

一种用于机械臂拟人化控制的学习框架

赵 亮^{1,2,3}, 杨 铁^{1,2}, 于 鹏^{1,2}, 杨 洋^{1,2}

(1. 中国科学院沈阳自动化研究所机器人学国家重点实验室, 辽宁 沈阳 110016;
2. 中国科学院机器人与智能制造创新研究院, 辽宁 沈阳 110169; 3. 中国科学院大学, 北京 100049)

摘 要: 赋予机器人拟人化的动作可以使机器人的行为更具可解释性和可预测性, 可以显著提升人机协作任务的质量和安全性。本文提出一个“人在回路”学习框架, 从人类遥操作示教中学习拟人化姿态特征, 并将学习到的特征模型应用于冗余机械臂的控制中产生拟人化动作。在模型训练过程中采用“人在回路”的在线再标注方法, 克服了协变量偏移问题, 将总示教时间缩减至 10 min 以内。人机姿态对比和动态轨迹跟踪任务实验验证了该框架训练得到的拟人化控制方法的有效性。用户评价测试表明, 使用拟人化姿态约束的机械臂在主观体验方面对用户更加友好, 作为共同作业的工具更容易被非专业用户所接受。

关键词: 拟人化控制; 行为克隆; 人机交互; 遥操作; 外骨骼

中图分类号: TP242

文献标识码: A

文章编号: 1002-0446(2023)-05-0513-10

A Learning Framework for Controlling Robotic Manipulators with Human-like Actions

ZHAO Liang^{1,2,3}, YANG Tie^{1,2}, YU Peng^{1,2}, YANG Yang^{1,2}

(1. State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China;
2. Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China;
3. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Endowing robotic manipulators with human-like actions can make their behaviours more explainable and predictable, which improves the quality and safety of human-robot collaboration (HRC). A human-in-the-loop learning framework is proposed to learn the human-like posture features from human teleoperation demonstration. The learnt feature model can be used to generate human-like actions on the redundant manipulators. Moreover, an online relabeling approach is adopted in the training session to address the covariate shift problem, which reduces the total demonstration time to 10 min or less. The effectiveness of the human-like control method trained by the proposed framework is validated through comparing the posture between the human and the robot, and tracking the dynamic trajectories. A user study shows that, manipulators controlled with the proposed human-like posture constraint method are more friendly to the users in subjective feelings, and more acceptable as co-operative tools for non-professional users.

Keywords: human-like control; behavioral cloning; human-robot interaction; teleoperation; exoskeleton

1 引言 (Introduction)

人类与机器人并肩协作的场景已经越来越常见, 这在增强人类能力的同时大幅提高了生产力。这种人机协作的工作方式创造了很多新型的机器人应用, 如机器人辅助手术^[1-2]、人机协作装配^[3-4]等。人机协作的一个重要特点就是人与机器人共享工作空间^[5]。在近距离、富交互场景中, 赋予机器人拟人化的属性可以使机器人的行为动作对于人类协作者展现出更好的可解释性和可预测性, 这对于提升人机交互的质量具有重要的意义。例如, 在机器人辅助手术任务中^[2], 无法预测机械臂的姿态可

能频繁导致机械臂阻挡医生的视线、操作空间, 甚至与正在进行手术操作的医生发生碰撞进而引发危险。具有拟人化特性的机械臂可以有效避免这些状况的发生, 在与机械臂共同作业时, 医生可以像和另一位人类助手协作一样, 完全把注意力的焦点放在末端执行器的位姿上, 并通过直觉和常识避免与机械臂发生干涉、碰撞等问题。再如, 大多数遥操作任务中, 操作者需要通过安装在远程环境中的摄像机对机械臂和被操作对象进行观测。与人类手臂形态、动作方式差异较大的机械臂会给遥操作作业带来较大的精神负担, 从而使操作者容易疲劳和出错^[6]。而具有拟人化特性的机械臂可以在遥操作中

产生更强的临场感，操作者对机械臂的控制更加直觉化。已有研究表明，类人机器人和拟人化的动作行为对提升人机交互的性能具有积极意义^[7]。

AVATAR 系统^[8-10] 可以将实时捕捉到的人体运动映射到一个虚拟化身或机器人中，从而在人工载体上实现拟人化动作，这是一种非常直觉化的方法。尤其在冗余结构的人形机械臂中，由于人一机运动学结构的近似性，得到拟人化的动作相对容易^[11-12]。在一些已有的研究工作中，曾使用外骨骼^[13-14] 或基于 IMU（惯性测量单元）^[15] 和光学原理^[16] 的运动捕捉设备来获取人体手臂的姿态。已有多项研究工作^[10,13-14,16] 基于遥操作接口，将实时捕捉到的人体手臂姿态融入到机械臂的控制量，实现机械臂的拟人化动作。然而，现实中大多数通用型机械臂并不具有类人化的结构；另一方面，机器人通常作为一种自动化工具参与人机协作，需要建立人类动作特征模型，使机器人能够自主产生拟人化的动作。

在部分 7 自由度机械臂上可以定义一种臂形角参数^[17]，在人体手臂上也可以找到相似的几何定义。文 [18] 研究了从人体手部位姿到臂形角的非线性回归模型。为了更精确地逼近这种回归关系，也有很多研究工作^[2,19-21] 采用基于神经网络的方法预

测臂形角来实现机械臂的拟人化姿态。然而，这个过程需要预先学习从人体手部姿态到臂形角的回归关系，并将基于人体手臂定义的臂形角映射到特定机械臂的定义中，建模和训练过程需要大量关于机械臂结构和机器人学的背景知识，对非专业用户并不友好。

本文采用模仿学习^[22-24] 的方法，通过可穿戴外骨骼精确捕捉人体上肢运动，并基于“人在回路”的示教、训练方法建立人类上肢“手—臂”运动协调的非线性关系模型来实现冗余机械臂的拟人化控制。本文研究内容的总体框架如图 1 所示。

2 具有姿态模仿特性的遥操作示教系统
(Posture-mimicry-featured teleoperation system for demonstration)

2.1 用于动作捕捉的外骨骼主端设备

所使用的训练数据集需要通过遥操作示教方式来获取。所采用的遥操作主端设备是在前期工作中研发的一种可穿戴外骨骼主手^[25]。在对人体上肢的生物机械系统建模时，常用图 2 所示的 7 自由度串联结构作为其运动学模型。该外骨骼设备在各个转动关节处安装有编码器，可以实时采集关节角度。当操作者穿戴外骨骼时，所获取的关节角度 q_i

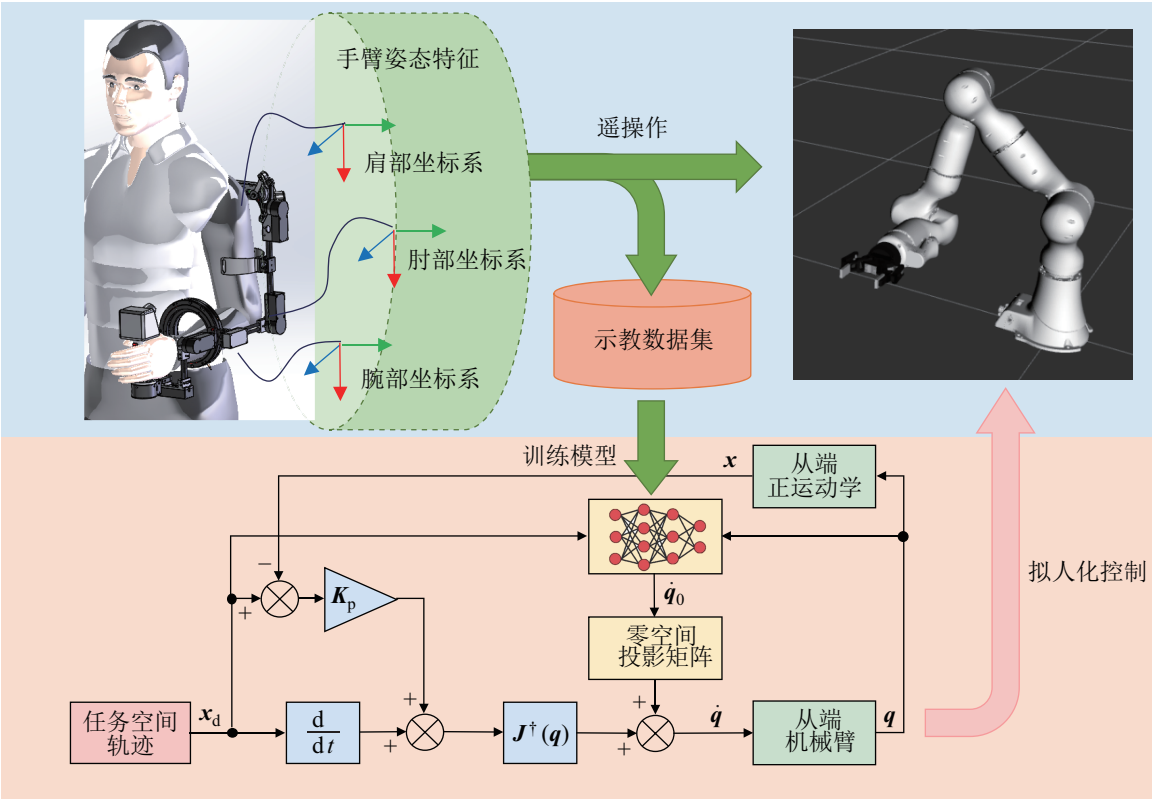


图 1 机械臂拟人化控制学习框架

Fig.1 Learning framework for controlling robotic manipulators with human-like actions

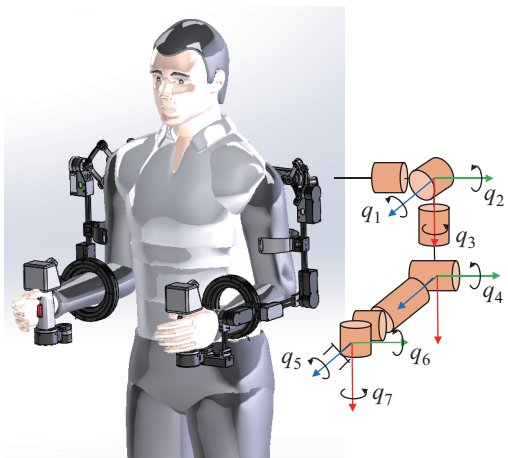


图 2 外骨骼主端及其运动学结构

Fig.2 Exoskeleton-based master device and its kinematic structure

($i = 1, 2, \dots, 7$) 被用作运动学模型的参数, 从而可以通过正运动学计算得到人体手臂运动链上任意坐标系相对于基坐标系的变换矩阵。关节 k 的坐标系相对于基坐标系的变换为

$${}^0T_k = {}^0T_1 T_2 \dots T_{k-1} T_k \quad (1)$$

其中, ${}^{k-1}T_k$ 为从坐标系 $\{k-1\}$ 到坐标系 $\{k\}$ 的变换矩阵, 并以关节 k 的旋转角度 q_k 为参数。其中变换矩阵 ${}^0T_7$ 和 ${}^0T_3$ 分别表示人体手部和肩部的位姿。

2.2 遥操作中的从端机械臂姿态约束方法

在具有姿态模仿特性的遥操作任务中, 从端机械臂末端执行器受控跟随操作者的手部位姿, 此外还有附加的姿态控制要求。对于冗余机械臂, 可以利用零空间控制量来驱动从端机械臂模仿主端操作者的手臂姿态。然而, 由于主、从端运动学结构上的差异, 上述 2 项控制要求在大多数情况下难以同时满足。因此, 本文采用了一种分级策略来协调遥操作中的 2 种控制要求。为了实现这一目标, 采用

基于 CLIK (闭环逆运动学) 算法^[26] 的控制策略, 在此基础上增加了一个可以模仿主端姿态的控制器, 并将其产生的控制量投影到机械臂末端的零空间, 从而在保证精确控制末端执行器位姿的同时对机械臂姿态加以约束。由这种控制策略构成的主从遥操作系统如图 3 所示。

CLIK 算法本质上是一种关节空间速度层面的控制策略。在冗余机械臂上, 可以产生关节空间速度的 CLIK 控制策略可以表示为^[26]

$$\dot{q} = J^\dagger(q) [\dot{x}_d + K_p(x_d - x)] + (I - J^\dagger(q)J(q)) \dot{q}_0 \quad (2)$$

其中 $J^\dagger(q)$ 是雅可比矩阵 $J(q)$ 的 Moore-Penrose 伪逆, x_d 是机械臂末端执行器的期望目标位姿, 可以通过主端的 ${}^0T_7$ 计算得到。 x 是机械臂末端执行器的实际位姿, 根据机械臂当前关节状态 q 可以通过正运动学方程计算得到。 K_p 是一个对称正定矩阵, 可以保证 $\dot{q}$ 是收敛的。矩阵 $I - J^\dagger(q)J(q)$ 可以将任意关节空间速度 $\dot{q}_0$ 映射到零空间。

式 (2) 右侧的 2 项分别作用于 2 个不同的任务: 第 1 项可以保证末端执行器始终跟随期望的目标位姿, 在该控制策略中作为主要任务; 第 2 项提供了额外的零空间约束, 作为次要任务。

次要任务可根据需要设定为不同的形式来帮助机械臂实现不同的控制特性, 例如奇异点规避、运动避障、关节限位规避等。在本系统中, 次要任务的目标是控制机械臂姿态, 使之模仿主端操作者的手臂姿态。以图 4(a) 所示的珞石 xMate3 型机械臂及其运动学结构为例, 机械臂的基座位姿是固定的, 而末端执行器的位姿由主要任务中的目标位姿确定。利用机械臂的冗余特性, 在保持基座和末端执行器位姿不变的前提下, 可以在末端执行器的零空间中搜索相应的关节控制量, 以使中间连杆达到最接近操作者手臂姿态的位置。如图 4(b) 所示, 根

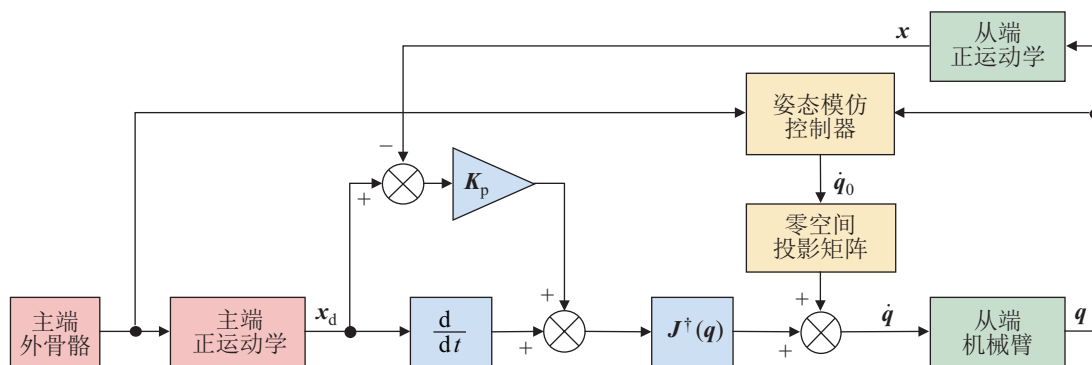
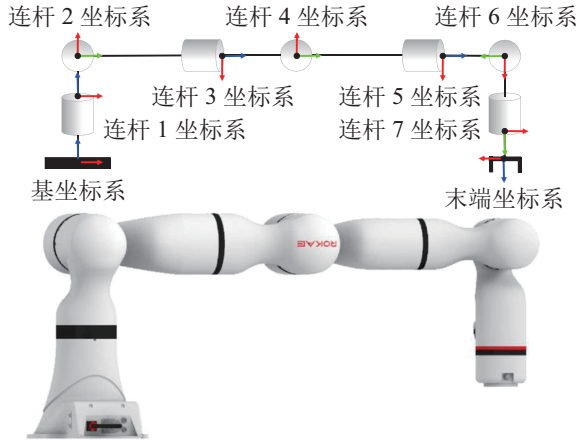
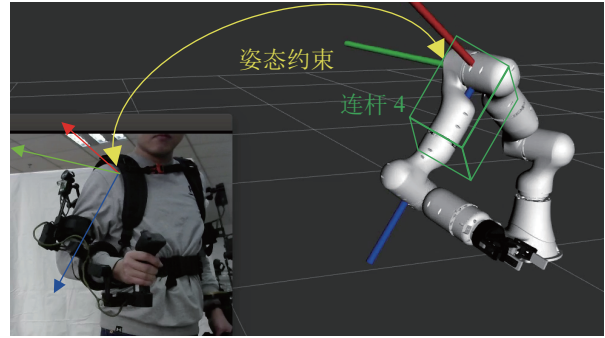


图 3 具有姿态模仿特性的 CLIK 控制框图

Fig.3 CLIK control scheme with posture-mimicry feature



(a) xMate3 机械臂的外形及运动学结构



(b) xMate3 机械臂的拟人化约束条件

图 4 从端运动学结构及其拟人化姿态约束条件

Fig.4 Kinematic structure of the slave device and its human-like posture constraints

据所使用的机械臂的运动学结构和安装方式, 连杆 4 在结构上相当于操作者的上臂, 二者方向越接近, 机械臂与操作者的手臂姿态越相似。因此, 本系统中选取连杆 4 的方向作为机械臂拟人化姿态的约束条件。以操作者上臂坐标系为参照, 将 2 个坐标系的方向对齐的问题可以通过优化方法进行求解。使 2 个坐标系之间的方向差异趋向最小的关节空间速度 $\dot{\mathbf{q}}_0$ 可以表示为

$$\dot{\mathbf{q}}_0 = -\eta \nabla H(\mathbf{q}), \quad \nabla H(\mathbf{q}) = \left(\frac{\partial H(\mathbf{q})}{\partial \mathbf{q}} \right)^T \quad (3)$$

其中, η 为正实系数, 梯度向量 $\nabla H(\mathbf{q})$ 保证了 $\dot{\mathbf{q}}_0$ 可以使目标函数 $H(\mathbf{q})$ 以最快速度趋向最小。在本文所涉及的问题中, 目标函数用来反映从端机械臂连杆 4 坐标系相对于操作者肩部坐标系的方向误差, 由此, 可将目标函数定义为

$$H(\mathbf{q}) = \delta_x^2 + \delta_y^2 + \delta_z^2 \quad (4)$$

用向量 $\boldsymbol{\delta}_\Theta = (\delta_x, \delta_y, \delta_z)^T$ 表示 2 个旋转矩阵之间的角度差异, 其各分量值可以从转角微分矩阵中得到:

$$\mathbf{S}(\boldsymbol{\delta}_\Theta) = \begin{bmatrix} 0 & -\delta_z & \delta_y \\ \delta_z & 0 & -\delta_x \\ -\delta_y & \delta_x & 0 \end{bmatrix} \quad (5)$$

对于给定的 2 个旋转矩阵, 二者间的角度微分矩阵可以表示为^[27]

$$\mathbf{S}(\boldsymbol{\delta}_\Theta) = {}^0\mathbf{R}_3 {}^0\mathbf{R}_4^T - \mathbf{I}_{3 \times 3} \quad (6)$$

其中, ${}^0\mathbf{R}_3$ 和 ${}^0\mathbf{R}_4^T$ 分别表示主端操作者肩部的方向和从端机械臂连杆 4 的方向, $\mathbf{I}_{3 \times 3}$ 是 3 阶单位矩阵。

3 机械臂拟人化动作学习框架 (Learning framework of human-like actions on manipulators)

3.1 基于行为克隆方法的模仿学习

提出了一种“人在回路”的学习框架来提取示教数据集集中潜在的拟人化动作约束特征。该框架使用行为克隆方法^[24]训练拟人化姿态生成策略, 目标是获得一个可以连续地从当前状态 $\mathbf{s}_t$ 产生期望动作 $\mathbf{a}_t$ 的映射模型 $\mathbf{a}_t \leftarrow \pi_\theta(\mathbf{s}_t)$ 。 π_θ 为以 θ 为参数的神经网络策略。示教数据集通常可表示为“状态-动作”二元组的集合, 即 $\mathcal{D} = \{(\mathbf{s}_t, \mathbf{a}_t)\}$ 。行为克隆算法训练回归模型的方法为

$$\arg \min_{\theta} E_{(\mathbf{s}, \mathbf{a}) \sim \mathcal{D}} (\pi_\theta - \mathbf{a})^2 \quad (7)$$

当策略模型训练收敛时, 对于 $\mathcal{D}$ 中的每个状态 $\mathbf{s}_t$, 策略 π_θ 产生的动作 $\hat{\mathbf{a}}_t$ 都可以非常接近 $\mathbf{a}_t$, 也就是说, 此时的策略可以模仿示教者的行为模式。

在本文的学习框架中, 定义 t 时刻的状态向量为 $\mathbf{s}_t = (\mathbf{Q}_t, \mathbf{p}_t, \mathbf{q}_t)$, 其中包括了末端执行器的位姿四元数 $\mathbf{Q}_t \in \mathbb{R}^4$, 位置向量 $\mathbf{p}_t \in \mathbb{R}^3$, 以及全部的关节角 $\mathbf{q}_t \in \mathbb{R}^7$ 。同时, 定义 t 时刻的动作向量 $\dot{\mathbf{q}}_{0t} \in \mathbb{R}^4$, 表示用于次要任务的部分关节空间速度。

在神经网络模型方面, 本文的训练框架采用了一种编码器-解码器的结构。网络结构和训练数据在 CLIK 控制回路中的流向如图 5(a) 所示。

网络中包括了一个从端约束生成器, 即编码器部分, 以及一个主端状态生成器, 即解码器部分。由于拟人化姿态约束是基于从端机械臂当前状态和目标位姿信息的, 编码器网络部分的输入形式为被

跟踪的目标位姿以及从端机械臂当前的关节角度的组合, 输出为用于产生拟人化姿态的关节空间角速度。而主端设备在控制回路中产生被跟踪的目标位姿和从端机械臂的姿态约束信息, 所以, 将编码器部分生成的拟人化姿态约束角速度向量 $\dot{q}_0$ 、从端机械臂当前状态 q 以及被跟踪的目标位姿 x_d 组合, 作为解码器部分的输入来还原主端外骨骼的状态。

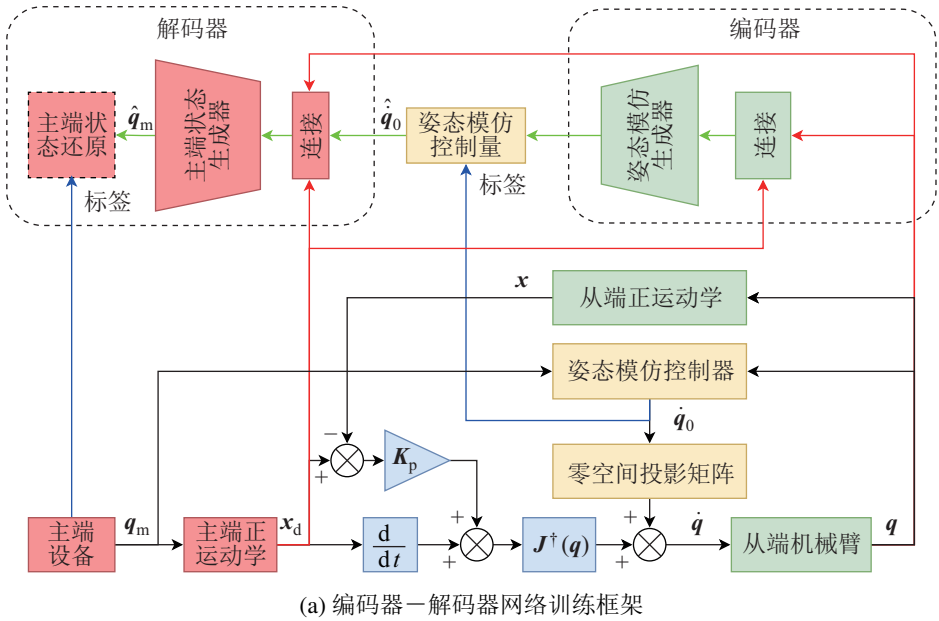
本文的学习框架使用编码器—解码器结构网络的目的在于建立一种闭环的自验证机制来判断所产生的拟人化姿态约束的正确性。在应用环节中, 可以通过解码器还原得到主端外骨骼的状态信息, 并根据主端外骨骼的运动学方程生成一个虚拟的目标位姿 x_{dv} , 进而根据虚拟目标位姿 x_{dv} 和真实的目标位姿 x_{dt} 之间的差异程度来判断编码器或解码器的输出结果中是否存在可能的错误。

该框架中还使用了阈值机制来控制一个门开关, 根据 x_{dv} 和 x_{dt} 之间的差异程度来决定是否将编码器部分生成的姿态约束动作应用于控制回路中。以上应用环节中, 编码器网络和解码器网络协同作用于控制回路的过程如图 5(b) 所示。

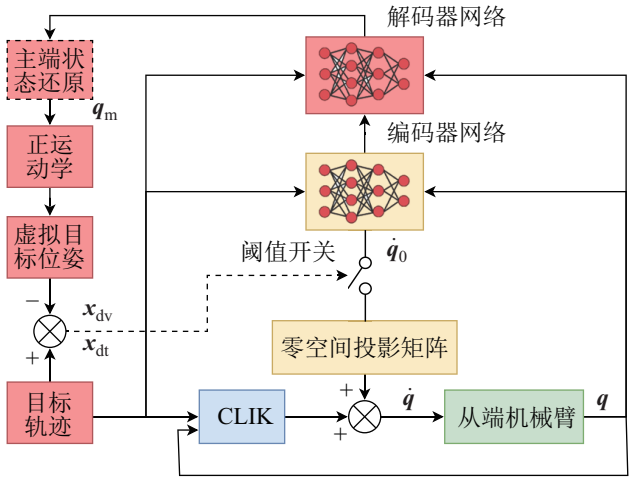
在本系统中, 编码器网络和解码器网络都使用多层感知机 (MLP), 且使用 $\tanh()$ 函数作为输入层和隐藏层的激活函数。2 个网络的层次结构和各层节点数如表 1 所示。

表 1 网络结构
Tab.1 Network structure

层数	编码器				解码器			
	L1	L2	L3	L4	L1	L2	L3	L4
节点数	14	28	56	4	18	36	72	7



(a) 编码器—解码器网络训练框架



(b) 编码器—解码器网络应用框架

图 5 编码器—解码器网络的训练和应用框架

Fig.5 Training and application frameworks of the encoder-decoder network

3.2 “人在回路”的在线再标注方法

行为克隆学习方法本质上属于监督学习, 通过有限的示教数据来训练拟人化姿态约束生成策略。这就意味着当未被示教数据集所包含的状态出现时, 策略难免会遇到特征漂移的问题, 导致不能输出正确的拟人化约束动作。另一方面, 采集示教数据的工作量也很大, 示教数据集不可能是无限大的、完整覆盖状态空间的。为了解决这个问题, 结合本文采用可穿戴设备进行示教数据采集的特点, 提出了一种“人在回路”的在线再标注方法, 通过多轮短时间的“再标注一再训练”过程对策略模型进行精细调整。算法 1 为在线再标注算法。

算法 1 在线再标注

前置条件: 预训练示教数据集 $\mathcal{D}$

使用数据集 $\mathcal{D}$ 预训练的策略 π_θ

```

1:  $\mathcal{D}_E \leftarrow \emptyset$ 
2: for 再标注轮次  $i = 1, 2, \dots, n_{\text{round}}$  do
3:   for 时间点  $t = 1, 2, \dots, n_{\text{steps}}$  do
4:     // 使用策略产生拟人化姿态约束控制量
      $\hat{\mathbf{a}}_t \leftarrow \pi_\theta(\mathbf{s}_t)$ 
     // 通过遥操作获取姿态约束控制量
      $\mathbf{a}_t \xleftarrow{\text{teleoperation}} \pi_h(\mathbf{s}_t)$ 
5:     if  $\|\mathbf{a}_t - \hat{\mathbf{a}}_t\|^2 \geq \varepsilon$  then
6:        $\mathcal{D}_E \leftarrow \mathcal{D}_E \cup (\mathbf{s}_t, \mathbf{a}_t)$ 
7:     end if
8:   end for
9:   // 合并再标注的数据集
10:   $\mathcal{D} \leftarrow \mathcal{D}_E \cup \mathcal{D}$ 
11:  // 执行再训练
12:  for each  $(\mathbf{s}_t, \mathbf{a}_t) \sim \mathcal{D}$  do
13:     $\theta_n \leftarrow \arg \min_{\theta} \|\pi_\theta(\mathbf{s}_t) - \mathbf{a}_t\|$ 
14:  end for
15:  // 更新策略网络参数
   $\pi_\theta \leftarrow \theta_n$ 
16: end for

```

再标注环节中, 基于主端外骨骼的姿态约束零空间控制器产生拟人化姿态约束控制量, 并作用于 CLIK 控制回路。同时, 已训练的拟人化约束生成策略网络也会得到主端外骨骼提供的操作者手部位姿信息, 以其作为输入来生成姿态约束控制量。2 种途径产生的约束控制量之间的均方误差可以用来评估当前状态下特征偏移的程度。DAgger^[28] 是一种行为克隆中应对特征漂移问题的常用的再标注技术手段, 这种方法需要人类专家对策略产生的动作进行逐帧的离线再标注后再对策略进行重新训练。当数据集很庞大时, DAgger 方法的人工再标注工

作负担就会变得难以接受。本文的遥操作系统可以通过外骨骼主端设备产生的约束控制量实时在线评估当前状态的特征漂移程度, 因而非常容易发现偏离分布的数据点, 然后可通过外骨骼主端产生的约束控制量进行再标注并存储到一个新的再标注数据集 $\mathcal{D}_E$ 中。这种在线再标注方法可以在很大程度上减轻操作者的工作负担, 并且几乎不需要操作者具有专业的数据甄别和再标注能力, 仅通过穿戴外骨骼主端设备对从端机械臂进行遥操作控制即可完成这一过程。

将新产生的再标注数据集 $\mathcal{D}_E$ 与原有的示教数据集进行合并形成扩增数据集。每次扩增后, 示教数据集的分布将更趋近于对状态空间的完整覆盖。使用扩增后的示教数据集对策略进行重新训练, 可以大幅度提高策略的泛化能力。

4 实验 (Experiments)

4.1 实验环境

在 ROS 系统的可视化仿真平台 RViz 中搭建仿真实验环境。在虚拟环境中设置了一套珞石 xMate3 型 7 自由度机械臂的 URDF (统一机器人描述格式) 模型作为实验从端, 并接收关节空间速度作为控制指令。外骨骼主端设备通过以太网连接到遥操作系统中, 并以 1 kHz 的频率发送 ROS 消息作为遥操作的动作指令。CLIK 算法的控制回路周期为 5 ms。

在示教环节中, 主端设备的状态对学习框架和控制回路完全开放。在测试环节中, 主端设备的关节数据将不再提供给从端, 从端仅可获取手部位姿作为目标指令。

4.2 在线再标注方法效果评价测试

为了评估本文的在线再标注方法的有效性和对数据采集效率的提升程度, 通过实验测试了多组不同时间参数的再标注过程, 并以单次示教、无再标注过程的方法作为参照进行对比。示教和再标注的时间参数如表 2 所示。

表 2 示教和再标注时间参数

Tab.2 Temporal parameters of demonstration and relabeling

单次长时间示教 /min	多轮再标注 /min		
	初始轮	轮次 1	轮次 2
5	3	1	1
8	4	2	2
10	5	2	2

在遥操作数据采集环节中, 操作者穿戴外骨骼进行自由的手臂运动, 尽可能完整地覆盖上肢可达

的运动范围。对于单次示教方法, 数据采集时间持续相对较长 (5 min、8 min 和 10 min), 且示教数据采集完成后, 策略模型仅进行 1 轮训练, 模型网络参数收敛后即完成训练过程。而对于再标注方法, 首次采集的示教数据仅用于对策略进行预训练, 而后还需要经历 2 轮不同长时的“再采集一再标注一再训练”的过程。此外, 实验中还单独采集了一段时间约为 2 min 的随机运动轨迹数据作为测试集, 用于对各个策略的性能进行评估。

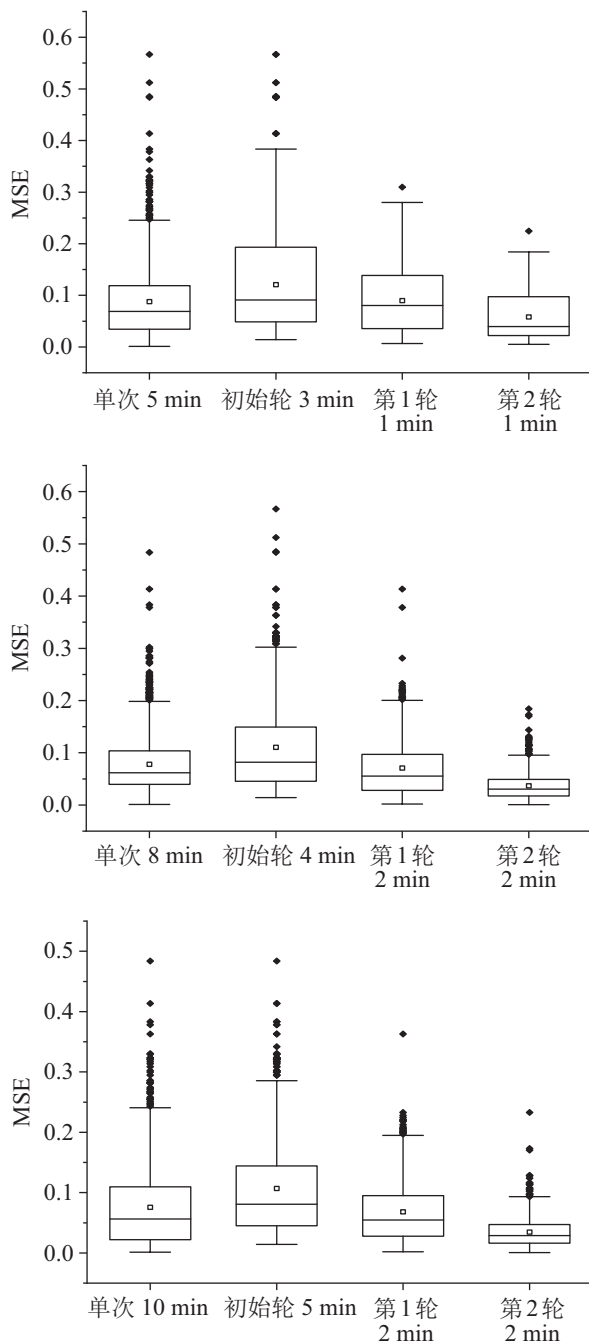


图 6 使用不同示教与再标注时间参数训练的模型误差
Fig.6 Error of models trained with different temporal parameters of demonstration and relabeling

将主端外骨骼产生的拟人化姿态约束控制量作为参照标准, 在测试集上使用不同示教时长参数的数据集训练得到的策略进行均方误差 (MSE) 评估, 得到的统计结果如图 6 所示。

结果表明, 对于单次示教数据的采集和训练过程, 更长时间的数据采集并不能显著提高模型的预测精度。这是因为操作者进行数据采集时虽然被要求示教范围尽可能广泛覆盖, 但是由于每个人固有的运动习惯不同, 即使延长数据采集时间, 也可能有很多区域会被遗漏, 导致训练得到的模型的泛化效果不佳。而相对应地, 在多轮在线再标注方法训练得到的模型中, 无论数据采集时间长短, 每一轮再标注和再训练都会对模型的预测精度有显著提升。该统计结果也揭示了单纯延长示教数据采集时间并不一定能提升模型预测精度。实验发现, 预训练数据采集达到 4 min, 再经过 2 轮时长为 2 min 的再标注过程, 就可以得到很高的模型精度。整个数据采集和再标注过程不超过 10 min, 这个过程的体力负担对大多数操作者都是可以接受的。

与在线再标注方法相对应的是在模仿学习中广为使用的 DAgger 离线再标注方法。而本文的在线再标注方法的特殊性在于使用了可穿戴设备进行数据采集, 并且通过遥操作的方式形成了一种“人在回路”的系统, 可以在遥操作的同时对数据集中偏差较大的数据点进行自动提取和再标注。而如果使用 DAgger 方法进行离线再标注, 即使采用两轮时间各为 2 min 的再标注数据采集过程, 也将产生约 48 000 个新的多维数据点, 依靠人工对这些数据进行离线再标注几乎不具可行性。

4.3 拟人化姿态约束有效性测试及用户评价

在遥操作控制中采用分级 CLIK 控制方法, 在 xMate3 机械臂模型上得到的拟人化姿态控制如图 7(a) 所示。图中用黄线标示出了操作者的手臂姿态, 黄色圆圈标示操作者手的位置; 绿线标示出了机械臂的姿态, 绿色圆圈标示出了机械臂末端的位置。在外骨骼主端提供了操作者手臂姿态信息的情况下, 从端机械臂始终保持与操作者近似的姿态。本文的分级 CLIK 控制方法可以在遥操作中实现从端对主端的姿态模仿特性。

为了验证本文的学习框架和方法训练得到的拟人化姿态约束动作生成策略是否能在脱离外骨骼主端设备提供的人体手臂姿态信息的情况下控制从端机械臂保持拟人化姿态运行, 实验中使用了 HTC Vive 的 VR (虚拟现实) 手柄来模拟产生实时的笛卡儿空间位姿指令, 从端机械臂执行对目标的跟随

任务，并通过已训练好的拟人化姿态约束策略产生约束动作。由于 HTC Vive 的 VR 手柄只能提供单点的笛卡儿空间位姿信息，在这样的实验设定下，从端控制器不会获取到任何关于操作者手臂姿态的信息，必须自主产生拟人化姿态约束。

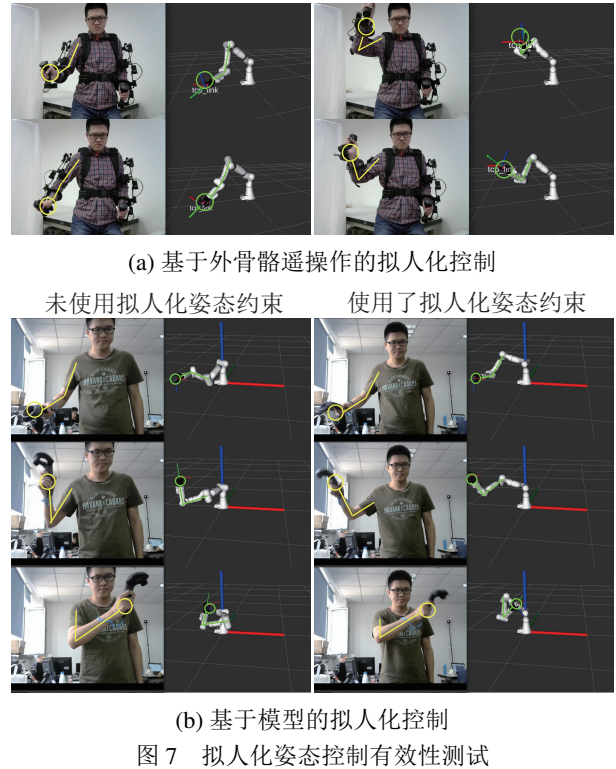


Fig.7 Effectiveness test of the human-like posture control

操作者手持 VR 手柄对仿真环境中的 xMate3 机械臂进行控制的结果如 7(b) 所示。标示方法与图 7(a) 相同。左侧的 3 组图片中，从端机械臂的控制未使用拟人化姿态约束策略进行约束，因此，机械臂仅跟踪了操作者手柄的位姿，姿态和操作者手臂完全不一样，并且机械臂的构形不可预测。而在右侧 3 组图片中，从端机械臂使用了已经过训练的拟人化姿态约束策略，可以明显看出，在保持追踪操作者手部位姿的同时，机械臂的姿态也与操作者手臂姿态保持了高度的相似性。

为了验证本文的机械臂拟人化控制方法是否能给予用户更高的可接受度和人机友好体验，参照文 [29] 中所使用的用户评价方法，邀请了 14 名不具备相关背景知识的志愿者（其中 10 名男性，4 名女性，年龄区间为 24~39 岁）对本文的机械臂拟人化控制方法进行评价。具体测试方法为：分别观看 30 段机械臂运动视频（其中包含使用拟人化姿态约束和未使用拟人化姿态约束控制的机械臂运动视频各 15 段，按随机顺序播放），请志愿者回答 3 个问题对每组视频的观看感受进行评价。问题包括：1)

哪一个视频中的机械臂看起来更像人类手臂；2) 哪一个视频的动作感觉更熟悉；3) 更愿意和哪一个视频中的机械臂在共享工作空间的情况下协作。评价统计结果如表 3 所示，体现了使用拟人化姿态约束的机械臂在感官上对用户更加友好，作为共同作业的工具更容易被非专业用户所接受。同时，评价结果中，使用拟人化控制的机械臂在 3 项测试中获得的优评率均超过 90%，表明本文方法在机械臂上产生的拟人化姿态能更加确定地给予用户较好的主观感受。

表 3 用户评价统计结果
Tab.3 Statistic results of user study

评价项目	使用拟人化控制	未使用拟人化控制
类人程度	207 (98.57%)	3 (1.43%)
熟悉程度	190 (90.48%)	20 (9.52%)
协作意愿	198 (94.29%)	12 (5.71%)

4.4 机械臂拟人化控制动态性能测试

实验中采用 4 段有代表性的轨迹，验证基于本文控制方法的冗余机械臂是否能在运动中保持拟人化的姿态。为了与人类姿态进行对比，以操作者穿戴外骨骼运动时录制的数据作为实验轨迹。录制的手部位姿轨迹作为轨迹指令，通过拟人化姿态控制方法在机械臂上重复该段轨迹指令。由于 xMate3 机械臂第 6 关节与人体肘部在形态上相近，因此，实验中选取机械臂的第 6 关节、末端执行器的位置作为关键点，用其轨迹与人体肘部、手部位姿轨迹进行对比。轨迹记录如图 8 所示。

在记录的全部轨迹中，机械臂末端执行器对手部轨迹的跟踪始终保持较高的精度。而机械臂的第 6 关节与肘部在空间轨迹上也始终保持较高的相似性。由于运动学结构上的差异，机械臂第 6 关节与人体肘部轨迹不可能完全一致，但相似的轨迹已经可以保证机械臂展现出拟人化的行为特征。实验结果也验证了本文所采用的分级 CLIK 控制策略的有效性，即末端执行器跟踪任务作为主要任务具有高优先级；而姿态约束任务作为次要任务，在保证主要任务精度的前提下尽可能提高从端机械臂的拟人化程度。

5 结论 (Conclusion)

针对人机交互任务中对机器人拟人化属性的需求，介绍了一种高度直觉化的方法赋予冗余机械臂拟人化姿态。目前，对于机器人的拟人化程度尚无公认的标准和精确模型，本文提出了直接通过人类示教来学习拟人化姿态特征的方法。

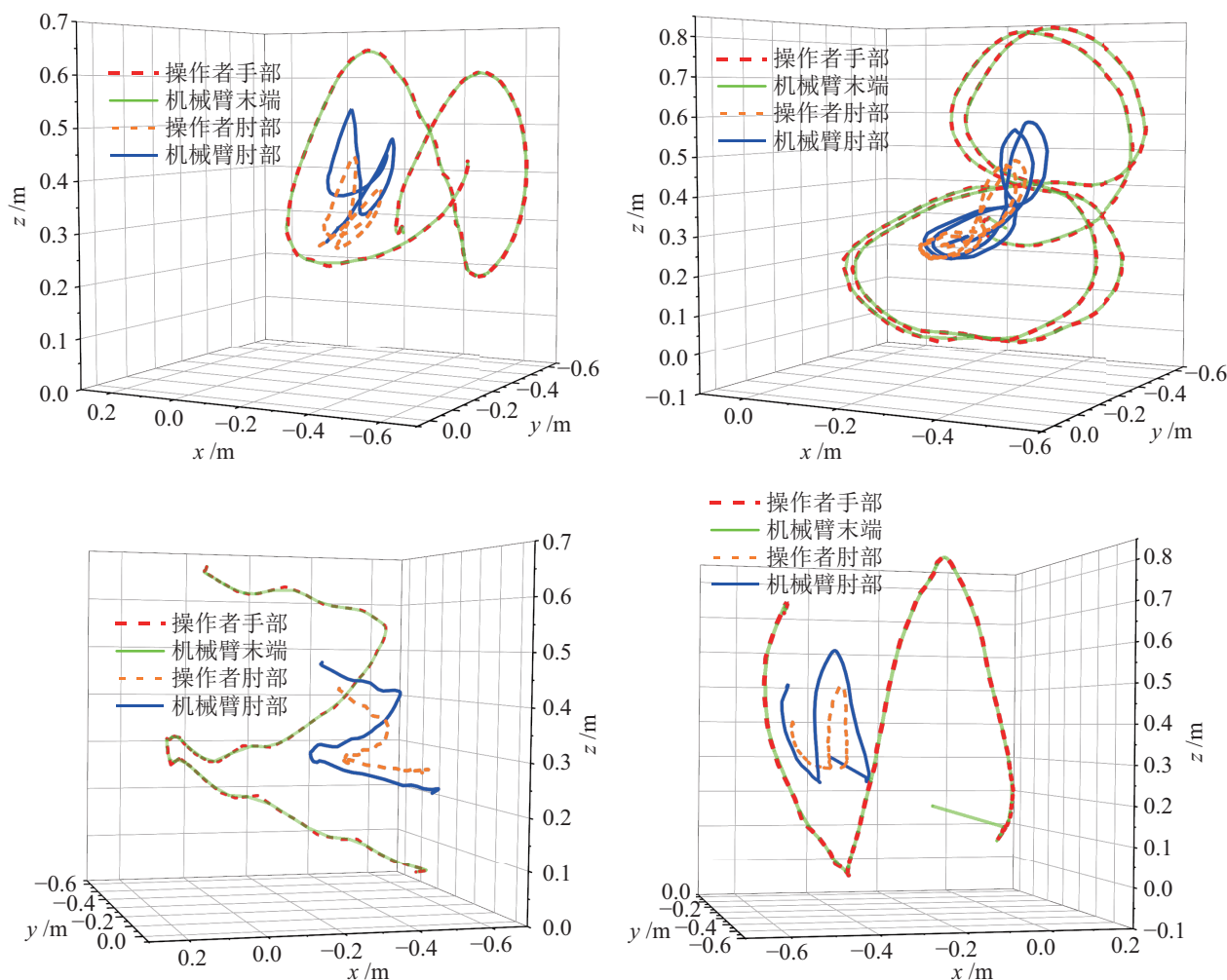


图 8 动态跟踪任务中关键点轨迹

Fig.8 Trajectories of key points in dynamic tracking tasks

基于可穿戴主端设备和分级 CLIK 控制方法, 建立了一种具有姿态模仿特性的主从遥操作系统, 并提供了便利直观的示教数据采集手段。通过基于行为克隆的学习方法, 从示教数据中提取拟人化姿态约束特征, 训练拟人化姿态约束生成网络。为了解决特征漂移的问题, 进一步利用遥操作系统进行“人在回路”的在线再标注, 对模型进行修正。一般情况下, 只需不超过 10 min 的示教就可以采集到足够的数据并成功训练拟人化姿态约束生成策略。

由于神经网络模型泛化误差的存在, 所生成的控制量仍然会有偶发错误, 这也是未在真实机械臂上开展本文实验的原因。在后续的工作中会在所生成的控制指令中增加滤波环节, 以避免危险动作的发生。

本文的生成模型可根据给定的人手位姿还原完整的人体形态, 从而估计操作者占用空间的情况。后续工作中将使用预测结果帮助机器人避免与人体在人机交互任务中发生碰撞, 提高操作安全性。

参考文献 (References)

- [1] Heunis C M, Barata B F, Furtado G P, et al. Collaborative surgical robots: Optical tracking during endovascular operations[J]. IEEE Robotics & Automation Magazine, 2020, 27(3): 29-44.
- [2] Su H, Sandoval J, Makhdoomi M, et al. Safety-enhanced human-robot interaction control of redundant robot for teleoperated minimally invasive surgery[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2018: 6611-6616.
- [3] Malik A A, Bilberg A. Collaborative robots in assembly: A practical approach for tasks distribution[J]. Procedia CIRP, 2019, 81: 665-670.
- [4] Maurtua I, Ibarburen A, Kildal J, et al. Human-robot collaboration in industrial applications: Safety, interaction and trust [J]. International Journal of Advanced Robotic Systems, 2017, 14(4). DOI: 10.1177/1729881417716010.
- [5] Kumar S, Savur C, Sahin F. Survey of human-robot collaboration in industrial settings: Awareness, intelligence, and compliance[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(1): 280-297.
- [6] Shao S, Zhou Q, Liu Z. Mental workload characteristics of manipulator teleoperators with different spatial cognitive abilities

- [J]. International Journal of Advanced Robotic Systems, 2019, 16(6). DOI: 10.1177/1729881419888042.
- [7] Coradeschi S, Ishiguro H, Asada M, et al. Human-inspired robots[J]. IEEE Intelligent Systems, 2006, 21(4): 74-85.
- [8] Lee J, Chai J, Reitsma P S A, et al. Interactive control of avatars animated with human motion data[J]. ACM Transactions on Graphics, 2002, 21(3): 491-500.
- [9] Lee J, Lee K H. Precomputing avatar behavior from human motion data[C]//ACM SIGGRAPH/Eurographics Symposium on Computer Animation. New York, USA: ACM, 2004: 79-87.
- [10] Spanlang B, Navarro X, Normand J M, et al. Real time whole body motion mapping for avatars and robots[C]//19th ACM Symposium on Virtual Reality Software and Technology. New York, USA: ACM, 2013: 175-178.
- [11] Alibeigi M, Rabiee S, Ahmadabadi M N. Inverse kinematics based human mimicking system using skeletal tracking technology[J]. Journal of Intelligent & Robotic Systems, 2017, 85(1): 27-45.
- [12] Pollard N S, Hodgins J K, Riley M J, et al. Adapting human motion for the control of a humanoid robot[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2002: 1390-1397.
- [13] Lee C H, Choi J, Lee H, et al. Exoskeletal master device for dual arm robot teaching[J]. Mechatronics, 2017, 43: 76-85.
- [14] Lee H, Kim J, Kim T. A robot teaching framework for a redundant dual arm manipulator with teleoperation from exoskeleton motion data[C]//IEEE-RAS International Conference on Humanoid Robots. Piscataway, USA: IEEE, 2014: 1057-1062.
- [15] Park S, Jung Y, Bae J. An interactive and intuitive control interface for a tele-operated robot (AVATAR) system[J]. Mechatronics, 2018, 55: 54-62.
- [16] Nunez M L, Dajles D, Siles F. Teleoperation of a humanoid robot using an optical motion capture system[C]//IEEE International Work Conference on Bioinspired Intelligence. Piscataway, USA: IEEE, 2018. DOI: 10.1109/IWOBI.2018.8464136.
- [17] Gonzalez G, Wachs J. Pose imitation constraints for collaborative robots[DB/OL]. (2020-10-12) [2022-11-01]. <https://arxiv.org/abs/2009.10947>.
- [18] Zanchettin A M, Rocco P, Bascetta L, et al. Kinematic analysis and synthesis of the human arm motion during a manipulation task[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2011: 2692-2697.
- [19] Zanchettin A M, Bascetta L, Rocco P. Acceptability of robotic manipulators in shared working environments through human-like redundancy resolution[J]. Applied Ergonomics, 2013, 44(6): 982-989.
- [20] Su H, Qi W, Yang C G, et al. Deep neural network approach in human-like redundancy optimization for anthropomorphic manipulators[J]. IEEE Access, 2019, 7: 124207-124216.
- [21] Su H, Qi W, Hu Y B, et al. An incremental learning framework for human-like redundancy optimization of anthropomorphic manipulators[J]. IEEE Transactions on Industrial Informatics, 2022, 18(3): 1864-1872.
- [22] Giusti A, Guzzi J, Cireşan D C, et al. A machine learning approach to visual perception of forest trails for mobile robots[J]. IEEE Robotics and Automation Letters, 2016, 1(2): 661-667.
- [23] Bojarski M, Del Testa D, Dworakowski D, et al. End to end learning for self-driving cars[DB/OL]. (2016-04-25) [2021-03-03]. <https://arxiv.org/abs/1604.07316>.
- [24] Ross S, Melik-Barkhudarov N, Shankar K S, et al. Learning monocular reactive UAV control in cluttered natural environments[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2013: 1765-1772.
- [25] Zhao L, Yang T, Yu P, et al. An exoskeleton-based master device for dual-arm robot teleoperation[C]//Chinese Automation Congress. Piscataway, USA: IEEE, 2020: 5316-5319.
- [26] Wang J G, Li Y M, Zhao X H. Inverse kinematics and control of a 7-DOF redundant manipulator based on the closed-loop algorithm[J]. International Journal of Advanced Robotic Systems, 2010, 7(4). DOI: 10.5772/10495.
- [27] Corke P. Robotics, vision and control: Fundamental algorithms in Matlab[M]. 2nd ed. Berlin, Germany: Springer-Verlag, 2011: 63.
- [28] Laskey M, Staszak S, Hsieh W Y S, et al. SHIV: Reducing supervisor burden in DAGger using support vectors for efficient learning from demonstrations in high dimensional state spaces [C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2016: 462-469.
- [29] Su H, Enayati N, Vantadori L, et al. Online human-like redundancy optimization for tele-operated anthropomorphic manipulators[J]. International Journal of Advanced Robotic Systems, 2018, 15(6). DOI: 10.1177/1729881418814695.

作者简介:

赵 亮 (1986-), 男, 博士生, 副研究员。研究领域: 机器人遥操作, 人机交互, 机器人智能控制。

杨 铁 (1990-), 男, 硕士, 助理研究员。研究领域: 机器人机构设计, 外骨骼, 精密测量系统。

于 鹏 (1981-), 男, 硕士, 研究员。研究领域: 外骨骼机器人, 微纳机器人。