

111Equation Chapter 2 Section 1中图分类号:

TP242

基于肌腱驱动的双足机器人跳跃运动控制研究

控制科学与工程

朱玉迪

李清都 教授

二〇二五年十二月

学校代码：10252
学 号：191550059

上海理工大学博士学位论文

基于肌腱驱动的双足机器人跳跃运动控制研究

姓 名 朱玉迪
系 别 光电信息与计算机工程学院
专 业 控制科学与工程
研究方向 模式识别与智能系统
指导教师 李清都 教授

学位论文完成日期 2025 年 12 月

RESEARCH ON JUMPING MOTION CONTROL OF TENDON-
DRIVEN BIPEDAL ROBOTS

by

Yudi Zhu

A Thesis Submitted to University of Shanghai for Science & Technology in

Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

Under the Supervision of
Professor Qingdu Li

University of Shanghai for Science & Technology

December 2025

学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学位论文保留并向国家有关部门或机构送交论文的复印件和电子版。允许论文被查阅和借阅。本人授权上海理工大学可以将本学位论文的全部内容或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于 保密 年
 不保密

学位论文作者签名： 指导教师签名：

年 月 日 年 月 日

声 明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已注明引用的内容外，本论文不包含任何其他个人或集体已经公开发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。

本声明的法律责任由本人承担。

学位论文作者签名：

年 月 日

摘 要

摘 要

关键词：双足机器人 控制策略

Abstract

Key Words: Bipedal robot, Control Strategy

目 录

中文摘要

ABSTRACT

第一章 绪论	1
1.1 研究背景与意义.....	1
1.2 人形机器人本体系统的发展历程	2
1.3 腿臂协同控制的研究综述	6
1.3.1 基于零空间分解的腿臂协同控制	7
1.3.2 基于分层二次规划的腿臂协同控制	10
1.3.3 基于强化学习的腿臂协同控制.....	12
1.3.4 基于模仿学习的腿臂协同控制.....	15
1.4 本文的主要问题.....	19
1.5 本文研究内容及章节安排	20
第二章 人形机器人全身动作重定向	22
2.1 引言	22
2.2 方法	24
2.2.1 动作模仿系统设计	24
2.3 改进重定向	24
2.3.1 人体肘关节建模	25
2.4 实验	28
2.4.1 单人动作模仿	28
2.4.2 多人动作模仿	30
2.4.3 任务操作.....	31
2.5 结果与讨论	32
2.5.1 数据分析.....	32
2.5.2 消融实验.....	35
2.5.3 局限	36
2.6 结论	37

第一章 绪论

1.1 研究背景与意义

近年来，人工智能技术在感知、认知与决策层面取得了突破性进展。现有系统已经能够在复杂场景中实现高精度的目标识别^[1]，合成高分辨率、逼真的照片级图像^[2-3]，根据自然语言描述自动生成具有实用价值的程序代码^[4-5]，并在多种具有挑战性的游戏环境中超越人类顶尖水平^[6-8]。然而，与其在感知与认知能力上的快速发展相比，人工智能体在物理世界中的运动能力仍显著落后于人类及其他生物所展现出的灵活性与适应性。

人类之所以能够在复杂、动态且高度不确定的环境中完成多样而精细的运动行为，得益于人类长期进化形成的高度协调的全身运动控制机制，尤其体现在上下肢之间的协同配合与任务分工上。相比之下，当前的机器人系统，无论是在仿真环境还是现实部署中，往往只能执行少量预定义或高度专用的动作，其行为形式僵化、鲁棒性有限，难以应对真实环境中的复杂扰动和突发状况。这种差距在双足人形机器人上尤为突出。

具备类人运动能力的双足人形机器人，被普遍认为是未来服务机器人、特种机器人以及人机协作系统的重要发展方向。若能够显著提升其全身运动控制能力，使其在非结构化环境中稳定完成诸如坐下、起立、跌倒与爬起等高风险、高自由度动作，将有助于机器人从实验室和受控工业场景中走向更加真实、复杂的应用环境。同时，具有自然、协调运动模式的虚拟人和机器人智能体，也将在计算机图形学、虚拟现实、生物力学以及康复医学等领域展现重要价值。

围绕上述目标，研究者提出了多种控制器设计方法。传统的基于模型的手工设计控制策略在复现人类敏捷性方面已取得一定成果^[9-17]。然而，尽管人类能够熟练完成各类全身运动技能，其内在控制机理却难以被清晰刻画，更难以将其系统性地编码进控制器之中。

在机器人领域，手工设计控制器同样取得了一些令人信服的成果^[18-22]。但此类方法通常高度依赖专家经验与大量精细调参，所得到的控制器往往针对特定任务或技能进行定制，难以自然扩展到更大规模、更丰富的技能集合。当研究对象转向难度更高、专业性更强的行为（如坐起或爬起）时，由于接触关系更加复杂、系统动力学更难精确建模，控制器的设计与调试过程将显著复杂化。总体而言，尽管手工设计方法取得了诸多成功，其所能达到的运动表现仍与人类所展现出的丰富性、灵活性与优雅性存在明显差距。

为降低控制器设计的人工成本,基于优化的方法,如模型预测控制(MPC)^[23-29]与强化学习(RL)^[30-37],逐渐成为合成复杂运动技能的重要工具。这类方法通过优化目标函数来自动搜索控制策略,将设计者的主要工作从“如何控制”转变为“希望智能体表现出何种行为”。其中,强化学习在合成高自由度、强非线性系统的运动控制策略方面表现出强大潜力,已被成功应用于多种仿人运动技能的学习^[38-43]。

然而,基于优化的方法在实际应用中仍面临显著挑战。目标函数的设计往往需要融合大量领域知识,并对任务高度敏感。例如,仅为了诱导智能体产生自然的人类步态,就需要在能量效率、对称性、冲击力抑制以及稳定性等多个维度之间进行精细权衡^[25,44-48]。当任务扩展至坐下起立、跌倒爬起等更具挑战性的动作时,这种基于目标函数的设计方式不仅调参成本高,而且难以系统性地刻画人类动作中隐含的腿部与上肢的协同控制规律。这不禁引人思考:是否有一种方法能够根据任务来实现高效的腿臂协同控制。

1.2 人形机器人本体系统的发展历程

人形机器人作为机器人领域最具综合性与挑战性的研究方向之一,其本体系统的发展始终围绕结构设计、驱动技术与全身运动控制能力展开。回顾近半个世纪的发展历程,人形机器人本体技术大致经历了从“结构验证与稳定行走”到“高动态运动控制”,再到“高功率密度电驱与具身智能融合”的阶段性演进。不同国家在不同时期形成了具有代表性的技术路线,其演进脉络呈现出较为清晰的时间顺序。

20世纪70年代,日本率先开展人形机器人系统研究。1976年早稻田大学研发的WABOT系列机器人首次实现双足行走,被认为是最早具有完整类人结构的双足机器人之一^[49]。该阶段的研究重点集中在双足结构构型设计与重心稳定控制问题,为人形机器人基本构型与自由度分配奠定了工程基础。进入1990年代后,日本逐步将研究重心转向电机驱动关节与系统工程实现。AIST与早稻田大学联合推出的HRP系列机器人在全身自由度配置、动力学建模与姿态控制方面形成了较为完整的技术体系^[50]。与此同时,本田公司推出的ASIMO系列(2000年首发)采用旋转电机与减速器结合的刚性电驱结构,实现了稳定行走、跑步与上下楼梯等动作^[51]。该阶段的核心贡献在于建立了电驱人形机器人标准化关节结构与控制框架。然而,由于采用较高减速比结构,关节反驱能力有限,动态性能受到一定约束。日本阶段的代表性平台如图1-1所示。

进入21世纪后,美国在人形机器人领域逐步将研究重心由稳定行走问题转向高动态能力与复杂环境适应性研究。MIT LegLab提出的弹跳型双足机器人系统为

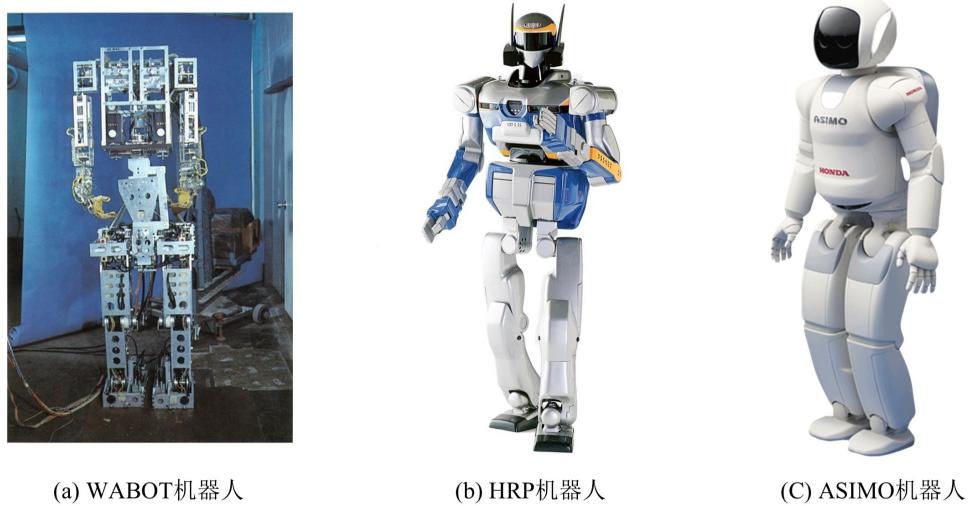


图 1-1 日本代表性人形机器人平台

高动态奔跑控制奠定理论基础^[52]。随后，波士顿动力公司发布 PETMAN 平台^[53]，并在 Atlas 机器人上采用基于优化的全身控制方法实现复杂动态动作^[54]。Atlas 采用集中式液压动力单元与高功率液压执行机构，实现奔跑、跳跃与翻滚等高动态动作，在全身动量控制与扰动恢复方面取得重要突破。然而，液压系统在能耗、噪声与维护复杂性方面存在不足，限制了其在工业与服务场景中的大规模部署。

近年来，美国人形机器人技术逐步向高功率密度电驱结构转型，强调工程化应用与场景落地能力。Agility Robotics 推出的 Digit 被认为是首个面向物流场景的人形机器人平台^[55]，其采用类鸟式腿部构型与全电驱系统，可在仓储环境中自主导航、搬运包裹并与人类协同作业。Digit 的结构设计强调稳定性与可部署性，使人形机器人首次在真实物流场景中实现持续运行。

Tesla 于 2022 年发布 Optimus，并在随后推出 Gen 2 版本^[56]。该平台面向规模化生产设计，结合 Tesla 在电池系统、制造工艺与自动驾驶算法方面的技术积累，采用改进型电驱执行器与 11 自由度灵巧手，实现更高步行速度与更强操作能力。其技术目标在于构建可大规模制造的人形机器人产品体系。

2024 年，Boston Dynamics 发布全电驱版本 Atlas，标志着高动态人形机器人由液压系统向高功率密度电驱系统的全面转型^[57]。全电驱 Atlas 采用新一代高性能电驱执行器与优化的结构构型设计，在保持高动态能力的同时，实现更高的能效与更低的系统复杂度。该平台强调全身协调控制与复杂操作能力，具备更强的机动性、力量输出与整机协同能力，为工业场景中的复杂搬运与装配任务提供了工程化基础。这一转型体现出人形机器人本体结构正由高功率液压驱动向高带宽电驱系统演进。

Figure 公司发布的 Figure 02 被定位为面向制造与物流场景的自主通用人形机器人平台，具备 40 余个自由度与高灵巧操作能力，并通过与大模型系统的融合实现语义理解与任务规划能力的提升^{2[58]}。该平台强调商业可行性与实际部署能力，标志着人形机器人由技术验证阶段进入工业落地阶段。

近年来，欧洲企业开始面向家庭场景探索工程化产品。挪威 1X 公司推出的 NEO Gamma 采用全电驱与整体柔顺结构设计^[59]，通过仿生构型与全身顺应性控制提升安全性，使机器人能够在无需物理隔离的情况下与人类共处于家庭环境之中。其设计理念强调低噪声、低刚度与安全交互能力，体现出人形机器人向消费级与家庭应用场景扩展的趋势。欧美代表性平台如图 1-2 所示。

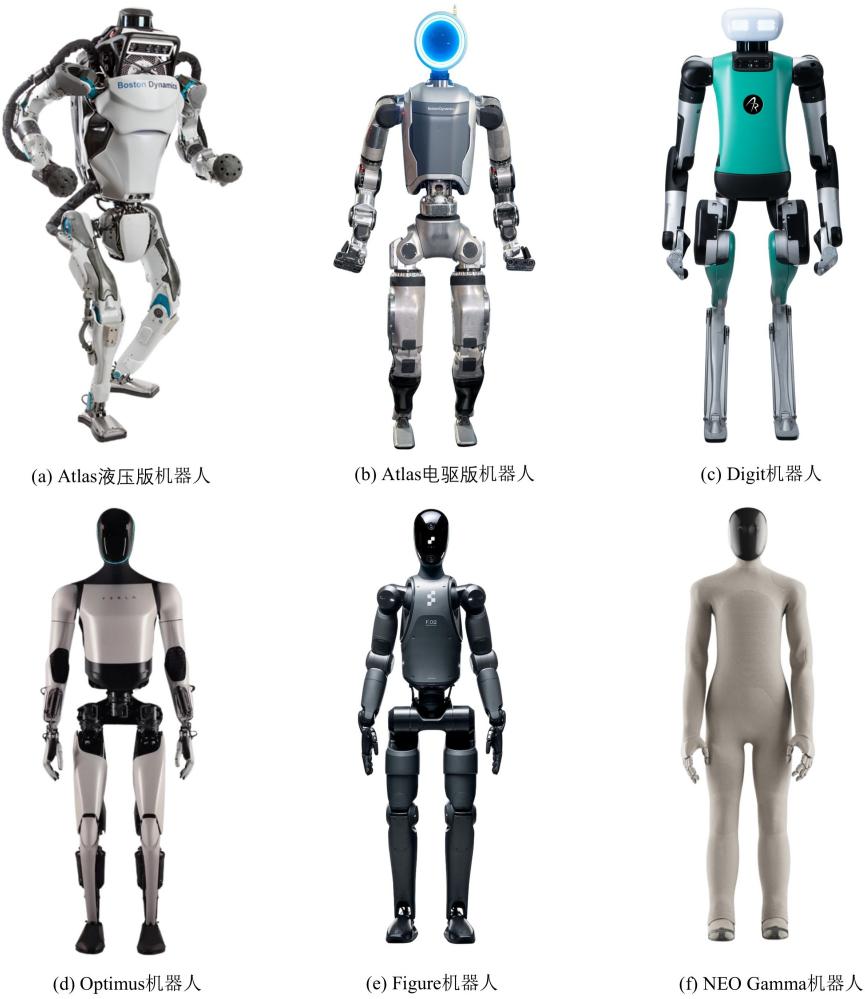


图 1-2 欧美代表性人形机器人平台

中国在人形机器人的研究始于 20 世纪 80 年代。哈尔滨工业大学、清华大学与国防科技大学等单位在双足结构设计、步态规划与动力学控制方面开展了系统探索^[60-61]。北京理工大学“汇童”系列实现了多代迭代升级，在高速视觉感

知与全身协调控制方面取得进展^[62]。上述研究为国内人形机器人本体结构设计与运动控制算法奠定了理论与工程基础。

进入 2020 年以后，在国家政策推动与产业资本支持下，中国人形机器人本体系统进入快速发展阶段。驱动技术由传统高减速比电驱逐步向准直驱与高功率密度关节模组演进，强调低惯量设计、高带宽力控能力以及整机一体化集成水平的提升。与早期以实验室验证为主的研究平台不同，近年来国内平台普遍面向工业、服务与具身智能应用场景，呈现出明显的工程化与规模化特征。

宇树科技推出的 H1 与 G1 系列平台代表了高功率密度电驱人形机器人的典型路线。H1 通过低减速比高扭矩电驱关节实现高速奔跑能力，刷新了全尺寸人形机器人的奔跑速度纪录。G1 平台则采用紧凑型一体化关节模组设计，整机质量约 35 kg，标准配置 23 个关节电机（可扩展至 43 自由度）^[63]，膝关节最大输出扭矩可达 90–120 N·m，配备低惯量高响应永磁同步电机与工业级轴承结构，实现较高的动态稳定性与重复运行可靠性。其稳定的结构设计与出色的运动性能，使其成为科研验证、公开展示以及大规模活动（如春晚演示）中广泛选用的平台，体现出工程可靠性与动态性能之间的平衡。

在高动态能力方面，EngineAI 推出的 SE01 平台实现了前空翻等高难度全身动作，展示出高功率密度驱动系统与全身协调控制算法的融合能力。该类平台通过优化关节输出密度与躯干惯量分布，使高动态动作在电驱结构下成为可能，标志着国内电驱人形机器人在动态性能方面取得突破^[64]。

优必选（UBTECH）推出的 Walker S2 面向工业应用场景^[65]，配置 41 个高性能伺服关节，强调稳定双足行走与精确操作能力，定位于制造与物流环境中的通用作业平台。与此同时，优必选与北京人形机器人创新中心联合研发的天工（Walker Tienkung）研究平台具备最高 12 km/h 的奔跑能力，并搭载 NVIDIA Jetson Orin 计算平台，支持强化学习与具身智能算法验证，在首届人形机器人马拉松赛事中取得优异成绩，体现出国内在高动态持续运行能力方面的系统集成水平^[66]。

智元发布的 A2 通用人形机器人采用模块化设计理念，强调工业装配与物流搬运能力，通过高带宽电驱关节与高自由度上肢系统实现复杂操作任务^[67]。星动纪元 Star1 则率先开源 humanoid_gym 框架，为双足机器人强化学习训练提供统一平台，推动算法与硬件平台的协同发展，体现出国内在“算法一本体一体化”方向上的探索趋势。

在医疗与康复技术积累基础上，傅利叶（Fourier）推出 GR-1 平台，采用自主研发的 FSA 执行器系统，在保证较高负载能力的同时实现较好的力控精度，形成医疗康复与通用人形机器人技术的融合路径。Adam 平台则采用透明仿生结构设计，

强调生物力学展示与教育研究功能，为人形机器人结构教学与运动机理研究提供实验载体。近年来中国代表性平台如图 1-3 所示。

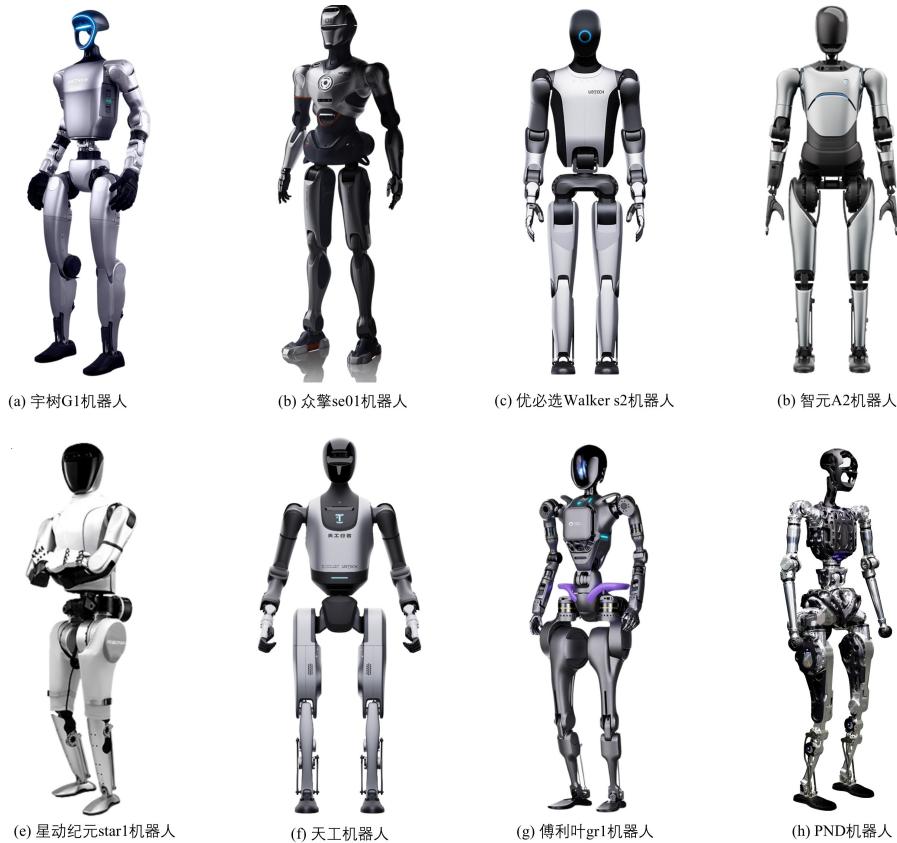


图 1-3 近年来中国代表性人形机器人平台

近年来中国人形机器人本体系统的发展呈现出以下特征：其一，形成了覆盖高动态运动、工业操作、教育科研与康复应用的多样化平台体系；其二，关节执行器由高减速比刚性结构向低减速比、高反驱能力与高功率密度方向演进；其三，强化学习与具身智能算法成为主流控制手段，对关节带宽、结构刚度与实时控制能力提出更高要求。整体技术水平已由“技术追赶”进入“并跑阶段”，在高动态电驱动平台与系统集成能力方面具备一定竞争优势。在此背景下，如何在高自由度电驱人形机器人本体上实现腿臂协同控制与任务优先级管理，成为支撑复杂场景部署的关键问题。

1.3 腿臂协同控制的研究综述

人形机器人通常具有高自由度、多关节串并联混合结构，其运动控制涉及躯干、双臂与双腿等多个子系统的协同配合。在执行站立、行走与操作等任务过程中，机器人不仅需要满足系统动力学约束和接触约束，还需同时兼顾平衡维持、末

端轨迹跟踪以及构形与能量优化等多重目标。因此，与传统机械臂或移动机器人控制问题不同，人形机器人的运动控制难以被分解为相互独立的子任务，而必须在统一框架下对全身多个关节和任务进行协调，这类方法通常被统称为全身控制（Whole-Body Control, WBC）。

在双足人形机器人中，腿部主要承担支撑与平衡维持的功能，而手臂除执行操作任务外，还可通过调节角动量、改变质心分布等方式辅助稳定性控制。尤其在外部扰动或复杂接触环境下，腿部与手臂之间往往存在显著的动力学耦合关系：腿部的支撑状态直接影响上肢可用的运动空间，而手臂的摆动和操作动作又会反作用于系统整体的稳定性。因此，如何在保证动态平衡与接触约束的前提下，实现腿臂之间的高效协同控制，是人形机器人全身控制研究中的关键问题之一。

从控制理论角度看，腿臂协同控制问题的本质可归结为多任务、多约束条件下的优先级分配与协调问题。在实际应用中，不同控制目标的重要性存在显著差异，例如保持接触约束、维持系统稳定性和避免关节超限通常具有更高优先级，而姿态优化、能量最小化或构形调整等目标则可在不破坏核心约束的前提下执行。围绕这一思想，研究者提出了以任务优先级为核心的全身控制范式，其关键在于确保高优先级任务得到满足，同时在其可行空间内协调低优先级任务。需要指出的是，任务优先级并非某一类特定控制方法所独有，而是一种贯穿人形机器人全身控制设计的基础思想，其具体实现形式随控制框架的不同而有所差异。

1.3.1 基于零空间分解的腿臂协同控制

在模型驱动的人形机器人全身控制方法中，任务优先级通常以显式层级结构进行刻画。早期研究主要基于解析方法，通过雅可比矩阵投影或零空间分解实现多任务协调，使低优先级任务在不干扰高优先级任务的前提下执行。这类方法结构清晰、物理意义明确，在人形机器人站立控制、姿态调节以及简单运动生成中得到了广泛应用。

基于零空间的多任务优先级控制思想最早起源于冗余机械臂的逆运动学求解问题。在具有冗余自由度的机械臂系统中，当关节自由度高于任务空间维度时，系统存在无限多组满足主任务的解。为在完成主要任务的同时执行次级目标（如避障、关节限位规避或构形优化），研究者提出利用雅可比矩阵的零空间结构进行分层控制。Lieggois^[68]首次系统性提出在主任务解的零空间中嵌入次级优化目标；随后 Nakamura^[69]对冗余系统的运动学分解与优先级结构进行了系统化分析，为多任务控制理论奠定了基础。

在此基础上，Siciliano 与 Slotine^[70]提出了较为完整的任务优先级控制框架，通过零空间递归投影实现多任务分层控制，该思想随后在“Stack-of-Tasks”结构中得

到形式化表达，并逐渐推广至复杂机器人系统。随着研究对象从单臂系统扩展至高自由度的人形机器人，零空间优先级思想被自然引入全身控制框架。Khatib 提出的操作空间控制理论（Operational Space Control）^[71-72] 将任务空间动力学与关节空间控制统一建模；Sentis 等进一步在动力学一致性框架下推广该思想^[73]，实现了多接触与多任务分层控制的系统化建模。由此，零空间分解逐渐成为人形机器人腿臂协同控制的基础理论框架。

其核心思想可概括为：通过零空间递归投影构建严格的任务层级关系。设机器人关节向量为 $q \in \mathbb{R}^n$ ，任务空间变量为 $x \in \mathbb{R}^m$ ，其速度映射关系为

$$\dot{x} = J(q)\dot{q}, \quad (1.1)$$

其中 $J(q)$ 为任务雅可比矩阵。当系统存在冗余 ($n > m$) 时，关节速度的一般解可表示为：

$$\dot{q} = J^\dagger \dot{x} + (I - J^\dagger J)\dot{q}_0, \quad (1.2)$$

其中 J^\dagger 为 Moore-Penrose 伪逆， $(I - J^\dagger J)$ 为零空间投影矩阵， \dot{q}_0 为任意零空间速度。

对于多任务系统，若任务 x_1 优先级高于 x_2 ，则其递归形式可表示为：

$$\dot{q} = J_1^\dagger \dot{x}_1 + N_1 J_2^\dagger \dot{x}_2 + N_1 N_2 \dot{q}_0, \quad (1.3)$$

其中 $N_i = I - J_i^\dagger J_i$ 为第 i 个任务的零空间投影矩阵。

为更直观地说明零空间递归投影在腿臂协同控制中的层级结构关系，图 1-4 给出了基于零空间分解的任务优先级全身控制框架示意图。

如图所示 1-4，该框架采用自上而下的层级结构。最上层为约束层（Constraint Layer），通常包括接触约束、质心控制或整体动量调节等稳定性相关任务。这一层对应公式中的最高优先级任务 x_1 ，其解通过雅可比伪逆 J_1^\dagger 求得，并生成零空间投影矩阵 N_1 。

在此基础上，任务层（Task Layer）中的手臂操作或轨迹跟踪任务被投影至约束层的零空间中执行，对应公式中的递归项 $N_1 J_2^\dagger \dot{x}_2$ 。这种结构确保操作任务不会破坏系统稳定性与接触一致性。

进一步地，姿态层（Posture Layer）利用剩余冗余自由度进行构形优化或冗余

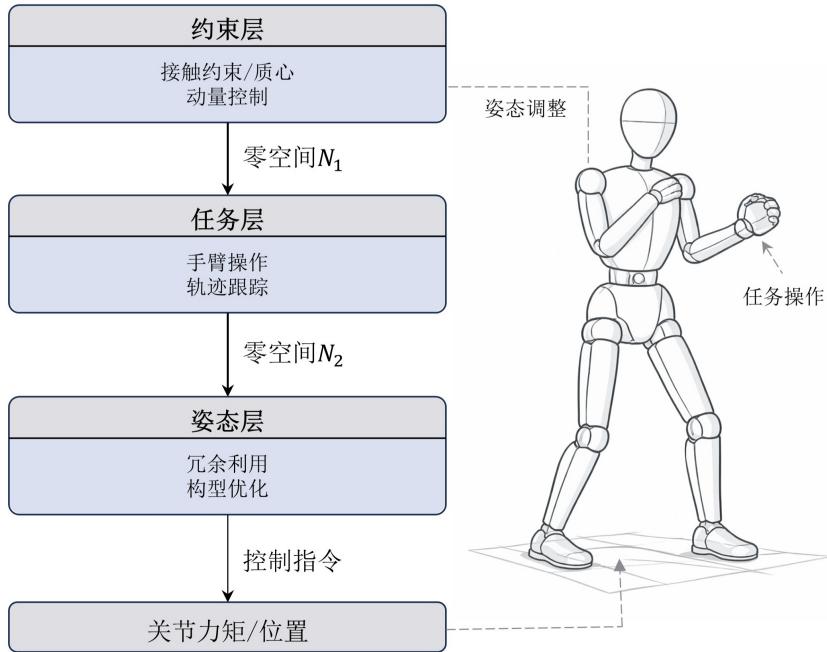


图 1-4 基于零空间分解的任务优先级全身控制框架示意图。

利用，其数学形式对应 $N_1 N_2 J_3^\dagger \dot{x}_3$ 。最终，各层任务的递归叠加形成关节空间控制指令（如关节速度、加速度或力矩命令）。

该算法的核心思想在于：通过零空间递归投影建立严格的任务层级关系——高优先级任务被优先满足，低优先级任务仅在其零空间内执行，而剩余冗余自由度用于姿态优化或次级性能提升。

在腿臂协同控制场景 1-4 中，这种层级结构通常体现为：支撑与质心稳定任务处于最高优先级，手臂操作任务嵌入稳定性零空间中执行，而构形与姿态调节则利用剩余冗余自由度进行优化。由此实现“稳定优先、操作嵌入、姿态优化”的分层协同控制机制。

从方法特性来看，基于零空间分解的任务优先级控制方法在腿臂协同问题中具有显著优势。首先，该方法能够通过严格的零空间投影机制保证高优先级任务的精确满足，使支撑稳定性、质心控制等关键任务在数学上得到明确保护。这种严格的层级结构具有清晰的物理含义，使不同任务之间的干扰关系可以通过投影矩阵形式直观刻画，因而具有良好的可解释性与结构透明性。其次，零空间方法通常基于解析形式推导，其计算结构相对紧凑，实时性较好，适用于早期计算资源受限的人形机器人控制系统。

然而，随着控制任务复杂度的提升，该方法的局限性逐渐显现。其一，零空间投影本质上基于等式约束构建，对于摩擦锥、关节力矩限制、接触不等式等物理约束的处理能力有限，往往需要额外近似或线性化处理。其二，在多接触与强非

线性动态场景下，连续零空间递归可能导致数值不稳定或解空间退化问题。其三，任务优先级通常在设计阶段固定，一旦高优先级任务过于严格，低优先级任务的可行空间可能被过度压缩，从而限制手臂与腿部之间的协调自由度。在高速动态运动或频繁接触切换过程中，这种刚性层级结构可能难以兼顾稳定性与灵活性。

因此，尽管零空间方法为多任务优先级控制提供了重要理论基础，并在人形机器人早期全身控制中发挥了关键作用，但在复杂动态腿臂协同场景下，其对不等式约束的处理能力、数值鲁棒性以及结构自适应能力仍存在一定局限。这也促使后续研究逐步向基于优化的分层控制框架发展。

1.3.2 基于分层二次规划的腿臂协同控制

随着人形机器人控制任务复杂度的不断提升，单纯依赖零空间投影的解析式优先级控制方法逐渐暴露出其在不等式约束表达、多接触建模以及数值稳定性方面的局限。为克服上述问题，研究者开始将多任务控制问题统一转化为带约束的优化问题，通过二次规划（Quadratic Programming, QP）框架在满足系统动力学方程的同时处理多种等式与不等式约束。

早期工作中，Mansard 等^[74]提出了将任务优先级嵌入二次规划结构的方法，通过构造分层优化问题实现严格优先级的保持，为多任务控制问题提供了统一的优化表达形式。随后，Escande 等^[75]系统化提出分层二次规划（Hierarchical Quadratic Programming, HQP）框架，通过递归构建一系列带约束的二次规划子问题，实现任务间的严格优先级分离，同时保证数值稳定性和计算可行性。这一工作在理论层面上将“Stack-of-Tasks”结构推广至优化框架，并为复杂约束处理提供了严谨的数学基础。在动力学层面，Righetti 等^[76]将逆动力学控制问题表述为二次规划问题，将关节加速度、力矩以及接触力作为决策变量，从而在满足动力学一致性的前提下显式引入摩擦锥约束和力矩限制，为全身动力学控制提供了可扩展的优化求解结构。Herzog 等^[77]在此基础上提出基于逆动力学的实时 QP 控制框架，强调在浮动基系统中直接将接触力作为优化变量处理，从而避免显式求解约束动力学逆问题，提高了实时性与数值鲁棒性。

在人形机器人系统中，HQP 方法逐渐成为主流全身控制实现方式。Kim 等^{Kim2019}提出 Whole-Body Locomotion Controller (WBLC)，采用分层优化结构统一处理质心控制、摆动足轨迹与接触约束，并通过软约束与权重调节机制实现平滑接触切换，为动态步行提供了实验验证。近年来，Paredes 等^{Paredes2023}在逆动力学 QP 框架中引入加速度形式的指数控制屏障函数（Acceleration-based Exponential Control Barrier Function, A-ECBF），在保持任务性能的同时显式保证安全约束不被破坏，进一步拓展了 HQP 在安全全身控制中的应用范围。

总体而言，分层二次规划方法在理论上继承了任务优先级控制的层级思想，在工程实现上则通过优化框架自然融合动力学约束、接触条件与不等式限制，使其成为现代人形机器人腿臂协同控制的重要技术路径。

在分层二次规划（HQP）框架中，多任务控制被统一表示为带约束的优化问题。对于单一任务，其基本形式为：

$$\min_z \|Az - b\|^2, \quad (1.4)$$

其中 z 为优化变量（如 \dot{q} 或 \ddot{q} ）， A 为任务雅可比矩阵， b 为期望任务量。

对于分层结构，设第 k 层任务为 (A_k, b_k) ，则 HQP 递归形式为：

$$z_k^* = \arg \min_z \|A_k z - b_k\|^2, \quad \text{s.t. } z \in \mathcal{S}_{k-1}, \quad (1.5)$$

其中 \mathcal{S}_{k-1} 表示满足第 1 至 $k-1$ 层任务最优解的可行解集合。该结构保证高优先级任务严格保持，而低优先级任务仅在其零干扰空间内优化。

在全身控制中，系统动力学方程作为等式约束引入：

$$M(q)\ddot{q} + h(q, \dot{q}) = S^\top \tau + J_c^\top f_c, \quad (1.6)$$

并可进一步加入接触与摩擦等不等式约束，从而形成完整的分层逆动力学优化框架。

综合近年来分层二次规划方法在人形机器人中的发展与应用，可以看到其在腿臂协同控制问题中形成了一种较为成熟的结构范式。该类方法通常以质心或动量控制、接触一致性作为最高优先级任务，在此基础上逐层嵌入躯干姿态调节与手臂末端操作任务，并通过不等式约束显式处理关节限位与力矩限制。这种分层逆动力学结构在多个扭矩控制人形机器人平台上得到验证，成为当前模型驱动全身控制的主流实现方式。

从方法特性来看，基于 HQP 的全身控制框架具有若干显著优势。首先，其优化结构能够在统一框架内同时处理等式与不等式约束，使接触力、摩擦锥、零力矩点（ZMP）以及关节力矩限制等物理约束得到显式表达，从而提高系统的安全性和物理一致性。其次，严格的层级优化机制保证高优先级稳定性任务不被低优先级操作任务破坏，使腿部支撑与手臂操作之间形成清晰的协调关系。此外，逆动力学形式下的优化变量通常包含关节加速度与接触力，使力控制与运动控制能

够在同一求解框架中实现统一协调，这对于多接触、高冗余的人形系统尤为重要。

然而，该方法同样存在一定局限性。其一，HQP 框架高度依赖精确的动力学模型与接触建模，当模型误差、结构柔性或环境不确定性较大时，控制性能可能显著下降。其二，多层优化问题的实时求解带来较高的计算开销，随着自由度与约束数量增加，计算复杂度迅速提升，对控制频率与硬件算力提出较高要求。其三，权重矩阵与松弛变量的设计往往依赖人工经验，不同任务组合与场景下需要反复调节参数，增加了工程实现复杂度。其四，优先级结构通常在设计阶段固定，缺乏在线自适应调整能力，在高度动态或接触频繁切换的场景中可能出现层级冲突或解空间收缩现象。尤其在腿臂强耦合的高速运动过程中，固定层级结构可能限制全身协调的灵活性。

因此，尽管 HQP 方法在结构化建模与物理约束表达方面具有显著优势，是当前模型驱动全身控制的重要技术路径，但其在模型依赖性、参数调节复杂性以及泛化能力方面仍存在改进空间。这些局限性也为后续结合数据驱动与学习方法的研究提供了重要动机，使控制策略能够在保持物理一致性的同时具备更强的自适应与泛化能力。

1.3.3 基于强化学习的腿臂协同控制

随着大规模并行仿真技术的发展，强化学习（Reinforcement Learning, RL）逐渐成为类人机器人全身控制的重要研究方向。不同于基于模型的解析优先级方法，强化学习通过与环境交互并最大化长期累积回报，在高维状态-动作空间中直接搜索控制策略，使复杂动力学耦合关系在训练过程中以涌现的形式形成。

基于强化学习的类人机器人控制最早主要集中于双足步态生成问题。针对 Cassie 等双足机器人，大量研究表明，仅通过精心设计的奖励函数，即可在无显式参考轨迹的情况下学习到站立、行走、跑步与跳跃等多种步态，并实现稳定的仿真到实机迁移^[78]。这一阶段的研究奠定了“基于奖励设计驱动步态生成”的技术路线，证明了在高度非线性动力学系统中，强化学习能够替代传统模型控制方法学习出稳定、鲁棒的运动行为。

随着类人机器人自由度的提升以及动力学仿真能力的增强，强化学习逐渐从下肢行走扩展至全身控制。Gu 等提出的 Humanoid-Gym 框架基于 Isaac Gym 构建大规模并行仿真环境，采用 PPO 算法训练高自由度类人机器人行走策略，并在多种机器人平台上实现零样本迁移^[79]。该框架展示了强化学习在高维全身控制问题中的可扩展性与工程可行性。

在此基础上，研究者进一步探索强化学习在更复杂全身运动任务中的应用，逐渐形成了从单一行走技能向多步态统一控制、高动态全身动作以及复杂多接触交

互任务扩展的发展趋势。Siekmann 等^[78] 提出统一奖励结构下的多步态学习方法，通过在同一策略中嵌入速度指令与步态调度变量，使机器人在无显式轨迹参考的情况下实现站立、行走、跑步与跳跃等多种步态的连续切换。该工作证明了强化学习策略能够在单一网络结构中内化多种动态模式，实现跨步态的统一控制。在高动态动作学习方面，Xie 等提出 PBHC (Phase-Based Hybrid Control) 框架^[80]，通过在强化学习中引入相位变量与混合控制结构，增强策略对高速摆动与冲击接触的建模能力，从而提升翻滚、腾空等高动态全身动作的稳定性。该方法在保留学习灵活性的同时，引入结构化先验以缓解训练难度，体现出“结构约束 + 数据驱动”结合的发展趋势。针对复杂多接触任务，Zhang 等提出 WoCoCo (Whole-Body Contact Control) 框架^[81]，通过在强化学习中显式引入接触模式编码与任务阶段划分，使策略能够在多接触序列（如攀爬、支撑转换等）中学习稳定的接触力分配与全身协调机制。该工作拓展了强化学习在多接触操作与复杂地形交互中的应用范围。Fu 等提出 HumanPlus 系统^[82]，通过域随机化、残差学习与控制器融合机制，实现高自由度类人机器人在真实环境中的稳定全身控制。该系统验证了强化学习策略在复杂机械结构与实际传感噪声条件下的可迁移性与工程可行性。这些工作共同表明，强化学习能够在不显式构建腿臂协调规则的情况下，自主学习复杂全身运动模式。

在强化学习框架下，类人机器人控制问题通常被建模为马尔可夫决策过程 (Markov Decision Process, MDP)：

$$\mathcal{M} = \langle \mathcal{S}, \mathcal{A}, \mathcal{T}, r, \gamma \rangle, \quad (1.7)$$

其中， \mathcal{S} 为状态空间， \mathcal{A} 为动作空间， \mathcal{T} 为状态转移概率函数， r 为即时奖励函数， $\gamma \in (0, 1)$ 为折扣因子。

在离散时间步 t ，机器人处于状态 $s_t \in \mathcal{S}$ ，根据策略 $\pi_\theta(a_t | s_t)$ 选择动作 $a_t \in \mathcal{A}$ ，环境根据转移概率 \mathcal{T} 演化至下一个状态 s_{t+1} ，并反馈奖励 $r_t = r(s_t, a_t)$ 。策略参数 θ 通过最大化期望累积折扣回报进行优化：

$$J(\theta) = \mathbb{E}_{\pi_\theta} \left[\sum_{t=0}^{\infty} \gamma^t r_t \right], \quad (1.8)$$

其中期望是对策略 π_θ 下产生的轨迹分布进行统计得到的。

图 1-5 所示为基于强化学习的腿臂协同控制框架。该框架可划分为“训练过程”和“执行过程”两个阶段。在训练阶段，机器人根据当前状态 s_t 通过策略 π_θ 产生动作 a_t ，环境反馈新的状态与奖励 r_t 。奖励函数通常由多个子项加权组合而成：

$$r_t = \sum_i w_i r_t^{(i)}, \quad (1.9)$$

其中 $r_t^{(i)}$ 表示第 i 个子奖励项，用于刻画特定控制目标（如速度跟踪、姿态稳定、能量消耗或接触约束等）， $w_i \in \mathbb{R}$ 为对应的权重系数，用于调节各子目标在总体奖励中的相对重要性。权重 w_i 通常由人工设计或通过经验调节确定。不同奖励项分别刻画稳定性、步态性能、角动量调节、摆臂协调以及能耗等因素。策略优化模块通过最大化累计折扣回报 $J(\theta)$ 更新参数。

在腿臂协同控制问题中，常见奖励构成包括：

- 质心速度或姿态跟踪奖励；
- 足端接触一致性或周期性奖励；
- 角动量抑制或躯干姿态稳定奖励；
- 手臂摆动对称性或相位协调奖励；
- 能耗与关节速度正则化项。

在执行阶段，训练得到的策略直接根据实时状态输出控制动作，实现闭环控制。

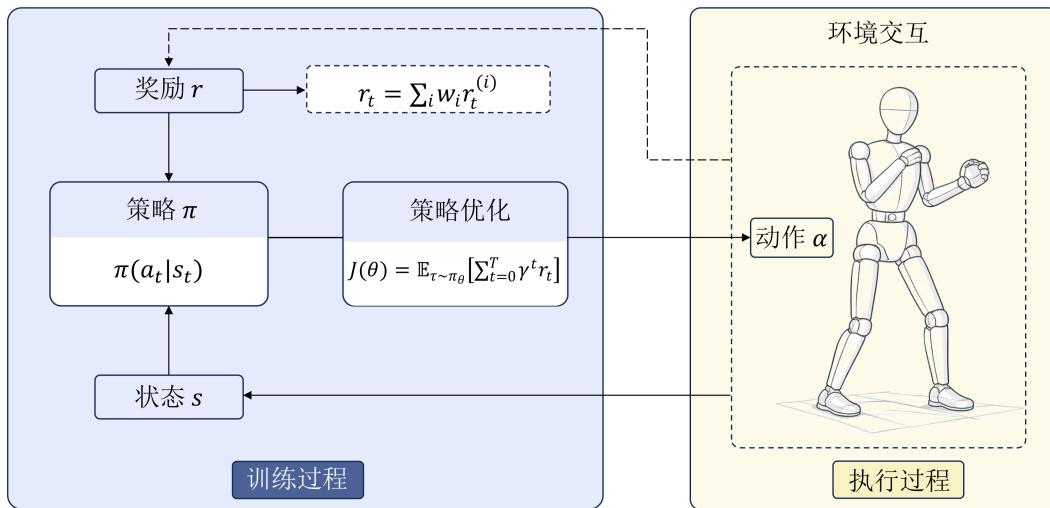


图 1-5 基于强化学习的腿臂协同控制框架示意图

综合现有研究可以发现，基于强化学习的腿臂协同控制方法在方法范式上呈现出明显区别于解析优先级控制框架的技术特征。其核心思想在于通过奖励函数的加权组合，在统一的优化目标下对稳定性、任务性能与能耗等因素进行整体建模，使不同控制目标在同一策略参数空间内协同优化。这种“隐式优先级”机制并非通过显式层级结构实现，而是通过长期回报最大化过程在高维动力学系统中自

动形成腿部支撑与手臂动作之间的协调关系。

从工程实践角度看，该范式具有若干显著优势。首先，强化学习能够直接在高维、强非线性的动力学系统中搜索控制策略，对复杂接触序列与频繁支撑切换具有较强适应能力。在多接触或高动态动作场景中，策略可以通过大量仿真交互学习到稳定的全身协调模式，而无需显式推导接触力分配规则。其次，依托大规模并行仿真与域随机化技术，策略在训练阶段可暴露于多种扰动与参数变化，从而获得一定程度的鲁棒性与跨平台迁移能力。此外，通过奖励结构设计与条件变量引入，单一策略网络可以统一编码多步态与多技能行为，使全身动作生成在形式上实现一体化。

然而，强化学习方法同样存在不可忽视的局限性。其一，不同控制目标被压缩为奖励加权结构后，其相对重要性依赖人工权重设计，缺乏类似零空间或 HQP 方法所具有的严格层级保障。当奖励项之间存在潜在冲突时，策略可能在训练过程中形成不可预测的折中行为。其二，强化学习训练通常需要大量仿真交互与计算资源，尤其在高自由度类人系统中，训练成本与调参复杂度较高。其三，硬约束（如接触力上界、关节力矩限制或安全边界）难以通过纯奖励形式严格保证，在安全性要求较高的实际部署场景中，往往仍需结合模型约束或安全过滤机制进行补偿。其四，由于协同关系以隐式形式存在于策略参数中，其结构解释性与理论可证明性相对较弱，难以像解析方法那样给出明确的优先级与稳定性保证。

因此，基于强化学习的腿臂协同控制方法在复杂动力学建模能力与自适应性方面具有显著优势，但在约束可控性、训练效率以及结构可解释性方面仍存在一定挑战。这些特性也推动近年来研究逐渐探索模型驱动方法与数据驱动方法的融合框架，以期在稳定性保障与自适应能力之间取得平衡。

1.3.4 基于模仿学习的腿臂协同控制

模仿学习（Imitation Learning, IL）的思想最早来源于机器人学习与控制中的示教学习问题^[83]。其核心目标是在不给定明确代价函数或系统模型的情况下，通过专家示范数据直接学习控制策略。早期方法主要基于监督学习框架，通过回归或分类模型逼近专家策略，即行为克隆^[84-85]。在该阶段，模仿学习主要应用于机械臂轨迹复现与简单运动控制问题，其优势在于实现简单、训练稳定，但在长时序控制与状态分布偏移问题上存在明显局限。

随着深度学习的发展以及大规模动作数据的获取，模仿学习逐渐从“轨迹复制”扩展至“运动分布建模”。Ho 与 Ermon 提出的生成对抗模仿学习（Generative Adversarial Imitation Learning, GAIL）^[86] 将策略学习转化为专家与策略轨迹分布之间的对抗匹配问题，为后续对抗式运动先验方法奠定了理论基础。在物理仿真角

色控制领域, Peng 等提出的 Adversarial Motion Priors (AMP)^[87] 首次将对抗式分布约束引入物理仿真控制框架, 通过训练判别器学习人类运动分布, 并将其输出转化为奖励信号, 使策略在强化学习过程中逐步逼近人类动作统计特征。该方法证明, 在不显式构造复杂奖励函数的情况下, 机器人可以生成具有自然节律与协调结构的全身运动。

随后, Adversarial Skill Embedding(ASE)^[88] 将对抗式运动先验扩展至多技能统一控制框架, 通过构建共享潜在技能空间, 使单一策略能够编码多种运动模式。这一阶段的重要进展在于将“模仿单一动作”拓展为“学习可组合的运动分布”, 为复杂全身行为的统一表达提供了结构化表示。

在人形机器人领域, 模仿学习逐渐与强化学习相结合, 形成“模仿 + 强化”的混合训练范式。一方面, Adaptive Humanoid Control (AHC) 等多行为蒸馏框架^[89] 通过多专家策略训练与蒸馏融合, 实现起身、行走等多技能统一控制, 并在蒸馏过程中保留不同动作中的腿臂协同模式; 另一方面, 基于生成模型的运动潜空间方法, 如 Motion Variational Autoencoder (MVAE)^[90], 通过变分自编码器 (Variational Autoencoder, VAE) 学习低维运动潜在空间, 使控制问题在结构化动作空间中进行优化, 从而在保证运动自然性的同时提高训练稳定性与可控性。

随着高精度动作捕捉数据集 (如 AMASS) 与高保真物理仿真平台的发展, 模仿学习方法逐渐能够在复杂接触、多阶段动作与高动态场景下实现全身控制策略训练, 并成功部署于真实人形机器人平台。这一阶段的核心贡献在于: 将人类运动分布作为结构先验引入控制问题, 使腿部推进、躯干稳定与手臂辅助之间的协同关系通过数据分布本身得到编码, 而无需显式构造耦合规则。

从发展脉络来看, 模仿学习在人形机器人腿臂协同控制中的演进可以概括为三个阶段: 第一阶段为基于监督回归的动作复制; 第二阶段为基于分布匹配的对抗式运动先验; 第三阶段为基于潜空间建模的生成式运动表示。在这一过程中, 腿臂协同结构逐渐从“被动复制的轨迹模式”演化为“可泛化的结构化运动分布表示”, 为复杂全身控制提供了重要的数据驱动支撑。

模仿学习的核心目标是使策略 $\pi_\theta(a|s)$ 逼近专家策略 $\pi_E(a|s)$, 从而在状态空间中复现专家行为。根据建模方式的不同, 现有方法大致可分为三类: 基于监督回归的行为克隆方法、基于分布匹配的对抗式模仿学习方法, 以及基于潜在空间建模的生成式运动先验方法。

1) 行为克隆 在监督学习框架下, 策略通过最小化专家动作回归误差进行训练, 其基本形式为:

$$\mathcal{L}_{BC}(\theta) = \mathbb{E}_{(s, a^*) \sim \mathcal{D}} [\|\pi_\theta(s) - a^*\|^2], \quad (1.10)$$

其中 \mathcal{D} 为专家数据集, a^* 为专家动作。行为克隆的核心思想是通过函数逼近直接复制专家策略。该方法实现简单、训练稳定, 但由于仅在专家状态分布上优化, 当执行过程中状态分布发生偏移时, 误差可能逐步累积, 影响长期控制性能。

2) 对抗式模仿学习 对抗式模仿学习通过匹配专家轨迹分布与策略生成轨迹分布进行训练。典型方法如 AMP^[87] 引入判别器 D_ϕ , 并构造如下对抗目标:

$$\mathcal{L}_D = \mathbb{E}_{s_E} [\log D_\phi(s_E)] + \mathbb{E}_{s_\pi} [\log(1 - D_\phi(s_\pi))]. \quad (1.11)$$

判别器用于区分专家状态 s_E 与策略生成状态 s_π , 并将判别结果转化为训练信号。其核心思想并非逐帧回归专家动作, 而是在分布层面对齐专家与策略行为, 使策略在统计意义上继承人类运动结构, 从而在物理仿真环境中自然形成腿臂协同模式。

3) 生成式运动先验 生成式运动先验方法通过学习低维潜在空间对复杂运动分布进行结构化建模。以基于变分自编码器的 MVAE^[90] 为例, 其运动生成过程可表示为:

$$p_\theta(x_{t+1} | x_t, z), \quad (1.12)$$

其中 z 为潜变量, 用于调节运动演化方向。通过在潜在空间中进行控制或优化, 原本高维复杂的动作生成问题被转化为对潜变量的调节过程。由于人类运动中的腿臂协同关系已在潜空间中被编码, 控制过程中无需显式构造耦合约束即可生成自然协调的全身动作。

为更加系统地分析模仿学习在腿臂协同控制中的技术特征, 有必要从结构继承能力、泛化性能以及约束表达能力等方面进行综合讨论。

首先, 在动作自然性与协同结构形成方面, 模仿学习具有显著优势。由于人类运动数据中天然包含腿部推进、躯干稳定与手臂辅助之间的动力学耦合关系, 策略在学习运动分布的过程中能够隐式继承这种协调结构, 而无需显式构建腿臂耦合规则或复杂奖励函数。相比纯强化学习依赖人工奖励调节的方式, 模仿学习在动作节律保持、动量分配一致性以及整体姿态连贯性方面通常表现更加自然与稳

定。此外，生成式潜空间方法通过构建结构化运动表示，使高维动作生成过程受限于低维潜变量调节，从而提升训练稳定性并减少策略搜索空间的复杂度。

其次，在训练效率与工程实现层面，模仿学习能够利用离线运动数据进行预训练，降低在线探索成本，并可作为强化学习的运动先验模块使用。这种“模仿+强化”的混合范式在复杂全身控制任务中具有较强实用价值，有助于缩短训练周期并提升策略初始稳定性。

然而，模仿学习方法同样存在若干局限。其一，方法性能高度依赖示范数据的质量与覆盖范围，当目标任务超出数据分布支持区域时，策略可能出现退化或失稳。其二，由于模仿学习主要关注运动分布匹配，动力学约束与接触一致性通常未被显式建模，在复杂多接触或强扰动环境中可能缺乏严格的稳定性保证。其三，虽然腿臂协同关系可以在数据中被继承，但其物理形成机制并未显式刻画，协同结构缺乏明确的理论可解释性与层级控制结构。

综合来看，模仿学习在腿臂协同控制中的核心价值在于结构继承，即通过数据驱动方式将人类运动中的协调模式直接编码于策略之中。然而，在泛化能力、约束显式建模以及可解释性方面仍存在改进空间。

表 1-1 四类腿臂协同控制方法的优先级建模机制与特性对比

方法类别	优先级建模方式	主要优势	主要局限
零空间分解方法	显式结构优先级	层级关系清晰；高优先级任务严格满足；物理意义明确；实时性较好	难处理不等式与复杂接触约束；优先级结构固定；高动态场景扩展性有限
分层 QP(HQP)方法	优化式优先级	可统一处理等式与不等式约束；可纳入动力学与接触力约束；数值稳定性较高；适合多接触控制	依赖精确模型；计算复杂度较高；权重与松弛变量需人工调节；层级结构缺乏自适应性
强化学习方法	隐式奖励优先级	统一优化框架；可处理强非线性与频繁接触切换；多技能可在单策略中涌现；环境适应能力强	高度依赖奖励设计；难以严格保证硬约束；训练成本高；协同结构缺乏显式可解释性
模仿学习方法	数据分布优先级	继承人类运动分布中的协同结构；动作自然性高；可作为运动先验提升训练稳定性	依赖数据质量与覆盖范围；泛化能力受限；动力学与安全约束缺乏显式保障

综合表 1-1 可见，从优先级建模机制的角度来看，模型驱动方法强调显式结构

分层与约束表达，强化学习强调奖励层面的整体优化，而模仿学习则通过数据分布继承协同结构。不同方法在优先级表达形式上的差异，直接影响其在自然性、泛化能力与可解释性方面的表现。

具体而言，模型驱动方法在任务优先级刻画与物理约束保障方面具有显著优势，但在复杂高动态场景与多阶段行为生成问题中扩展性相对有限；强化学习方法能够在统一优化框架下处理强非线性动力学与复杂接触切换问题，但协同关系主要通过奖励加权隐式形成，其泛化能力与稳定性往往依赖大量训练与精细调参；模仿学习方法则在结构继承方面表现突出，但对数据分布依赖较强，在未见初始姿态或复杂环境下的泛化能力仍存在局限。

因此，在面向人形机器人全身控制与腿臂协同控制的实际需求时，单一技术路线往往难以同时兼顾自然性、泛化能力与稳定性保障。

基于上述分析，本文采用强化学习与模仿学习相结合的混合训练框架。首先，在具备参考轨迹的条件下，通过强化学习训练策略，使机器人在物理约束与环境交互过程中学习逼近目标动作，从而获得动力学一致且自然协调的全身行为。然而，该阶段所得策略通常对初始姿态较为敏感，泛化能力有限。

在此基础上，引入模仿学习进行策略融合与结构扩展。例如，将任意姿态到达起身预备姿态的策略与标准起身动作策略进行融合，使控制器能够覆盖更广泛的初始状态空间，实现任意姿态下的稳定起身。同时，在融合阶段引入地形随机化与扰动增强机制，可进一步提升策略在复杂环境中的鲁棒性。

该两阶段框架结合了强化学习在动力学一致性与复杂接触处理方面的优势，以及模仿学习在结构继承与行为扩展方面的优势，为实现自然、稳定且具备泛化能力的腿臂协同控制提供了一种有效路径。

1.4 本文的主要问题

在构建基于任务优先级的腿臂协同控制总体框架之后，本文进一步围绕影响全身协调控制性能与工程落地能力的关键因素，凝练出以下几个核心问题：

(1) 如何获取高质量且物理可行的全身参考动作数据？

腿臂协同控制的基础在于高质量的全身动作参考。然而，人类动作捕捉数据与机器人结构之间在自由度配置、关节限位、质量分布以及驱动能力等方面存在显著差异，导致动作无法直接迁移。因此，如何在保持人类运动自然性与协调性的同时，满足机器人动力学约束与物理可实现性要求，构建高精度、稳定的重定向算法，是实现高质量全身控制的首要问题。若重定向质量不足，将直接影响后续策略学习的效果与收敛稳定性。

(2) 如何在强化学习过程中避免训练偏置，提升复杂动作的学习能力？

在基于强化学习的全身控制策略训练中，简单动作往往更容易获得稳定奖励反馈，从而被策略反复强化；而复杂或高难度动作由于奖励稀疏或梯度波动较大，容易被忽略，导致策略出现动作单一、模式塌陷或泛化能力不足等问题。因此，如何构建合理的课程学习机制，对动作进行难度分级与分布调控，并在风格保持与任务目标之间取得平衡，是保证整体运动质量与表达能力的关键挑战。

(3) 如何减小绳驱动人形机器人中的 Sim-to-Real 差异？

绳驱动系统具有柔顺性强、传动耦合复杂以及非线性特性显著等特点。绳索弹性、摩擦损耗、多级滑轮传动以及电机—关节映射关系，都会导致仿真模型与真实系统之间存在差异。若缺乏准确建模与补偿机制，强化学习策略在仿真环境中的性能难以稳定迁移至真实平台。因此，如何建立合理的驱动映射模型，刻画耦合动力学特性，并通过系统辨识与域随机化等方法提升策略的迁移能力，是实现工程落地的核心问题之一。

(4) 如何实现复杂功能性任务在多场景下的鲁棒泛化？

对于起身、坐起等功能性任务，其运动过程通常涉及多阶段接触切换、大范围姿态变化以及高稳定性要求。传统方法往往针对特定初始姿态单独训练策略，缺乏跨姿态与跨地形的泛化能力。因此，如何结合模仿学习提供的动作先验与强化学习的自适应优化能力，实现复杂任务在不同初始条件、地形扰动以及接触不确定性下的稳定执行，是实现高可靠性全身控制的重要问题。

综上，本文围绕“动作数据构建—策略训练优化——驱动系统一致性建模—任务鲁棒泛化”这一完整技术链条展开研究，旨在构建兼具运动自然性、控制稳定性与工程可部署性的类人机器人腿臂协同控制方法。

1.5 本文研究内容及章节安排

本文围绕人形机器人在复杂全身运动任务中的腿臂协同控制问题展开研究，重点探讨如何在人体与机器人存在显著结构与动力学差异的条件下，通过运动重定向与模仿强化学习方法，实现具有隐式任务优先级特性的全身控制策略。全文的研究内容主要分为三个层次：理论基础与问题建模、基于模仿学习的全身控制方法设计，以及仿真与真实机器人系统上的实验验证。

第2章研究人形机器人全身动作重定向问题，重点解决人体与机器人在结构、自由度及运动约束方面存在差异所导致的动作失真问题。针对仅在关节空间或笛卡尔空间进行重定向各自存在的局限性，本章提出了一种关节空间与笛卡尔空间相融合的动作表示方式，在保证动作自然性与可行性的同时，提高末端执行精度

与接触一致性。在此基础上，构建了融合多种约束条件的动作优化模型，并通过定量与定性分析对重定向动作的质量进行了系统评估。本章为后续模仿学习阶段提供了高质量、可控的参考动作。

第3章提出了一种基于强化学习的全身控制框架，是全文的核心内容之一。本章首先介绍了整体学习流程，包括参考动作数据的重定向、状态与动作空间设计以及控制频率设置。随后，提出了一种极简但信息密度较高的奖励设计方法，仅依赖少量关键奖励项即可引导智能体学习稳定且自然的全身运动。针对复杂动作中不同阶段难度分布不均的问题，本章进一步提出了一种难点采样的参考动作调度策略，通过动态调整参考动作中关键阶段的采样概率，有效提升了训练稳定性与收敛速度，并通过对比实验进行了验证。

第4章围绕论文题目，系统分析了基于任务优先级的腿臂协同控制机制。首先对人形机器人在全身运动任务中的腿臂协同任务进行建模，明确腿部支撑、上肢支撑与平衡调节在不同动作阶段中的功能分工。随后，从学习结果的角度分析隐式任务优先级在策略中的体现方式，包括奖励函数中隐含的优先级关系以及动作阶段切换过程中的协调行为。通过对摔倒后爬起、坐下与站起等典型全身任务的分析，对比是否使用上肢参与支撑的控制策略，展示了所提出方法在复杂任务中实现自然腿臂协同控制的能力。

第5章进一步在真实双足人形机器人平台上对所提出方法进行实验验证。首先介绍实验硬件系统与控制架构，并说明学习到的策略在真实系统中的部署方式。随后，通过爬起、坐站等典型高风险全身动作实验，从成功率、鲁棒性及动作协调性等方面评估策略在真实环境中的表现，验证所提方法的实际可行性。

第6章对全文工作进行总结，并讨论了本文方法的局限性以及未来在复杂交互任务、自主技能扩展和人机协作等方向上的进一步研究展望。

第二章 人形机器人全身动作重定向

2.1 引言

人机交互（Human-Robot Interaction, HRI）是仿人机器人研究中的重要方向，而运动模仿则是实现自然、高效交互行为的核心手段之一。由于人体与机器人在运动学结构、关节自由度配置以及运动范围等方面存在显著差异，如何将人体动作合理映射到机器人平台上，仍然是运动模仿与动作重定向中的关键挑战。

现有的动作重定向方法通常可分为基于关节空间和基于笛卡尔空间的两类。前者侧重于保持人体与机器人在关节构型层面的相似性，但往往难以保证末端执行器在任务空间中的精确运动；后者则以末端轨迹跟踪为主要目标，但容易忽略整体肢体构型的一致性，从而导致机器人运动不自然。这种构型一致性与任务空间精度之间的权衡，限制了重定向结果在复杂交互场景中的适用性。

针对上述问题，本章在已有动作重定向研究的基础上，对传统重定向框架进行了改进，通过在关节空间构型约束与笛卡尔空间末端跟踪之间建立协调机制，实现了对人体上肢动作的有效映射。在保证机器人手臂整体构型与人体动作保持一致的同时，该方法能够显著提升末端执行器位置与姿态的跟踪精度，从而兼顾运动自然性与任务可执行性。实验结果表明，该重定向方法在不同动作类型与操作场景下均表现出良好的稳定性与适应性，为后续结合学习方法进行全身动作生成与控制奠定了基础。

近年来，利用人体运动数据对机器人进行示教与运动生成，已成为简化机器人运动编程与学习过程的一种重要途径^[91-94]。通过观测和映射人类动作，机器人能够在无需显式建模复杂任务逻辑的情况下获得具有较强表现力和适应性的运动行为，这对于提升仿人机器人的自然性、自主性以及人机交互能力具有重要意义^[95-96]。尽管该方向已取得一定进展^[97-98]，但在结构受限的真实机器人平台上实现稳定且精确的人体运动模仿，仍然面临诸多挑战^[99]。

从运动映射的角度来看，人体到机器人的动作模仿通常可抽象为一个动作重定向（*motion retargeting*）问题，即在满足机器人自身运动学与动力学约束的前提下，将人体动作合理映射到机器人关节空间或任务空间中。由于人体与机器人在自由度数量、关节拓扑结构、运动范围以及力矩能力等方面存在显著差异，人体动作往往无法被机器人直接复现，这使得动作重定向成为运动模仿中的核心问题之一。

现有的动作重定向方法大致可以分为基于学习的方法和基于模型的方法。在

动画与虚拟角色控制领域，强化学习被广泛用于从人体动作数据中学习复杂运动策略，并在多种任务中展现出良好的泛化能力与风格一致性^[100-103]。然而，将此类方法直接应用于真实仿人机器人仍然面临仿真到现实差距（sim-to-real gap）的问题，尤其是在关节力矩限制、硬件安全性以及高精度末端控制等方面存在明显挑战^[104-105]。

相比之下，基于模型的动作重定向方法由于其可解释性强、对物理约束处理明确，在真实机器人系统中仍被广泛采用。根据映射空间的不同，这类方法通常可进一步分为基于笛卡尔空间和基于关节空间的重定向策略。基于笛卡尔空间的方法通常以末端执行器位姿跟踪为主要目标，通过逆运动学或优化方法实现任务空间误差最小化^[106-107]。这类方法在任务执行精度方面具有优势，但由于缺乏对整体肢体构型的约束，容易导致机器人运动不自然，且在存在避障或关节限制时可行性较差。相对而言，基于关节空间的方法更关注机器人与人体在构型层面的相似性，能够生成较为自然的运动，但往往难以保证末端执行器在任务空间中的精确跟踪，从而限制了其在精细操作任务中的应用。

因此，如何在动作重定向过程中同时兼顾关节构型一致性与末端执行器的任务空间精度，成为当前类人机器人运动模仿研究中的一个关键问题。现有方法通常在这两者之间进行权衡，尚难以在统一框架下有效协调二者。

基于上述背景，本章在已有动作重定向研究的基础上，对传统的重定向流程进行了系统梳理与分析，并引入一种结合关节空间构型约束与笛卡尔空间任务优化的重定向建模思路。该思路以人体手臂运动为示教源，通过建立人体与机器人之间的几何对应关系，在满足机器人关节空间约束的前提下，引入基于笛卡尔空间的优化机制，对末端执行器运动进行精确调节，从而在构型一致性与任务精度之间取得更为平衡的解。

需要强调的是，本章关注的重点在于动作重定向问题本身的建模方式与约束处理机制，其目的在于为后续章节中基于学习的全身动作生成与控制策略提供稳定、合理的低层动作映射基础。

2.2 方法

2.2.1 动作模仿系统设计

图 ?? 展示了本文所构建的机器人机械臂动作模仿系统整体框架。系统主要由视觉采集系统、上位处理器、下位处理器以及机械运动系统四个部分组成，各模块协同工作以实现人体上肢动作向机器人机械臂的实时重定向。

视觉采集系统采用标准的 RGB-D 摄像设备，本文实验中使用 Intel RealSense

D435 深度相机。摄像头安装于被试者正前方，距地面约 1.5 m，用于实时获取人体运动图像及对应的深度信息。采集到的数据通过 USB 接口传输至上位处理器进行处理。

上位处理器为一台 PC，负责完成人体姿态信息的提取与预处理。具体而言，上位处理器对输入的 RGB 图像进行人体姿态估计，提取人体手臂关键点的二维信息，并进一步结合深度数据恢复三维关键点位置。为降低传感噪声对后续动作重定向的影响，本文对三维关键点序列引入卡尔曼滤波进行平滑处理。处理后的关键点数据通过 ROS 通信机制发送至下位处理器。

下位处理器用于执行动作重定向计算与机器人控制指令生成。在接收到人体三维关键点数据后，系统首先基于人体手臂的几何模型计算对应的关节角度，该关节角度作为机器人动作重定向问题的初始参考。随后，通过迭代优化求解满足机器人关节约束并同时逼近人体末端运动目标的关节角解。最终得到的目标关节角通过 CAN 总线发送至机械运动系统，用于驱动机器人机械臂执行相应动作。

机械运动系统由机器人双臂、电机及其驱动模块构成。本文所使用的仿人机器人平台为自主研发系统，如图 2-1 所示。每条机械臂具有六个自由度，关节配置包括肩关节屈伸、肩关节外展内收、肩关节旋内外、肘关节屈伸、前臂旋前旋后以及腕关节屈伸。该结构能够满足机械臂在三维空间内完成复杂动作的运动学需求，为人体动作的有效重定向提供了硬件基础。

2.3 改进重定向

由于肩关节和肘关节主要影响手臂的运动，而手腕主要影响手的运动，本研究重点对肩关节和肘关节进行几何建模。为清楚说明肩关节和肘关节的建模方法，下面以左臂的建模过程进行说明。

如图 2-2 所示，点 S 表示左臂的肩关节，点 E 表示左臂的肘关节，点 M 则位于 Y 轴上，距离为 a 的一点。

根据上述定义，可以证明 $|SM| = a$ 。通过姿态估计，可以获得各关键点在相机坐标系下的坐标，分别为 $S(x_0, y_0, z_0)$ 、 $E(x_1, y_1, z_1)$ 以及 $M(x_0, y_0 - a, z_0)$ 。以点 S 为原点建立肩关节局部坐标系后，各点在该坐标系下的坐标可表示为 $S(0, 0, 0)$ 、 $E(x_1 - x_0, y_1 - y_0, z_1 - z_0)$ 以及 $M(0, -a, 0)$ 。

点 E 在平面 (ZSM) 上的正交投影记为点 F ，其坐标为 $F(0, y_0 - y_1, z_1 - z_0)$ 。向量 SE 在平面 (ZSM) 上的投影与 Y 轴负方向之间的夹角定义为 θ_1 ，即 $\angle FSM = \theta_1$ ，其中 θ_1 表示肩关节的屈伸角；向量 SE 与其在平面 (ZSM) 上投影之间的夹角定义为 θ_2 ，即 $\angle FSE = \theta_2$ ，其中 θ_2 表示肩关节的外展-内收角。

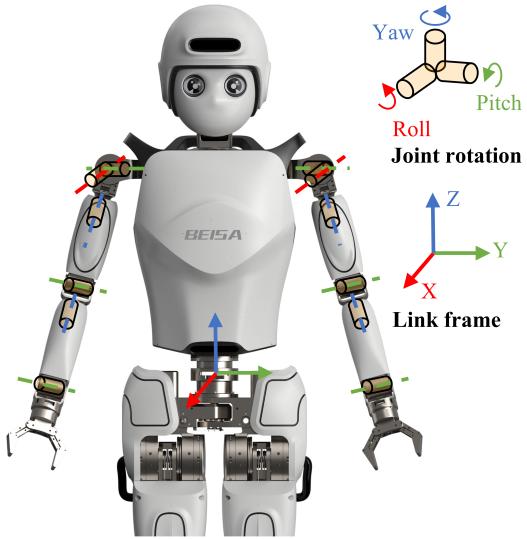


图 2-1 Joint structures and link frames of the robot's arm.

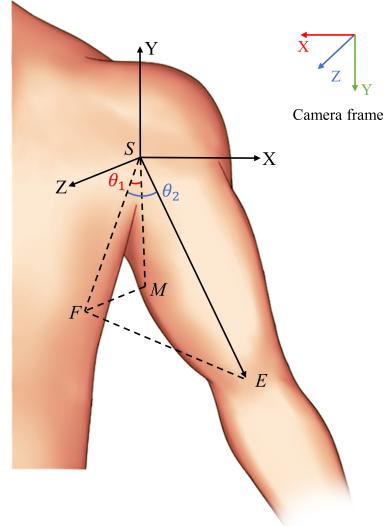


图 2-2 左手肩关节建模。

在上述几何关系下，可在三角形 $\triangle FSM$ 中求解角度 θ_1 ，根据余弦定理有

$$\theta_1 = \arccos\left(\frac{|SF|^2 + |SM|^2 - |FM|^2}{2|SF||SM|}\right) = \arccos\left(\frac{y_0 - y_1}{|SF|}\right), \quad (2.1)$$

其中角度 θ_1 的符号可由 $z_1 - z_0$ 的符号确定：当 $z_1 - z_0 > 0$ 时，表示手臂向前摆动；当 $z_1 - z_0 < 0$ 时，表示手臂向后摆动。同理，在三角形 $\triangle FSE$ 中可求解肩关节外展-内收角 θ_2 ，其表达式为

$$\theta_2 = \arccos\left(\frac{|SF|^2 + |SE|^2 - |FE|^2}{2|SF||SE|}\right). \quad (2.2)$$

2.3.1 人体肘关节建模

如图 2-3(a) 所示， W 表示手腕的实际位置， W' 表示手腕的假定初始位置（此时 $\theta_3 = 0$ ）。角度 θ_3 定义为从肘部指向手腕的向量 EW 与其初始位置向量 EW' 之间的夹角；角度 θ_4 定义为向量 EW 与向量 SE 的延长线之间的夹角。

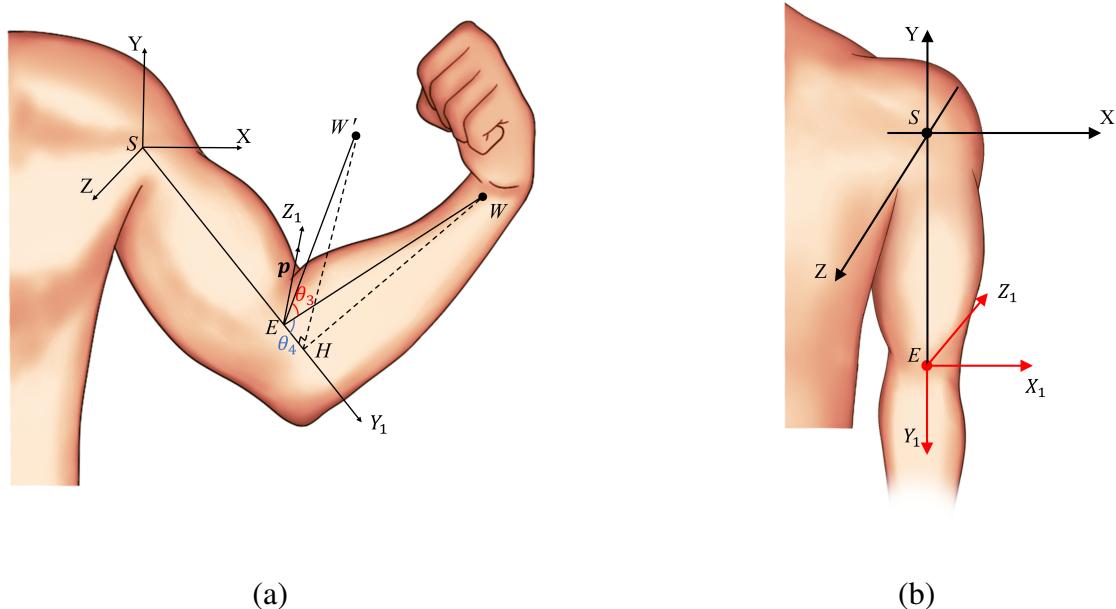


图 2-3 肘关节建模。(a) 当手臂关节角不为零时的肘关节坐标系, 其中 θ_3 表示肩关节外-内旋角, θ_4 表示肘关节屈伸角的补角。(b) 当手臂所有关节角均为零时的肘关节坐标系。

当手腕位于位置 W' 时, 肘关节的局部坐标系记为 $EX_1Y_1Z_1$ 。假设手臂的初始状态为 θ_1 、 θ_2 、 θ_3 和 θ_4 均为零, 则此时肘关节坐标系的定义如图 2-3(b) 所示。

通过姿态估计, 可以获得点 S 、 E 和 W 在相机坐标系下的绝对坐标, 分别为 $S(x_0, y_0, z_0)$ 、 $E(x_1, y_1, z_1)$ 和 $W(x_2, y_2, z_2)$ 。在以肩关节为原点建立的相对坐标系下, 手腕点 W 的坐标可表示为 $W(x_2 - x_0, y_2 - y_0, z_2 - z_0)$ 。

如图 2-3(a) 所示, 在以肩关节为原点建立的相对坐标系中, 肩关节外旋角 θ_4 可通过向量 SE 与向量 EW 的点积计算得到。

$$\theta_4 = \arccos \left[\frac{\mathbf{SE} \cdot \mathbf{EW}}{|\mathbf{SE}| |\mathbf{EW}|} \right] \quad (2.3)$$

人体肘关节的活动范围为 $\pi - \theta_4$, 其中向量 EW 的坐标为 $(x_1 - x_2, y_1 - y_2, z_2 - z_1)$ 。

为求解 θ_3 , 已知实际手腕位置 W , 首先需要计算向量 HW 。如图 2-3(a) 所示, Y_1 轴定义为始终指向肩关节到肘关节的连线, 即沿向量 SE 的方向。由于 θ_4 已确定, 无论 θ_3 如何变化, 向量 EW 、 EW' 与 EH (即 Y_1 轴) 之间的夹角均保持为 θ_4 , 换言之, 连接肘关节与手腕的线 EW' 围绕 Y_1 轴旋转形成向量 EW 。

$$|EW'| = |EW| \quad (2.4)$$

$$|EH| = |EW'| * \cos\theta_4 \quad (2.5)$$

$$EH = |EH| * \frac{SE}{|SE|} \quad (2.6)$$

因此, 通过 $HW = EW - EH$ 可以得到向量 HW 。需要注意的是, 虽然肩关节坐标系在手臂运动过程中保持固定, 但肘关节坐标系会随姿态变化而改变方向, 如图 2-4 所示。在手臂运动过程中, 肘关节坐标系会旋转到新的状态。在图 2-4 中, Z_1 轴的方向会随着 θ_1 的值而变化, Z'_1 轴表示 Z_1 轴从肘关节平移到肩关节的位置。沿 Z'_1 轴的方向向量记为 p , 且满足 $p \perp m$ 且 $|p| = 1$, 因此 $p = (0, \sin\theta_1, -\cos\theta_1)$ 。如图所示, p 表示当手腕达到位置 W' 时, 肘关节处 Z_1 轴的方向。已知向量 HW 以及与 HW' 平行的方向向量 p , 即可求得 θ_3 。

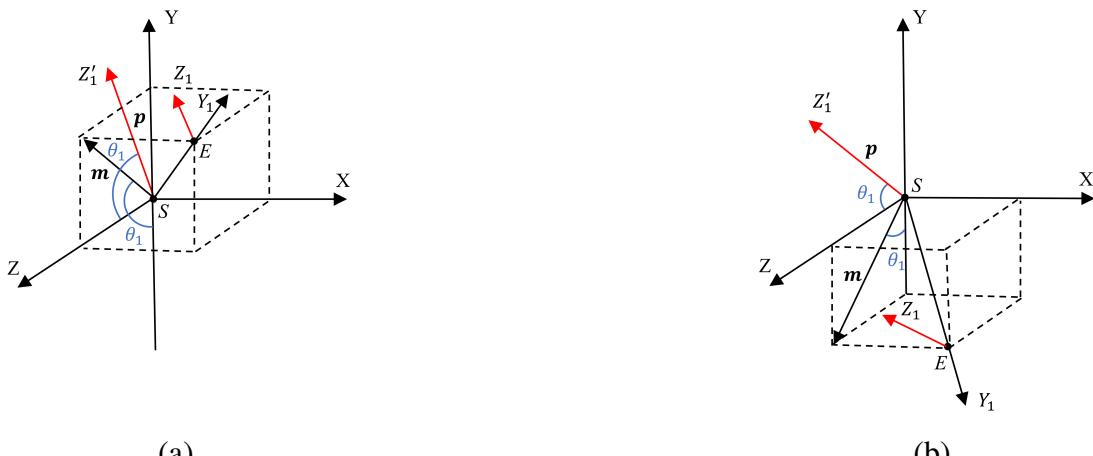


图 2-4 手臂运动过程中肩关节坐标系的旋转示意图。(a) 肩关节屈伸运动为锐角示意, (b) 肩关节屈伸运动为钝角示意。

在通过人体几何建模得到手臂关节角度后, 可以将这些角度直接作为目标角度来控制机械臂的运动。虽然这种方法能够模仿动作, 但无法实现精确的末端执行器位置跟踪, 从而难以完成高精度任务。为解决这一问题, 我们引入了末端执行器位置跟踪约束。

考虑到机器人与人体模型在尺寸、自由度和物理参数等方面的差异, 即使关节角度相同, 末端执行器的位置也可能不同。为了在任务执行过程中满足笛卡尔空间位置跟踪的要求, 我们将末端执行器位置的变化 (相对于设定的初始手臂位置 $P_{human init}$) 作为控制目标。

通过监测人体手臂末端执行器位置的变化, 可以实现机器人手臂在任务空间中的精确跟踪。在此基础上, 从手臂几何建模中得到的四个关节角被用作二次规划 (QP) 优化的初始条件。该方法确保在跟踪笛卡尔目标位置的同时, 机器人手

臂的几何构型与人体手臂保持一致。

关节空间与任务空间之间的速度映射关系可以通过微分运动学获得。

根据 OSQP^[108], 其中 Hessian 矩阵 $H = J^T J$, 梯度向量 $g = -J^T v$, 约束上下界则根据任务需求进行定义。整体算法流程如算法 2.1 所示, 最终求解得到满足几何构型约束和末端执行器位置约束的手臂关节角 q 。

2.4 实验

2.4.1 单人动作模仿

为了验证人体手臂运动学建模与动作重定向算法的有效性, 我们首先进行了单人动作模仿实验。实验步骤如下:

步骤 1: 初始化机器人, 确保所有关节移动至零位。

步骤 2: 示范者站在摄像头前, 保证身体处于视野范围内并开始执行动作。上位机处理器实时获取示范者肩部、肘部和手腕的关键点坐标, 用于后续手臂运动计算。

步骤 3: 在下位机处理中, 利用人体手臂几何模型计算肩关节屈伸角 (θ_1)、肩关节外展-内收角 (θ_2)、肘关节屈伸角 (θ_4) 以及肩关节内外旋角 (θ_3)。随后, 通过末端执行器位置跟踪方法对关节角进行优化, 得到调整后的机器人目标关节角 q 。

步骤 4: 将计算得到的手臂目标关节角直接发送至机器人, 完成动作模仿过程。

如图 2-5 所示, 从示范者的连续动作中选取了八帧, 展示机器人模仿人体手臂动作的过程。前两组展示了机器人模仿人体左臂动作的情况, 第三、四组展示了机器人模仿人体右臂动作的情况, 最后四组展示了机器人同时模仿双臂动作的情况。

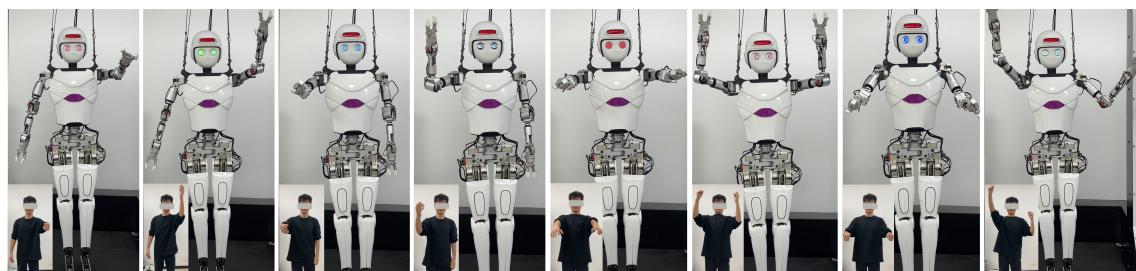


图 2-5 单人手臂动作模仿.

实验结果表明, 机器人能够通过动作重定向方法准确复制人体手臂动作, 验证了基于手臂几何模型的动作计算和末端执行器跟踪算法的有效性。

算法 2.1 几何建模与二次规划融合的动作重定向算法

Require: 通过几何建模得到的参考关节角 $q_{\text{init}} = \{\theta_1, \theta_2, \theta_3, \theta_4\}$

Ensure: 最终优化后的机器人关节角 q

1: 通过正向运动学计算初始末端位置:

$$P_{\text{init}} = f(q_{\text{init}})$$

2: 设置仿真参数:

3: 时间步长 $t = 0.05$, 迭代步数 $\text{step} = 20$

4: 上一时刻末端位置 $P_{t-1} = P_{\text{init}}$

5: 计算末端位置增量:

$$\Delta P = \frac{(P_{\text{robot}} - P_{\text{init}})}{\text{step}}$$

6: **for** $i = 1 : \text{step}$ **do**

7: **步骤 1:** 计算当前期望笛卡尔速度及二次规划的 Hessian 矩阵:

$$v = \frac{(P_{\text{init}} + i \cdot \Delta P) - P_{t-1}}{t}$$

8: **步骤 2:** 通过二次规划求解关节速度:

$$\dot{q} = \text{osqp}(H, g, l, u)$$

9: **步骤 3:** 状态更新:

10: 更新关节角度:

$$q = q + \dot{q} \cdot t$$

11: 更新当前末端位置:

$$P_{t-1} = f(q)$$

12: **end for**

13: **return** q

2.4.2 多人动作模仿

在双人动作模仿实验中，尽管两台机器人具有相同的结构尺寸，但两位示范者在身高比例和手臂关节长度上存在差异。图 2-6 展示了四组不同动作状态，每组均包含不同的手臂动作。实验结果表明，两台机器人能够平滑地复现示范者的手臂动作及末端执行器位置，验证了改进重定向方法的通用性。

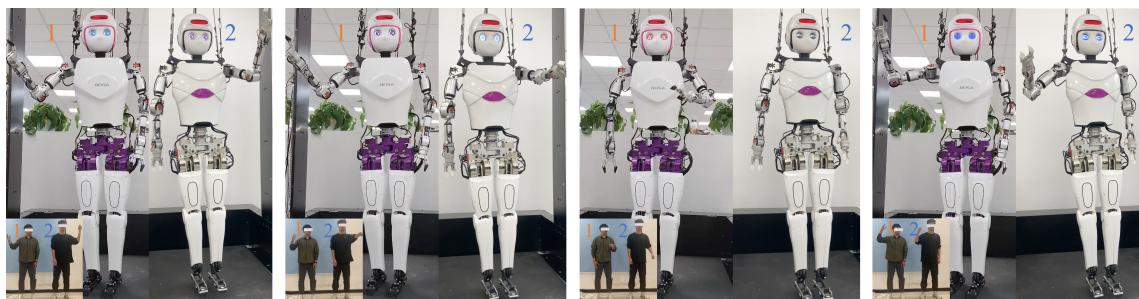


图 2-6 双人手臂动作模仿：机器人 1 模仿示范者 1 的动作，机器人 2 模仿示范者 2 的动作。

为了进一步评估动作重定向算法在多人场景下的泛化能力，我们在仿真环境中进行了多人动作模仿实验。如图 2-7 所示，在 Mujoco 仿真环境中创建了四个机器人模型。四位示范者的动作通过动作捕捉系统采集并生成手臂关键点，这些关键点随后输入改进重定向算法，计算四组目标手臂关节角度，用于驱动仿真机器人关节的运动。关节控制系统采用 500 Hz 的 PD 位置控制器运行，以确保仿真与实体机器人的控制频率一致，实现同步运动。

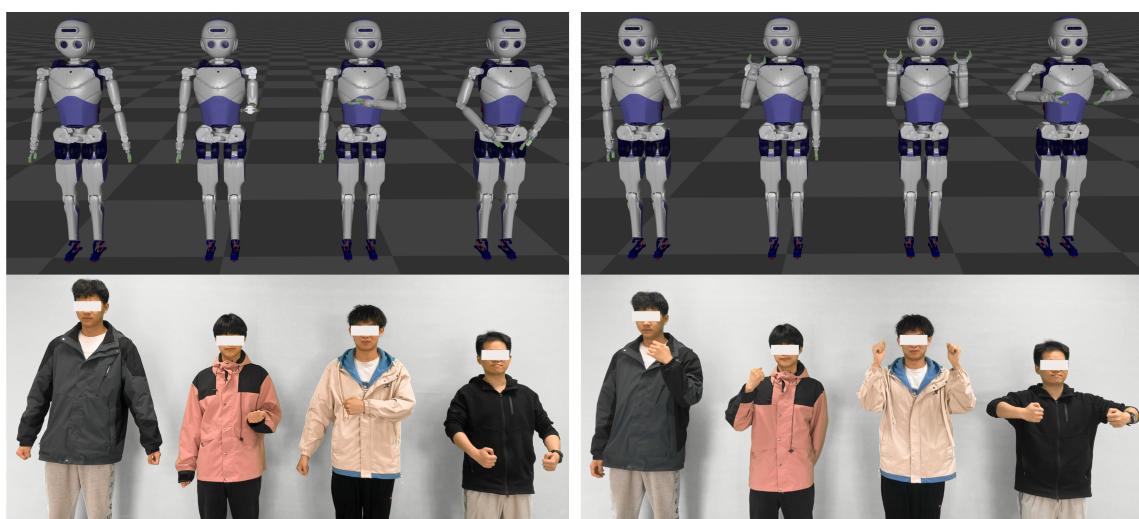


图 2-7 多人动作模仿：选择两组动作。

实验结果显示，随着示范者数量增加，机器人仍能保持动作模仿的平滑性和

精确性。所采用的改进重定向方法能够有效保证手臂运动相似性，同时在不同体型和动作模式之间具有良好的泛化能力。

2.4.3 任务操作

为了评估我们多人动作模仿系统在实际任务场景中的可行性，我们选择了四项任务：两项单臂模仿任务（倒水和分类）以及两项双臂协作任务（交接和搬运）。这些任务涵盖了在实际应用中可能出现的多种场景。为了便于任务执行，我们使用手持控制器操作机器人的前臂旋前-旋后关节、手腕屈伸关节以及末端执行器的夹爪。

在倒水任务中，如图 2-8 所示，机器人需要将水从瓶子倒入指定的杯子中。该任务对机器人具有一定的操作风险，因此需要保证动作的稳定性和准确性。

对于物体分类任务，如图 2-9 所示，桌面上有四类物品：芒果、枣子、猕猴桃和玩具。机器人需要将这些物品移动到各自对应的区域。该任务类似于日常的分类操作任务，需要双手协作进行操作。由于单臂机器人的作业空间有限，即便完成任务也需要较长的时间。

对于协作移交任务，如图 2-10 所示，桌面上放置了一个工具箱，里面有一把六角螺丝刀。机器人 B 需要从工具箱中取出螺丝刀并递给机器人 A，随后机器人 A 将螺丝刀放置到桌面上。该任务需要机器人之间的相互协作。

对于大型物体搬运任务，如图 2-11 所示，桌面上放置了一副金属框架（长度 180 cm，宽度 40 cm，重量 3.1 kg）。两台机器人协作完成该框架的搬运任务。由于目标物体的尺寸和重量较大，单台机器人在执行任务时会面临诸如重心难以保持稳定、手臂电机力矩不足等问题。通过协作，两台机器人可以共同抓取并抬起物体，从而降低任务难度。

需要注意的是，虽然执行任务的双足机器人外观与前述有所不同，但其机械臂的结构尺寸保持一致。

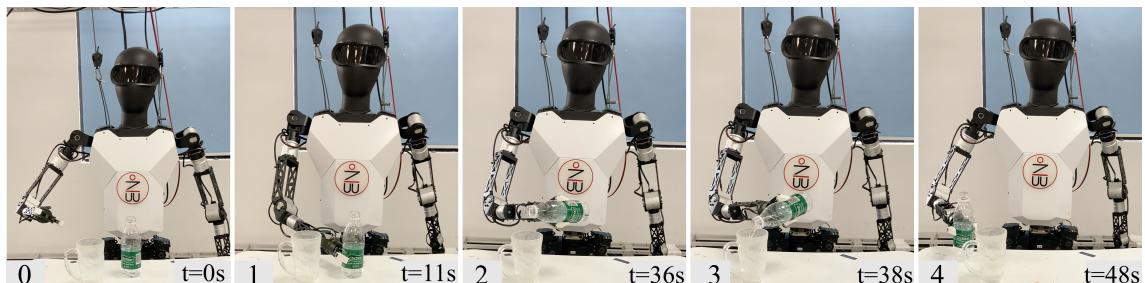


图 2-8 倒水。(0) 机器人在工作台前初始化；(1) 首先抓取桌面上的水瓶；(2) 将水瓶移动至目标杯子上方；(3) 旋转水瓶，将水倒入杯中；(4) 最后放开水瓶。整个过程耗时 48 秒。

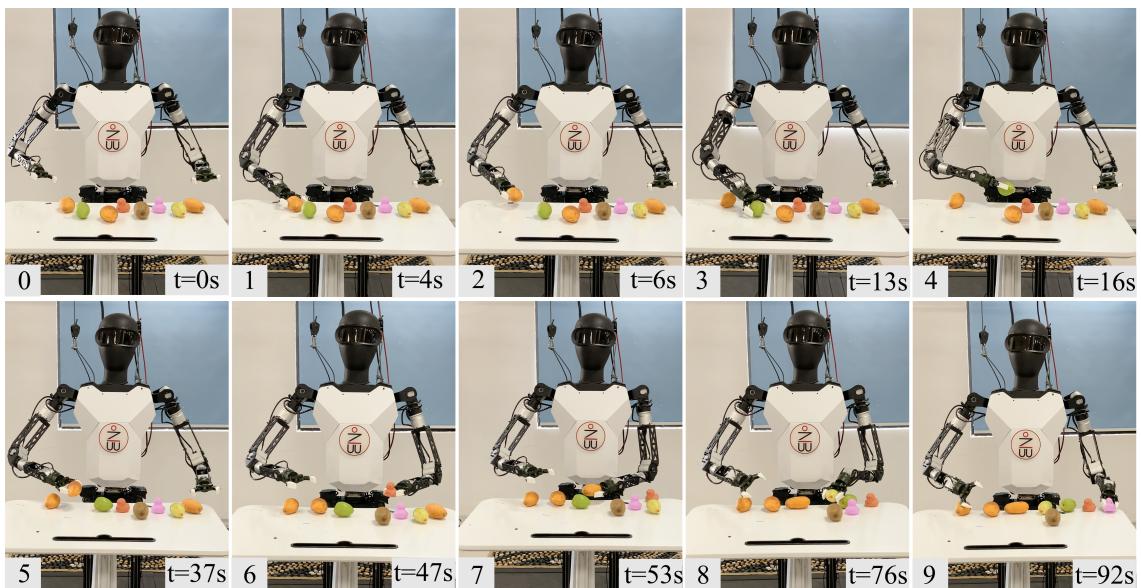


图 2-9 物体分类。(0) 机器人在工作台前初始化; (1) 首先选择右侧的物体进行分类, 抓取最右侧的芒果; (2) 提起并放置到目标位置; (3) 然后抓取枣子; (4) 提起并放置到目标位置。随后的步骤 (5-9) 重复 (1)(2) 的过程, 依次完成芒果 (7)、枣子 (8) 和玩具 (9) 的分类。最终完成所有物体的分类, 整个过程耗时 92 秒。



图 2-10 协作交接: 两台机器人在桌子前初始化。(1) 机器人 B 抓取六角螺丝刀; (2) 机器人 A 接收机器人 B 递来的螺丝刀; (3) 机器人 A 准备放置螺丝刀; (4) 机器人 A 松开夹爪完成放置。整个过程耗时 17 秒。

2.5 结果与讨论

2.5.1 数据分析

通过观察图 2-12 可知, 在双人动作模仿过程中, 机器人能够有效跟踪人体左肩的屈伸关节和外展内收关节的运动轨迹。如图 2-13 所示, 肩部旋转关节 (外旋-内旋) 和肘关节也表现出良好的跟踪性能。然而, 在图中的 (b) 和 (d) 中肘关节未能跟踪到动作, 这是由于关节角度受限, 无法超过机械极限所致。



图 2-11 协作搬运: (0) 两台机器人在金属框架前初始化; (1) 机器人 A 与 B 抓住框架; (2) 两台机器人同时抬起框架; (3) 两台机器人同时放下框架。整个过程耗时 15 秒。

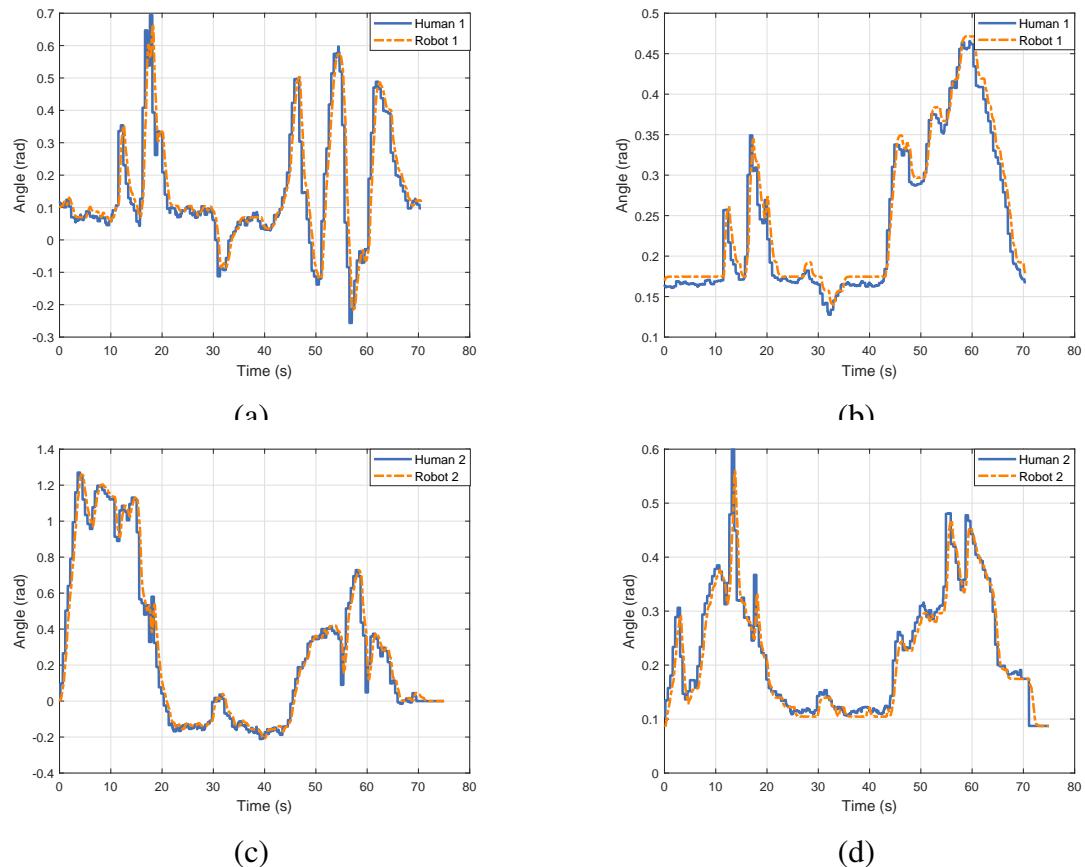


图 2-12 双人动作模仿中左肩屈伸关节和外展-内收关节的轨迹跟踪: (a) 和 (b) 机器人 1 跟踪表演者 1 左肩屈伸关节和外展-内收关节的关节轨迹; (c) 和 (d) 机器人 2 跟踪表演者 2 左肩屈伸关节和外展-内收关节的关节轨迹。

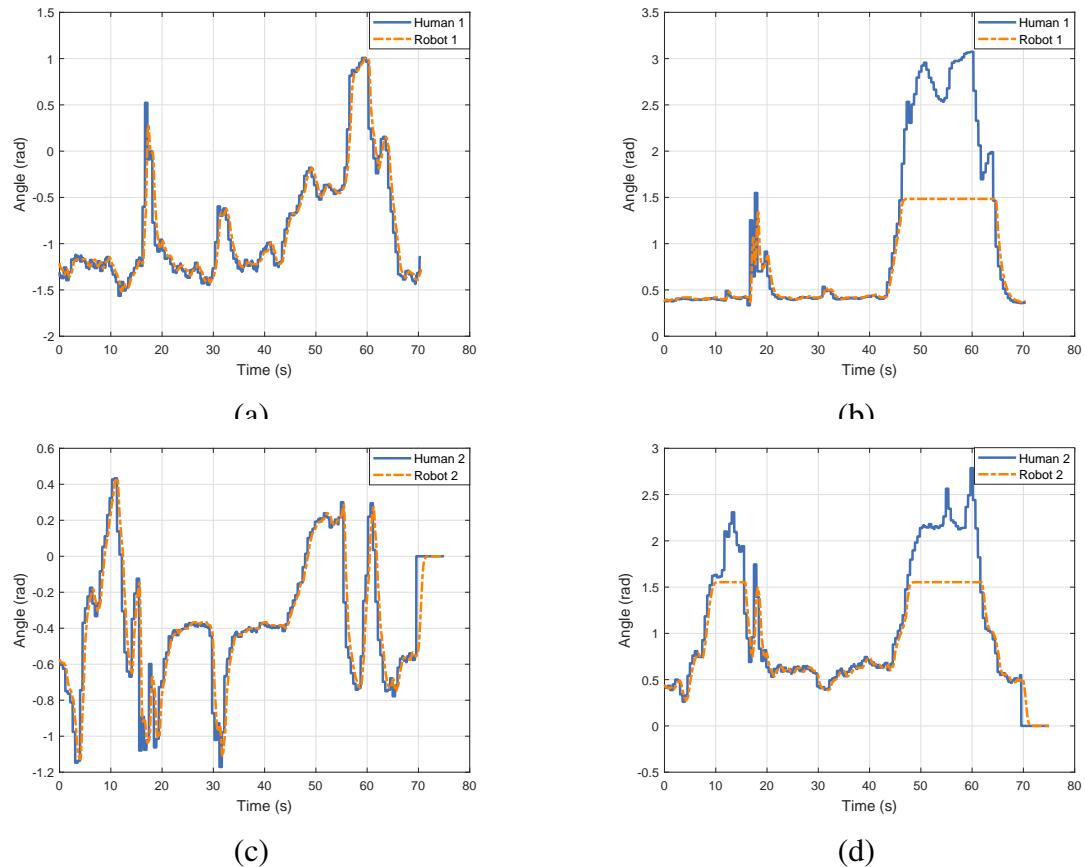


图 2-13 双人动作模仿中左肩关节外旋-内旋及肘关节的轨迹跟踪: (a) 和 (b) 表示机器人 1 跟踪示范者 1 左肩关节外旋-内旋及肘关节的轨迹; (c) 和 (d) 表示机器人 2 跟踪示范者 2 左肩关节外旋-内旋及肘关节的轨迹。

为了定量评估动作模仿的跟踪性能, 本文采用均方误差 (Mean Square Error, MSE) 作为关节轨迹跟踪精度的指标。MSE 值越小, 表示机器人的关节运动越接近目标轨迹。

图 2-14 展示了单人和双人动作模仿过程中各关节的均方误差 (MSE)。由于关节电机存在固有延迟, 跟踪过程中不可避免地引入了误差, 从而影响了 MSE 值。结果表明, 与单人动作模仿相比, 双人动作模仿时各关节的 MSE 并未显著增加, 这验证了系统在多智能体动作模仿中的可行性和有效性。此外, 肩部外展-内收关节的跟踪 MSE 最小, 而肩部内外旋关节的跟踪 MSE 最大。这主要是因为肩部外展-内收的运动特征更为显著, 使得从姿态估计输出的人体关键点数据更稳定。因此, 通过几何建模计算得到的肩部外展-内收关节参考角波动较小。相比之下, 手腕深度信息的波动较大, 使得生成的关键点数据变化更明显, 导致肩部内外旋关节参考角的波动较大。总之, 关节跟踪 MSE 主要来源于姿态估计过程中的数据波动, 同时电机响应延迟也会对跟踪精度产生一定影响。

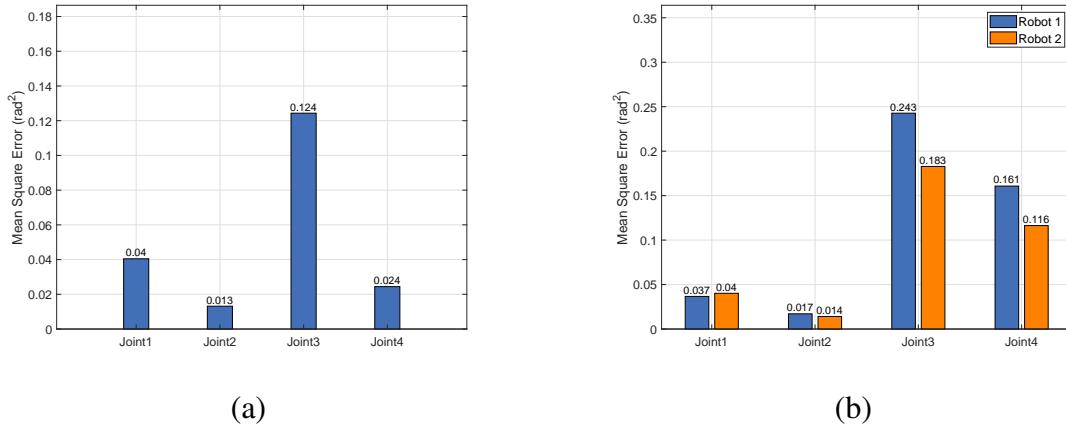


图 2-14 手臂动作模仿过程中关节角度的均方误差 (MSE): (a) 单人手臂动作模仿; (b) 双人手臂动作模仿。关节 1 对应肩屈伸关节, 关节 2 对应肩外展-内收关节, 关节 3 对应肩外旋-内旋关节, 关节 4 对应肘屈伸关节。

2.5.2 消融实验

通过消融实验, 我们比较了仅使用几何方法以及基于优化的二次规划 (QP) 方法在末端执行器位置跟踪任务中的性能。如表 2-1 所示, 当采用纯几何方法时, 末端执行器的跟踪误差为 9.35×10^{-2} ; 而采用 QP 优化方法时, 该误差显著降低至 3.75×10^{-4} 。相比之下, 本文提出的方法将末端执行器的跟踪误差进一步降低至 2.065×10^{-5} , 略高于解析逆向运动学方法, 但明显优于几何方法和 QP 方法。

在机器人任务执行中, 末端执行器的跟踪精度固然重要, 但保持合理的手臂构型同样关键。良好的手臂构型不仅可以避免与环境及机器人自身发生碰撞, 还能增强运动的稳定性与安全性。如表 2-1 所示, 只有本文提出的改进重定向方法在保持可行手臂构型的同时, 实现了较小的末端执行器误差。

表 2-1 不同重定向方法的性能比较

方法	关节限制	构型约束	末端执行器误差	迭代次数
几何方法	×	✓	9.35×10^{-2}	×
解析逆运动学	×	✗	2×10^{-6}	1
QP 方法	✓	✗	3.75×10^{-4}	≈ 100
改进重定向	✓	✓	2.065×10^{-5}	≈ 20

此外, 在计算效率方面, 所提出的方法仅需约 20 次迭代即可收敛至最优解, 相较于传统的基于 QP 的方法, 更适合实时应用。

为了直观比较不同重定向方法在动作模仿任务中的表现, 我们提供了本文所提出方法的实验视频 (实验视频), 并在同一关键帧下展示了各重定向方法的差异。

如图 ?? 所示，实验结果表明：

基于几何的方法通过追踪肩关节和肘关节角度来保证手臂的配置相似性。然而，由于它并未显式地追踪末端执行器的位置，因此末端执行器的跟踪误差相对较大。逆向运动学 (IK) 方法能够实现末端执行器位置的跟踪，但由于在求解过程中未考虑关节限制且忽略手臂配置，导致肘关节位置偏差明显，手臂配置相似性降低。QP 方法在优化关节限制的同时能够成功跟踪末端执行器位置，并改善手臂配置。然而，该方法增加了计算复杂度，且手臂配置仍存在一定误差。我们提出的改进重定向方法在保持低末端执行器跟踪误差的同时，能够维持手臂的相似配置，从而提升运动的稳定性和自然性。

为了评估多机器人动作模仿的自然性和可接受性，我们邀请了六名参与者进行示范动作，随后机器人执行模仿及任务操作。评价指标如下：1. ** 动作平滑性 (Motion Smoothness, MS)**：评估机器人执行过程中是否存在不连续或突兀的动作。2. ** 姿态可行性 (Pose Feasibility, PF)**：评估机器人手臂姿态是否与人类示范者保持相似。3. ** 任务完成度 (Task Achievement, TA)**：评估机器人是否准确完成指定的模仿任务。4. ** 总体用户偏好 (Overall User Preference)**：反映参与者对整体模仿质量的主观评价。

用户评分采用 1 (差) 到 5 (优) 的五分制，统计结果如表 ?? 所示。实验结果表明，在所有评价指标上，我们的方法均优于基于几何方法 (geometry-based)、逆向运动学 (IK) 和二次规划 (QP) 方法。尤其在动作平滑性和姿态可行性方面，用户主观评分明显更高。这表明，改进的重定向方法不仅提高了模仿精度，同时也增强了动作的自然性和用户可接受性。

2.5.3 局限

所提出的改进重定向算法能够有效地计算出最优关节角，同时满足手臂几何结构一致性和末端执行器位置跟踪的要求。通过结合多人人体姿态估计技术，该框架实现了多机械臂的动作模仿与协作任务，增强了系统的适应性和多任务处理能力。尽管人体运动运动学长期以来一直是研究前沿，但本研究在重定向算法方面相较于现有方法取得了进展。尤其是，将改进的重定向算法与多机器人协作任务相结合，并利用实时人体姿态估计进行机器人动作模仿，本方法在复杂环境中显著提升了动作模仿的精度与协调性。然而，所提出的框架仍存在一定局限性，尤其是在姿态方向跟踪方面。由于我们的姿态估计系统主要通过单目相机估计末端执行器的位置，它只能获取位置数据，而无法完整提取关节的全姿态信息（包括旋转分量）。这一限制显著影响了关节运动的灵活性，尤其是在涉及前臂旋前-旋后和手腕屈伸等复杂动作时，机器人无法充分调整这些关节以满足任务要求。未来

的研究可以引入完整的姿态估计，实现位置与方向的综合跟踪，从而提升机器人在复杂任务中的灵活性与精度。关于关节限位约束，如图 2-13 (b) 和 (d) 所示，无法实现精确跟踪主要是由于机械结构限制了关节运动范围，这一问题在传统控制方法中同样存在。为解决该局限性，未来工作可引入更加灵活的自适应关节限位控制策略，使机器人在不超出机械极限的前提下调整运动范围，从而减轻关节限位带来的影响。

表 2-2 不同重定向方法的用户主观评价

方法	MS	PF	TA	综合评分
几何方法	4	5	2	4
逆运动学方法	4	2	3	2
QP 方法	3	3	3	3
本文方法	5	5	4	5

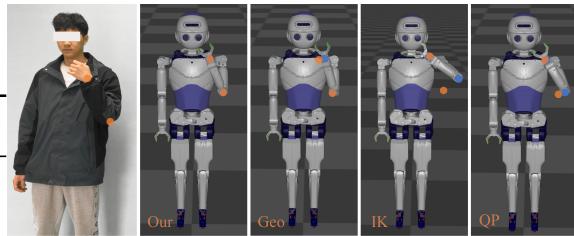


图 2-15 相同动作关键帧下不同重定向方法的对比结果。橙色点表示目标位置，蓝色点表示实际到达位置。

2.6 结论

本文提出了一个综合性的实时多仿人机器人手臂动作模仿框架，该框架采用改进的重定向方法，有效实现了手臂运动学配置与末端执行器精度之间的最佳平衡。通过整合多人人体姿态估计技术，该框架不仅支持多仿人机器人之间的协作运动，还为复杂任务中的多机器人协作提供了有效解决方案。

实验结果表明，机器人能够成功复现各种人类手臂动作。尽管随着示范者数量的增加，姿态估计误差有所上升，但整体动作模仿仍保持较高的跟踪精度和末端执行器准确性。此外，在倒水、物品分类及协作等任务中，我们进一步验证了改进重定向方法在保持手臂运动学结构完整性和末端执行器精度方面的优越性能。此外，随着多人人体姿态估计技术的整合，即使检测对象数量增加，系统仍能保持良好的实时性能，有效避免了计算负载的大幅增加。

尽管该框架在高精度与高效率方面表现出色，但在姿态跟踪上仍存在一定局限性。未来的研究将重点进一步优化姿态估计技术，实现位置与姿态的综合跟踪，从而提升机器人在复杂任务中的灵活性与精度，并使其在多智能体交互与协作任务中发挥更大作用。

参考文献

- [1] Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge[J]. International journal of computer vision, 2015, 115(3): 211-252.
- [2] Brock A, Donahue J, Simonyan K. Large scale gan training for high fidelity natural image synthesis[J]. arXiv preprint arXiv:1809.11096, 2018.
- [3] Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks[C]. //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 4401-4410.
- [4] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners[J]. Advances in neural information processing systems, 2020, 33: 1877-1901.
- [5] Chen M. Evaluating large language models trained on code[J]. arXiv preprint arXiv:2107.03374, 2021.
- [6] Berner C, Brockman G, Chan B, et al. Dota 2 with large scale deep reinforcement learning[J]. arXiv preprint arXiv:1912.06680, 2019.
- [7] Arulkumaran K, Cully A, Togelius J. Alphastar: an evolutionary computation perspective[C]. //Proceedings of the genetic and evolutionary computation conference companion. 2019: 314-315.
- [8] Silver D, Huang A, Maddison C J, et al. Mastering the game of go with deep neural networks and tree search[J]. nature, 2016, 529(7587): 484-489.
- [9] Raibert M H. Hopping in legged systems—modeling and simulation for the two-dimensional one-legged case[J]. IEEE Transactions on Systems, Man, and Cybernetics, 2012, (3): 451-463.
- [10] McGeer T. Passive bipedal running[J]. Proceedings of the Royal Society of London. B. Biological Sciences, 1990, 240(1297): 107-134.
- [11] Raibert M H, Hodgins J K. Animation of dynamic legged locomotion[C]. //Proceedings of the 18th annual conference on Computer graphics and interactive techniques. 1991: 349-358.
- [12] Schwind W J. Spring loaded inverted pendulum running: a plant model[B]. University of Michigan, 1998.

- [13] Geyer H, Seyfarth A, Blickhan R. Positive force feedback in bouncing gaits?[J]. Proceedings of the Royal Society of London. Series B: Biological Sciences, 2003, 270(1529): 2173-2183.
- [14] Vukobratović M, Borovac B. Zero-moment point—thirty five years of its life[J]. International journal of humanoid robotics, 2004, 1(01): 157-173.
- [15] Yin K, Loken K, Van de Panne M. Simbicon: simple biped locomotion control[J]. ACM Transactions on Graphics (TOG), 2007, 26(3): 105-es.
- [16] Da Silva M, Abe Y, Popović J. Simulation of human motion data using short-horizon model-predictive control[C]. //Computer Graphics Forum: vol. 27: 2. 2008: 371-380.
- [17] Bledt G, Powell M J, Katz B, et al. Mit cheetah 3: design and control of a robust, dynamic quadruped robot[C]. //2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2018: 2245-2252.
- [18] Raibert M H, Brown Jr H B, Cheponis M. Experiments in balance with a 3d one-legged hopping machine[J]. The International Journal of Robotics Research, 1984, 3(2): 75-92.
- [19] Miura H, Shimoyama I. Dynamic walk of a biped[J]. The International Journal of Robotics Research, 1984, 3(2): 60-74.
- [20] Hodgins J K, Wooten W L, Brogan D C, et al. Animating human athletics[C]. //Proceedings of the 22nd annual conference on Computer graphics and interactive techniques. 1995: 71-78.
- [21] Coros S, Beaudoin P, Van de Panne M. Generalized biped walking control[J]. ACM Transactions On Graphics (TOG), 2010, 29(4): 1-9.
- [22] Hutter M, Gehring C, Jud D, et al. Anymal-a highly mobile and dynamic quadrupedal robot[C]. //2016 IEEE/RSJ international conference on intelligent robots and systems (IROS). 2016: 38-44.
- [23] De Lasas M, Mordatch I, Hertzmann A. Feature-based locomotion controllers[J]. ACM transactions on graphics (TOG), 2010, 29(4): 1-10.
- [24] Sreenath K, Park H W, Poulakakis I, et al. A compliant hybrid zero dynamics controller for stable, efficient and fast bipedal walking on mabel[J]. The International Journal of Robotics Research, 2011, 30(9): 1170-1193.

- [25] Mordatch I, Todorov E, Popović Z. Discovery of complex behaviors through contact-invariant optimization[J]. ACM Transactions on Graphics (ToG), 2012, 31(4): 1-8.
- [26] Al Borno M, De Lasas M, Hertzmann A. Trajectory optimization for full-body movements with complex contacts[J]. IEEE transactions on visualization and computer graphics, 2012, 19(8): 1405-1414.
- [27] Gehring C, Coros S, Hutter M, et al. Practice makes perfect: an optimization-based approach to controlling agile motions for a quadruped robot[J]. IEEE Robotics & Automation Magazine, 2016, 23(1): 34-43.
- [28] Dogar M, Srinivasa S. A framework for push-grasping in clutter[J]. Robotics: Science and systems VII, 2011, 1: 65-72.
- [29] Apgar T, Clary P, Green K, et al. Fast online trajectory optimization for the bipedal robot cassie.[C]. //Robotics: Science and Systems: vol. 101. 2018: 14.
- [30] Van de Panne M, Kim R, Fiume E. Virtual wind-up toys for animation[C]. //Graphics Interface. 1994: 208-208.
- [31] Kohl N, Stone P. Policy gradient reinforcement learning for fast quadrupedal locomotion[C]. //IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004: vol. 3. 2004: 2619-2624.
- [32] Tedrake R, Zhang T W, Seung H S. Stochastic policy gradient reinforcement learning on a simple 3d biped[C]. //2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566): vol. 3. 2004: 2849-2854.
- [33] Endo G, Morimoto J, Matsubara T, et al. Learning cpg sensory feedback with policy gradient for biped locomotion for a full-body humanoid[C]. //AAAI. 2005: 1267-1273.
- [34] Coros S, Beaudoin P, Van de Panne M. Robust task-based control policies for physics-based characters[B]. //ACM SIGGRAPH Asia 2009 papers. 2009: 1-9.
- [35] Tan J, Zhang T, Coumans E, et al. Sim-to-real: learning agile locomotion for quadruped robots[J]. arXiv preprint arXiv:1804.10332, 2018.
- [36] Haarnoja T, Ha S, Zhou A, et al. Learning to walk via deep reinforcement learning[J]. arXiv preprint arXiv:1812.11103, 2018.

- [37] Hwangbo J, Lee J, Dosovitskiy A, et al. Learning agile and dynamic motor skills for legged robots[J]. *Science Robotics*, 2019, 4(26): eaau5872.
- [38] Wang J M, Hamner S R, Delp S L, et al. Optimizing locomotion controllers using biologically-based actuators and objectives[J]. *ACM Transactions on Graphics (TOG)*, 2012, 31(4): 1-11.
- [39] Tan J, Gu Y, Liu C K, et al. Learning bicycle stunts[J]. *ACM Transactions on Graphics (TOG)*, 2014, 33(4): 1-12.
- [40] Levine S, Finn C, Darrell T, et al. End-to-end training of deep visuomotor policies[J]. *Journal of Machine Learning Research*, 2016, 17(39): 1-40.
- [41] Peng X B, Berseth G, Van de Panne M. Terrain-adaptive locomotion skills using deep reinforcement learning[J]. *ACM Transactions on Graphics (TOG)*, 2016, 35(4): 1-12.
- [42] Gu S, Holly E, Lillicrap T, et al. Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates[C]. //2017 IEEE international conference on robotics and automation (ICRA). 2017: 3389-3396.
- [43] Andrychowicz O M, Baker B, Chociej M, et al. Learning dexterous in-hand manipulation[J]. *The International Journal of Robotics Research*, 2020, 39(1): 3-20.
- [44] Wang J M, Fleet D J, Hertzmann A. Optimizing walking controllers[B]. //ACM SIGGRAPH Asia 2009 papers. 2009: 1-8.
- [45] Geijtenbeek T, Van De Panne M, Van Der Stappen A F. Flexible muscle-based locomotion for bipedal creatures[J]. *ACM Transactions on Graphics (TOG)*, 2013, 32(6): 1-11.
- [46] Clegg A, Yu W, Tan J, et al. Learning to dress: synthesizing human dressing motion via deep reinforcement learning[J]. *ACM Transactions on Graphics (TOG)*, 2018, 37(6): 1-10.
- [47] Yu W, Turk G, Liu C K. Learning symmetry and low-energy locomotion[J]. ArXiv e-prints.
- [48] Abdolhosseini F, Ling H Y, Xie Z, et al. On learning symmetric locomotion[C]. //Proceedings of the 12th ACM SIGGRAPH Conference on Motion, Interaction and Games. 2019: 1-10.

- [49] Kato I, Takanishi A. Development of wabot-1: the first anthropomorphic robot[C]. //Proceedings of the 6th International Symposium on Industrial Robots. 1976.
- [50] Kaneko K, Kanehiro F, Kajita S, et al. Humanoid robot hrp-2[C]. //Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). 2004: 1083-1090.
- [51] Sakagami Y, Watanabe R, Aoyama C, et al. The intelligent asimo: system overview and integration[C]. //Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2002: 2478-2483.
- [52] Raibert M H. Legged robots that balance[B]. MIT Press, 1986.
- [53] Nelson G, Saunders A, Playter R. Petman: a humanoid robot for testing chemical protective clothing[C]. //Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). 2012: 2209-2216.
- [54] Kuindersma S, Deits R, Fallon M, et al. Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot[J]. Autonomous Robots, 2016, 40(3): 429-455.
- [55] Agility Robotics. Digit: human-centric, multi-purpose robot for logistics work. <https://agilityrobotics.com>. Accessed 2026-03-01. 2023.
- [56] Tesla Inc. Tesla optimus gen 2 humanoid robot. <https://www.tesla.com/AI>. Accessed 2026-03-01. 2023.
- [57] Boston Dynamics. Atlas: the fully electric humanoid robot. <https://www.bostondynamics.com>. Accessed 2026-03-01. 2024.
- [58] Figure AI. Figure 02 autonomous general-purpose humanoid robot. <https://figure.ai>. Accessed 2026-03-01. 2024.
- [59] 1X Technologies. Neo gamma: home-safe humanoid robot. <https://www.1x.tech>. Accessed 2026-03-01. 2024.
- [60] Fu C, Liu S, Wang J, et al. Sensory reflex control for a pneumatic biped robot[C]. //Proceedings of the 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems (CYBER). Hong Kong, China: IEEE, 2014: 401-406.
- [61] 侯月阳. 挠性驱动单元及其在仿人双足步行机器人中的应用研究 [D]. 哈尔滨: 哈尔滨工业大学, 2014.

- [62] Xu W, Huang Q, Li J, et al. An improved zmp trajectory design for the biped robot bhr[C]. //Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). Shanghai, China: IEEE, 2011: 569-574.
- [63] Unitree Robotics. Unitree g1 humanoid robot. <https://www.unitree.com>. Accessed 2026-03-01. 2024.
- [64] EngineAI. Se01 high-dynamic humanoid robot. <https://www.engineai.com>. Accessed 2026-03-01. 2024.
- [65] UBTECH Robotics. Walker s2 industrial humanoid robot. <https://www.ubtrobot.com>. Accessed 2026-03-01. 2024.
- [66] UBTECH Robotics and Beijing Humanoid Innovation Center. Walker tienkung research platform. <https://www.ubtrobot.com>. Accessed 2026-03-01. 2024.
- [67] Fourier Intelligence. Gr-1 general-purpose humanoid robot. <https://www.fftai.com>. Accessed 2026-03-01. 2024.
- [68] Liegeois A. Automatic supervisory control of the configuration and behavior of multibody mechanisms[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1977, 7(12): 868-871.
- [69] Nakamura Y. Advanced robotics: redundancy and optimization[B]. Addison-Wesley, 1987.
- [70] Siciliano B, Slotine J J E. A general framework for managing multiple tasks in highly redundant robotic systems[C]. //Proceedings of the 5th International Conference on Advanced Robotics. 1991: 1211-1216.
- [71] Khatib O. A unified approach for motion and force control of robot manipulators: the operational space formulation[J]. IEEE Journal on Robotics and Automation, 1987, 3(1): 43-53.
- [72] Khatib O, Sentis L, Park J Y, et al. Whole-body dynamic behavior and control of human-like robots[J]. International Journal of Humanoid Robotics, 2004, 1(1): 29-43.
- [73] Sentis L. Synthesis and control of whole-body behaviors in humanoid systems[D]. Stanford University, 2007.
- [74] Mansard N, Khatib O, Kheddar A. A unified approach to integrate unilateral constraints in the stack of tasks[J]. IEEE Transactions on Robotics, 2009, 25(3):

670-685.

- [75] Escande A, Mansard N, Wieber P B. Hierarchical quadratic programming: fast on-line humanoid-robot motion generation[J]. *The International Journal of Robotics Research*, 2014, 33(7): 1006-1028.
- [76] Righetti L, Buchli J, Mistry M, et al. Inverse dynamics control of floating-base robots with external constraints: a unified view[C]. //Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). 2011: 1085-1090.
- [77] Herzog A, Rotella N, Mason S, et al. Momentum control with hierarchical inverse dynamics on a torque-controlled humanoid[J]. *Autonomous Robots*, 2016, 40(3): 473-491.
- [78] Siekmann J, et al. Learning diverse bipedal gaits from scratch[J]. *Science Robotics*, 2021.
- [79] Gu X, et al. Humanoid-gym: reinforcement learning for humanoid robots[J], 2024.
- [80] Xie Z, et al. Pbhc: physics-based humanoid control[J], 2025.
- [81] Zhang Y, et al. Wococo: whole-body contact control via reinforcement learning[J], 2024.
- [82] Fu Y, et al. Humanplus: humanoid whole-body control via reinforcement learning[J], 2024.
- [83] Argall B D, Chernova S, Veloso M, et al. A survey of robot learning from demonstration[J]. *Robotics and Autonomous Systems*, 2009, 57(5): 469-483.
- [84] Pomerleau D A. Alvinn: an autonomous land vehicle in a neural network[C]. //Advances in Neural Information Processing Systems (NeurIPS). 1989.
- [85] Bain M. Experiments in imitation learning[D]. University of Cambridge, 1995.
- [86] Ho J, Ermon S. Generative adversarial imitation learning[C]. //Advances in Neural Information Processing Systems (NeurIPS). 2016.
- [87] Peng X B, Kumar A, Zhang G, et al. Amp: adversarial motion priors for stylized physics-based character control[J]. *ACM Transactions on Graphics (SIGGRAPH)*, 2021, 40(4): 1-15.
- [88] Peng X B, Zhang Z, Yu W, et al. Ase: large-scale reusable adversarial skills[J]. *ACM Transactions on Graphics (SIGGRAPH)*, 2022, 41(4): 1-15.

-
- [89] Zhao Y, Wang X, Wang D, et al. Towards adaptive humanoid control via multi-behavior distillation and reinforced fine-tuning[J]. arXiv preprint arXiv:2511.06371, 2025.
 - [90] Ling H Y, Zinno F, Cheng G, et al. Character controllers using motion vaes[J]. ACM Transactions on Graphics, 2020, 39(4): 40:1-40:12.
 - [91] Fu Z, Zhao T Z, Finn C. Mobile aloha: learning bimanual mobile manipulation with low-cost whole-body teleoperation[J]. arXiv preprint arXiv:2401.02117, 2024.
 - [92] He T, Luo Z, Xiao W, et al. Learning human-to-humanoid real-time whole-body teleoperation[J]. arXiv preprint arXiv:2403.04436, 2024.
 - [93] Bütepage J, Ghadirzadeh A, Öztimur Karadağ Ö, et al. Imitating by generating: deep generative models for imitation of interactive tasks[J]. Frontiers in Robotics and AI, 2020, 7: 47.
 - [94] Kulić D, Venture G, Yamane K, et al. Anthropomorphic movement analysis and synthesis: a survey of methods and applications[J]. IEEE Transactions on Robotics, 2016, 32(4): 776-795.
 - [95] Chi C, Xu Z, Pan C, et al. Universal manipulation interface: in-the-wild robot teaching without in-the-wild robots[J]. arXiv preprint arXiv:2402.10329, 2024.
 - [96] Bahl S, Gupta A, Pathak D. Human-to-robot imitation in the wild[J]. arXiv preprint arXiv:2207.09450, 2022.
 - [97] Ishiguro Y, Makabe T, Nagamatsu Y, et al. Bilateral humanoid teleoperation system using whole-body exoskeleton cockpit tablis[J]. IEEE Robotics and Automation Letters, 2020, 5(4): 6419-6426.
 - [98] Ramos J, Kim S. Humanoid dynamic synchronization through whole-body bilateral feedback teleoperation[J]. IEEE Transactions on Robotics, 2018, 34(4): 953-965.
 - [99] Kim S, Kim C, You B, et al. Stable whole-body motion generation for humanoid robots to imitate human motions[C]. //IEEE/RSJ International Conference on Intelligent Robots and Systems. 2009: 2518-2524.
 - [100] Peng X B, Abbeel P, Levine S, et al. Deepmimic: example-guided deep reinforcement learning of physics-based character skills[J]. ACM Transactions On Graph-

- ics (TOG), 2018, 37(4): 1-14.
- [101] Peng X B, Ma Z, Abbeel P, et al. Amp: adversarial motion priors for stylized physics-based character control[J]. ACM Transactions on Graphics (ToG), 2021, 40(4): 1-20.
- [102] Won J, Gopinath D, Hodgins J. A scalable approach to control diverse behaviors for physically simulated characters[J]. ACM Transactions on Graphics (TOG), 2020, 39(4): 33-1.
- [103] Luo Z, Cao J, Kitani K, et al. Perpetual humanoid control for real-time simulated avatars[C]. //Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 10895-10904.
- [104] Bohez S, Tunyasuvunakool S, Brakel P, et al. Imitate and repurpose: learning reusable robot movement skills from human and animal behaviors[J]. arXiv preprint arXiv:2203.17138, 2022.
- [105] Tang A, Hiraoka T, Hiraoka N, et al. Humanmimic: learning natural locomotion and transitions for humanoid robot via wasserstein adversarial imitation[C]. //IEEE International Conference on Robotics and Automation (ICRA). 2024: 13107-13114.
- [106] Koenemann J, Burget F, Bennewitz M. Real-time imitation of human whole-body motions by humanoids[C]. //IEEE International Conference on Robotics and Automation (ICRA). 2014: 2806-2812.
- [107] Arduengo M, Arduengo A, Colomé A, et al. Human to robot whole-body motion transfer[C]. //IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids). 2021: 299-305.
- [108] Stellato B, Banjac G, Goulart P, et al. OSQP: an operator splitting solver for quadratic programs[J]. Mathematical Programming Computation, 2020, 12(4): 637-672. <https://doi.org/10.1007/s12532-020-00179-2>.