

<http://bhxb.buaa.edu.cn> jbuaa@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2023.0314

基于迭代最优阻抗的人机物理交互控制方法

刘微容*, 魏子丰, 晋振兵, 孟家豪, 王星琨, 张浩琛

(兰州理工大学 电气工程与信息工程学院, 兰州 730050)

摘 要: 为提高人机物理交互的准确性和柔顺性, 实现最优交互性能, 针对基于迭代学习的阻抗控制方法需要多次重复同一任务的问题, 借鉴迭代最优控制无需系统矩阵信息即可优化代价函数确定系统最优控制输入的机制, 提出了基于迭代最优阻抗的人机物理交互控制方法。方法采用双环控制结构。面向任务的外环设计了迭代最优阻抗控制器 (IOIC), 将求取最优阻抗参数的问题描述成线性二次型调节器问题, 利用迭代最优控制, 求取最优反馈增益, 使包括轨迹跟踪误差和交互力在内的代价函数最小化; 同时引入软辅助函数, 避免参数突变可能带来的机器人抖动问题。面向机器人的内环设计了非奇异终端滑模轨迹跟踪控制器 (NTSMTC), 使机器人实际轨迹跟踪外环输出的阻抗轨迹, 通过饱和函数消减控制律的抖振。仿真结果证明: 所提方法在人机协作任务中, 仅利用一次任务初始阶段的交互信息即可求得最优阻抗参数, 使任务过程中的轨迹跟踪误差和交互人所消耗的力最小化。

关键词: 人机物理交互; 阻抗控制; 轨迹跟踪控制; 最优控制; 滑模控制

中图分类号: TP242

文献标志码: A

文章编号: 1001-5965(2025)06-1843-09

近年来, 工业机器人已经成功应用于执行各类高精度重复性任务^[1]。然而, 仍有一些任务由于工序复杂、交互环境结构性差, 导致机器人无法单独完成, 需要人对机器人运动进行指导和调整^[2]。在这种情况下, 机器人和人不可避免地相互接触, 形成人机耦合的动力学系统来协同完成任务^[3]。因此, 人机物理交互系统受到越来越多的研究人员关注。

为保证交互过程的安全、高效, 机器人与人接触时, 不仅需要控制交互位置, 还需要控制交互力。阻抗控制因其能够调节交互位置和力之间的关系^[4], 在人机物理交互系统中受到了广泛的研究与应用。但是, 在执行任务的过程中, 由于交互人的随机性, 难以用固定的阻抗参数满足所有的交互需求, 特别是面对具有变化和不确定性的交互对象^[5]。

通过模拟人类特有的、可根据交互对象不同而主动调节肢体阻抗的属性, 研究人员提出了多种变阻抗控制方法^[6-9], 通过实时改变虚拟阻抗参数, 提高人机交互系统的性能和安全性。Liu 等^[10]提出了一种基于速度的无传感器变阻抗控制方法, 利用末端执行器速度和交互力在线调整阻抗参数, 使末端执行器速度更符合人的行为, 以提高人机物理交互过程中的顺应性。Grafakos 等^[11]提出了一种变导纳控制器, 通过测量人体肌电信号, 实时调节阻抗控制器的阻尼参数, 以提高交互性能。然而, 上述方法并不能在阻抗参数的变化过程中保证交互系统的稳定性, 因此, 基于稳定性的变阻抗方法受到广泛关注。Dong 等^[12]针对模型不确定的机器人提出了基于不确定性和扰动估计器 (uncertainty and disturbance estimator, UDE) 的变阻抗方法, 利用

收稿日期: 2023-06-05; 录用日期: 2023-10-16; 网络出版时间: 2023-11-24 14:33

网络出版地址: link.cnki.net/urlid/11.2625.V.20231123.1034.002

基金项目: 国家自然科学基金 (62261032); 中央引导地方科技发展资金项目 (25ZYJA026); 甘肃省重点人才项目; 甘肃省高等学校创新基金 (2022A-021)

* 通信作者. E-mail: liuwr@lut.edu.cn

引用格式: 刘微容, 魏子丰, 晋振兵, 等. 基于迭代最优阻抗的人机物理交互控制方法 [J]. 北京航空航天大学学报, 2025, 51(6): 1843-1851. LIU W R, WEI Z F, JIN Z B, et al. Human-robot physical interaction control method based on iterative optimal impedance [J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(6): 1843-1851 (in Chinese).

UDE逼近不确定的被控模型,同时为阻抗参数的变化提供了稳定性条件,且该稳定性条件可在执行任务前离线验证。虽然基于稳定性的变阻抗控制方法可以保证交互过程的稳定性,但基于此设计的变阻抗策略往往比较保守,柔顺交互性能受到了较大的限制。总体而言,各类变阻抗控制方法虽然能够通过实时改变阻抗参数实现机器人与人的高效交互,提高了交互性能,但不能保证实现最优的交互性能^[13]。

受到人类自主学习时通过积累先前失败经验,反复迭代练习从而提高对此技能掌控能力的启发,基于迭代学习的阻抗控制方法被提出。在每次执行任务过程中,利用先验信息,迭代更新阻抗参数,直至最优。Yang等^[14]受到点对点运动的人机交互实验启发,基于迭代学习方法,设计包括前馈控制律和阻抗增益在内的阻抗参数迭代更新规则。Yang等^[15]针对人体动力学时变难以建模的问题,利用期望轨迹和人体阻抗的空间周期性特征,提出一种空间重复阻抗学习控制策略,用于提高人机交互性能。然而,基于迭代学习的阻抗控制方法需要在相同的持续时间和初始条件下多次重复执行同一任务,在实际应用中,难以满足此严格条件^[16]。

为实现最优交互性能,基于最优的阻抗控制同样受到研究人员的广泛关注。因为阻抗控制既需要控制力,又需要控制位置,所以,最优阻抗的思想是:权衡力和位置跟踪二者之间的关系^[17],定义包括力和位置跟踪误差在内的代价函数或奖励函数描述实际交互性能,通过最小化代价函数或最大化奖励函数求取最优的阻抗参数。Johansson等^[18]针对环境模型已知情况,利用线性二次型调节器(linear quadratic regulator, LQR),通过求解代数Riccati方程,确定最优阻抗参数。在此基础上,Matinfar等^[19]考虑机器人与环境系统的动态特性,提出一类新的代价函数,不同于上述方法,采用在线求解LQR的方法求取最优阻抗参数。但是,上述方法均假设环境模型已知,因为LQR方法无法对参数未知的线性模型进行求取。而自适应动态规划(adaptive dynamic programming, ADP)和强化学习(reinforcement learning, RL)方法的提出,为解决上述问题提供了思路。Huang等^[20]将机器人和环境的交互系统视为线性未知模型,利用策略迭代的方法,在线求取最优阻抗参数,实现了机器人与人的最优交互,使轨迹跟踪误差和交互力最小化。Guo等^[21]利用相同的思路,设计了一种基于Q-学习算法的自适应阻抗控制器,实现阻抗自适应,保证在人肢体模型动态完全未知的情况下,阻抗参数收

敛到最优值,但在每次阻抗参数更新后,仍需重新收集数据进行计算,并未从根本上解决需要重复执行任务的问题。

基于上述讨论,本文提出了基于迭代最优阻抗的人机物理交互控制方法,仅利用执行1次任务时初始阶段的部分信息,通过迭代求取最优反馈增益,获得最优阻抗模型,明显减少了数据收集次数,节省了数据收集与计算的时间;在实际应用中,提高了系统运行效率和响应速度。同时,引入软辅助函数消减参数突变问题。利用非奇异终端滑模控制使机器人实际轨迹跟踪阻抗轨迹,并通过饱和函数消减控制力矩的抖振问题,实现了确保轨迹跟踪精度的同时最小化交互力。

1 控制系统方案

基于迭代最优阻抗的人机物理交互控制方法的结构如图1所示。该方法采用双环控制结构,由阻抗控制外环和轨迹跟踪内环构成。

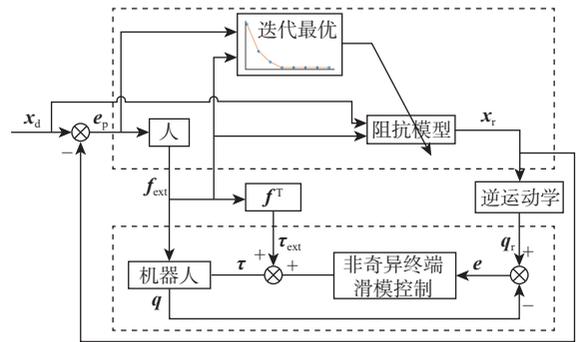


图1 基于迭代最优阻抗的人机物理交互控制方法结构
Fig. 1 Structure of human-robot physical interaction control method based on iterative optimal impedance

外环面向任务,利用期望轨迹跟踪误差 e_p 、人与机器人的交互力 f_{ext} 作为迭代最优阻抗控制器(iterative optimal impedance controller, IOIC)的输入,求取最优阻抗参数。根据期望轨迹 x_d 和交互力 f_{ext} 得到阻抗轨迹 x_r ,使得机器人的运动可以确保人以最小的工作量协助机器人跟踪期望轨迹,从而实现最优交互性能。

内环面向机器人,利用非奇异终端滑模轨迹跟踪控制器(nonsingular terminal sliding mode trajectory tracking controller, NTSMTC),使机器人的运动响应与所设计阻抗模型一致。

2 外环: 迭代最优阻抗控制器

机器人与人交互和与环境交互不同。在交互力方面,环境只能根据机器人的运动被动改变交互力,而人能根据经验和感官对机器人当前所执行的

任务进行指导, 主动改变交互力的大小。因此, 与人交互的机器人需要调节自身阻抗参数, 以适应不同人员的技能水平差异, 与人共同完成指定任务。本节设计了 IOIC 以保证期望轨迹跟踪性能的同时减少人在任务过程中消耗的力。

2.1 人机物理交互系统

为构建轨迹跟踪误差和交互力之间的关系, 建立如下机器人阻抗模型:

$$\mathbf{M}_d \ddot{\mathbf{e}}_p + \mathbf{B}_d \dot{\mathbf{e}}_p + \mathbf{K}_d \mathbf{e}_p = -\mathbf{f}_{\text{ext}} \quad (1)$$

式中: \mathbf{M}_d 、 \mathbf{B}_d 和 \mathbf{K}_d 分别为期望质量、阻尼和刚度参数; \mathbf{f}_{ext} 为交互力; \mathbf{e}_p 为期望轨迹跟踪误差。

人机协作型任务中, 人的动力学模型可建立为^[22]

$$\mathbf{K}_r \dot{\mathbf{f}}_{\text{ext}} + \mathbf{K}_f \mathbf{f}_{\text{ext}} = \mathbf{K}_c \mathbf{e}_p \quad (2)$$

式中: \mathbf{K}_r 、 \mathbf{K}_f 和 \mathbf{K}_c 为未知增益, 表示人对所执行任务的熟练度。

从式(2)可以看出, 人对机器人施加的力是基于对任务完成度的判断。

为构建人机物理交互系统的状态空间方程, 基于上述模型和控制目标, 选取系统的状态 \mathbf{X} 为

$$\mathbf{X} = [\mathbf{e}_p \quad \dot{\mathbf{e}}_p \quad \mathbf{f}_{\text{ext}}]^T \quad (3)$$

则人机物理交互系统的状态方程可写为

$$\dot{\mathbf{X}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{K}_r^{-1} \mathbf{K}_c & \mathbf{0} & -\mathbf{K}_r^{-1} \mathbf{K}_f \end{bmatrix} \mathbf{X} + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \\ \mathbf{0} \end{bmatrix} \mathbf{u}_c = \mathbf{A} \mathbf{X} + \mathbf{B} \mathbf{u}_c \quad (4)$$

式中: \mathbf{I} 为单位矩阵; $\mathbf{A} \in \mathbf{R}^{p \times p}$ 为状态矩阵; $\mathbf{B} \in \mathbf{R}^{p \times r}$ 为输入矩阵; $\mathbf{u}_c \in \mathbf{R}^r$ 为交互系统的控制输入, 表示为

$$\mathbf{u}_c = \ddot{\mathbf{e}}_p = -\mathbf{M}_d^{-1} \mathbf{B}_d \dot{\mathbf{e}}_p - \mathbf{M}_d^{-1} \mathbf{K}_d \mathbf{e}_p - \mathbf{M}_d^{-1} \mathbf{f}_{\text{ext}} = -\mathbf{K} \mathbf{X} \quad (5)$$

其中: $\mathbf{K} = [\mathbf{M}_d^{-1} \mathbf{K}_d \quad \mathbf{M}_d^{-1} \mathbf{B} \quad \mathbf{M}_d^{-1}]$ 为反馈增益。

因此, 对于式(5)描述的人机物理交互系统, 求取最优阻抗参数 \mathbf{M}_d^* 、 \mathbf{B}_d^* 和 \mathbf{K}_d^* , 即求取最优反馈增益 \mathbf{K}^* 。

该问题可看作是一个 LQR 问题, 但系统中的状态矩阵 \mathbf{A} 包含了人体模型中的未知参数 \mathbf{K}_r 、 \mathbf{K}_f 和 \mathbf{K}_c , 考虑到 LQR 方法无法求解含有未知参数的线性模型, 借鉴 Jiang 等^[23] 所提迭代最优控制方法, 可以无需系统模型, 仅利用部分状态信息和输入信息, 即可优化代价函数, 确定系统最优控制输入的机制, 设计 IOIC。

2.2 迭代最优阻抗控制器设计

为实现最小化期望轨迹跟踪误差和交互力的控制目标, 构建代价函数 L 为

$$L = \int_{t_0}^{\infty} \left([\mathbf{e}_p \quad \dot{\mathbf{e}}_p \quad \mathbf{f}_{\text{ext}}] \mathbf{Q} \begin{bmatrix} \mathbf{e}_p \\ \dot{\mathbf{e}}_p \\ \mathbf{f}_{\text{ext}} \end{bmatrix} + \mathbf{u}_c^T \mathbf{R} \mathbf{u}_c \right) dt = \int_{t_0}^{\infty} (\mathbf{X}^T \mathbf{Q} \mathbf{X} + \mathbf{u}_c^T \mathbf{R} \mathbf{u}_c) dt \quad (6)$$

式中: $\mathbf{Q} = \mathbf{Q}^T > \mathbf{0}$ 为期望轨迹跟踪误差和交互力的权重; $\mathbf{R} = \mathbf{R}^T > \mathbf{0}$ 为与期望轨迹跟踪误差和交互力呈线性依赖关系的控制输入的权重。

通过设计最优的控制输入 \mathbf{u}_c^* , 使代价函数 L 最小, 即可同时使轨迹跟踪误差和交互力最小, 实现最优交互性能。

Jiang 等^[23] 所提迭代最优控制方法通过收集一段时间内的状态和输入信息, 并经过采样, 构建足够多的数据矩阵, 实现仅利用在该时间段内收集的数据, 迭代计算得出近似最优反馈增益矩阵。

因此, 定义如式(7)所示的 3 个算子, 用于在人机物理交互初始阶段 $[0, t]$ 时间内, 收集状态 \mathbf{X} 和控制输入 \mathbf{u}_c 的信息。

$$\begin{cases} \kappa_{\mathbf{X}\mathbf{X}} = [\mathbf{X}^T \otimes \mathbf{X}^T|_0^1, \mathbf{X}^T \otimes \mathbf{X}^T|_1^2, \dots, \mathbf{X}^T \otimes \mathbf{X}^T|_{t_{h-1}}^t] \\ \delta_{\mathbf{X}\mathbf{X}} = [\int_0^{t_1} \mathbf{X}^T \otimes \mathbf{X}^T dt, \int_{t_1}^{t_2} \mathbf{X}^T \otimes \mathbf{X}^T dt, \dots, \int_{t_{h-1}}^t \mathbf{X}^T \otimes \mathbf{X}^T dt] \\ \delta_{\mathbf{X}\mathbf{u}} = [\int_0^{t_1} \mathbf{X}^T \otimes \mathbf{u}_c^T dt, \int_{t_1}^{t_2} \mathbf{X}^T \otimes \mathbf{u}_c^T dt, \dots, \int_{t_{h-1}}^t \mathbf{X}^T \otimes \mathbf{u}_c^T dt] \end{cases} \quad (7)$$

式中: $\kappa_{\mathbf{X}\mathbf{X}}$ 、 $\delta_{\mathbf{X}\mathbf{X}}$ 、 $\delta_{\mathbf{X}\mathbf{u}}$ 为数据储存矩阵; “ \otimes ” 表示克罗内克乘积; h 为 $[0, t_0]$ 时间段的采样个数。

在此数据收集阶段, 为保证收集到的数据丰富且非奇异, 设定输入为

$$\mathbf{u}_c = -\mathbf{K}_0 \mathbf{X} + \mathbf{v}_n \quad (8)$$

式中: \mathbf{K}_0 为使系统(4)稳定的初始反馈增益; \mathbf{v}_n 为满足持续激励的噪声函数。

计算 $\kappa_{\mathbf{X}\mathbf{X}}$ 、 $\delta_{\mathbf{X}\mathbf{X}}$ 和 $\delta_{\mathbf{X}\mathbf{u}}$, 直到满足 $\text{rank}([\delta_{\mathbf{X}\mathbf{X}}, \delta_{\mathbf{X}\mathbf{u}}]) = p(p+1)/2 + pr$ 。

求解式(9)得到 \mathbf{P}_k 和 \mathbf{K}_{k+1} :

$$\begin{bmatrix} \text{vec}(\mathbf{P}_k) \\ \text{vec}(\mathbf{K}_{k+1}) \end{bmatrix} = (\mathbf{\Gamma}_k^T \mathbf{\Gamma}_k)^{-1} \mathbf{\Gamma}_k^T \boldsymbol{\theta}_k \quad (9)$$

式中: $\text{vec}(\mathbf{P}_k)$ 和 $\text{vec}(\mathbf{K}_{k+1})$ 表示待求解项, $\text{vec}(\cdot)$ 表示矩阵向量化, \mathbf{P}_k 为李雅普诺夫方程的对称正定解; $\mathbf{\Gamma}_k = [\kappa_{\mathbf{X}\mathbf{X}} \quad -2\delta_{\mathbf{X}\mathbf{X}}(\mathbf{I} \otimes \mathbf{K}_k^T \mathbf{R}) - 2\delta_{\mathbf{X}\mathbf{u}}(\mathbf{I} \otimes \mathbf{R})]$ 为可在线测量的状态信息; $\boldsymbol{\theta}_k = \delta_{\mathbf{X}\mathbf{X}} \text{vec}(\mathbf{Q}_k)$ 为输入信息, $\mathbf{Q}_k = \mathbf{Q} + \mathbf{K}_k^T \mathbf{R} \mathbf{K}_k$ 。

通过不断重复迭代计算式(9), 直到 $\|\mathbf{P}_k - \mathbf{P}_{k-1}\| \leq \varepsilon$, ε 为很小的正实数, 即可得到最优反馈增益的近似值 \mathbf{K}_{k+1} , 从而计算得到最优阻抗参数。

特别注意的是, 当迭代结束得到最优反馈增益 \mathbf{K}_{k+1} 后, 系统的反馈增益会从初始的 \mathbf{K}_0 突变至

K_{k+1} 。这种突变会对系统控制输入 u_e 产生影响,使得加速度发生突变,可能导致机器人在运动过程中发生抖动,这种抖动在机器人的控制过程中是不允许存在的。

因此,为解决由于参数突变带来的问题,设计如下软辅助函数:

$$K' = \frac{K_{k+1} + K_0}{2} - \frac{K_{k+1} - K_0}{2} \cos\left(\frac{\pi t}{T_s}\right) \quad (10)$$

式中: $T_s = T_1 - T_2$ 为软辅助函数的周期。

软辅助函数的变化曲线如图2所示。可以看到,为避免参数突变,设计的软辅助函数利用三角函数平滑连续的曲线特性,使初始反馈增益在时间 $[T_1, T_2]$ 间隔内平滑、单调地变化到近似最优反馈增益 K_{k+1} 。

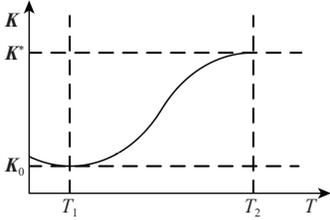


图2 软辅助函数曲线

Fig. 2 Soft auxiliary function curves

进一步,根据上述设计,本文所提 IOIC 的控制流程如下。

输入: 初始反馈增益 K_0 , 权重 Q 和 R , 阈值 ε , 探测噪声 v_n 。

1. 根据选定的初始反馈增益 K_0 和探测噪声 v_n , 设置机器人的初始阻抗增益, 同时设置初始的系统输入 u_e 。当人与机器人开始交互后的 $[0, t]$ 时间内, 收集系统状态 e_p 、 \dot{e}_p 和 f_{ext} 信息, 以及系统输入 u_e 信息。利用收集到的信息建立矩阵 κ_{XX} 、 δ_{XX} 和 δ_{Xu} , 直到 Γ_k 列满秩。

2. 利用式(9), 求解 P_k 和 K_{k+1} 。

3. 令 $k \rightarrow k+1$, 重复步骤 2, 直到 $\|P_k - P_{k-1}\| \leq \varepsilon$ 为止。

4. 取 K_{k+1} 为近似最优控制策略。

5. 在时间 $[t, t+T_s]$ 内, 利用式(10), 求取由软辅助函数过渡的反馈增益 K' , 使得反馈增益平滑地变化至最优。

输出: 近似最优反馈增益 K_{k+1} , 求取最优的阻抗增益 M_d^* 、 B_d^* 和 K_d^* 。

3 内环: 非奇异终端滑模轨迹跟踪控制器

在第2节已完成外环交互控制设计的基础上, 还需要进一步完成内环机器人的轨迹跟踪控制设

计。因此, 本节设计 NTSMTC 以保证机器人能跟踪外环输出的阻抗轨迹, 从而使机器人的实际运动表现出外环阻抗控制器预设的阻抗行为。

3.1 非奇异终端滑模轨迹跟踪控制器设计

机器人和人的交互任务发生在笛卡儿空间, 而机器人的控制发生在关节空间中, 因此, 机器人在关节空间的动力学方程可写为

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \tau + \tau_{ext} + \tau_d \quad (11)$$

式中: q 、 \dot{q} 、 $\ddot{q} \in \mathbf{R}^n$ 分别为机器人的实际关节角、关节角速度和关节角加速度; $M(q) \in \mathbf{R}^{n \times n}$ 为正定对称的惯性矩阵; $C(q, \dot{q}) \in \mathbf{R}^{n \times n}$ 为哥氏离心力矩阵; $G(q) \in \mathbf{R}^n$ 为重力向量; $\tau \in \mathbf{R}^n$ 为控制输入力矩; $\tau_{ext} = J^T f_{ext} \in \mathbf{R}^n$ 为操作人员与机器人在关节空间中的交互力; $\tau_d \in \mathbf{R}^n$ 为外部干扰力矩向量, 且假设 $\|\tau_d\| \leq \eta$ 。

通过逆运动学将外环的阻抗轨迹 x_r 转换为关节空间中的阻抗轨迹 q_r , 设计滑模控制律使得机器人的关节位置 q 跟踪 q_r , 再通过正运动学实现机器人的实际位置 x 跟踪 x_r 。

为实现轨迹跟踪控制, 定义阻抗轨迹位置跟踪误差为 $e = q_r - q$, 速度跟踪误差可定义为 $\dot{e} = \dot{q}_r - \dot{q}$ 。根据机器人的动力学方程(11), 可以得到

$$\ddot{e} = M_0^{-1}(q)(\tau - C_0(q, \dot{q})\dot{q} - G_0(q) + \tau_{ext} + \tau_d) - \ddot{q}_r \quad (12)$$

根据所定义的位置跟踪误差及速度跟踪误差, 可设计如下非奇异终端滑模面:

$$s = \dot{e} + \lambda_1 e + \lambda_2 \phi(e) \quad (13)$$

式中: $\lambda_j = \text{diag}(\lambda_{j1}, \dots, \lambda_{jn})$, $j = 1, 2$, 为正定的对角矩阵; $\phi(e) = [\phi(e(1)), \dots, \phi(e(n))]^T$ 为关于跟踪误差的函数, 具体可以表示为

$$\phi(e(i)) = \begin{cases} e(i)^{b/a} & R_i = 0 \text{ 或 } R_i \neq 0, |e(i)| \geq \mu \\ \sigma_1 e(i) + \sigma_2 e(i)^2 \text{sign}(e(i)) & R_i \neq 0, |e(i)| < \mu \end{cases} \quad (14)$$

式中: $R_i = \dot{e}(i) + \lambda_{1i} e(i) + \lambda_{2i} e(i)^{b/a}$ ($i = 1, \dots, n$); μ 为足够小的正常数; a 和 b 为正奇数, 且 $1/2 < b/a < 1$; $\sigma_1 = (2 - b/a)\mu^{b/a-1}$, $\sigma_2 = (b/a - 1)\mu^{b/a-2}$; $\text{sign}(\cdot)$ 表示符号函数。

为设计控制力矩, 对滑模面(13)进行求导, 可得

$$\dot{s} = \ddot{e} + \lambda_1 \dot{e} + \lambda_2 \dot{\phi}(e, \dot{e}) \quad (15)$$

将式(12)代入到式(15)得

$$\dot{s} = M_0^{-1}(q)(\tau - C_0(q, \dot{q})\dot{q} - G_0(q) + \tau_{ext} + \tau_d) - \ddot{q}_r + \lambda_1 \dot{e} + \lambda_2 \dot{\phi}(e, \dot{e}) \quad (16)$$

因此, 设计控制力矩为

$$\tau = \tau_{eq} + \tau_{sw} = \mathbf{M}(\ddot{\mathbf{q}}_r - \lambda_1 \dot{\mathbf{e}} - \lambda_2 \phi(\mathbf{e}, \dot{\mathbf{e}})) + \mathbf{C}_0(\mathbf{q}, \dot{\mathbf{q}})\mathbf{q} + \mathbf{G}_0(\mathbf{q}) - \tau_{ext} - \mathbf{M}(\eta \text{sign}(\mathbf{s}) + \beta \mathbf{s}) \quad (17)$$

式中: τ_{eq} 为等效控制力矩; τ_{sw} 为切换控制力矩; β 为控制增益。

3.2 稳定性证明

定理 1 考虑非线性系统 (11), 在滑模面 (13) 和控制律 (17) 的作用下, 机器人系统阻抗轨迹跟踪误差能在有限时间内收敛至零。

证明 定义如下李雅普诺夫函数:

$$V = \frac{1}{2} \mathbf{s}^T \mathbf{s} \quad (18)$$

对式 (18) 求导得

$$\dot{V}(\mathbf{s}) = \mathbf{s}[\ddot{\mathbf{e}} + \lambda_1 \dot{\mathbf{e}} + \lambda_2 \phi(\mathbf{e}, \dot{\mathbf{e}})] = \mathbf{s}[\mathbf{M}_0^{-1}(\mathbf{q})(\tau - \mathbf{C}_0(\mathbf{q}, \dot{\mathbf{q}})\mathbf{q} - \mathbf{G}_0(\mathbf{q}) + \tau_{ext} + \tau_d) - \ddot{\mathbf{q}}_r + \lambda_1 \dot{\mathbf{e}} + \lambda_2 \phi(\mathbf{e}, \dot{\mathbf{e}})] \quad (19)$$

将控制律 (17) 代入式 (19) 得

$$\dot{V} = \mathbf{s}^T(-\eta \text{sign}(\mathbf{s}) - \beta \mathbf{s} + \tau_d) \leq \mathbf{s}^T(-\eta \text{sign}(\mathbf{s}) + \tau_d) \leq -\rho \|\mathbf{s}\| \leq -\rho V^{1/2} \quad (20)$$

式中: $\rho = \eta - \|\tau_d\| > 0$ 。

V 初始时刻状态记为 $V(0)$, 设 t_r 为收敛时间, 则 $V(t_r) = 0$, 对式 (20) 从 0 时刻到 t_r 时刻积分得

$$\int_0^{t_r} dt \leq \int_{V(0)}^{V(t_r)} \frac{-dV}{\rho V^{1/2}} \quad (21)$$

可以计算出收敛时间为

$$t_r \leq \frac{1}{\rho} V^{1/2}(0) \quad (22)$$

假设 t_r 时刻 $V(t_r) \neq 0$, \mathbf{e} 沿着非奇异终端滑模收敛到零的时间为 $t_r + t_s$, 为方便计算, 令 $t_r = 0$ 。

当 $R_i = 0$ 或 $R_i \neq 0, |e(i)| \geq \mu$ 时, 令

$$\mathbf{s} = \dot{\mathbf{e}} + \lambda_1 \mathbf{e} + \lambda_2 \mathbf{e}^{b/a} = \mathbf{0} \quad (23)$$

令 $\mathbf{y} = \mathbf{e}^{1-b/a}$, 把式 (23) 写成一阶线性微分方程标准形式为

$$\frac{d\mathbf{y}}{dt} + \frac{a-b}{a} \lambda_1 \mathbf{y} = -\frac{a-b}{a} \lambda_2 \quad (24)$$

通过求解一阶微分方程得

$$t_s = \frac{a}{\lambda_1(a-b)} \ln \left(\frac{\lambda_1 V(0)^{a-b/a} + \lambda_2}{\lambda_2} \right) \quad (25)$$

当 $|e(i)| \geq \mu$ 时, \mathbf{e} 将沿着非奇异终端滑模面有限时间内收敛到零; 当 $|e(i)| < \mu$ 时, \mathbf{e} 将在一般滑模面上做滑模运动。通过选取一个充分小的 μ , 使 $|e(i)| \geq \mu$ 总能成立, 从而可以实现 \mathbf{e} 在非奇异终端滑模面上有限时间收敛到零, 即 \mathbf{e} 的有限时间稳定性能够得到保证。

由式 (22) 和式 (25) 可知, 通过设计的滑模面和控制律, 机器人能在有限时间内跟踪阻抗轨迹。

注 1 为减少控制律的抖振, 可以利用饱和函

数 $\text{sat}(\cdot)$ 代替符号函数 $\text{sign}(\cdot)$ 。饱和函数可表示为

$$\text{sat}(z) = \begin{cases} 1 & z \geq \theta \\ z/\theta & -\theta < z < \theta \\ -1 & z \leq -\theta \end{cases} \quad (26)$$

4 仿真实验

本节搭建了二连杆机器人与人协同跟踪目标轨迹的场景, 机器人在操作人员外力的作用下完成跟踪期望轨迹的任务, 人机交互场景示意图如图 3 所示。为验证本文所提基于迭代最优阻抗的人机物理交互控制方法的有效性, 分别对外环阻抗控制器和内环轨迹跟踪控制器进行验证。外环控制目标为求取最优阻抗参数, 使阻抗轨迹跟踪期望轨迹, 同时最小化操作人员工作量。内环控制目标为机器人实际轨迹跟踪阻抗轨迹。在 MATLAB/Simulink 中进行仿真, 采用龙格库塔求解器, 仿真步长设置为 0.01 s, 仿真时长为 15 s。

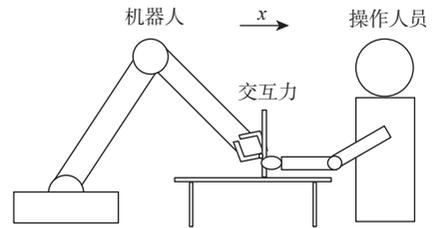


图 3 人机物理交互场景示意图

Fig. 3 Human-robot physical interaction scenario

4.1 控制外环仿真

为验证外环 IOIC 的有效性, 设置期望轨迹为沿着 x 方向上的点对点运动, 具体轨迹设置为幅值 0.4、周期 10 s 的方波信号。机器人的阻抗模型 (1) 中, 初始阻抗参数分别设置为 \mathbf{M}_0 、 \mathbf{B}_0 和 \mathbf{K}_0 。人的运动学模型 (2) 中, 增益分别设置为 $K_r = 0.5$ 、 $K_f = 5$ 和 $K_e = 20$, 这些增益对于 IOIC 来说是未知的。代价函数 (6) 中的权重 \mathbf{Q} 和 R 分别设置为

$$\begin{cases} \mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ R = 0.01 \end{cases} \quad (27)$$

实验开始的前 0.5 s 是数据收集阶段, 设置此阶段控制输入 (8) 中的探测噪声 v_n 为

$$v_n = \sum_{j=1}^{10} (0.05/j) \sin(100t) \quad (28)$$

1) IOIC 性能验证。为证明本文所提外环 IOIC 的性能, 与定阻抗控制 (fix impedance control, FIC) 和利用 LQR 法直接求取的最优阻抗控制 (optimal impedance control, OIC) 进行对比。其中, FIC 通常

需要在设计阶段手动选择适当的阻抗参数。在未知的环境和任务变化中,由于环境多变、任务多变,难以得到普适的阻抗参数,不具备自适应性;OIC在已知环境下被视为一种理想的控制策略,虽然性能最优,但难以在实际中取得应用。实验结果如图4~图6所示。

内收集到的数据迭代求解出近似最优阻抗参数,从而实现与OIC几乎相同的性能。而FIC由于阻抗参数设置不合适,且无法改变,导致经过超调振荡后才能跟踪期望轨迹。由此可得,IOIC能够无需重复实验,仅利用任务初始阶段的交互信息,实现最优交互性能。

为验证IOIC能实现最小化期望轨迹跟踪误差和交互力的控制目标,在图5和图6分别展示了IOIC、FIC和OIC的期望轨迹跟踪误差 e_{p0} 和交互力 f 变化曲线。由图5可以看到,IOIC的期望轨迹跟踪误差在1s后快速收敛至0,且在5~10s和10~15s这2个阶段中,误差曲线与OIC基本重合。由图6可以看到,0~5s阶段,IOIC可以使交互力快速收敛至0,而FIC由于跟踪误差的存在,导致交互力经过几次振荡后才逐渐收敛至0。且通过5~10s和10~15s这2个阶段可以看出,IOIC的交互力峰值明显低于FIC,且能快速无超调地收敛至0。因此,结合图5和图6的仿真实验结果可以得出结论,IOIC可以实现最小化期望轨迹跟踪误差和交互力的控制目标。

2) 软辅助函数有效性验证。为验证所提软辅助函数的作用,图7展示了人机物理交互系统的控制输入,即轨迹跟踪误差加速度。从图7中红色虚线和蓝色虚线可以看到,当不添加软辅助函数时,在0.5s时刻,由于阻抗参数从初始值变为迭代得到的最优值,导致跟踪误差加速度有一个突变,取不同的阻抗参数时,其幅值突变大小不同,当所选的初始阻抗参数与迭代得到的最优阻抗参数差距过大时,会出现显著的突变。因此,可以得出结论:突变的大小与初始阻抗参数的选择有关。这种突变在实际应用中可能会造成机器人系统的抖动。如图7中红色实线和蓝色实线所示,当添加软辅助函数

后,跟踪误差加速度的突变被平滑处理,且收敛速度更快。图7展示了期望轨迹跟踪误差加速度的变化曲线。图中包含四条曲线:红色虚线表示无软辅助函数且初始阻抗参数为[5, 20, 50]的情况,在0.5s处有明显的负向尖峰;蓝色虚线表示无软辅助函数且初始阻抗参数为[10, 20, 50]的情况,尖峰幅值较小;红色实线表示有软辅助函数且初始阻抗参数为[5, 20, 50]的情况,曲线平滑且无尖峰;蓝色实线表示有软辅助函数且初始阻抗参数为[10, 20, 50]的情况,曲线同样平滑且无尖峰。

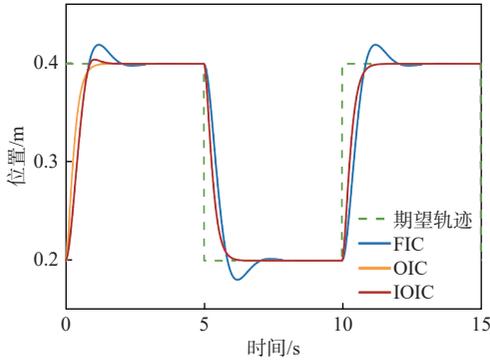


图4 期望轨迹和阻抗轨迹

Fig. 4 Desired trajectory and impedance trajectory

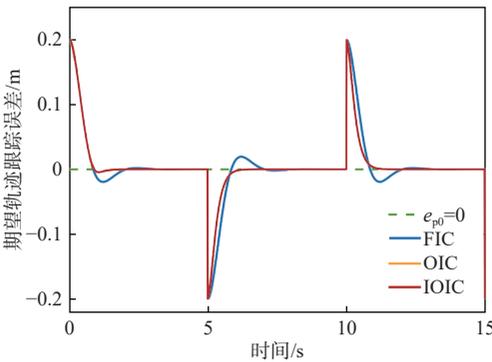


图5 期望轨迹跟踪误差

Fig. 5 Desired trajectory tracking error

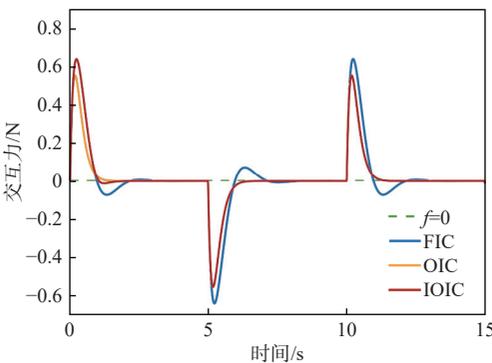


图6 机器人与人的实际交互力

Fig. 6 Actual interaction force between robots and humans

图4展示了IOIC、FIC和OIC在所搭建人机物理交互仿真实验中得到的阻抗轨迹结果。可以看到,0~5s阶段,IOIC的阻抗轨迹在1s后快速跟踪上期望轨迹,且在5~10s和10~15s这两个阶段中与OIC的最优阻抗轨迹基本重合,可以无超调地跟踪期望轨迹。这是因为IOIC能利用0~0.5s

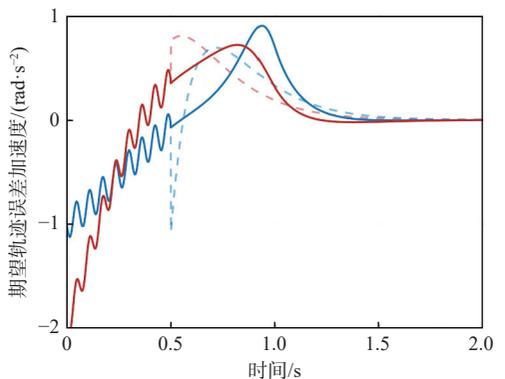


图7 期望轨迹跟踪误差加速度

Fig. 7 Desired trajectory tracking error acceleration

后, 可以看到, 控制输入曲线变化较为平缓 and 连续。由此验证了所提软辅助函数的有效性。同时, 图 7 中, 前 0.5 s 内有一定的抖振, 这是由于在系统输入中加入了探测噪声, 使收集到的数据丰富且非奇异。

3) IOIC 数值验证。为进一步展示本文所提 IOIC 能得到最优阻抗参数, 将 IOIC 与 OIC 结果进行比较。根据选择操作人员的模型参数, 可以得到系统 (4) 的状态矩阵 A 和输入矩阵 B 分别为

$$\begin{cases} A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 40 & 0 & -10 \end{bmatrix} \\ B = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{cases} \quad (29)$$

OIC 利用 LQR 法直接求解 Riccati 方程, 可以得到最优的 P^* 和 K^* 分别为

$$\begin{cases} P^* = \begin{bmatrix} 4.7958 & 0.3830 & 0.1708 \\ 0.3830 & 0.1329 & 0.0073 \\ 0.1708 & 0.0073 & 0.0497 \end{bmatrix} \\ K^* = [38.2969 \quad 13.2889 \quad 0.7335] \end{cases} \quad (30)$$

基于 IOIC 的控制流程, 根据选取的初始阻抗参数, 设置初始的反馈增益为

$$K_0 = [10 \quad 4 \quad 0.2] \quad (31)$$

根据所设置的代价函数权重, 利用前 0.5 s 收集到的数据进行迭代。图 8 展示了矩阵 P 和反馈增益 K 的迭代过程。可以看到, 经过 7 次迭代, 迭代值与最优值的欧氏距离逐渐趋于 0, 即迭代反馈增益逐渐趋于最优反馈增益值。

最终得到的反馈增益为

$$K_7 = [38.2966 \quad 13.2889 \quad 0.7336] \quad (32)$$

可以得到基于迭代求解的最优阻抗参数为

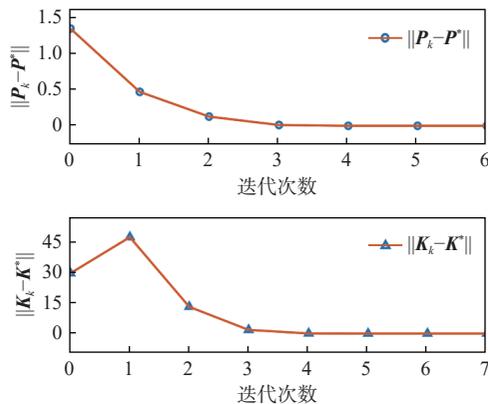


图 8 P 和 K 矩阵迭代过程
Fig. 8 P and K matrix iteration

$$\begin{cases} M^* = 1.3632 \\ B^* = 18.1150 \\ K^* = 52.2048 \end{cases} \quad (33)$$

通过比较式 (30) 和式 (32) 可以看出, IOIC 求解得到的反馈增益和 OIC 直接求解 Riccati 方程得到的最优反馈增益数值基本相同, 从数值上验证了 IOIC 的有效性。

4.2 控制内环仿真

为验证所提内环 NTSMTC 的有效性, 设置二连杆机器人的物理参数如表 1 所示。

表 1 二连杆机器人物理参数
Table 1 Physical parameters of two-link robot

参数	数值
连杆 1 长度 l_1/m	1.0
连杆 2 长度 l_2/m	1.0
连杆 1 质量 m_1/kg	1.5
连杆 2 质量 m_2/kg	1.5
连杆 1 惯性张量 $I_1/(kg \cdot m^2)$	5.0
连杆 2 惯性张量 $I_2/(kg \cdot m^2)$	5.0
重力常数 $g/(m \cdot s^{-2})$	9.81

设置机器人运动学模型 (11) 中的转矩扰动向量为

$$\tau_d = \begin{bmatrix} \sin t + 0.2 \sin(200\pi t) \\ \cos t + 0.2 \cos(200\pi t) \end{bmatrix} \quad (34)$$

滑模面中的参数设置为: $\lambda_1 = \lambda_2 = \text{diag}(3, 3)$ 。式 (13) 中的参数分别设置为: $a = 5, b = 3, \mu = 0.005$ 。控制器中的参数分别设置为: $\beta = \text{diag}(0.5, 1), \eta = 1.5$ 。

仿真结果如图 9~图 11 所示。

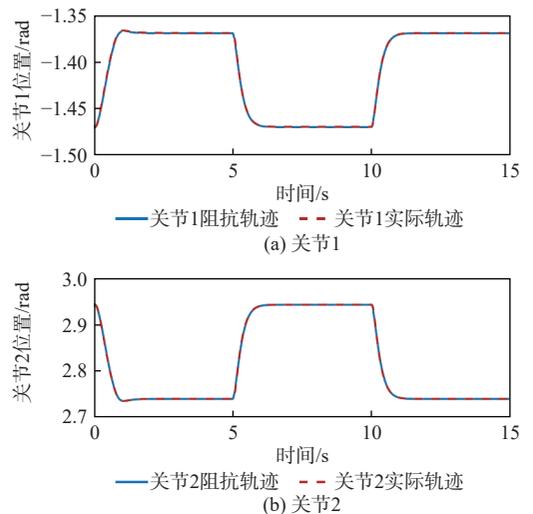


图 9 关节空间中的阻抗轨迹和实际轨迹
Fig. 9 Impedance trajectory and actual trajectory in joint space

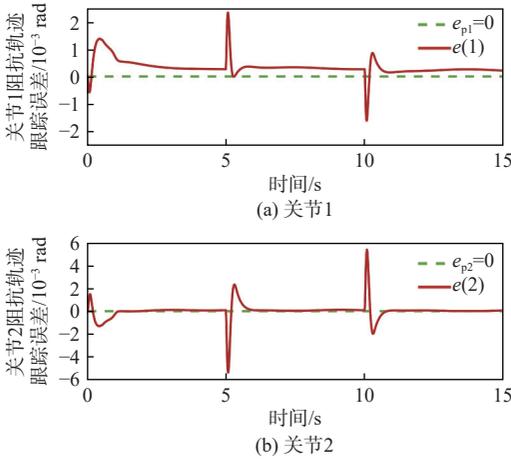


图 10 关节空间中的阻抗轨迹跟踪误差

Fig. 10 Impedance trajectory tracking error in joint space

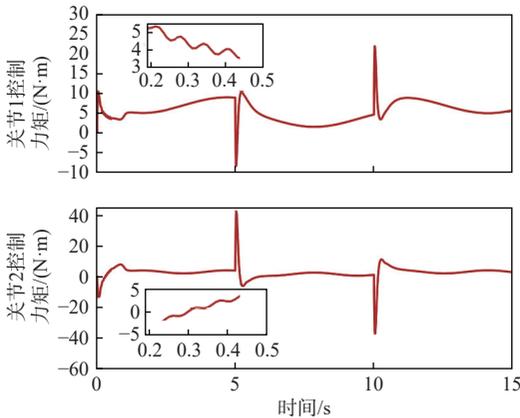


图 11 关节控制力矩

Fig. 11 Joint control torque

图 9 展示了关节空间中的阻抗轨迹和实际轨迹跟踪曲线。图 10 展示了机器人 2 个关节的轨迹跟踪误差,可以看出,关节 1 的轨迹跟踪误差始终在 3×10^{-3} rad 范围内,关节 2 的轨迹跟踪误差始终在 6×10^{-3} rad 范围内,证明了所提轨迹跟踪方法的有效性。图 11 展示了 2 个关节的控制力矩曲线,可以看到,关节 1 的控制力矩最大不超过 25 N·m,关节 2 的控制力矩最大不超过 42 N·m,在 5 s 和 10 s 时,关节力矩发生较大的变化,这是因为此刻阻抗轨迹突然变化所导致。0.5 s 前的抖振是由于加入的探测噪声所引起。整体来看,控制力矩平稳无明显抖振,验证了所提内环轨迹跟踪方法控制力矩的平稳性。

5 结论

1) 本文提出的基于迭代最优阻抗的人机物理交互控制方法,在人机协作型任务中,无需多次重复实验,能够实现机器人与人的最优交互性能,确保机器人轨迹跟踪精度的同时减少了人在执行任务过程中消耗的力。

2) 外环 IOIC 无需交互人的模型信息,利用可在线测量的轨迹跟踪误差和交互力信息,求取近似最优阻抗参数,与 OIC 求取的最优阻抗参数基本一致。

3) 内环 NTSMTC 在交互力和内部干扰的条件下具有很好的轨迹跟踪能力,控制力矩整体平滑且无明显抖振,并通过李雅普诺夫能量函数对控制律进行收敛性证明。

参考文献 (References)

- [1] 李琳,肖佳栋,张铁,等.基于自适应迭代的机器人曲面恒力跟踪[J].北京航空航天大学学报,2019,45(4):641-649.
LI L, XIAO J D, ZHANG T, et al. Constant-force curved-surface-tracking with robotic manipulator based on adaptive iterative algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(4): 641-649(in Chinese).
- [2] ZAHEDI F, CHANG D, LEE H. User-adaptive variable damping control using Bayesian optimization to enhance physical human-robot interaction[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 2724-2731.
- [3] LIU X, GE S S, ZHAO F, et al. A dynamic behavior control framework for physical human-robot interaction[J]. Journal of Intelligent & Robotic Systems, 2020, 101(1): 14.
- [4] HOGAN N. Impedance control: an approach to manipulation[J]. Journal of Dynamic Systems, Measurement, and Control, 1985, 107(1): 1-24.
- [5] YU X B, LI B, HE W, et al. Adaptive-constrained impedance control for human-robot co-transportation[J]. IEEE Transactions on Cybernetics, 2022, 52(12): 13237-13249.
- [6] SHARIFI M, ZAKERIMANESH A, MEHR J K, et al. Impedance variation and learning strategies in human-robot interaction[J]. IEEE Transactions on Cybernetics, 2022, 52(7): 6462-6475.
- [7] ABU-DAKKA F J, SAVERIANO M. Variable impedance control and learning-a review[J]. Frontiers in Robotics and AI, 2020, 7: 590681.
- [8] SUN T R, PENG L, CHENG L, et al. Stability-guaranteed variable impedance control of robots based on approximate dynamic inversion[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(7): 4193-4200.
- [9] JIN Z H, QIN D D, LIU A D, et al. Model predictive variable impedance control of manipulators for adaptive precision-compliance tradeoff[J]. IEEE/ASME Transactions on Mechatronics, 2023, 28(2): 1174-1186.
- [10] LIU X, WU J G, WANG L. Sensorless variable impedance compliant control for human-robot interaction[J]. AIP Advances, 2022, 12(11): 115121.
- [11] GRAFAKOS S, DIMEAS F, ASPRAGATHOS N. Variable admittance control in pHRI using EMG-based arm muscles co-activation[C]//Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. Piscataway: IEEE Press, 2016: 1900-1905.
- [12] DONG Y T, REN B B. UDE-based variable impedance control of uncertain robot systems[J]. IEEE Transactions on Systems, Man,

- and *Cybernetics: Systems*, 2019, 49(12): 2487-2498.
- [13] 梅雪松, 刘星, 赵飞, 等. 协作机器人外力感知与交互控制研究现状及展望[J]. *航空制造技术*, 2020, 63(9): 22-32.
- MEI X S, LIU X, ZHAO F, et al. Research status and prospects of external force perception and interaction control methods for collaborative robots[J]. *Aeronautical Manufacturing Technology*, 2020, 63(9): 22-32(in Chinese).
- [14] YANG C G, GANESH G, HADDADIN S, et al. Human-like adaptation of force and impedance in stable and unstable interactions[J]. *IEEE Transactions on Robotics*, 2011, 27(5): 918-930.
- [15] YANG J T, SUN T R, CHENG L, et al. Spatial repetitive impedance learning control for robot-assisted rehabilitation[J]. *IEEE/ASME Transactions on Mechatronics*, 2023, 28(3): 1280-1290.
- [16] RAHIMI H N, HOWARD I, CUI L. Neural impedance adaption for assistive human-robot interaction[J]. *Neurocomputing*, 2018, 290: 50-59.
- [17] XU Z H, LI X X, LI S, et al. Dynamic neural networks based adaptive optimal impedance control for redundant manipulators under physical constraints[J]. *Neurocomputing*, 2022, 471: 149-160.
- [18] JOHANSSON R, SPONG M W. Quadratic optimization of impedance control[C]//Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 1994: 1: 616-621.
- [19] MATINFAR M, HASHTRUDI-ZAAD K. Optimization-based robot compliance control: geometric and linear quadratic approaches[J]. *International Journal of Robotics Research*, 2005, 24(8): 645-656.
- [20] HUANG H H, YANG C G, CHEN C L P. Optimal robot-environment interaction under broad fuzzy neural adaptive control [J]. *IEEE Transactions on Cybernetics*, 2021, 51(7): 3824-3835.
- [21] GUO Y D, TIAN Y, WANG H P. Learning-based adaptive optimal impedance control to enhance physical human-robot interaction performance[J]. *International Journal of Control, Automation and Systems*, 2022, 20(9): 3053-3062.
- [22] SUZUKI S, FURUTA K. Adaptive impedance control to enhance human skill on a haptic interface system[J]. *Journal of Control Science and Engineering*, 2012, 2012: 365067.
- [23] JIANG Y, JIANG Z P. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics [J]. *Automatica*, 2012, 48(10): 2699-2704.

Human-robot physical interaction control method based on iterative optimal impedance

LIU Weirong*, WEI Zifeng, JIN Zhenbing, MENG Jiahao, WANG Xingkun, ZHANG Haochen

(College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: In order to improve the accuracy and compliance of human-robot physical interaction and achieve optimal interaction performance, a human-robot physical interaction control method based on iterative optimal impedance was proposed to solve the problem that iterative learning-based impedance control method needs to repeat the same task many times. The proposed method draws on the mechanism by which iterative optimal control can optimize cost function to determine optimal control input to the system without information of the system matrix. A double-loop control structure was used for the proposed control method. An iterative optimal impedance controller (IOIC) was designed for a task-oriented outer loop. The problem of determining optimal impedance parameters was described as a linear quadratic regulator problem, which utilized iterative optimal control to find optimal feedback gain and minimize cost function including tracking error and interaction force. Robot jitter caused by parameter mutations was avoided by introducing soft auxiliary functions. A nonsingular terminal sliding mode trajectory tracking controller (NTSMTC) was used in the inner loop of the robot to make the actual trajectory of the robot track impedance trajectory output by the outer loop, and the chattering of control law was eliminated by saturation function. Simulation results prove that the proposed method can obtain optimal impedance parameters only by using interactive information in the initial stage of the task once in a human-robot collaborative task, so as to minimize the trajectory tracking error and the force consumed by the human during the task.

Keywords: human-robot physical interaction; impedance control; trajectory tracking control; optimal control; sliding mode control

Received: 2023-06-05; Accepted: 2023-10-16; Published Online: 2023-11-24 14:33

URL: link.cnki.net/urlid/11.2625.V.20231123.1034.002

Foundation items: National Nature Science Foundation of China (62261032); Central Government Guiding Funds for Local Science and Technology Development Program (25ZYJA026); Key Talent Project of Gansu Province; Gansu Education Science and Technology Innovation Fund (2022A-021)

* Corresponding author. E-mail: liuwr@lut.edu.cn