

DOI: 10.19650/j.cnki.cjsi.J2614887

基于视触融合及刚度自适应的机械臂抓取方法*

李佳钰^{1,2}, 王宏斌¹, 陈新宇¹, 杨怀广³, 尤 波^{1,2}

(1. 哈尔滨理工大学黑龙江省复杂智能系统与集成重点实验室 哈尔滨 150080; 2. 哈尔滨理工大学先进制造智能化技术教育部重点实验室 哈尔滨 150080; 3. 哈尔滨工业大学机电工程学院 哈尔滨 150080)

摘 要:针对非结构化环境中单一视觉几何定位残余误差与物体物理属性未知相耦合,且现有视触系统缺乏底层动态闭环补偿导致接触力失控的问题,提出一种基于视觉先验引导与触觉反馈修正的视触协同机械臂抓取方法。首先,以 ResUNet 替换 GraspNet 主干网络并引入基于抓取点与物体重心欧氏距离的重心约束机制,增强点云特征提取完整性,提高抓取位姿检测的物理稳定性与精确率;其次,利用触觉数据集训练卷积神经网络(CNN)-Transformer 刚度估计器,并将在线辨识出的软、中、硬 3 类刚度等级,依据分类概率与阻抗系数映射的导纳参数转换规则输入控制器,实现抓取接触力动态调整,增强交互柔顺性,降低物品损坏风险;最后,在深度确定性策略梯度(DDPG)基础上融合快速随机搜索树算法(RRT)与事后经验回放机制(HER),构建融合路径导向与经验重构的 RHER-DDPG 算法(RHER-DDPG),通过 RRT 专家轨迹缩小初期动作探索空间,结合 HER 将失败目标重新标记为成功样本,共同加速了 Actor 网络高维状态到动作映射的收敛过程,生成高效抓取策略。对比实验表明:相较于原始 GraspNet 网络,抓取检测精确率提升 12.8%;相较于传统 DDPG 算法,模型收敛迭代次数减少约 30%,并在真实场景下机械臂抓取测试成功率达 90.3%,物品损坏率仅为 7.1%,验证了所提方法的有效性。

关键词: 柔顺抓取;视-触觉融合;刚度估计;CNN-Transformer

中图分类号: TP24 TH29 **文献标识码:** A **国家标准学科分类代码:** 510.80

Mechanical arm grasping method based on vision-tactile fusion and stiffness self-adaptation

Li Jiayu^{1,2}, Wang Hongbin¹, Chen Xinyu¹, Yang Huaiguang³, You Bo^{1,2}

(1. Heilongjiang Provincial Key Laboratory of Complex Intelligent System and Integration, Harbin University of Science and Technology, Harbin 150080, China; 2. Key Laboratory of Advanced Manufacturing Intelligent Technology, Ministry of Education, Harbin University of Science and Technology, Harbin 150080, China; 3. School of Mechatronics Engineering, Harbin Institute of Technology, Harbin 150080, China)

Abstract: To address the coupling of residual errors of single-visual geometric localization are coupled with the unknown physical properties of objects in unstructured environments, as well as the lack of low-level dynamic closed-loop compensation in existing visuo-tactile systems leading to uncontrolled contact forces, a visuo-tactile collaborative robotic grasping method based on visual prior guidance and tactile feedback correction is proposed. First, the GraspNet backbone network is replaced by ResUNet, and a center-of-mass constraint mechanism based on the Euclidean distance between the grasp point and the object's centroid is introduced. This enhances the completeness of point cloud feature extraction, thereby improving the physical stability and accuracy of grasp pose detection. Second, a convolutional neural network (CNN)-Transformer stiffness estimator is trained using a tactile dataset. The online-identified soft, medium, and hard stiffness levels are input into the controller according to the admittance parameter conversion rules mapped by classification probabilities and impedance coefficients. This realizes the dynamic adjustment of grasping contact force, enhances interaction compliance, and reduces the risk of object damage. Finally, based on the deep deterministic policy gradient (DDPG), the rapidly-exploring random tree (RRT) algorithm and hindsight experience replay (HER) mechanism are integrated to construct the RRT-

收稿日期: 2026-01-22 Received Date: 2026-01-22

* 基金项目: 黑龙江省自然科学基金(PL2025E050)、黑龙江省省属本科高校优秀青年教师基础研究支持计划(YQJH2025073)项目资助

guided hindsight experience replay deep deterministic policy gradient (RHER-DDPG) algorithm, which fuses path guidance and experience reconstruction. By utilizing RRT expert trajectories to narrow the initial action exploration space and combining HER to relabel failed targets as successful samples, the convergence process of the Actor network mapping from high-dimensional states to actions is jointly accelerated, generating efficient grasping strategies. Comparative experiments show that, compared with the original GraspNet, the grasp detection accuracy is improved by 12.8%. Compared with the traditional DDPG algorithm, the model convergence iterations are reduced by approximately 30%. In real-world robotic grasping tests, the success rate reaches 90.3% and the object damage rate is only 7.1%, verifying the effectiveness of the proposed method.

Keywords: flexible grasping; vision-haptic fusion; stiffness estimation; CNN-Transformer

0 引言

随着人工智能与控制技术的发展,机械臂已在航空航天、医疗等领域得到了广泛应用^[1]。柔顺抓取控制是解决接触力失控的关键技术^[2-4],可有效应对物理属性未知与定位误差导致的抓取失效问题^[5-6]。因此,设计能够融合视-触多模态感知信息、自适应物体刚度变化且具备高效规划能力的抓取方法,对于提升机械臂在复杂动态环境下的作业性能具有重要意义。

目前,机械臂抓取技术主要基于视觉感知与力位协同控制算法^[7-9]。在视觉检测方面,基于深度学习的抓取位姿估计已成为主导方法^[10-11],代表性方法包括 GraspNet^[12]、PointNet++^[13] 及各类基于卷积神经网络的检测模型^[14-15]。GraspNet 虽然在大规模数据集上表现优异,但其原始架构采用 PointNet++ 作为主干网络,该网络在建模物体局部几何细节时存在感受野有限、细粒度特征表达能力不足等问题。针对该问题,Chen 等^[16]指出,现有基于点云的深度学习方法通常难以建模数据内部的高阶非线性拓扑关系,导致在复杂背景下难以充分提取具有判别性的局部几何特征;王立辉等^[17]进一步研究发现,在存在环境干扰的实际应用中,仅依赖基础点云特征难以保证位姿估计的鲁棒性与准确性。这些工作共同揭示了单一视觉模态在特征提取完整性与复杂场景适应能力方面的局限性。为此,融合视觉与触觉可有效提升抓取系统在复杂环境中的适应能力^[18-20]。

在运动控制与规划方面,传统阻抗控制和导纳控制^[21-23]通常需预先设定环境刚度参数,无法适应未知刚度的易碎物体。为克服模型参数对环境变化的敏感性,基于数据驱动的强化学习方法被广泛应用于接触力与位置的协同调节^[24]。其中,深度确定性策略梯度算法(deep deterministic policy gradient, DDPG)虽被视为解决此类连续动作空间控制问题的有效手段,但 Ren 等^[25]在最新研究中指出,该算法在处理高维复杂的接触任务时,仍具有奖励稀疏与收敛速度慢的问题。解决这一问题,通常引入事后经验回放机制(hindsight experience replay,

HER)进行改进,然而 Sanchez 等^[26]进一步分析发现,HER 机制在初期探索阶段依然存在盲目性,导致无效采样过多,无需引入外部引导机制以优化策略更新。作为一种潜在的引导手段,快速随机搜索树算法(rapidly-exploring random tree, RRT)虽具备概率完备性,但 Xun 等^[27]指出其在非结构化狭窄环境下的搜索效率通常难以满足实时性要求;同时,Zhang 等^[28]研究强调,基于 RRT 的几何采样路径在动态重规划中易导致动作延迟,难以满足精密抓取对末端稳定跟踪的要求。

综上所述,针对现有多模态抓取策略多将视觉检测与力觉控制独立设计或仅进行简单拼接,难以应对非结构化环境中视觉几何残余误差与未知物理属性的耦合干扰,且缺乏系统级底层动力学闭环补偿的问题,本文提出了一种基于视觉先验引导与触觉闭环补偿的视触协同机械臂抓取方法。本文的主要内容为:

1) 基于物体重心约束机制的 ResUNet 改进方法。此处重心约束机制界定为:构建一个与物体重心距离成反比的稳定得分函数,以惩罚几何可行但物理容易失衡的抓取点。引入 ResUNet 增强特征提取能力,从几何特征与物理约束双重维度提高了抓取位姿的预测精度。

2) 设计了卷积神经网络(convolutional neural network, CNN)结合 Transformer 架构的刚度在线辨识与多级导纳控制策略。本文中,刚度等级被定义为基于触觉时序信号分类出的软、中、硬 3 类力学响应特征;导纳参数转换规则界定为将上述分类概率动态转化为质量-弹簧-阻尼系统中具体刚度系数与阻尼系数的数学映射逻辑,从而在物理层面吸收并修正视觉几何误差,实现柔顺无损抓取。

3) 提出了融合路径导向与经验重构的 RHER-DDPG 算法(RRT-guided hindsight experience replay deep deterministic policy gradient, RHER-DDPG)。其具体作用机制为:在训练初期利用 RRT 生成的无碰撞专家轨迹作为高优样本引导智能体跨越盲目探索陷阱起到路径导向作用,同时利用 HER 机制通过目标重置手段将无效记录转化为有效学习信号起到经验重构作用,二者协同提高了高维连续动作空间下的映射求解效率。

1 基于视触觉融合感知与特征提取方法

1.1 视触协同控制机制

本文构建的视触协同抓取系统包含3个在时间维度上严格递进的交互阶段,形成了从视觉开环引导到触觉闭环修正的完整协作回路。第1阶段为视觉引导期,抓取启动前由深度相机获取场景点云,改进的GraspNet结合几何与重心信息输出目标期望位姿,此阶段作为前馈开环控制为机械臂提供宏观运动目标。第2阶段为轨迹规划期,RHER-DDPG决策模块接收目标期望位姿并输出关节角度增量指令,其中RRT算法负责生成避障的专家轨迹以引导初期探索,DDPG负责末端从预抓取点向接触点的精细趋近操作。第3阶段为触觉修正与闭环期,夹爪末端接触物体瞬间触发触觉感知,CNN-Transformer网络在线辨识物体刚度等级并将其下发至导纳控制器;当视觉定位存在残余偏差导致过早或异常接触时,导纳控制器通过监测实际接触力与期望抓取力的差值,根据动力学方程生成位置修正量,实时修改并下发新的位置指令。该时序机制实现了感知、决策与控制的深度融合,确保系统在存在视觉残余误差的环境下仍能维持安全、稳定的柔顺接触。

1.2 改进GraspNet的抓取位姿检测

针对原始GraspNet由于缺乏高阶特征提取能力与物理约束,导致抓取点易发生偏心侧滑的问题,本文从特征

提取与物理约束两方面进行改进:

1) 引入ResUNet特征提取,采用ResUNet替换原主干网络PointNet++。借助其编码器-解码器架构与跨层残差融合机制,有效提升了点云高阶拓扑信息的丢失与复杂背景下局部几何特征的提取完整性与位姿预测精度问题。

2) 构建物体重心约束机制。现有方法生成的候选点虽满足几何力闭合,但实际操作中常因受力不均导致滑脱。为从物理力学层面滤除此类易引发偏心翻转的边缘抓取点,本文在抓取得分阶段引入重心约束机制。其稳定得分函数基于物体重心与抓取点的欧氏距离构建,距离越小得分越高,以此实现对候选抓取点的物理稳定性惩罚与评估。其稳定得分 s_{stable} 如式(1)所示。

$$s_{\text{stable}} = 1 - \frac{d_{\text{COC}}}{d_{\text{max}}} \quad (1)$$

式中: d_{COC} 表示物体重心到抓取点的距离; d_{max} 为所有抓取点与物体重心距离中的最大值。稳定得分通过归一化计算,值域为 $[0, 1]$,与物体重心距离成反比,即距离越小,得分越高。

最终评估分数由原始力闭合得分与稳定得分的乘积构成。该机制从物理力学层面滤除了几何可行但易引发偏心翻转的边缘抓取点,最终抓取得分如式(2)所示。

$$s_{\text{final}} = s_{\text{grasp}} \times s_{\text{stable}} \quad (2)$$

改进后的GraspNet网络结构如图1所示。

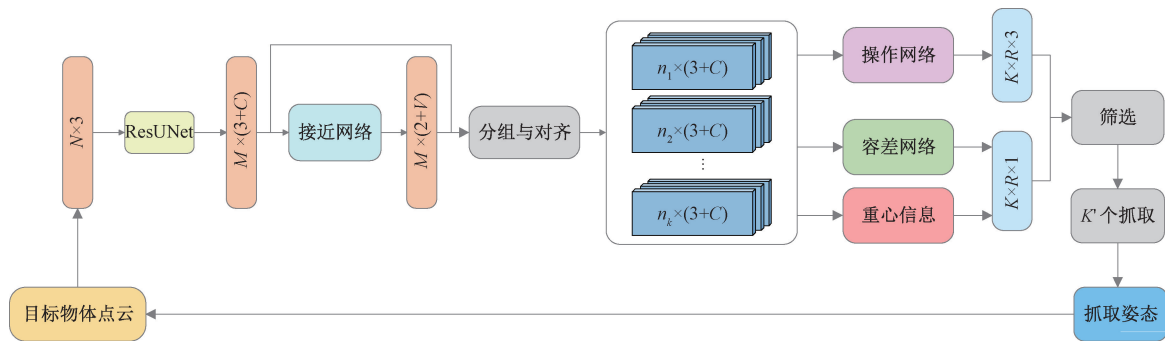


图1 改进GraspNet网络结构

Fig. 1 Structure of the improved GraspNet network

1.3 基于CNN-Transformer的刚度在线辨识方法

在抓取易碎物体时,仅依赖视觉的位姿估计难以适应接触过程中的物理柔顺性变化,即使抓取位姿准确,仍可能因接触力失控导致物体损坏。为实现安全交互,本文引入触觉感知,并提出一种CNN-Transformer混合架构用于物体刚度辨识。该模型利用CNN提取触觉信号的局部力响应特征,同时通过Transformer捕捉力-形变响应的长时序动态,表明物体的动态力学特性。CNN-Transformer结构如图2所示。

为了明确模型训练的数据基础并保证算法的可重复性,本文构建了机械臂柔顺抓取触觉数据集。该数据集涵盖了非结构化场景下典型的3类刚度物品,其具体的硬件配置、信号采集标准、样本规模及网络输入维度等详细信息如表1所示。

基于上述构建的触觉数据集与如图2所示的网络结构,模型具体运行流程为:1)输入层接收批次大小为32、包含力与时间特征双维度且单次时间窗口长度为60点的触觉数据;2)数据调整维度后输入CNN模块,经两层

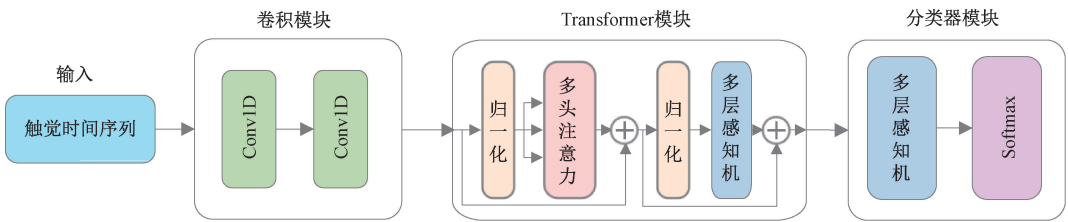


图 2 CNN-Transformer 结构

Fig. 2 Architecture of the CNN-Transformer model

表 1 机械臂柔顺抓取触觉数据集参数详情

Table 1 Details of tactile dataset parameters for compliant grasping

数据集特征维度	详细说明
硬件与传感器	Robotiq 2F-85 二指夹爪,贴附接触式力传感器(量程 0~50 N,分辨率 0.01 N)
信号采集配置	采样频率 1 000 Hz,内置低通滤波处理
数据集样本规模	采集 450 组双指触觉时间序列,形成 900 条独立力学曲线;划分 600 条为训练集,300 条为验证集
网络输入维度	Tensor 大小设定为[32,2,60](批次大小 32,包含力和时间特征 2 维,单次时间窗口长度 60 点)
物体类型及划分	选取 12 种典型物品,由 4 名测试者独立盲测后划分为 3 类:1)软质/易碎品(Soft):牛奶盒、药盒、抽纸、毛巾卷; 2)中等硬度(Mid):矿泉水瓶、洗面奶瓶、杂物筐、包装纸盒;3)硬质(Hard):陶瓷杯、铁盒、牙签盒、玻璃水晶球

一维卷积等操作提取力的局部特征,随后将输出再次调整维度并输入两层 Transformer 编码器以捕捉长时序关联;3)分类头经过全局平均池化与全连接层,通过 Softmax 函数输出对应软、中、硬 3 个类别的概率分布,该分类结果将直接用于驱动导纳控制器进行动力学参数的动态切换,以实现与物体物理柔顺特性的精准匹配。

2 基于 RHER-DDPG 的柔顺抓取控制策略

2.1 多导纳柔顺控制模型

针对机械臂与不同刚度物体的安全柔顺交互,本文提出一种分阶段的多级导纳控制策略。该策略首先通过慢速闭合获取触觉响应序列,其在 1 000 Hz 传感器采样率下采集 60 个时间点所带来的 0.06 s 感知延迟,系统将此阶段作为安全缓冲期。通过约束夹爪以 0.05 rad/s 等极低初速度闭合,将延迟盲区内的瞬态接触力严格限制在 2 N 以内,有效规避了软质物品的破损风险。软质易碎品瞬态接触力学响应如图 3 所示。

完成刚度在线辨识后,系统据此动态切换导纳控制参数,最终实现与物体物理特性的自适应匹配。在该策略中,导纳参数转换规则是指将刚度等级映射为多级导纳控制器动态参数的逻辑。导纳控制通过构建质量-弹簧-阻尼系统来描述夹爪与外部环境的接触。忽略惯性参数后,设定期望抓取力为 F_0 ,实际接触力为 F_e ,视觉引导的初始期望轨迹(即刚接触时的位置)为 x_0 ,导纳控制器修正后的目标位置指令为 x_d ,其底层误差补偿控制律

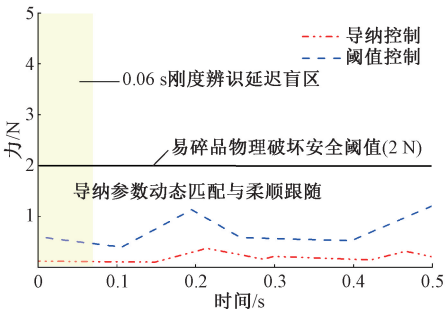


图 3 软质易碎品接触瞬态力学响应曲线

Fig. 3 Transient force response curve during contact

定义如式(3)所示。

$$D_d(\dot{x}_d - \dot{x}_0) + K_d(x_d - x_0) = F_e - F_0 \tag{3}$$

式中:阻尼参数与刚度参数均由 CNN-Transformer 输出的刚度等级依据转换规则动态赋予。当视觉定位存在过切误差导致实际接触力 F_e 大于期望抓取力 F_0 时,该控制律会自动生成反向的位置修正量 $x_d - x_0$,实时修改下发给机械臂的位置指令,从而实现对视觉误差的物理吸收与柔顺补偿。其多级导纳闭环控制系统框图如图 4 所示。



图 4 导纳闭环控制系统

Fig. 4 Admittance closed-loop control system

2.2 RHER-DDPG 路径规划算法

在实际抓取任务中,部分物体的刚度呈连续分布或处于类别边界。为避免分类器硬判决输出导致导纳参数 D_d 和 K_d 频繁跳变并引发机械臂末端震荡,本文进一步设计了基于模糊概率的参数平滑过渡机制。CNN-Transformer 网络的分类层输出结果本质为对应软、中、硬 3 个维度的概率分布 P_{soft} 、 P_{mid} 与 P_{hard} ,且满足 $P_{\text{soft}} + P_{\text{mid}} + P_{\text{hard}} = 1$ 。因此,系统将离散的控制参数连续化,实际下发的动态导纳参数通过概率加权计算获得如式(4)所示。

$$\begin{cases} K_{\text{actual}} = P_{\text{soft}} K_{\text{soft}} + P_{\text{mid}} K_{\text{mid}} + P_{\text{hard}} K_{\text{hard}} \\ D_{\text{actual}} = P_{\text{soft}} D_{\text{soft}} + P_{\text{mid}} D_{\text{mid}} + P_{\text{hard}} D_{\text{hard}} \end{cases} \quad (4)$$

式中: K_{actual} 和 D_{actual} 分别为平滑过渡后实际下发的动态刚度与阻尼参数; K_{soft} 、 K_{mid} 、 K_{hard} 与 D_{soft} 、 D_{mid} 、 D_{hard} 分别为系统预设的软、中、硬 3 个等级对应的基准刚度与阻尼参数。通过引入该平滑过渡机制,不仅保证了典型物体参数的精确映射,也确保了处于刚度边界的物体能够获得平滑连续的导纳阻抗特性,有效吸收了分类抖动,提升了系统的控制鲁棒性。

为提升机械臂柔顺抓取的自主路径规划能力,本文采用 RHER-DDPG 算法在 UR3e 机械臂的关节空间中直接生成运动策略。控制变量选取第 1~6 关节的角位移, RHER-DDPG 算法的状态输入量是当前机械臂的 6 个关节的角度和目标抓取位姿,如式(5)所示。

$$I = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, x, y, z, \text{roll}, \text{pitch}, \text{yaw}\} \quad (5)$$

式中: $\theta_1 \sim \theta_6$ 为当前状态机械臂的 6 个关节角度。

选择机械臂各关节角度改变量为 RHER-DDPG 算法输出状态量,如式(6)所示。

$$O = \{o_1, o_2, o_3, o_4, o_5, o_6\} \quad (6)$$

式中: $o_1 \sim o_6$ 为 RHER-DDPG 算法输出的机械臂的 6 个关节角度改变量。

面向高维状态空间下机械臂抓取规划存在的收敛慢、探索效率低等问题,本文提出一种融合路径导向与经验重构的 RHER-DDPG 算法,如图 5 所示。在其具体作用机制中,路径导向利用 RRT 专家轨迹引导初期探索以避免高维空间无效采样,同时经验重构通过 HER 机制将失败轨迹重置为到达虚拟目标的成功样本来克服奖励稀疏问题。该算法在传统 DDPG 基础上引入 HER 机制,并利用 RRT 规划器在训练初期生成专家路径引导智能体学习。Actor 与 Critic 网络均采用多层全连接架构配合修正线性单元(rectified linear unit, ReLU)激活函数,提升建模状态与动作间的非线性关系。

如图 5 所示, RHER-DDPG 算法在原始 DDPG 的 Actor-Critic 架构基础上,引入了 RRT 与 HER 重构机制。主要改进策略为:

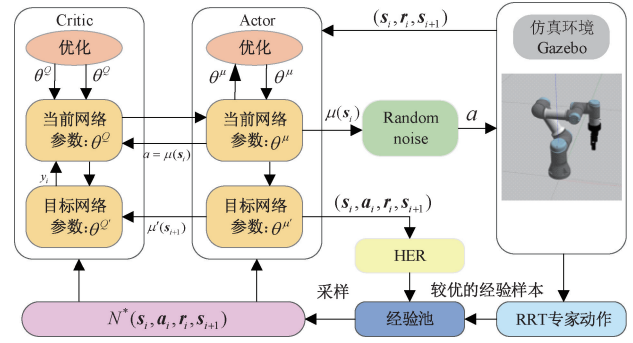


图 5 RHER-DDPG 算法流程

Fig. 5 Flowchart of the RHER-DDPG algorithm

1) RRT 引导的初始探索策略

训练初期,DDPG 因机械臂探索空间大、训练无序而效率低。为加快其初期学习速度,用 RRT 算法生成路径作示例引导。先以 RRT 算法生成 UR3e 六自由度协作机械臂末端到目标抓取位置的可行路径,再经逆运动学解算出各关节角度改变量即 RRT 专家动作如式(7)所示。

$$a_{\text{RRT}} = (\Delta\theta_1, \Delta\theta_2, \Delta\theta_3, \Delta\theta_4, \Delta\theta_5, \Delta\theta_6) \quad (7)$$

式中: $\Delta\theta_1 \sim \Delta\theta_6$ 为经过逆运动学解算得出的机械臂 6 个关节的角度改变量。

接下来在当前状态 s_t 下让机械臂执行每一个 RRT 专家动作 a_{RRT} ,获得奖励 r_t ,更新状态 s_{t+1} 。最后,将得到的状态转移元组 $(s_t, a_{\text{RRT}}, r_t, s_{t+1})$ 存入经验回放池 R 中,如式(8)所示。

$$R \leftarrow R \cup \{(s_t, a_{\text{RRT}}, r_t, s_{t+1}) | T_{\text{RRT}}\} \quad (8)$$

式中: T_{RRT} 为 RRT 生成的可行路径(专家路径),本文中 DDPG 算法的输出动作 a_t 和 RRT 专家动作均为机械臂各个关节角度改变量。

2) 奖励函数设计

分析了训练初期盲目探索导致收敛慢的问题,引入 RRT 专家动作作为先验引导,构建包含目标接近度与路径引导项的复合奖励函数。该函数综合考量位置与姿态误差,在保证目标导向性的同时提升路径可行性,当末端误差满足阈值条件时判定成功并给予奖励。整合后的奖励函数 r 定义如式(9)所示。

$$r = -(\lambda_1 d_p^2 + \lambda_2 d_o^2) + \lambda_3 e^{-\eta \|a - a_{\text{RRT}}\|_2} + \xi \cdot \cap_{\text{succ}} \quad (9)$$

式中:第 1 项为位姿接近度惩罚, λ_1, λ_2 为位置与姿态的权重系数, $d_p = \|\mathbf{P}_{\text{ee}} - \mathbf{P}_{\text{target}}\|_2$, $d_o = \|\Phi_{\text{ee}} - \Phi_{\text{target}}\|_2$ 分别表示末端与目标的欧氏距离;第 2 项为 RRT 引导奖励, λ_3 为权重系数, η 为指数衰减率, $\|a - a_{\text{RRT}}\|_2$ 表示当前动作与专家动作的差异,该项随训练进行逐渐衰减以减少对先验知识的依赖;第 3 项为成功奖励, $\cap_{\text{succ}}$ 为指示函数(满足阈值条件时为 1,否则为 0), ξ 为到达奖励常数。

3) HER 算法集成

HER 通过重定义经验池中的失败样本,将其转化为

面向新目标的成功经验,从而提升学习效率与收敛速度。对于每一个样本 $(s_t, \mathbf{a}_t, \mathbf{r}_t, s_{t+1})$, 引入一个新的目标状态 $\mathbf{g}'$, 并计算新的奖励, 如式 (10) 所示。

$$\mathbf{r}_{\text{HER}} = \mathbf{r}(s_t, \mathbf{a}_t, \mathbf{g}')$$

(10)

这些重新定义的经验样本是一个五元组 $(s_t, \mathbf{a}_t, \mathbf{r}_{\text{HER}}, s_{t+1}, \mathbf{g}')$, 该五元组也被加入经验回放池, 如式 (11) 所示。

$$R \leftarrow R \cup \{(s_t, \mathbf{a}_t, \mathbf{r}_{\text{HER}}, s_{t+1}, \mathbf{g}')\}$$

(11)

4) 混合更新策略及策略校正

针对探索广度与样本效率的平衡问题, 构建了混合经验回放机制。设定经验回放池总容量为 N , 其中 RRT 专家数据占比 ρ_1 , HER 重构造数据占比 ρ_2 , 其余为 DDPG 自探索数据。训练时按比例 $(\rho_1:\rho_2:1-\rho_1-\rho_2)$: 采样以

平衡路径引导与稀疏奖励下的学习效率。此外, 利用 RRT 规划路径对策略网络进行辅助监督更新, 定义策略校正规则如式 (12) 所示。

$$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} - \lambda \nabla_{\boldsymbol{\theta}} \|Q(s_t, \mathbf{a}_t | \boldsymbol{\theta}) - Q(s_t, \mathbf{a}_{\text{RRT}} | \boldsymbol{\theta})\|$$

(12)

式中: $\boldsymbol{\theta}$ 为策略网络的参数; λ 为策略校正更新梯度学习率。

3 实验结果与分析

3.1 实验平台与参数设置

算法的有效性通过一套集成物理实验平台与仿真的验证系统进行评估, 如图 6 所示。

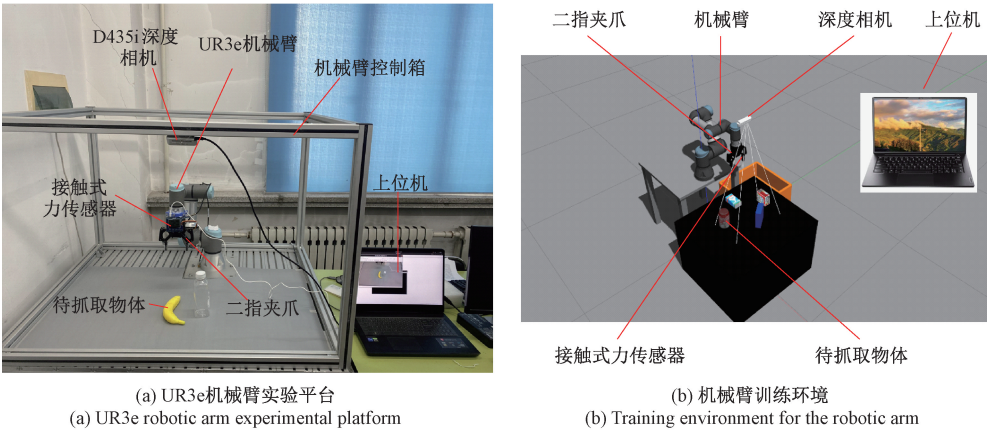


图 6 实验平台搭建
Fig. 6 Experimental platform

其中, 硬件系统包含 UR3e 协作机械臂、Robotiq 2F-85 自适应夹爪、Intel RealSense D435i 深度相机及指尖接触力传感器; 软件系统基于 Ubuntu 20.04 与 ROS Noetic 构建, 并采用 MoveIt 进行运动规划。此外, 构建了与物理样机几何结构与运动学参数严格一致的 Gazebo 仿真环境, 用于模型训练与算法验证。实验的可复现性及所提 RHER-DDPG 算法的性能验证基于 PyTorch 深度学习框架实现。训练过程中采用 Adam 优化器更新网络参数, 具体的网络超参数与训练环境配置如表 2 所示。

3.2 算法性能对比与验证

1) 视觉抓取检测性能

改进 GraspNet 模型训练使用标准数据集 GraspNet-1Billion, 其包含多样物体及详细抓取位姿标注。训练时, 模型用 Adam 优化器更新参数, 初始学习率设为 0.001, 结合余弦退火策略动态调整, 使模型初期快速收敛、后期精细优化。评估采用多指标, 包括抓取点识别的精确率、召回率, 接近向量方向性误差, 以及衡量实际抓取能力的抓取成功率。改进 GraspNet 网络的训练结果如图 7 所示。

表 2 训练参数表
Table 2 Training parameters

参数	数值
Actor 网络学习率	10^{-4}
Critic 网络学习率	10^{-3}
批处理大小	64
经验回放池大小	10^6
折扣因子	0.99
软更新参数	0.005
噪声幅度	0.1
RRT 专家经验占比	0.2
HER 样本占比	0.2
额外奖励影响因子	0.5
额外奖励指数衰减速率	0.05
网络更新间隔	100
开始训练最小样本数	1 000
RRT 路径规划长度限制	200
策略校正更新梯度学习率	0.01

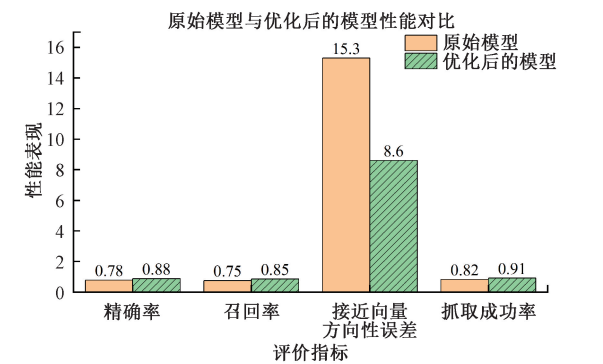


图 7 改进 GraspNet 网络训练结果

Fig. 7 Training results of the improved GraspNet network

通过实验数据对比可以清晰地看出,优化后的模型在各项性能指标上均优于原始模型,具体表现为抓取点分类的精确率和召回率提升,接近向量方向误差有效降低,抓取成功率明显提高。

2) 刚度辨识与力控性能

在触觉感知实验中,抓取软、中硬、硬物体形成 3 类可分时序特征。二指夹爪对每类硬度各抓 150 次,得 450 组双指序列、900 条曲线,标注后前 600 条作训练集,后 300 条作验证集。

其中,CNN-Transformer 模型的软硬度识别性能通过夹爪触觉数据集进行验证。该模型基于 PyTorch 实现,采用 Adam 优化器,共训练 50 轮。如图 8 所示,学习曲线显示训练初期损失迅速下降并趋于稳定,训练与验证损失同步收敛且差距较小,未出现过拟合现象。同时,训练集与验证集的准确率均稳步提升,表明该模型能够有效建模触觉时序特征,具备良好的泛化能力。

夹爪多级导纳控制器与阈值控制在抓取不同软硬度

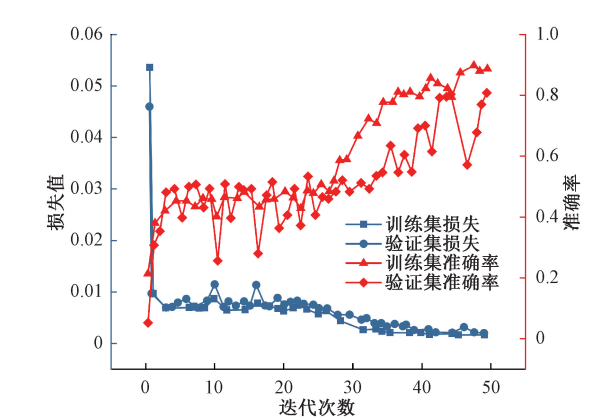


图 8 CNN-Transformer 模型训练损失与准确率变化曲线

Fig. 8 Learning curve of the CNN-Transformer network model

物体时的性能通过对比实验进行评估。实验中,导纳控制需配置阻尼等参数,阈值控制则设定初始闭合速度与期望抓取力;两种策略均在达到期望抓取力后停止闭合并进入稳定抓取阶段,具体参数如表 3 所示。抓取过程中抓取力与夹爪关节位移的变化曲线如图 9 所示,用于直观反映两种控制方法的动态响应特性。

表 3 实验参数表

刚度等级	阈值控制	导纳控制		期望抓取力/N
	初速度/ (rad·s ⁻¹)	阻尼/ (N·s·rad ⁻¹)	刚度/N	
软	0.05	80	150	3
中	0.03	150	350	5
硬	0.01	80	800	8

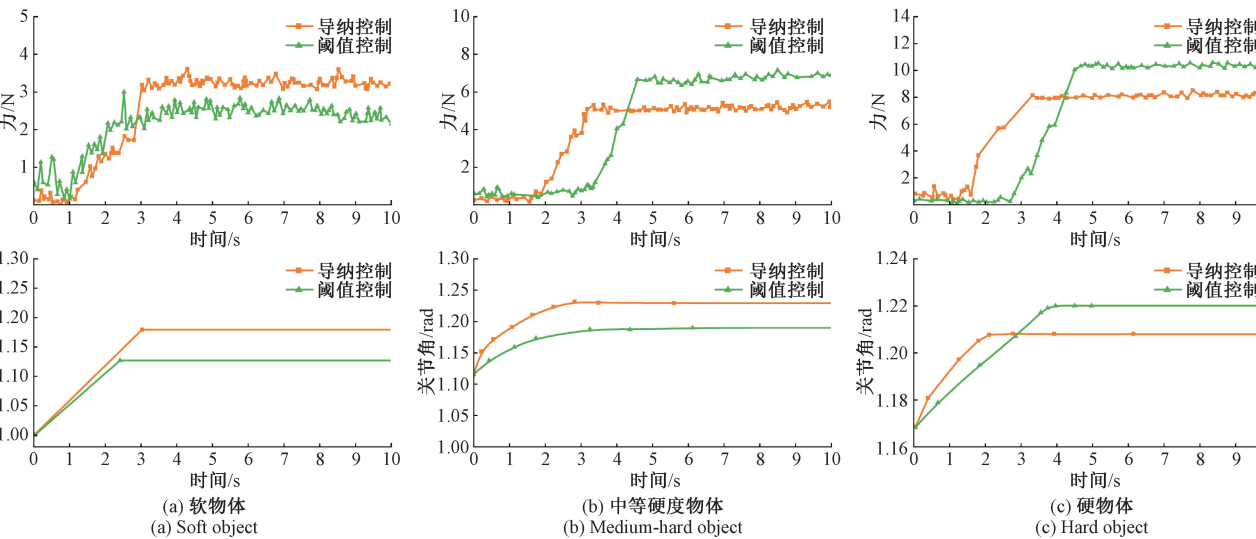


图 9 不同刚度物体抓取实验力与位移变化曲线

Fig. 9 Force and displacement curves during hard object grasping

由图 9 可知,抓取不同软硬度物体时,导纳与阈值控制差异大。软物抓取,导纳力稳至 3 N,阈值仅 2~2.5 N;中硬物抓取,导纳力准达 5 N,阈值超调 7 N;硬物抓取,导纳力稳至 8 N 且柔顺,阈值超调超 10 N、刚性接触强。

3) 路径规划收敛性对比

本文选取传统 DDPG 算法、只加入 RRT 专家经验的 DDPG 算法 (RRT-guided DDPG, R-DDPG) 和所提出的 RHER-DDPG 算法分别进行训练,并从奖励值、损失值和成功率 3 方面来进行对比分析。

如图 10 所示,在奖励值对比中,RHER-DDPG 在约 700 epoch 达到最高奖励并稳定收敛,而 R-DDPG 需至 900 epoch 才收敛,传统 DDPG 奖励值始终处于低位且波动较大;在损失值对比中,RHER-DDPG 的策略网络损失下降最快,并在训练早期趋于稳定,表现出最优的训练稳定性;在成功率对比中,RHER-DDPG 仅需 500 epoch 即达到 80%~90% 的稳定成功率,R-DDPG 在 600 epoch 达到 80%,传统 DDPG 则始终未能突破 60%。结果表明,融合 RRT 引导与事后经验回放机制的 RHER-DDPG 显著提升了样本利用效率与策略收敛速度。

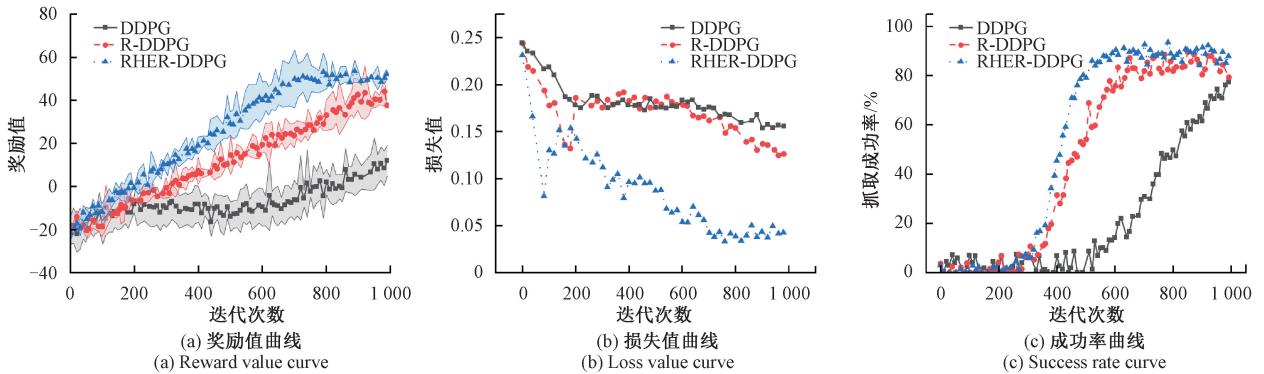


图 10 3 种算法对比曲线
Fig. 10 Reward curves of the three algorithms

值得注意的是,本文选取经典的 DDPG 算法作为基础框架,核心目的在于控制变量,以独立验证 RRT 专家引导与 HER 机制的作用。该架构优化策略具有模块的正交性,未来可作为通用插件迁移至软 Actor-Critic (soft actor-critic, SAC) 或双延迟深度确定性策略梯度 (twin delayed deep deterministic policy gradient, TD3)。

3.3 机械臂柔顺抓取实验

实验开始,先启动硬件设备及程序,机械臂复位,深度相机采集抓取区域的 RGB 和深度图像并转为点云数据。接着运行改进 GraspNet 网络,接收信息并预测目标物体抓取位姿。每次预测后,上位机选取置信度最高的位姿,转换至机械臂基座坐标系。各场景首次抓取时,目标物体的抓取位姿预测情况如图 11 所示。

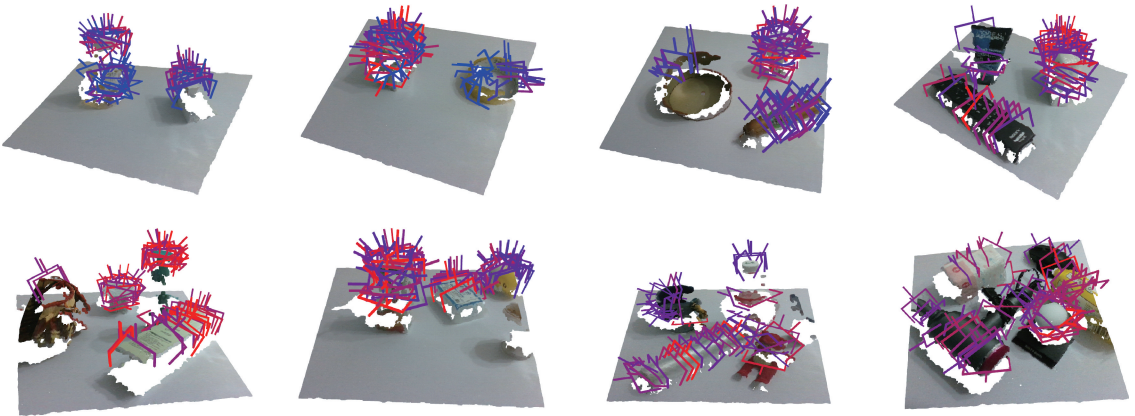


图 11 各场景抓取位姿预测情况
Fig. 11 Grasp pose predictions in various scenes

随后,RHER-DDPG 将目标位姿转为关节增量指令,驱动机械臂逼近目标,二指夹爪闭合。接触瞬间,CNN-Transformer 在线辨识物体刚度,导纳控制器实时切换参数,将稳态抓取力自适应调节至 3、5 或 8 N。如图 12 的力响应曲线所示,该过程平稳无超调,有效避免了软物体损伤与硬物体滑脱,实现了柔顺抓取。

抓取首个物品后,控制机械臂移至放置区,张开夹爪放下物品,随后机械臂复位,继续预测并抓取,直至所有物品抓取完毕,实验结束。

机械臂抓取实验过程示意图如图 13 所示。

抓取成功率计算公式如式(13)所示。

$$S = \frac{I_a}{I_b} \times 100\%$$

(13)

式中: I_a 为场景中待抓取物品的总数; I_b 为抓取次数。

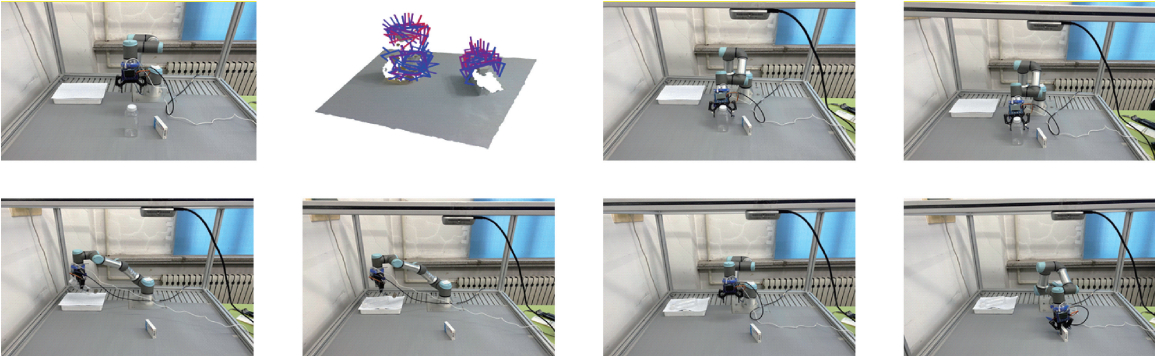


图 13 机械臂抓取实验过程示意图

Fig. 13 Schematic of the robotic arm grasping experiment process

物品损坏率计算公式如式(14)所示。

$$D = \frac{I_c}{I_a} \times 100\%$$

(14)

式中: I_c 为场景中损坏的物品数。只要夹爪抓取物品过程中物品发生形变或破裂均视为物品损坏。

受限于视觉点云获取、强化学习推理与高频触觉闭环等多模块全系统硬件联调的时间成本,本文在 8 个高度杂乱的非结构化场景中进行了 31 次高难度抓取。这 31 次测试全面覆盖了 12 种不同几何形状、材质与未知硬度的典型物品。在此高维物理变量耦合的复杂工况下,取得了 90.3% 的综合抓取成功率,损坏率低至 7.1%。实验表明,本文所提方法不仅在保证高成功率的同时,还能有效实现对各类刚度物体柔顺且无损的抓取。

3.4 消融实验与分析

针对视觉定位误差导致的滑脱问题,本文对比了原始 GraspNet 与引入 ResUNet 及物理重心约束的改进模型。实验表明,改进模型的抓取点检测精确率由 78% 提升至 88%,表征抓取方向偏差的接近向量方向性误差由 15.3 降至 8.6。且方向性误差的降低,从物理力学角度减少了由偏心效应引起的滑脱风险,使模块独立抓取成

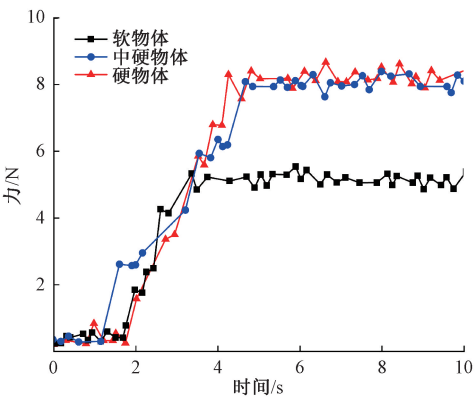


图 12 不同刚度物体的抓取力自适应响应曲线

Fig. 12 Grasp force variation curves in different scenes

功率由 81.7% 提升至 90.3%。

为验证多级导纳控制与 RHER-DDPG 强化学习模块的综合贡献,本文设计了不同系统配置进行消融测试,结果如表 4 所示。

表 4 消融实验系统性能对比

Table 4 Ablation experiment system performance comparison

实验组别	视觉重心约束	动态导纳控制	RHER-DDPG	平均接触力超调量/N	抓取成功率/%
基线系统	×	×	×	> 2.5	68.4
仅视觉改进	✓	×	×	> 2.0	76.5
视觉+触觉	✓	✓	×	< 0.5	83.2
本文框架	✓	✓	✓	< 0.5	90.3

如表 4 可知,视觉重心约束提升了宏观定位准确性;动态导纳控制消除了力超调风险;最终引入 RHER-DDPG 算法弥补了视觉残余误差,使系统综合成功率达到 90.3%。

4 结 论

针对机械臂在非结构化环境下因视觉特征提取不充分导致的抓取滑脱、物体刚度未知引起的接触损伤以及传统强化学习算法收敛速度缓慢等问题,提出一种基于视触融合及刚度自适应的机械臂抓取方法。通过引入 ResUNet 主干网络与物体重心约束机制优化位姿检测精度,利用 CNN-Transformer 架构建立刚度-导纳参数的自适应映射,并采用 RRT 引导与 HER 经验回放策略对 DDPG 算法进行优化,实现了复杂场景下精准的柔顺抓取。实验结果表明,相比于原始 GraspNet 算法,在复杂场景下的抓取成功率明显提升,有效降低了因特征提取不足导致的抓取滑脱问题;与传统 DDPG 算法相比,训练收敛步数大幅减少,策略探索的盲目性明显降低;与固定参数导纳控制相比,在抓取易碎物体时的无损率有效提高,实现了接触力的自适应调节。解决了机械臂在动态作业中位姿估计鲁棒性差、接触力控制僵硬及路径规划效率低的问题。

参考文献

- [1] 尤波,刘嘉琦,程晨晨,等. 基于 sEMG 的人机交互随动控制研究[J]. 仪器仪表学报, 2025, 46(3): 123-142.
YOU B, LIU J Q, CHENG CH CH, et al. Human-computer interaction follow-up control based on sEMG[J]. Chinese Journal of Scientific Instrument, 2025, 46(3): 123-142.
- [2] YANG H, LIU J Q, LIU W B, et al. Compliant grasping control for a tactile self-sensing soft gripper[J]. Soft Robotics, 2024, 11(2): 230-243.
- [3] LIANG Q K, XIAO W X, LONG J Y, et al. A multimodal robust recognition method for grasping objects with robot flexible grippers[J]. IEEE/ASME Transactions on Mechatronics, 2024, 30(2): 1154-1165.
- [4] 褚文敏,周蒯,滕利臣. 基于分布式三维力传感器的装配接触力测量方法[J]. 仪器仪表学报, 2024, 45(11): 30-41.
CHU W M, ZHOU K, TENG L CH. Assembly contact force measurement method based on distributed threedimensional force sensor[J]. Chinese Journal of Scientific Instrument, 2024, 45(11): 30-41.
- [5] VOIGT F, NACERI A, HADDADIN S. Learning wrist policies for anthropomorphic soft power grasping in handle and door manipulation[J]. IEEE Transactions on Robotics, 2025, 41: 3738-3759.
- [6] CHEN G, WU J S, ZHANG Z J, et al. A soft robotic gripper with a belt loop actuated adhesion design for gentle handling of fragile object[J]. IEEE Robotics and Automation Letters, 2024, 9(8): 6784-6791.
- [7] TANG K, HUANG X ZH, SEINO A, et al. Fixture-free automated sewing system using dual-arm manipulator and high-speed fabric edge detection[J]. IEEE Robotics and Automation Letters, 2025, 10(9): 8962-8969.
- [8] ZAKERI E, XIE W F. Robust AI-driven target-object-free hybrid vision/force control of industrial robotic systems[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2025, 55(7): 4899-4914.
- [9] 胡瑀晖,吴立刚,陈建国,等. 空间机械臂多模态视觉感知与操作技术综述[J]. 哈尔滨工业大学学报, 2025, 57(12): 1-21.
HU Y H, WU L G, CHEN J G, et al. Review of multi-modal visual perception and manipulation technologies for space manipulator[J]. Journal of Harbin Institute of Technology, 2025, 57(12): 1-21.
- [10] WANG L, ZHANG Y ZH, ZHU SH D, et al. Improving grasp pose detection by implicitly utilizing geometric information and spatial relations of objects in clutter[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 1-11.
- [11] HAN H Y, WANG W J, HAN X, et al. 6-DoF grasp pose estimation based on instance reconstruction[J]. Intelligent Service Robotics, 2024, 17(2): 251-264.
- [12] GAO H J, ZHAO J J, HU J Q, et al. A real-time grasping detection network architecture for various grasping scenarios[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 36(5): 8215-8226.
- [13] YU SH, ZHAI D H, XIA Y Q, et al. Robotic grasp detection with 6-D pose estimation based on graph convolution and refinement[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2024, 54(6): 3783-3795.
- [14] ZUO G Y, SHEN ZH H, YU SH Y, et al. HBGNet: Robotic grasp detection using a hybrid network[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 74: 1-9.
- [15] 吴海滨,左云逸,王爱丽,等. CNN 结合 Transformer 的高光谱图像和 LiDAR 数据协同地物分类方法[J]. 仪器仪表学报, 2025, 46(8): 286-301.

- WU H B, ZUO Y Y, WANG A I L, et al. Collaborative land classification method using CNN combined with Transformer for hyperspectral images and LiDAR data[J]. Chinese Journal of Scientific Instrument, 2025, 46(8): 286-301.
- [16] CHEN W, JIANG Q P, ZHOU W, et al. Dynamic hypergraph convolutional network for no-reference point cloud Quality assessment [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024, 34(10): 10479-10493.
- [17] 王立辉, 陈勇吉, 韩华春, 等. 基于点云分割与主成分分配的换电机器人位姿估计方法[J]. 仪器仪表学报, 2024, 45(10): 123-132.
- WANG L H, CHEN Y J, HAN H CH, et al. Pose estimation method for battery swapping robot based on point cloud segmentation and principal component registration[J]. Chinese Journal of Scientific Instrument, 2024, 45(10): 123-132.
- [18] YAN H X, FANG H SH, LU C W. AViT_a: Adaptive visual-tactile dexterous grasping[J]. IEEE Robotics and Automation Letters, 2024, 9(11): 9462-9469.
- [19] CHEN K CH, CHEN G, LIN J R, et al. Temporal cross-modal fusion method of tactile, 2D visual, and point cloud data for slip detection in robotic grasping [J]. Measurement Science and Technology, 2025, 36(10): 106220.
- [20] AN J, JIA Q X, LI T, et al. Dexterous grasping of robotic hand posture and force generation through visual-tactile fusion[J]. Measurement Science and Technology, 2024, 36(2): 026202.
- [21] FU J L, BURZO I, IOVENE E, et al. Optimization-based variable impedance control of robotic manipulator for medical contact tasks [J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 4004608.
- [22] ZHANG Y H, ZHAO X W, CHEN Y M, et al. Observer-based variable impedance control using moving horizon estimation for robot machining thin-walled workpieces[J]. IEEE Transactions on Industrial Electronics, 2023, 71(6): 5972-5982.
- [23] 郝正阳, 张帆, 张帆扬. 自动装配的导纳控制系统中关键影响机制研究[J]. 仪器仪表学报, 2024, 45(7): 292-300.
- HAO ZH Y, ZHANG F, ZHANG F Y. Research on key influencing mechanisms in admittance control systems for automated assembly [J]. Chinese Journal of Scientific Instrument, 2024, 45(7): 292-300.
- [24] 宁睿, 刘志晨, 刘毅, 等. 强化学习引导变导纳控制的机械臂精密轴孔装配模糊位姿估计与精确调整[J]. 仪器仪表报, 2024, 45(11): 170-177.
- NING R, LIU ZH CH, LIU Y, et al. Reinforcement learning guided variable admittance control for fuzzy pose estimation and precise adjustment of robotic precise peg-in-hole assembly [J]. Journal of Scientific Instrument, 2024, 45(11): 170-177.
- [25] REN L, FANG Z L, LIU W ZH, et al. Deep reinforcement learning for UAV indoor navigation through task decomposition [J]. IEEE/CAA Journal of Automatica Sinica, 2025, 12(12): 2627-2629.
- [26] SANCHEZ F R, WANG Q, BULENS D C, et al. Learning and reusing primitive behaviours to improve hindsight experience replay sample efficiency[C]. 2024 10th International Conference on Automation, Robotics and Applications, 2024: 328-333.
- [27] XUN X T, ZHANG R D, CHAI S CH, et al. Light reflection-guided RRT*: Efficient path planning in narrow passages [J]. IEEE Robotics and Automation Letters, 2025, 10(11): 11474-11481.
- [28] ZHANG Y, WEN L, HUANG Y H, et al. Meta-Learning-based safety-critical control in multi-obstacles environments [J]. IEEE Transactions on Automation Science and Engineering, 2025, 22: 15299-15313.

作者简介



李佳钰(通信作者), 2009年、2013年和2018年于哈尔滨理工大学分别获得学士学位、硕士学位和博士学位, 现为哈尔滨理工大学副教授, 主要研究方向为足式机器人控制与人机交互技术。

E-mail: lijiayu@hrbust.edu.cn

Li Jiayu (Corresponding author) received his B. Sc., M. Sc., and Ph. D. degrees all from Harbin University of Science and Technology in 2009, 2013, and 2018, respectively. He is currently an associate professor at Harbin University of Science and Technology. His main research interest includes control of the legged robot and human machine interaction.



王宏斌, 2024年于哈尔滨理工大学获得学士学位, 现为哈尔滨理工大学研究生, 主要研究方向为机器人多模态感知与智能控制。

E-mail: 2420510141@stu.hrbust.edu.cn

Wang Hongbin received his B.Sc. degree from Harbin University of Science and Technology in 2024. He is currently a master's graduate of Harbin University of Science and Technology. His main research interests include robot multimodal perception and intelligent control.



陈新宇,2021 年于哈尔滨理工大学获学士学位,2025 年于哈尔滨理工大学获硕士学位,现为哈尔滨理工大学硕士毕业生,主要研究方向为机器视觉、机械臂抓取规划及控制。

E-mail:18045189903@ 163. com

Chen Xinyu received his B. Sc. degree from Harbin University of Science and Technology in 2021, and his M. Sc. degree from Harbin University of Science and Technology in 2025. He is currently a master's graduate of Harbin University of Science and Technology. His main research interests include machine vision, and robotic arm grasp planning and control.



杨怀广,2009 年、2013 年、2015 年于哈尔滨工业大学分别获得工学学士、硕士、博士学位。现为哈尔滨工业大学副教授,主要研究方向为机器人地面力学与数字孪生。

E-mail:yanghuaiguang@ hit. edu. cn

Yang Huaiguang received his B. Sc. , M. Sc. , and Ph. D. degrees all from Harbin Institute of Technology in 2009, 2013, and 2015, respectively. He is currently an associate professor at Harbin Institute of Technology. His main research interests include robot ground mechanics and digital twins.



尤波,1982 年于哈尔滨工业大学获得学士学位,1987 年于哈尔滨科技大学获得硕士学位,1995 年于哈尔滨工业大学获得博士学位,现为哈尔滨理工大学教授,主要研究方向为智能机器人与机电控制。

E-mail:youbu@ hrbust. edu. cn

You Bo received his B. Sc. degree from Harbin Institute of Technology in 1982, his M. Sc. degree from Harbin University of Science and Technology in 1987, and his Ph. D. degree from Harbin Institute of Technology in 1995. He is currently a professor at Harbin University of Science and Technology. His main research interest includes intelligent robot and mechatronic control.