines with circle markers (line width increases as iteration number increases from 14 to 16).

the interaction force converge after about 14 iterations. This is further confirmed by Fig. 12, where the cost and the trajectory parameter are illustrated. The above experimental results are similar to that in the simulation studies. The proposed method achieves the desired interaction performance by adapting the reference trajectory, without the requirement of the knowledge of the environment. Different desired interaction performances can be achieved by choosing different cost functions, as in the simulation studies.

B. Discussions

During the experiments, we note that the calculated gradient (25) or (26) may get near to singularities, since the measurement noise exists or the values of V at two adjacent iterations are too close. To address this issue, the adaptation rate σ^j can be reset as a constant when the difference of V at two adjacent iterations is smaller than a prescribed threshold.

Humans adapt both impedance and reference trajectory simultaneously during the interaction with environments. How to integrate the proposed reference adaptation with impedance

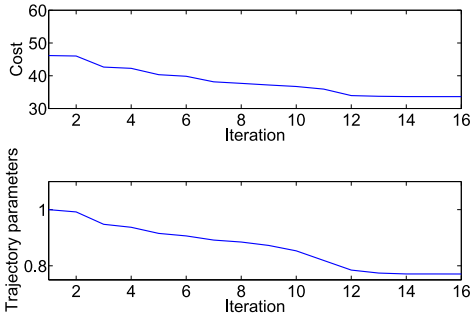


Fig. 12. Cost function and trajectory parameter.

learning/adaptation in a unified framework needs to be further investigated.

It is worth noting that in the proposed reference adaptation, the interaction performance cost is minimized using iterative learning. In this regard, the proposed method is inevitably subject to some drawbacks of iterative learning such as requirement of iterative searching and task repeatability. We will investigate how to address this issue in our future works.

Moreover, the interaction performance relies on the selection of the cost function, which has been shown to be nontrivial [18]. *A priori* partial knowledge of the environment can be used to cope with this problem in some cases, while how to address it in a general case is still an open problem.

VII. Conclusion

In this paper, reference adaptation has been developed to refine the reference trajectory of the robot arm, such that the desired interaction performance can be achieved subject to unknown environments. The desired interaction performance has been defined by minimizing a certain cost function which describes a tradeoff of trajectory tracking and force minimization. This cost function has been parameterized and the trajectory parameters have been updated to minimize it. The validity of the proposed method has been verified through simulation and experimental studies.

Appendix

Proof of Theorem 1

We consider the following Lyapunov function candidate:

$$W(t) = \frac{1}{2}Z^T(t)M_R(q(t))Z(t) + \frac{1}{2}\tilde{\Psi}^T\Gamma\tilde{\Psi} \tag{44}$$

where $\tilde{\Psi} = \hat{\Psi} - \Psi$. The derivative of $W(t)$ with respect to time is

$$\dot{W}(t) = \tilde{\Psi}^T\Gamma\dot{\tilde{\Psi}} + Z^T(t)M_R(q(t))\dot{Z}(t) + \frac{1}{2}Z^T(t)\dot{M}_R(q(t))Z(t). \tag{45}$$

According to (33), we obtain

$$\begin{aligned}\dot{W}(t) = \tilde{\Psi}^T\Gamma\dot{\tilde{\Psi}} + Z^T(t)\big[u + F(t) - M_R(q)\dot{X}_v(t) \\ - C_R(q(t),\dot{q}(t))X_v(t) - G_R(q(t))\big] \\ + Z^T(t)\left[-C_R(q(t),\dot{q}(t)) + \frac{1}{2}\dot{M}_R(q(t))\right]Z(t).\end{aligned} \tag{46}$$

Considering Property 2, we have

$$\dot{W}(t) = Z^T(t)\big[u + F(t) - M_R(q)\dot{X}_v(t) - C_R(q(t), \dot{q}(t))X_v(t) - G_R(q(t))\big] + \tilde{\Psi}^T\Gamma\dot{\tilde{\Psi}}. \tag{47}$$

According to Property 3, we have

$$M_R(q(t))\dot{X}_v(t) + C_R(q(t), \dot{q}(t))X_v(t) + G_R(q(t)) = Y(q(t), \dot{q}(t), \dot{X}_v(t), X_v(t))\Psi. \tag{48}$$

Thus, we obtain

$$\dot{W}(t) = Z^T(t)\big[u + F(t) - Y\big(q(t), \dot{q}(t), \dot{X}_v(t), X_v(t)\big)\Psi\big] + \tilde{\Psi}^T\Gamma\dot{\tilde{\Psi}}. \tag{49}$$

By substituting the control input (36) and the update law (37), we obtain

$$\begin{aligned}\dot{W}(t) &= Z^T(t)\Big[Y\big(q(t), \dot{q}(t), \dot{X}_v(t), X_v(t)\big)\tilde{\Psi} - KZ(t)\Big] \\ &\quad - \tilde{\Psi}^T Y^T(q(t), \dot{q}(t), \dot{X}_v(t), X_v(t))Z(t) \\ &= -Z^T(t)KZ(t) \le 0.\end{aligned} \tag{50}$$

By integrating $\dot{W}(t)$, we have

$$-\int_0^T Z^T(t)KZ(t)dt = W(t) - W(0). \tag{51}$$

Since K is positive definite, we have

$$\lambda_{\min}(K)\int_0^T Z^T(t)Z(t)dt \le W(0). \tag{52}$$

Then, we can obtain $Z(t) \in L_2^n$. Since $\dot{W}(t) \le 0$, we have $0 \le W(t) \le W(0)$, for $\forall t \ge 0$, leading to $W(t) \in L_\infty^n$. Suppose that $Z(t)$ is uniformly continuous. Then, we can conclude that $Z(t) \to 0$ as $t \to \infty$, which completes the proof.

REFERENCES

[1] G. B. Avanzini, N. M. Ceriani, A. M. Zanchettin, P. Rocco, and L. Bascetta, "Safety control of industrial robots based on a distributed distance sensor," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 6, pp. 2127–2140, Nov. 2014.

[2] J. J. Craig and M. H. Raibert, "A systematic method of hybrid position/force control of a manipulator," in *Proc. IEEE Comput. Soc. Comput. Softw. Appl. Conf.*, Chicago, IL, USA, 1979, pp. 446–451.

[3] N. Hogan, "Impedance control: An approach to manipulation—Part I: Theory; Part II: Implementation; Part III: Applications," *Trans. ASME J. Dyn. Syst. Meas. Control*, vol. 107, no. 1, pp. 1–24, Mar. 1985.

[4] J. E. Colgate and N. Hogan, "Robust control of dynamically interacting systems," *Int. J. Control*, vol. 48, no. 1, pp. 65–88, 1988.

[5] M. S. Erden and A. Billard, "Robotic assistance by impedance compensation for hand movements while manual welding," *IEEE Trans. Cybern.*, to be published.

[6] C.-C. Cheah and D. Wang, "Learning impedance control for robotic manipulators," *IEEE Trans. Robot. Autom.*, vol. 14, no. 3, pp. 452–465, Jun. 1998.

[7] S. Jung, T. C. Hsia, and R. G. Bonitz, "Force tracking impedance control of robot manipulators under unknown environment," *IEEE Trans. Control Syst. Technol.*, vol. 12, no. 3, pp. 474–483, May 2004.

[8] S. Hussain, S. Q. Xie, and P. K. Jamwal, "Adaptive impedance control of a robotic orthosis for gait rehabilitation," *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 1025–1034, Jun. 2013.

[9] S. P. Buerger and N. Hogan, "Complementary stability and loop shaping for improved human–robot interaction," *IEEE Trans. Robot.*, vol. 23, no. 2, pp. 232–244, Apr. 2007.

[10] S. S. Ge, Y. Li, and C. Wang, "Impedance adaptation for optimal robot–environment interaction," *Int. J. Control*, vol. 87, no. 2, pp. 249–263, 2014.

[11] M. S. De Queiroz, J. Hu, D. M. Dawson, T. Burg, and S. R. Donepudi, "Adaptive position/force control of robot manipulators without velocity measurements: Theory and experimentation," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 27, no. 5, pp. 796–809, Sep. 1997.

[12] 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.

[13] V. S. Chib, J. L. Patton, K. M. Lynch, and F. A. Mussa-Ivaldi, "Haptic identification of surfaces as fields of force," *J. Neurophysiol.*, vol. 95, no. 2, pp. 1068–1077, 2006.

[14] D. W. Franklin *et al.*, "Endpoint stiffness of the arm is directionally tuned to instability in the environment," *J. Neurosci.*, vol. 27, no. 29, pp. 7705–7716, 2007.

[15] B. Kim, J. Park, S. Park, and S. Kang, "Impedance learning for robotic contact tasks using natural actor-critic algorithm," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 2, pp. 433–443, Apr. 2010.

[16] J. Buchli, F. Stulp, E. Theodorou, and S. Schaal, "Learning variable impedance control," *Int. J. Robot. Res.*, vol. 30, no. 7, pp. 820–833, 2011.

[17] C. Yang *et al.*, "Human-like adaptation of force and impedance in stable and unstable interactions," *IEEE Trans. Robot.*, vol. 27, no. 5, pp. 918–930, Oct. 2011.

[18] Y. Li and S. S. Ge, "Impedance learning for robots interacting with unknown environments," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 4, pp. 1422–1432, Jul. 2014.

[19] P. Dizio and J. R. Lackner, "Motor adaptation to coriolis force perturbations of reaching movements: Endpoint but not trajectory adaptation transfers to the nonexposed arm," *J. Neurophysiol.*, vol. 74, no. 4, pp. 1787–1792, 1995.

[20] J. Barraquand and J.-C. Latombe, "Robot motion planning: A distributed representation approach," *Int. J. Robot. Res.*, vol. 10, no. 6, pp. 628–649, 1991.

[21] S. M. LaValle, *Planning Algorithms*. Cambridge, NY, USA: Cambridge Univ. Press, 2006.

[22] S. S. Ge and F. L. Lewis, *Autonomous Mobile Robots: Sensing, Control, Decision-Making, and Applications*. Boca Raton, FL, USA: CRC Press, 2006.

[23] A. Saccon, J. Hauser, and A. Beghi, "Trajectory exploration of a rigid motorcycle model," *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 2, pp. 424–437, Mar. 2012.

[24] S. Ito, M. Darainy, M. Sasaki, and D. J. Ostry, "Computational model of motor learning and perceptual change," *Biol. Cybern.*, vol. 107, no. 6, pp. 653–667, 2013.

[25] B. Corteville, E. Aertbelien, H. Bruyninckx, J. De Schutter, and H. V. Brussel, "Human-inspired robot assistant for fast point-to-point movements," in *Proc. IEEE Int. Conf. Robot. Autom.*, Rome, Italy, 2007, pp. 3639–3644.

[26] M. S. Erden and T. Tomiyama, "Human-intent detection and physically interactive control of a robot without force sensors," *IEEE Trans. Robot.*, vol. 26, no. 2, pp. 370–382, Apr. 2010.

[27] Z. Wang, A. Peer, and M. Buss, "An HMM approach to realistic haptic human–robot interaction," in *Proc. 3rd Joint Eurohaptics Conf. Symp. Haptic Interf. Virtual Environ. Teleoper. Syst.*, Salt Lake City, UT, USA, 2009, pp. 374–379.

[28] Z. Wang, E. Giannopoulos, M. Slater, A. Peer, and M. Buss, "Handshake: Realistic human-robot interaction in haptic enhanced virtual reality," *Presence Teleoper. Virtual Environ.*, vol. 20, no. 4, pp. 371–392, 2011.

[29] 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 Trans. Mechatronics*, vol. 18, no. 1, pp. 285–296, Feb. 2013.

[30] Y. Li and S. S. Ge, "Human–robot collaboration based on motion intention estimation," *IEEE/ASME Trans. Mechatronics*, vol. 19, no. 3, pp. 1007–1014, Jun. 2014.

[31] Y. Li and S. S. Ge, "Force tracking control for motion synchronization in human–robot collaboration," *Robotica*, vol. 34, no. 6, pp. 1260–1281, 2016.

[32] K. P. Tee, R. Yan, and H. Li, "Adaptive admittance control of a robot manipulator under task space constraint," in *Proc. IEEE Int. Conf. Robot. Autom.*, Anchorage, AK, USA, 2010, pp. 5181–5186.

[33] C. Yang and E. Burdet, "A model of reference trajectory adaptation for interaction with objects of arbitrary shape and impedance," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, San Francisco, CA, USA, 2011, pp. 4121–4126.
[34] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of robots by learning," *J. Robot. Syst.*, vol. 1, no. 2, pp. 123–140, 1984.
[35] S. S. Ge, T. H. Lee, and C. J. Harris, *Adaptive Neural Network Control of Robotic Manipulators*. River Edge, NJ, USA: World Sci., 1998.
[36] F. L. Lewis, S. Jagannathan, and A. Yesildirak, *Neural Network Control of Robot Manipulators and Nonlinear Systems*. London, U.K.: CRC Press, 1998.
[37] J.-J. E. Slotine and W. Li, "On the adaptive control of robot manipulators," *Int. J. Robot. Res.*, vol. 6, no. 3, pp. 49–59, 1987.
[38] W. He and S. S. Ge, "Vibration control of a flexible beam with output constraint," *IEEE Trans. Ind. Electron.*, vol. 62, no. 8, pp. 5023–5030, Aug. 2015.
[39] W. He, Y. Chen, and Z. Yin, "Adaptive neural network control of an uncertain robot with full-state constraints," *IEEE Trans. Cybern.*, vol. 46, no. 3, pp. 620–629, Mar. 2016.
[40] W. He, S. Zhang, and S. S. Ge, "Adaptive control of a flexible crane system with the boundary output constraint," *IEEE Trans. Ind. Electron.*, vol. 61, no. 8, pp. 4126–4133, Aug. 2014.
[41] R. Z. Stanisic and A. V. Fernandez, "Adjusting the parameters of the mechanical impedance for velocity, impact and force control," *Robotica*, vol. 30, no. 4, pp. 583–597, 2012.
[42] 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," *Robot. Auton. Syst.*, vol. 57, no. 1, pp. 23–33, 2009.
[43] S. Liu and D. Sun, "Minimizing energy consumption of wheeled mobile robots via optimal motion planning," *IEEE/ASME Trans. Mechatronics*, vol. 19, no. 2, pp. 401–411, Apr. 2014.
[44] Z. Xu, S. Wei, N. Wang, and X. Zhang, "Trajectory planning with Bezier curve in Cartesian space for industrial gluing robot," in *Intelligent Robotics and Applications*. Cham, Switzerland: Springer, 2014, pp. 146–154.
[45] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of robots by learning," *J. Robot. Syst.*, vol. 1, no. 2, pp. 123–140, 1984.
[46] J.-X. Xu, D. Huang, and S. Pindi, "Optimal tuning of PID parameters using iterative learning approach," *SICE J. Control Meas. Syst. Integr.*, vol. 1, no. 2, pp. 143–154, 2011.
[47] Y. Li, S. S. Ge, and C. Yang, "Learning impedance control for physical robot–environment interaction," *Int. J. Control*, vol. 85, no. 2, pp. 182–193, 2012.
[48] P. I. Corke, "A robotics toolbox for MATLAB," *IEEE Robot. Autom. Mag.*, vol. 3, no. 1, pp. 24–32, Mar. 1996.

Chen Wang received the B.Eng. degree in control science engineering from Zhejiang University, Hangzhou, China, in 2011. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore.

His current research interests include artificial intelligence and learning in human–robot interaction and human–robot collaboration.

Yanan Li (M'14) received the B.Eng. degree in control science and engineering and the M.Eng. degree in control and mechatronics engineering from the Harbin Institute of Technology, Harbin, China, in 2006 and 2009, respectively, and the Ph.D. degree from the NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore, in 2013.

He has been a Research Scientist with the Institute for Infocomm Research, Agency for Science, Technology, and Research, Singapore, from 2013 to 2015. He is currently a Research Associate with the Department of Bioengineering, Imperial College London, London, U.K. His current research interests include physical human–robot interaction and human–robot collaboration.

Shuzhi Sam Ge (S'90–M'92–SM'99–F'06) received the B.Sc. degree from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1986, and the Ph.D. degree from the Imperial College of Science, Technology, and Medicine, University of London, London, U.K., in 1993.

He is the Founding Director of the Social Robotics Laboratory, Interactive Digital Media Institute, National University of Singapore, Singapore, and on leave with the University of Electronic Science and Technology of China, Chengdu, China. He has authored or co-authored six books and over 400 international journal and conference papers. His current research interests include social robotics, multimedia fusion, medical robots, and intelligent systems.

Dr. Ge is an Editor-in-Chief of the *International Journal of Social Robotics*. He has served/been serving as an Associate Editor for a number of flagship journals. He also serves as an Editor of the Taylor & Francis Automation and Control Engineering Series. He also served as the Vice President of Technical Activities from 2009 to 2010, and the Vice President for Membership Activities from 2011 to 2012 and the IEEE Control Systems Society.

Tong Heng Lee (M'90) received the B.A. degree (First Class Hons.) in engineering tripos from Cambridge University, Cambridge, U.K., in 1980, and the Ph.D. degree in electrical engineering from Yale University, New Haven, CT, USA, in 1987.

He is a Professor with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, where he is currently the Head of the Drives, Power, and Control Systems Group. He has also co-authored three research monographs and holds four patents (two of which are in the technology area of adaptive systems, and the other two are in the area of intelligent mechatronics). His current research interests include adaptive systems, knowledge-based control, intelligent mechatronics, and computational intelligence.

Dr. Lee was a recipient of the Cambridge University Charles Baker Prize in Engineering. He is the Deputy Editor-in-Chief of the *International Federation of Automatic Control* and *Mechatronics International Journal* and serves as an Associate Editor of many other flagship journals.