

融合腰部虚拟关节的人形机器人仿人全身运动控制

刘煜程¹ 田羽锋¹ 杨玥² 苏晓杰¹

摘 要 复杂运动的稳定性与社会环境的适应性是人形机器人实现自然交互与任务执行的核心能力, 而高质量、高协调的仿人运动是保障其高效稳定运行的基础. 腰部自由度作为上下肢协调控制的关键枢纽, 对整体动态稳定性与动作流畅性具有显著影响. 然而, 现有研究在腰部自由度建模与控制方面关注不足, 导致机器人在仿人行走中易出现姿态畸变甚至运动失稳. 为此, 提出一种融合腰部虚拟关节的人形机器人仿人全身运动控制框架. 该框架首先利用动作捕捉系统采集人体多维关节数据, 并通过关节空间映射实现运动风格的高保真迁移. 随后, 针对人体腰部与机器人结构间的自由度差异, 构建融合俯仰与滚转虚拟关节的上下肢动力学模型, 提出多关节主动补偿机制以实现虚拟腰部自由度的协同补偿. 进一步地, 将虚拟关节引入特权观测空间, 构建对抗式动作先验风格奖励与任务奖励的可变融合策略, 从而实现全身协调的仿人运动控制. 实验基于 Isaac Gym 训练平台与 MuJoCo 仿真环境, 在多种行走速率 (0.2 ~ 1.6 m/s) 下开展测试, 结果表明所提方法能显著提升人形机器人的仿人运动自然度与泛化能力.

关键词 仿人运动控制; 多关节主动补偿; 虚拟关节控制; 全身协调运动

引用格式 刘煜程, 田羽锋, 杨玥, 苏晓杰. 融合腰部虚拟关节的人形机器人仿人全身运动控制. 自动化学报, 2026, 52(9): 1992–2004

DOI 10.16383/j.aas.c250667 **CSTR** 32138.14.j.aas.c250667

Human-like Whole-body Motion Control for Humanoid Robot With Integrated Virtual Waist Joints

LIU Yu-Cheng¹ TIAN Yu-Feng¹ YANG Yue² SU Xiao-Jie¹

Abstract The stability of complex motion and the adaptability to social environments are core capabilities for humanoid robots to achieve natural interaction and task execution, where high-quality and highly coordinated human-like motion serves as the foundation for their efficient and stable operation. As the key hub for upper-lower limb coordination control, the waist degrees of freedom (DOFs) play a critical role in maintaining overall dynamic stability and motion fluency. However, limited attention to waist DOF modeling and control in existing studies often leads to posture distortion or instability during human-like walking. To address this issue, this paper proposes a human-like whole-body motion control framework for humanoid robot with integrated virtual waist joints. The framework first employs a motion capture system to acquire multi-dimensional human joint data, which are mapped into the humanoid robot's joint space to achieve high-fidelity transfer of motion style. Then, considering the DOF discrepancy between the human waist and the robot structure, a coupled upper-lower limb dynamic model incorporating virtual pitch and roll waist joints is constructed, and a multi-joint active compensation mechanism is developed to realize coordinated compensation for the virtual waist DOFs. Furthermore, the virtual joints are incorporated into the privileged observation space, and a variable fusion strategy combining adversarial motion prior style rewards and task rewards is designed to achieve whole-body coordinated human-like motion control. Experiments conducted in the Isaac Gym training platform and MuJoCo simulation environment demonstrate that the proposed method significantly improves the naturalness and generalization of humanoid robots' human-like motion under various walking speeds (0.2 ~ 1.6 m/s).

Keywords human-like motion control; multi-joint active compensation; virtual joint control; whole-body coordinated motion

Citation Liu Yu-Cheng, Tian Yu-Feng, Yang Yue, Su Xiao-Jie. Human-like whole-body motion control for humanoid robot with integrated virtual waist joints. *Acta Automatica Sinica*, 2026, 52(9): 1992–2004

收稿日期 2025-11-24 录用日期 2026-03-31

Manuscript received November 24, 2025; accepted March 31, 2026

国家自然科学基金 (62403084, 62173051), 中央高校基本科研业务费专项资金 (2025CDJJKZCQ-08), 重庆市科技创新重大研发项目 (CSTB2023TIAD-STX0037) 资助

Supported by National Natural Science Foundation of China (62403084, 62173051), Fundamental Research Funds for the Central Universities (2025CDJJKZCQ-08), and the Science and

Technology Innovation Key Research and Development Program of Chongqing under Grant (CSTB2023TIAD-STX0037)

本文责任编辑 李繁斌

Recommended by Associate Editor LI Fan-Biao

1. 重庆大学自动化学院 重庆 401331 2. 西安建筑科技大学信息与控制工程学院 西安 710311

1. School of Automation, Chongqing University, Chongqing 401331 2. College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710311

近年来, 人形机器人凭借类人多自由度的肢体结构, 在工厂装配、货物搬运、家庭服务等场景中表现出独特的灵活性与通用性^[1]. 稳定、高效且可扩展的全身运动控制策略, 成为其在复杂工况下可靠作业的核心前提^[2]. 然而, 人形机器人上下肢耦合机理不明晰, 导致全身运动过程中保持上下肢协同稳定十分困难. 因此, 全身协同控制作为人形机器人执行复杂任务的关键基础, 受到学术界的广泛关注.

当前研究可分为基于模型与基于学习两条主要技术路线. 基于模型的方法通常将全身平衡维持和任务跟踪问题统一描述为带约束的优化命题^[3], 包括模型预测控制 (model predictive control, MPC) 以及基于倒立摆或中枢发生器等简化模型控制. 模型预测控制因其能显式处理实时多约束且易被转化为二次规划问题进行快速计算, 被广泛用于提升机体在扰动下的稳定性与敏捷性^[4-5]. 为进一步降低计算复杂度, 研究者把质心动力学解耦为凸子问题, 或引入倒立摆^[6]、中枢发生器^[7]等简化模型, 实现快速反馈实时求解. 然而, 人形机器人固有的多关节且强非线性属性导致其动力学模型难以精准建立^[8]. 此外, 随着环境复杂度的提升, 如地面摩擦力、环境形变以及传感器噪声等不确定因素也会影响模型控制器的性能表现^[9]. 强化学习 (reinforcement learning, RL) 为应对难以精确建模的接触非线性与环境不确定性提供了新思路^[10]. 与传统的基于模型的控制方法不同, RL 利用大规模仿真训练与仿真迁移实机 (sim-to-real, S2R) 技术架构, 在无需大量在线调参或模型辨识的情况下即可实现多地形鲁棒行走^[11]. 此外, 也有研究提出“教师-学生”协同框架^[12], 显著降低了样本需求并提升了泛化能力. 总体来看, 两类方法在单一步态与站立平衡任务上已表现出不错的效果. 但是, 简化模型以及缺少样本数据的学习策略均难以实现对全身关节动作的合理分配, 导致其能量效率及复杂环境适应能力难以提升. 因此, 研究者们进一步探索提升人形机器人执行任务时动作的仿人协调性, 从而改善机体能量分配及运动平稳性^[13]. 相关研究可以归类为机理模型引导学习^[14-16]和动捕数据支持的模仿学习^[17] (imitation learning, IL) 两条主要路线. 一方面, 机理模型引导策略虽能通过二次规划或简化动力学模型显式约束质心位置与零力矩^[15], 但过于保守的可行域会导致动作僵硬及上肢可达空间受限, 难以满足高拟人的需求^[16]. 另一方面, 以融合对抗式动作先验 (adversarial motion prior, AMP) 的深度强化学习为代表的驱动策略虽能拟合动捕数据, 但受人-机器人腰部结构物理差异的影响, 躯干动作产生的

惯性扭矩通过浮动基座动力学直接耦合至下肢. 在质量分布不均与动力学约束失配的双重影响下, 系统常通过下肢关节的非自然补偿维持平衡, 进而诱发动作畸形或失稳摔倒^[17]. 因此, 如何通过腰部关节耦合减弱上肢运动惯量对下肢产生的扰动显得尤为重要^[18]. 研究者们针对腰部耦合方式的改进相继开展了两类研究: 一是设计特殊的多自由度腰部机械结构, 并进行针对性多模态控制^[19-21], 但由于腰部关节使动作空间复杂度提高, 其对复杂环境的适应能力也大打折扣; 二是聚焦下肢稳定独立优化, 尝试完全代替腰部运动^[22], 但由于缺乏对腰-腿协同机制的系统性建模与实时补偿, 始终存在髋关节负载过大的风险.

针对上述挑战, 本文提出一种融合腰部虚拟关节的人形机器人仿人全身运动控制框架 VW-MAC (virtual waist-guided multi-joint active compensation), 主要创新包括:

1) 针对因腰部物理自由度缺失导致的仿人运动平稳性差的问题, 提出一种基于虚拟腰部关节的耦合动力学建模方法及多关节主动补偿策略. 不同于用髋关节优化代替腰部自由度的策略^[22], 本文所提方法通过发掘下肢运动链的冗余自由度空间, 将虚拟腰部的运动需求更合理地重映射至下肢所有关节, 实现了多关节协同的主动补偿. 在提升协调性的同时保证能量的均匀分担, 有效避免了髋关节负载过大的问题.

2) 针对不同类型奖励函数权重固化导致的多目标策略难调和的问题, 基于融合 AMP 的深度强化学习控制框架, 提出对抗式动作先验风格奖励与任务奖励的可变融合机制, 动态调节奖励权重, 实现人形机器人在任务执行过程中全身稳定性与仿人协调性的最优平衡.

本文整体研究架构如图 1 所示. 第 1 节将人体运动的动捕数据全维映射至机器人关节空间, 并提取全身运动学信息; 为完整利用全维信息, 第 2 节针对腰部自由度失配的问题, 提出虚拟腰部建模方法及多关节主动补偿策略; 在 AMP 及深度强化学习的基础上, 第 3 节结合虚拟腰部关节耦合动力学模型设计任务奖励函数, 并提出一种任务-风格奖励可变融合机制, 兼顾机器人仿人运动的稳定与协调; 第 4 节在 Isaac Gym 及 MuJoCo 仿真环境下进行实验验证.

1 仿人运动关节空间全维映射

仿人运动控制的重要前提在于从人体动作中提取结构化运动特征, 并将其有效地映射至机器人的

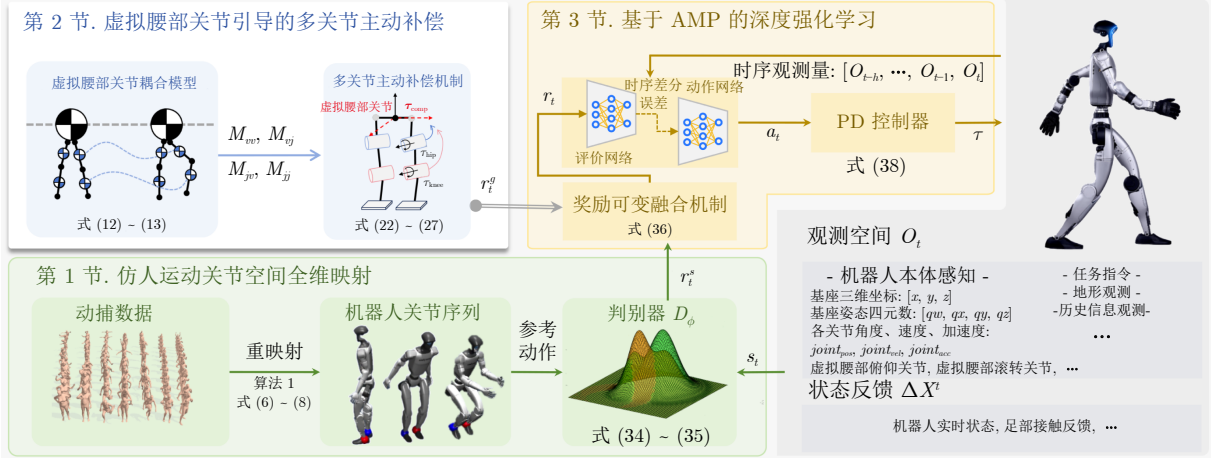


图1 本文主要研究内容结构

Fig.1 Structure of the main research content in this work

关节空间^[23]. 本节首先建立带约束的全身运动学模型, 用于描述机器人多关节耦合运动的可行解空间. 随后提出人-机器人动作空间重定向方法, 实现人体骨骼时序数据向机器人关节序列的全维映射, 为虚拟腰部关节建模与仿人运动学习提供基础数据支持.

1.1 带约束的全身运动学模型构建

人形机器人动作表现出仿人特性, 根本原因在于其关节及末端在笛卡尔空间中的运动轨迹符合人体运动学规律^[24-25]. 人形机器人运动学建模旨在描述各关节运动与末端执行器位姿之间的映射关系. 机器人四肢可以用多关节刚体连杆进行描述, 其基本正向运动学可以用一系列转换矩阵的乘积表示:

$$\mathbf{T}_e(\mathbf{q}) = \prod_{i=1}^n \mathbf{T}_{i-1, i}(\mathbf{q}_i) \quad (1)$$

其中 $\mathbf{T}_e(\mathbf{q})$ 为手臂或腿部连杆的齐次变换矩阵; $\mathbf{q} = [q_1, q_2, \dots, q_n]^T$ 为四肢的关节角向量; $\mathbf{T}_{i-1, i}(\mathbf{q}_i)$ 表示相邻关节间的位姿变换.

在仿人运动控制中, 机器人需同时满足多种需求, 包括任务精度、姿态稳定性、关节物理限制及运动连续性等^[26], 而机器人上肢、腰部及下肢在运动过程中优化的目标存在差异. 为此, 本文将全身运动学约束分为下肢、腰部与上肢三个层次, 通过分层建模实现整体协调.

下肢约束用于维持机器人在动态动作中的稳定性, 并保证踝部位姿与支撑面约束条件一致, 该约束主要包括踝部位姿误差约束和质心稳定约束. 在控制过程中, 踝部位姿需保持与期望位姿 $\mathbf{T}_{\text{foot, target}}$ 的差异最小化, 同时要求质心投影实时位于支撑多边形内部, 以规避姿态失稳风险. 下肢部分的优化目标函数可描述为:

$$\mathcal{L}_{\text{leg}} = (1 - \omega_{\text{com}})\mathcal{L}_{\text{foot}} + \omega_{\text{com}}\mathcal{L}_{\text{CoM}} \quad (2)$$

其中, 踝部位姿误差项定义为 $\mathcal{L}_{\text{foot}} = \|\mathbf{T}_{\text{foot}}(\mathbf{q}) - \mathbf{T}_{\text{foot, target}}\|^2$, $\mathbf{T}_{\text{foot}}$ 为踝部实际位姿; 质心稳定项定义为 $\mathcal{L}_{\text{CoM}} = \|\mathbf{p}_{\text{CoM}}(\mathbf{q}) - \mathbf{p}_{\text{support}}\|^2$, $\mathbf{p}_{\text{CoM}}$ 为质心实际位置, $\mathbf{p}_{\text{support}}$ 为期望支撑域; $\omega_{\text{com}} \in [0, 1]$ 为平衡约束权重系数, 用于调节稳定性在整体多目标优化中的优先级.

为了尽可能保留人体动捕数据在各个维度上的信息, 与降维映射方式^[23]不同, 本文方法在动捕数据重定向阶段将机器人与人体的腰部关节自由度对齐. 由关节角 $\mathbf{q}$ 计算得到的机器人腰部姿态为 $\mathbf{R}_{\text{waist}}(\mathbf{q})$, 由人体动作捕捉数据得到的目标腰部姿态为 $\mathbf{R}_{\text{waist, target}}$, 则腰部姿态误差项可表示为:

$$\mathcal{L}_{\text{waist}} = \|\mathbf{R}_{\text{waist}}(\mathbf{q}) - \mathbf{R}_{\text{waist, target}}\|^2 \quad (3)$$

该项用于直接量化机器人腰部与目标姿态之间的偏差. 在整体优化过程中, $\mathcal{L}_{\text{waist}}$ 与下肢和上肢的约束项共同作用, 通过调整各关节角实现腰部方向与目标姿态的一致性, 达到姿态协调与运动连续性的要求.

上肢约束主要保证操作及摆臂动作的准确性和轨迹的平滑性. 手部末端的期望位姿为 $\mathbf{T}_{\text{hand, target}}$, 实际位姿 $\mathbf{T}_{\text{hand}}$, 上肢约束项可写为:

$$\mathcal{L}_{\text{arm}} = \|\mathbf{T}_{\text{hand}}(\mathbf{q}) - \mathbf{T}_{\text{hand, target}}\|^2 \quad (4)$$

机器人全身运动的优化目标可表示为:

$$\min_{\mathbf{q}} \mathcal{L}_{\text{wb}} = \omega_{\text{leg}}\mathcal{L}_{\text{leg}} + \omega_{\text{waist}}\mathcal{L}_{\text{waist}} + \omega_{\text{arm}}\mathcal{L}_{\text{arm}} \quad (5)$$

其中 ω_{leg} , ω_{waist} , ω_{arm} 为各层约束的权重系数. 不同任务阶段可动态调整各权重, 以在平衡性与任务精度之间取得最优协调. 带约束的全身分层建模方式在运动学层面实现了机器人整体姿态的约束协调, 既保证了足部与质心的稳定控制, 又通过全维腰部映

射保持人体动作的自然特征. 机器人全身运动的优化目标 $\mathcal{L}_{wb}$ 为奖励函数的设计提供了全身运动学基础.

1.2 人-机器人关节空间重定向

人-机器人关节空间重定向旨在将人体动作捕捉数据映射到机器人的可行运动空间, 在保留人体运动风格的同时满足机器人物理约束^[27]. 该过程包括坐标系对齐、尺度适配、关节对应映射及多层约束优化四个主要步骤.

首先, 为统一参考坐标系, 将人体骨骼点对应的齐次变换矩阵 T_h 通过刚体变换映射至机器人坐标系, 得到机器人坐标系下的人体姿态矩阵:

$$T_r = R \cdot T_h + t \quad (6)$$

其中 R 为旋转矩阵; t 为平移向量, 用于补偿人体与机器人坐标系间的姿态差异, 确保后续映射与 $T_{\text{foot}}(q)$ 、 $T_{\text{hand}}(q)$ 等机器人运动学向量的坐标一致性. 考虑人的肢体长度 L_h 与机器人肢体长度 L_r 存在差异, 引入比例因子 $s_p = L_r/L_h$ 进行尺度适配. 基于适配后的人体骨骼数据, 分别生成机器人踝部、腕部的目标齐次变换矩阵及腰部目标姿态矩阵:

$$\begin{cases} T_{\text{foot, target}} = T_{\text{base}} + s_p \cdot (T_{h, \text{foot}} - T_{\text{base}}) \\ T_{\text{hand, target}} = T_{\text{base}} + s_p \cdot (T_{h, \text{hand}} - T_{\text{base}}) \\ R_{\text{waist, target}} = R \cdot R_{h, \text{waist}} \end{cases}$$

其中, $T_{h, \text{foot}}$ 、 $T_{h, \text{hand}}$ 分别为人体踝部、腕部的齐次变换矩阵; $R_{h, \text{waist}}$ 为人体腰部姿态矩阵; T_{base} 为机器人基坐标系矩阵, 确保目标位姿与运动学模型中的约束目标完全对齐. 为表示人体与机器人关节的对应关系, 将关节划分为上肢、腰部和下肢三大模块, 如图2所示.

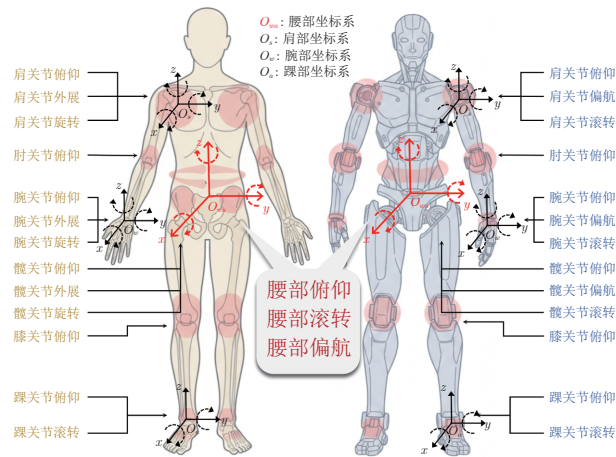


图2 人体关节与机器人关节对应关系

Fig.2 Correspondence between human joints and robot joints

基于全维关节映射及运动学模型, 采用雅可比迭代法求解机器人关节角向量 q 的增量. 定义综合位姿误差向量 ΔX 包含踝部位姿误差、腕部位姿误差及腰部姿态误差:

$$\Delta X = [\omega_{\text{leg}} \Delta x_{\text{foot}}^T, \omega_{\text{arm}} \Delta x_{\text{hand}}^T, \omega_{\text{waist}} \Delta r_{\text{waist}}^T]^T \quad (7)$$

其中

$$\Delta x_{\text{foot}} = \text{vec}(T_{\text{foot, target}} - T_{\text{foot}}(q))$$

$$\Delta x_{\text{hand}} = \text{vec}(T_{\text{hand, target}} - T_{\text{hand}}(q))$$

分别为踝部、腕部的位姿误差向量; $\Delta r_{\text{waist}} = \text{vec}(R_{\text{waist, target}} - R_{\text{waist}}(q))$ 为腰部姿态误差向量. 由此, 关节角增量可以表示为:

$$\Delta q = J^+ \cdot \Delta X \quad (8)$$

其中 J^+ 为机器人全身雅可比矩阵的伪逆. 在求解过程中引入阻尼项增强数值稳定性, 避免奇异构型, 并通过权重分配保证各约束优先级与运动学模型一致.

帧间运动连续性通过关节角速度约束实现, 同时关联关节物理限制:

$$\left| \frac{q_t - q_{t-1}}{\Delta t} \right| \leq v_{\text{max}} \quad (9)$$

其中, q_t 与 q_{t-1} 分别表示当前帧与上一帧的机器人关节角向量; Δt 为相邻帧之间的时间间隔; v_{max} 表示各关节允许的最大角速度向量. 该约束确保动作在时间上的平滑过渡, 同时维护上肢、腰部和下肢之间的协调性. 不同任务阶段可通过动态调整 ω_{leg} , ω_{waist} , ω_{arm} 实现多目标优化平衡, 在操作任务中提高 ω_{arm} , 而在动态行走任务中提高 ω_{leg} .

算法1. 人-机器人关节空间重定向算法

输入. 人体运动序列 $\{T_h^i\}_{i=1}^N$, 机器人模型参数, 约束权重 ω_{leg} 、 ω_{waist} 、 ω_{arm} .

输出. 机器人关节轨迹 $\{q_t\}_{t=1}^T$.

- 1: 初始化机器人关节姿态 q_0
- 2: for $t = 1$ to T do
- 3: 读取当前帧人体骨骼姿态 T_h^t (含上肢、腰部、下肢子矩阵)
- 4: 计算坐标系对齐矩阵 $T_{\text{align}} = [R, t]$
- 5: 将人体数据映射到机器人坐标系 $T_r^t \leftarrow R \cdot T_h^t + t$
- 6: 计算肢体比例因子 $s_p = L_r/L_h$
- 7: 生成目标约束量 $T_{\text{foot, target}}^t, T_{\text{hand, target}}^t, R_{\text{waist, target}}^t$
- 8: 构建综合误差 ΔX^t , 求解关节角 $q_t = q_{t-1} + J^+ \cdot \Delta X^t$
- 9: 检查关节角速度约束, 若不满足则修正 q_t
- 10: 输出当前帧机器人关节姿态 q_t
- 11: end for

值得注意的是, 基于本节构建的全维运动学模型实现的重定向过程, 虽然能最大限度保留运动数据中的关节信息, 但由于人体与实际机器人腰部结构存在差异, 不可避免地会出现自由度不匹配的问题. 在此情形下, 如何让机器人成功学习仿人运动数据显得尤为重要.

2 虚拟腰部关节引导的多关节主动补偿

现实中人形机器人由于缺少腰部俯仰、滚转自由度, 难以从全维重映射数据中有效学习到这两个自由度的信息, 导致上下肢运动失配, 进而出现仿人运动动作畸形的问题^[9, 28]. 本节基于人体腰部真实自由度与机器人腰部关节自由度的差异, 构建虚拟腰部关节耦合模型, 提出腿部多关节主动补偿机制, 实现下肢关节对欠自由度运动的联合补偿. 基于此, 本节设计特权观测空间, 并引入虚拟关节信息, 为深度强化学习控制框架的搭建提供理论支撑.

2.1 虚拟腰部关节耦合模型构建

为了更精确地刻画虚拟腰部自由度与下肢运动之间的动力学耦合关系, 本节基于浮动基座建模思想^[28], 在系统的广义坐标中引入虚拟腰部的滚转与俯仰角自由度, 从而形成一个扩展的系统动力学模型. 系统的广义坐标向量为:

$$\mathbf{q}_v = \begin{bmatrix} \mathbf{q}_b \\ \phi_v \\ \mathbf{q}_j \end{bmatrix} = \begin{bmatrix} \mathbf{q}_b \\ \mathbf{q}_{aug} \end{bmatrix} \quad (10)$$

其中, $\mathbf{q}_b = [x, y, z, \phi_r, \phi_p, \phi_y]^T$ 表示基座在世界坐标系中的位置与姿态, x, y, z 表示基座空间位置, 基座空间姿态 ϕ_r, ϕ_p, ϕ_y 分别为滚转、俯仰、偏航的欧拉角表示; $\phi_v = [\phi_{v,r}, \phi_{v,p}]^T$ 分别为虚拟腰部的滚转角与俯仰角; $\mathbf{q}_j = [q_1, q_2, \dots, q_n]^T$ 表示下肢关节角度向量, 维度与机器人实际驱动关节数一致. 为便于表示, 将虚拟腰部与下肢联合定义为扩展坐标 $\mathbf{q}_{aug} = [\phi_v^T, \mathbf{q}_j^T]^T$. 根据拉格朗日方程, 系统的动力学模型可写为:

$$M(\mathbf{q})\ddot{\mathbf{q}} + C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + G(\mathbf{q}) = S^T \boldsymbol{\tau} + J_c^T \mathbf{f}_c \quad (11)$$

其中, $M(\mathbf{q})$ 为系统的质量矩阵, 反映机器人整体的惯性特性; $C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}}$ 表示科氏力与离心力项; $G(\mathbf{q})$ 为重力项; S 为选择矩阵, 用于将驱动扭矩 $\boldsymbol{\tau}$ 映射到广义坐标空间中, 其中虚拟腰部对应列均为零; J_c 为接触雅可比矩阵; $\mathbf{f}_c$ 为地面对机器人产生的接触反作用力. 将式 (11) 按照坐标类型进行分块, 可得系统的分层动力学结构:

$$\begin{bmatrix} M_{bb} & M_{ba} \\ M_{ab} & M_{aa} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{q}}_b \\ \ddot{\mathbf{q}}_{aug} \end{bmatrix} + \mathbf{h} = \begin{bmatrix} \mathbf{0} \\ S_{aug}^T \boldsymbol{\tau} \end{bmatrix} + J_c^T \mathbf{f}_c \quad (12)$$

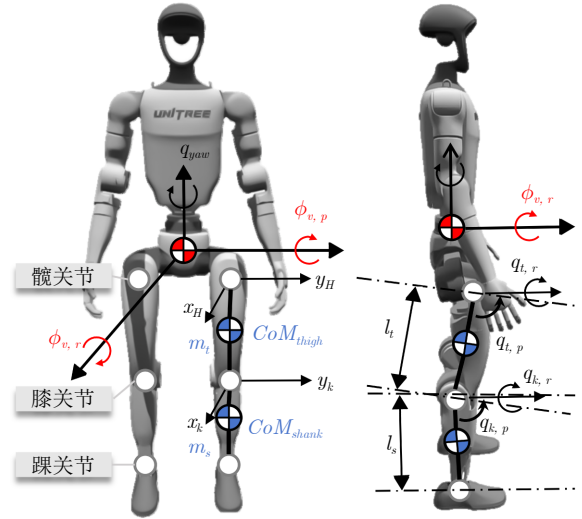
其中, S_{aug} 为扩展选择矩阵; M_{bb} 表示基座部分的惯性矩阵, 反映机身整体的惯性分布特性; M_{aa} 即为 $M_{aug, aug}$, 代表虚拟腰部及下肢组合系统的惯性矩阵, 描述两者之间的惯性耦合关系; M_{ba} 与 M_{ab} 分别为 $M_{b, aug}$ 与 $M_{aug, b}$, 描述基座与虚拟腰部、下肢之间的耦合惯性项, 揭示身体姿态变化对下肢运动的影响; $\mathbf{h} = [h_b, h_{aug}]^T$ 为科氏力与重力项的组合, 包含基座 h_b 与虚拟腰部及下肢组合系统 h_{aug} 的动力学效应.

进一步地, 将扩展坐标 $\mathbf{q}_{aug}$ 细化为虚拟腰部与下肢关节两个部分, 可得耦合动力学形式:

$$\begin{bmatrix} M_{vv} & M_{vj} \\ M_{jv} & M_{jj} \end{bmatrix} \begin{bmatrix} \ddot{\phi}_v \\ \ddot{\mathbf{q}}_j \end{bmatrix} + \begin{bmatrix} h_v \\ h_j \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\tau}_j \end{bmatrix} + J_{aug, c}^T \mathbf{f}_c \quad (13)$$

其中, M_{vv} 为虚拟腰部惯性矩阵, 反映其滚转与俯仰自由度的惯性特征; M_{vj} 与 M_{jv} 表示虚拟腰部与下肢关节之间的惯性耦合项; M_{jj} 为下肢各关节自身的惯性矩阵; h_v 和 h_j 分别为虚拟腰部与下肢部分的非线性动力学项, 包括科氏力、离心力以及重力补偿项; $\boldsymbol{\tau}_j$ 为下肢驱动扭矩向量; $J_{aug, c}$ 为虚拟腰部与下肢联合在接触约束下的雅可比矩阵.

为了更具体地刻画虚拟腰部与腿部连杆之间的耦合关系, 考虑单腿由大腿与小腿两段组成, 如图 3 所示. 机器人大腿质量为 m_t , 小腿质量为 m_s ; 大腿



注: q_{yaw} 为腰部关节偏航角度; x_H, y_H 为髋关节坐标系; x_k, y_k 为膝关节坐标系; $q_{k,r}$ 为膝关节滚转角, 对应图 2 中膝关节俯仰角.

图 3 带有虚拟腰部关节的机器人动力学模型

Fig.3 Robot dynamic model with a virtual waist joint

长度为 l_t , 小腿长度为 l_s ; 大腿滚转角为 $q_{t,r}$ 、俯仰角为 $q_{t,p}$, 膝关节俯仰角为 $q_{k,p}$; 大腿质心为 CoM_{thigh} , 小腿质心为 CoM_{shank} . 则腿部的动能为:

$$T_{leg} = \frac{1}{2} m_t \dot{\mathbf{p}}_t^T \dot{\mathbf{p}}_t + \frac{1}{2} m_s \dot{\mathbf{p}}_s^T \dot{\mathbf{p}}_s \quad (14)$$

其中, $\dot{\mathbf{p}}_t$ 与 $\dot{\mathbf{p}}_s$ 分别为大腿与小腿质心的线速度. 根据几何关系, 大腿质心位置可表示为:

$$\mathbf{p}_t = \mathbf{p}_h + \frac{l_t}{2} \begin{bmatrix} \cos q_{t,p} \cos q_{t,r} \\ \sin q_{t,r} \\ \sin q_{t,p} \cos q_{t,r} \end{bmatrix} \quad (15)$$

小腿质心位置为:

$$\mathbf{p}_s = \mathbf{p}_h + l_t \begin{bmatrix} \cos q_{t,p} \cos q_{t,r} \\ \sin q_{t,r} \\ \sin q_{t,p} \cos q_{t,r} \end{bmatrix} + \frac{l_s}{2} \begin{bmatrix} \cos(q_{t,p} + q_{k,p}) \\ 0 \\ \sin(q_{t,p} + q_{k,p}) \end{bmatrix} \quad (16)$$

其中 $\mathbf{p}_h$ 表示髋关节在基座坐标系中的位置.

对式 (15) 和 (16) 求导, 可得质心线速度:

$$\begin{cases} \dot{\mathbf{p}}_t = \mathbf{J}_t(\phi_{v,r}, \phi_{v,p}, q_{t,r}, q_{t,p}) \dot{\mathbf{q}}_t \\ \dot{\mathbf{p}}_s = \mathbf{J}_s(\phi_{v,r}, \phi_{v,p}, q_{t,r}, q_{t,p}, q_{k,p}) \dot{\mathbf{q}}_s \end{cases} \quad (17)$$

其中 $\mathbf{J}_t$ 和 $\mathbf{J}_s$ 分别为大腿和小腿的雅可比矩阵; $\dot{\mathbf{q}}_t = [\dot{\phi}_{v,r}, \dot{\phi}_{v,p}, \dot{q}_{t,r}, \dot{q}_{t,p}]^T$; $\dot{\mathbf{q}}_s = [\dot{\phi}_{v,r}, \dot{\phi}_{v,p}, \dot{q}_{t,r}, \dot{q}_{t,p}, \dot{q}_{k,p}]^T$. 虚拟腰部角速度 $(\dot{\phi}_{v,r}, \dot{\phi}_{v,p})$ 与各关节角速度共同决定了腿部质心的线速度和动能分布.

将式 (14) 代入拉格朗日方程, 可得关于虚拟腰部角度的局部惯性项:

$$M_{vv} = m_t \mathbf{J}_t^T \mathbf{J}_t + m_s \mathbf{J}_s^T \mathbf{J}_s \quad (18)$$

从而构成系统虚拟腰部惯性矩阵的一部分. 进一步, 虚拟腰部与下肢关节间的耦合项为:

$$M_{vj} = m_t \mathbf{J}_t^T \mathbf{J}_{t,j} + m_s \mathbf{J}_s^T \mathbf{J}_{s,j} \quad (19)$$

其中 $\mathbf{J}_{t,j}$ 和 $\mathbf{J}_{s,j}$ 分别为大腿与小腿质心线速度对各关节角速度的雅可比矩阵. 可见, 大腿和小腿的质量与长度 (m_t, m_s, l_t, l_s) 直接影响虚拟腰部的惯性耦合强度, 而其姿态角度决定了这种耦合随运动状态的变化规律. 因此, 虚拟腰部的等效扭矩可表示为:

$$\boldsymbol{\tau}_v = M_{vv} \ddot{\boldsymbol{\phi}}_v + M_{vj} \ddot{\mathbf{q}}_j + \mathbf{h}_v - \mathbf{J}_c^T \mathbf{f}_c \quad (20)$$

其中, 第一项体现了虚拟腰部姿态变化对整条腿的惯性影响, 第二项反映了各关节加速度对虚拟腰部运动的反馈. 特别地, 当大腿俯仰角 $q_{t,p}$ 与膝关节俯仰角 $q_{k,p}$ 发生较大变化时, 惯性耦合项 M_{vj} 会随之显著调整, 从而影响虚拟腰部扭矩的动态特性. 另一方面, 虚拟腰部的角加速度还通过质心雅可

比矩阵影响整体稳定性. 定义质心雅可比矩阵 $\mathbf{J}_{CoM}(\mathbf{q})$, 则质心加速度满足:

$$\ddot{\mathbf{p}}_{CoM} = \mathbf{J}_{CoM}(\mathbf{q}) \ddot{\mathbf{q}} + \dot{\mathbf{J}}_{CoM}(\mathbf{q}, \dot{\mathbf{q}}) \dot{\mathbf{q}} \quad (21)$$

由此可知, 虚拟腰部角加速度 $\ddot{\boldsymbol{\phi}}_v$ 的变化通过惯性耦合与质心动力学共同作用于全身姿态平衡, 尤其在考虑大腿与小腿的长度、质量分布后, 这种动态效应在需要全身协调的仿人运动中尤为显著. 综上, 该模型不仅揭示了虚拟腰部与下肢之间的物理耦合机理, 也为后续主动补偿控制与深度强化学习策略的设计提供了动力学基础.

注 1. 不同于传统的全身运动控制仅在现有硬件拓扑下求解受限的零空间优化^[8], 本文提出的 VW-MAC 框架通过引入虚拟广义坐标 $\mathbf{q}_v$, 在算法层面构建了跨拓扑的动量平衡补偿器. 在上述建模过程中, 虚拟关节不仅提供了运动学上的位姿参考, 而且参与了系统整体动量 $\mathbf{L}(\mathbf{q}, \dot{\mathbf{q}})$ 的重新分配. 该机制将人体动捕数据中由脊柱产生的角动量特征, 通过等效动力学映射矩阵 $\mathbf{J}_v$ 解耦至下肢冗余自由度空间. 这种建模方式在逻辑上实现了躯干惯性扰动与浮动基座动力学的深度解耦, 将原本可能导致失稳的强耦合扭矩转化为下肢各关节的主动协同运动, 从而在本质上解决了由于硬件自由度缺失引发的动力学失配问题.

2.2 多关节主动补偿机制设计

虚拟腰部关节在实体机器人上并不存在, 因此需要通过下肢其他关节的联合运动对两个虚拟自由度进行主动补偿. 在虚拟腰部与下肢的动力学耦合模型的基础上, 本节结合扭矩等效映射思想提出主动补偿机制. 在缺乏真实腰部驱动器的情况下, 通过协调下肢关节运动动态地合成虚拟腰部的扭矩响应, 从而在动力学层面实现欠自由度补偿. 主动补偿控制的核心思想是利用下肢关节产生等效的扭矩作用, 从而补偿腰部物理自由度缺失造成的姿态误差. 下肢扭矩由任务扭矩、补偿扭矩组成:

$$\boldsymbol{\tau}_j = \boldsymbol{\tau}_{task} + \boldsymbol{\tau}_{comp} \quad (22)$$

其中, $\boldsymbol{\tau}_{task}$ 为实现行走、平衡等主要任务的控制输入, $\boldsymbol{\tau}_{comp}$ 为用于虚拟腰部扭矩补偿的附加项. 为了在动力学上实现等效, 补偿扭矩应满足如下条件:

$$\mathbf{J}_j^T \boldsymbol{\tau}_{comp} \approx \mathbf{J}_v^T \boldsymbol{\tau}_v \quad (23)$$

其中 $\mathbf{J}_j$ 为下肢扭矩雅可比矩阵. 下肢关节产生的广义扭矩应在动量空间中等效于虚拟腰部所需的扭矩. 考虑到实际关节数有限, 为避免过约束或冗余分配, 引入正则化最小二乘形式:

$$\min_{\tau_{\text{comp}}} \|J_j^T \tau_{\text{comp}} - J_v^T \tau_v\|^2 + \lambda \|\tau_{\text{comp}}\|^2 \quad (24)$$

其中 $\lambda > 0$ 为正则化系数, 用于平衡扭矩匹配精度与能量消耗. τ_{comp} 解析解可写为:

$$\tau_{\text{comp}}^* = (J_j J_j^T + \lambda I)^{-1} J_j J_v^T \tau_v \quad (25)$$

进一步地, 为体现补偿作用在关节层面的具体分布, 将补偿扭矩展开为:

$$\tau_{\text{comp}}^* = [\tau_{\text{hip}}^T \quad \tau_{\text{knee}}^T \quad \tau_{\text{ankle}}^T]^T = B_{\text{dist}} \tau_v \quad (26)$$

其中, τ_{hip} , τ_{knee} , τ_{ankle} 分别为髋关节扭矩、膝关节扭矩、踝关节扭矩; $B_{\text{dist}} \in \mathbf{R}^{n \times 2}$ 为补偿分配矩阵, 其元素反映不同关节对虚拟腰部滚转与俯仰扭矩的响应能力和动态耦合程度. 一般而言, 髋关节在姿态补偿中起主导作用, 主要承担虚拟腰部俯仰与滚转分量; 膝关节提供辅助调节, 以实现角度过渡与冲击吸收; 踝关节则主要负责末端扭矩的细微平衡和接触稳定.

综合任务控制与补偿控制后, 为保证系统输出的物理可实现性及安全性, 需要对最终输出扭矩进行可行域约束, 得到:

$$\tau_j^{\text{safe}} = \text{proj}_{\mathcal{T}}(\tau_{\text{task}} + \tau_{\text{comp}}^*) \quad (27)$$

其中, $\text{proj}_{\mathcal{T}}(\cdot)$ 表示将扭矩投影到可行域 $\mathcal{T} = \{\tau : \tau_{\min} \leq \tau \leq \tau_{\max}\}$ 内, 以保证各关节扭矩不超过执行器能力范围.

2.3 观测空间及动作空间设计

为兼顾任务执行与补偿机制的耦合需求, 除了包含机体各关节角度、角速度等信息的一般性观测空间外^[23], 本节设计包含虚拟腰部信息的特权观测空间, 并将动作空间限制在实际可控关节范围内. 在训练阶段, 观测向量划分为两类: 普通观测数据与特权观测数据. 普通观测数据为机器人在实际运行中可直接获得的传感信息:

$$\mathbf{o}_{\text{obs}} = [\mathbf{q}_j, \dot{\mathbf{q}}_j, \tau_j, \tau_j^{\text{cmd}}, \mathbf{q}_b, \dot{\mathbf{q}}_b] \quad (28)$$

其中, $\mathbf{q}_j$, $\dot{\mathbf{q}}_j$, τ_j , τ_j^{cmd} 分别表示下肢各关节角度、角速度、扭矩及命令输入; $\mathbf{q}_b$, $\dot{\mathbf{q}}_b$ 为由 IMU 提供的基座姿态与角速度信息. 特权观测数据仅在训练阶段可用, 用于增强学习的表示能力:

$$\mathbf{o}_{\text{priv}} = [\mathbf{v}_b, \phi_v, \dot{\phi}_v, \tau_v] \quad (29)$$

其中, $\mathbf{v}_b$ 为机身线速度; ϕ_v , $\dot{\phi}_v$, τ_v 分别为虚拟腰部的关节角度、角速度与扭矩.

动作空间 $\mathcal{A}$ 包含机器人全身各关节的角度命令, 虚拟腰部自由度不直接进入动作空间, 其作用通过补偿机制间接体现.

综上, 本节建立了包含虚拟腰部滚转与俯仰自由度的扩展浮动基座动力学模型, 揭示了虚拟腰部与下肢关节之间的惯性耦合关系, 并提出通过双腿主动补偿实现等效扭矩的控制方法. 特权观测空间的设计使系统在学习与控制过程中能够显式利用虚拟关节信息, 为后续基于深度强化学习的全身协调控制提供了动力学基础与理论支撑.

3 基于 AMP 的深度强化学习

基于第 1 节中提取的人形机器人关节空间多维运动数据及第 2 节中构建的虚拟腰部关节机理, 本节面向人形机器人仿人运动控制问题, 构建一种融合任务约束与风格先验的深度强化学习框架. 该框架在近端策略优化 (proximal policy optimization, PPO) 算法的基础上, 引入 AMP 判别器, 实现对机器人行为风格与任务目标的联合优化. 通过设计任务-风格奖励可变融合机制, 有效缓解多种奖励项之间的相互干扰, 提升所学习运动控制策略的自然度与稳定性.

3.1 深度强化学习控制框架及任务奖励设计

人形机器人的运动控制问题可形式化为一个马尔科夫决策过程 (Markov decision process, MDP)^[29], 记为 $(\mathcal{S}, \mathcal{A}, f, r_t, p_0, \gamma)$. 其中, $\mathcal{S}$ 表示状态空间, 包含机器人当前的关节角度、角速度、基座姿态与线速度等观测量; $\mathcal{A}$ 为动作空间, 对应四肢驱动扭矩或期望关节加速度; $f(s, a)$ 表示系统的动力学转移函数, 描述状态转移规律, s 为状态空间, a 为动作空间; p_0 为初始状态分布; $\gamma \in (0, 1)$ 为折扣因子, 用于平衡长期与短期收益.

深度强化学习的目标是寻找最优策略 π_{θ} : $\mathcal{S} \mapsto \mathcal{A}$, 使期望折扣回报最大化:

$$J(\theta) = \mathbb{E}_{\pi_{\theta}} \left[\sum_{t=0}^{T-1} \gamma^t r_t(s_t, a_t, s_{t+1}) \right] \quad (30)$$

在本框架中, 策略优化基于 PPO 算法. 通过策略更新引入截断概率比, 以保证策略更新的稳定性, 即优化以下目标函数:

$$L(\theta) = \mathbb{E}_t \left[\min \left(r_t^{\text{clip}} \hat{A}_t, r_t \hat{A}_t \right) \right] \quad (31)$$

其中 $r_t = \pi_{\theta}(a_t|s_t)/\pi_{\theta_{\text{old}}}(a_t|s_t)$ 为策略概率比, $\pi_{\theta_{\text{old}}}$ 为上一轮策略; $r_t^{\text{clip}} = \text{clip}(r_t, 1 - \epsilon, 1 + \epsilon)$ 为截断概率比; $\hat{A}_t$ 为优势函数估计; ϵ 为截断阈值. 该设计通过在更新过程中引入截断概率比以限制策略变化幅度, 从而在保证优化稳定性的同时实现高效的策略改进, 适用于人形机器人仿人运动控制任务.

虚拟腰部姿态一致性奖励设计. 虚拟腰部的滚转与俯仰角 $\phi_{v,r}, \phi_{v,p}$ 表征了机身与下肢之间的动态协调状态, 其期望角度 $\tilde{\phi}_v = [\tilde{\phi}_{v,r}, \tilde{\phi}_{v,p}]^T$ 对应于平衡姿态或目标运动轨迹. 设计如下姿态一致性奖励函数:

$$r_{tt}^v = \exp\left(-(\phi_v - \tilde{\phi}_v)^T W_v (\phi_v - \tilde{\phi}_v)\right) \quad (32)$$

其中 $W_v = \text{diag}\{w^{vr}, w^{vp}\}$ 为姿态误差加权矩阵, 分别调节滚转与俯仰方向上的误差敏感度. 该奖励项通过加权最小化虚拟腰部姿态偏差, 引导策略在动力学耦合约束下学习隐式的姿态补偿行为, 使机器人在行走与外部扰动条件下保持平衡与自然的全身协调运动. 仅依赖任务奖励可能导致运动在动力学层面上不自然, 例如出现周期性振荡或姿态偏差, 因此需要引入风格约束以规范其动作分布.

3.2 AMP 判别器设计

为使学习到的控制策略在实现任务目标的同时具备与人体参考数据一致的运动风格, 本节引入基于对抗学习的动作先验机制. 该方法通过判别器网络学习真实人类动作的分布特征, 并以此指导策略生成符合人类风格的运动轨迹^[30]. 在对抗式模仿学习框架中, 总奖励函数由任务分量与风格分量构成:

$$r_{tt} = w^g r_{tt}^g + w^s r_{tt}^s \quad (33)$$

其中 r_{tt}^g 为任务奖励; r_{tt}^s 为风格奖励; w^g 与 w^s 为对应的权重系数. 判别器 D_ϕ 由参数 ϕ 表示, 用于判断状态转移 (s_t, s_{t+1}) 是否来自自动捕数据 $\mathcal{D}$. 判别器的训练目标函数为:

$$\min_{\phi} \mathcal{L}_{\text{disc}} = \mathcal{L}_{\text{real}} + \mathcal{L}_{\text{fake}} + \frac{w^{\text{gp}}}{2} \mathcal{L}_{\text{gp}} \quad (34)$$

其中 $\mathcal{L}_{\text{real}} = \mathbb{E}_{(s, s') \sim \mathcal{D}} [(D_\phi(s, s') - 1)^2]$ 表示判别器在真实数据 (s, s') 上的重建误差, 用于鼓励判别器对真实样本输出接近 1 的值; $\mathcal{L}_{\text{fake}} = \mathbb{E}_{(s, s') \sim \pi_\theta} [(D_\phi(s, s') + 1)^2]$ 表示在策略生成数据上的判别误差, 使得判别器对生成样本输出接近 -1 的值; w^{gp} 为梯度惩罚项权重系数; $\mathcal{L}_{\text{gp}} = \mathbb{E}_{(s, s') \sim \mathcal{D}} [\|\nabla_\phi D_\phi(s, s')\|^2]$ 为梯度惩罚项, 用于抑制判别器在真实数据流形上的非零梯度, 从而提升对抗训练的稳定性.

在策略更新阶段, 判别器输出被用作风格奖励的生成依据:

$$r_{tt}^s = \max[0, 1 - 0.25(D_\phi(s_t, s_{t+1}) - 1)^2] \quad (35)$$

通过对 D_ϕ 输出的缩放与偏移, 风格奖励被限制在 $[0, 1]$ 的范围内, 使其在不同任务阶段均具有平滑的影响. 高质量的动捕数据 $\mathcal{D}$ 可通过人类动作捕捉系统或高精度仿真模型生成. 最终, 策略与判

别器的优化目标分别为最大化期望回报与最小化判别误差, 实现生成策略与真实运动分布的一致性.

3.3 任务-风格奖励可变融合机制设计

任务奖励与风格奖励在本质上具有不同的优化方向: 前者关注动作结果的任务达成度, 后者关注动作过程的风格一致性. 若直接线性叠加两者, 可能导致学习过程中不同奖励项之间的竞争效应, 进而引起收敛不稳定或策略退化. 为此, 本节设计一种基于策略表现的可变融合机制, 根据当前策略在任务与风格方面的表现动态调整奖励权重. 定义融合系数 $\alpha_t \in [0, 1]$, 则总奖励函数可表示为:

$$r_{tt} = (1 - \alpha_t) r_{tt}^g + \alpha_t r_{tt}^s \quad (36)$$

在训练早期, 当策略尚未掌握基本运动结构时, 设置较小的 α_t , 使策略优先优化任务可行性; 随着训练稳定, 逐步增大 α_t , 加强风格约束以提升动作自然度. 融合系数的更新可与 PPO 中的策略性能指标结合, 实现自适应调节:

$$\alpha_{t+1} = \text{clip}\left(\alpha_t + \eta \frac{\Delta J_t}{J_{\text{max}}}, 0, 1\right) \quad (37)$$

其中, η 为更新步长; ΔJ_t 为最近若干步的期望回报增量; J_{max} 为历史最大奖励值. 该设计确保奖励在不同训练阶段的动态平衡, 使策略能够从任务驱动逐步过渡到风格驱动, 从而在任务完成度和动作自然度之间获得合理权衡.

在实际训练过程中, 深度强化学习策略输出为期望关节角度 $\mathbf{q}_j^{\text{ref}}$, 而非直接的驱动扭矩. 实际执行扭矩由各关节的比例-微分控制器产生, 即:

$$\boldsymbol{\tau}_j = K_p(\mathbf{q}_j^{\text{ref}} - \mathbf{q}_j) - K_d \dot{\mathbf{q}}_j \quad (38)$$

其中 K_p 与 K_d 分别为比例与微分增益矩阵. 为了在动力学层面实现虚拟腰部与多关节之间的主动补偿, 虚拟腰部扭矩 $\boldsymbol{\tau}_v$ 由式 (20) 计算得到, 并通过补偿分配矩阵 $B_{\text{dist}} \in \mathbf{R}^{n \times 2}$ 映射到实际关节空间:

$$\boldsymbol{\tau}_{\text{comp}} = B_{\text{dist}} \boldsymbol{\tau}_v \quad (39)$$

综合任务与补偿项, PD 控制器的期望关节角度更新为:

$$\mathbf{q}_j^{\text{ref}} = \mathbf{q}_{j, \text{task}}^{\text{ref}} + K_{\text{map}} \boldsymbol{\tau}_{\text{comp}} \quad (40)$$

其中 K_{map} 为角度-扭矩映射增益矩阵, 用于在训练中将虚拟腰部补偿扭矩转化为等效的关节角度修正. 通过该映射机制, 虚拟腰部姿态误差在姿态一致性奖励函数 (32) 的反馈下, 通过 PD 控制与分配矩阵 B_{dist} 的作用传递至各个实际关节动作层, 实现联合主动补偿.

4 实验验证与分析

为了验证本文所提 VW-MAC 方法在腰部物理自由度缺失人形机器人仿人运动控制中的有效性, 本文基于 Isaac Gym 仿真平台开展训练实验. 通过 MuJoCo 物理引擎进行高保真迁移验证, 从而测试 VW-MAC 框架对不同物理环境的泛化性及鲁棒性. 此外, 本文设计基准 AMP^[30] 与 VW-MAC 两套控制框架的对比及消融实验, 系统评估所提方法在仿人协调性及目标任务执行稳定性方面的综合性能.

4.1 实验平台与参数设置

实验平台基于宇树 G1 人形机器人构建. 为满足对抗式动作先验框架对全维仿人运动学习的自由度需求, 选用具备 29 个主动自由度的机器人模型作为基准硬件. 实验中保留上肢双臂及头部的完整自由度配置, 旨在实现人体运动风格的高保真迁移. 针对所提虚拟腰部关节框架在欠自由度约束下的补偿效能验证, 实验对机器人的腰部自由度进行了特殊限定: 锁定腰部俯仰与滚转关节, 仅保留偏航关节为激活状态. 此类构型模拟了典型的人形机器人腰部物理自由度缺失场景, 能够有效评估下肢冗余自由度对虚拟腰部运动轨迹的主动补偿能力. 在系统控制架构中, 核心运动策略的更新频率设定为 50 Hz, 关节层低级控制器的运行频率为 200 Hz, ω_{com} 设置为 0.5. 运动训练数据采自卡内基梅隆大学运动捕捉数据库, 涵盖了多种速度特征下的典型仿人行走动作序列. 实验设计主要包括以下三个维度:

1) 多速率运动测试. 将机器人运动速度指令设置为 0.2 m/s 至 1.6 m/s 范围内的多个离散值, 以验证所提控制框架在不同运动强度下的动态稳定性.

2) 奖励融合机制消融实验. 通过设计可变融合机制与固定权重融合机制的对照组, 评估基于特权

观测的奖励调节策略对运动协调性的影响.

3) 能量效率评估. 针对能量效率问题设计对比实验, 量化分析所提框架与常规简化模型控制器的能耗特征. 通过实时提取物理引擎中的执行器输出扭矩与角速度数据, 计算平均输运代价 (cost of transport, CoT), 旨在验证虚拟腰部关节在提升物理系统能量利用率方面的优势.

4.2 仿真实验结果与分析

4.2.1 离散指令下整体协调性测试

人形机器人在不同速度指令下保持全身动作协调与速度跟踪稳定, 是评价仿人运动控制方法有效性的关键指标. 在 Isaac Gym 并行训练中采用 0 m/s 至 1.5 m/s 的课程学习速度指令, 并将策略迁移至 MuJoCo, 以 10 Hz 的频率采集机器人质心及左右髋、踝关节点轨迹. 如图 4 所示, VW-MAC 在 0.2 m/s 至 1.6 m/s 离散速度指令下均保持平滑连续的质心轨迹, y 、 z 方向偏移较小, 全身协调性良好. 尽管动捕数据中人体参考速度仅为 1.2 m/s, 当指令提高至 1.6 m/s 时机器人仍保持较小的偏航与滚转波动, 体现了一定的超范围泛化能力. 左右髋、踝轨迹与质心轨迹的相对关系表明, 机器人在各速度下保持了运动的周期性与对称性, 有效缓解了腰部物理自由度缺失引起的下肢失配. 从图 5 ~ 7 的质心轨迹投影可见: 在 0.5 m/s 下沿 x 方向行进 2 m 时, y 向偏移为 0.27 m; 速度增至 1.0 m/s 时行进 4 m, 偏移缩至 0.12 m; 1.5 m/s 时行进 6 m, 偏移仅为 0.10 m, 相较 1.0 m/s 略有缩小. 图 8 和图 9 给出了 1.5 m/s 和 0.5 m/s 下的逐帧动作切片, 其中原色机器人表示 VW-MAC 生成的动作, 红色机器人表示相同速度指令和时刻下基准 AMP 的动作, cmd 为机器人运动速度指令. 当速度为 0.5 m/s 时, AMP 在 3.5 s 已出现左腿髋关节侧偏, 到 6.0 s 因

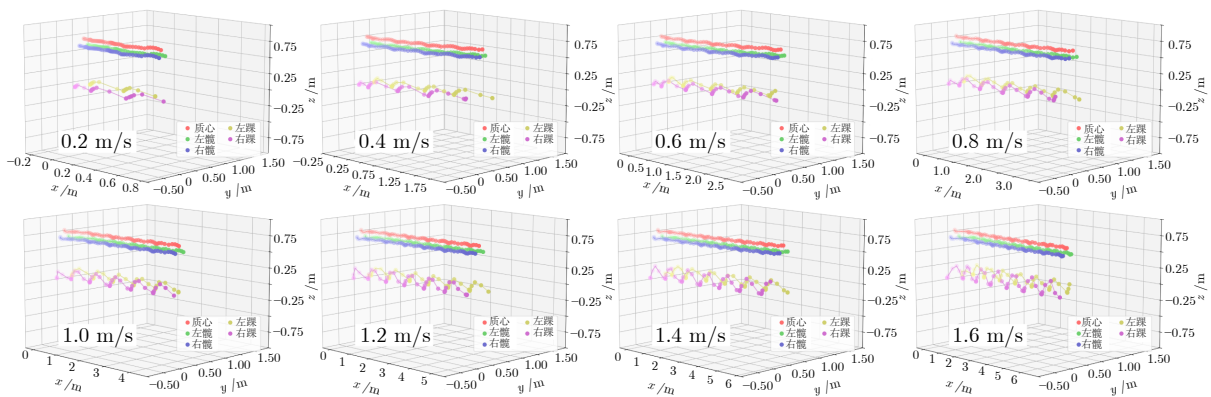


图 4 离散速度指令下机器人质心、髋关节及踝关节空间轨迹

Fig. 4 Spatial trajectories of the robot's center of mass, hip joints, and ankle joints under discrete velocity commands

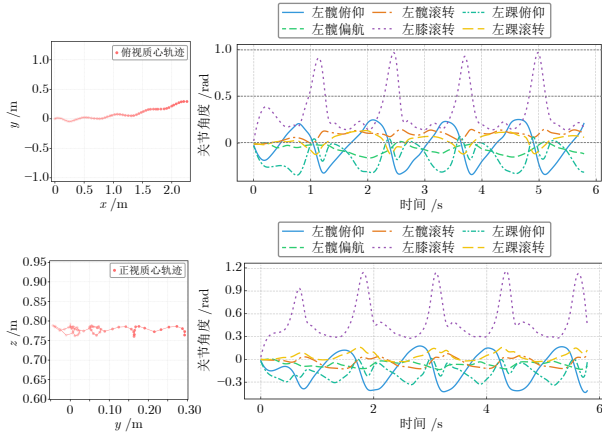


图5 速度指令 0.5 m/s 下的性能

Fig.5 Performance under 0.5 m/s velocity command

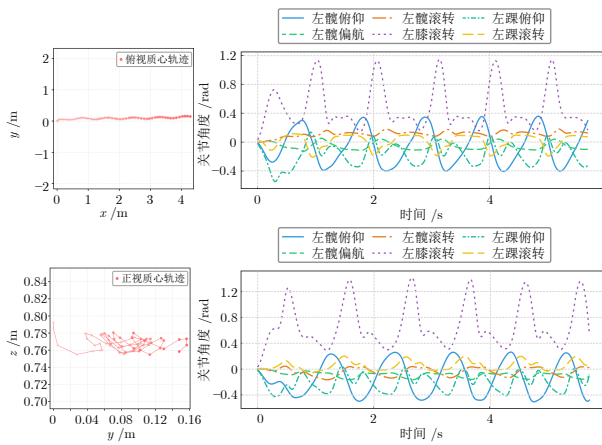


图6 速度指令 1.0 m/s 下的性能

Fig.6 Performance under 1.0 m/s velocity command

误差累积产生明显相位滞后; 当速度提升至 1.5 m/s 时, 该问题进一步放大, AMP 在 4.0 s 右腿抬起阶段即出现较大侧偏, 到 6.0 s 已难以完成跟踪, 而 VW-MAC 仍能通过腿部关节协同补偿偏移, 保持较稳定的行进轨迹. 此外, 训练中每 5.0 s 施加幅值 0.5 m/s 的随机脉冲扰动以模拟推拉干扰, 结果表明 VW-MAC 在 1.6 m/s 下的质心轨迹偏差较基准分别降低约 12.0% 和 15.0%, 说明该框架在复杂扰动下具有更强的恢复能力.

4.2.2 典型速度指令下的基线对比与消融实验

为分析虚拟腰部关节机理及任务-风格奖励融合机制对仿人运动控制性能的影响, 本文开展了基准 AMP 与 VW-MAC 的对比及奖励融合机制消融实验. 如表 1 所示, 根据不同运动速度指令下左右腿主要关节角度转动范围的差异, 评估下肢运动对称性. 在 0.5 m/s 下, 基准 AMP 的髋关节与膝关节对称性误差均不低于 0.167 267 rad, 而 VW-MAC

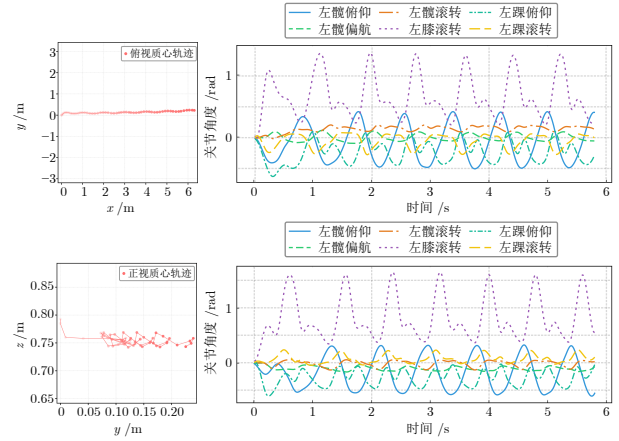


图7 速度指令 1.5 m/s 下的性能

Fig.7 Performance under 1.5 m/s velocity command

将膝关节对称性误差控制在 0.028 924 rad 以内; 速度提升至 1.0 m/s 和 1.5 m/s 时, AMP 的最大对称性误差增至 0.248 462 rad, VW-MAC 始终不超过 0.175 321 rad, 表现出更稳定的对称控制性能. 在消融实验中, 去除任务-风格奖励的可变融合机制后, AMP 在髋关节和膝关节上的对称性误差分别增加 0.003 067 rad 和 0.021 458 rad (分别对比最小的差值), 而 VW-MAC 退化更为明显, 在 1.0 m/s 下膝关节误差由 0.119 843 rad 增至 0.249 911 rad. 这表明相较于固定奖励融合方式^[25, 30], 可变融合机制能更有效地平衡任务执行与姿态对称性, 是提升多速度仿人运动平稳性的关键因素.

4.2.3 能量效率评估

仿真数据从定量指标和动态分布两方面验证了 VW-MAC 的节能效果. 在 1.2 m/s 稳定行走条件下, 引入虚拟腰部主动补偿机制后, CoT 由基准组的 0.556 降至 0.353, 能量利用效率提升 36.5%. 对比图 10 和图 11 的关节能耗热力图可见, 基准组能耗主要集中于髋、膝关节, 这是由于腰部锁定后上身惯性扰动需由支撑腿强力代偿, 从而产生高幅值脉冲扭矩; VW-MAC 则通过下肢冗余自由度协同抵消动力学干扰, 使关节负荷分布更为平顺均衡. 从图 12 和图 13 的系统总功耗曲线可进一步看出, 相比基准控制器的剧烈瞬时功率波动, VW-MAC 显著抑制了峰值功率, 且平均功率基准线明显下移, 说明多关节主动补偿机制在优化扭矩分配和削减物理功耗方面具有一定优势.

5 结束语

本文提出一种新型人形机器人仿人运动控制方法 VW-MAC, 通过融合虚拟关节机理与基于 AMP 的深度强化学习, 实现机器人全身协调的仿人运动

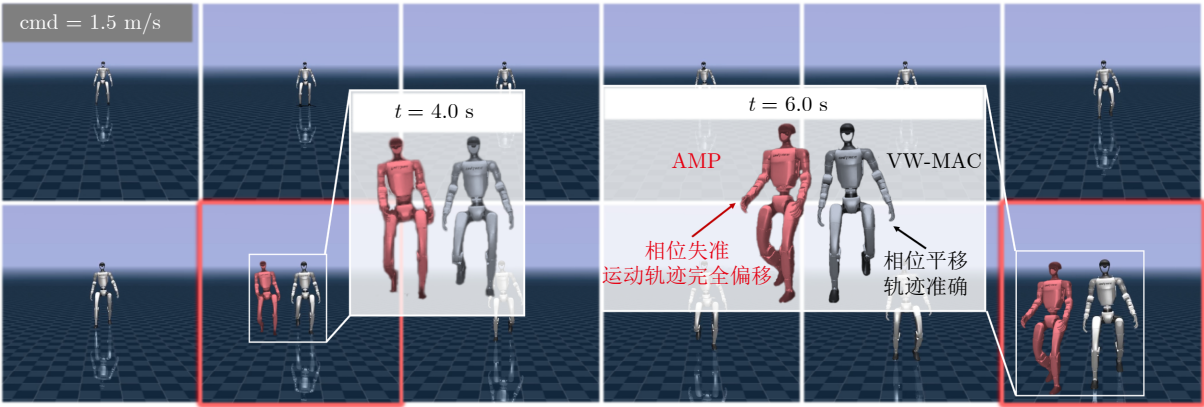


图 8 速度指令 1.5 m/s 下人形机器人行走关键帧切片

Fig.8 Keyframe slices of humanoid robot walking at a 1.5 m/s velocity command

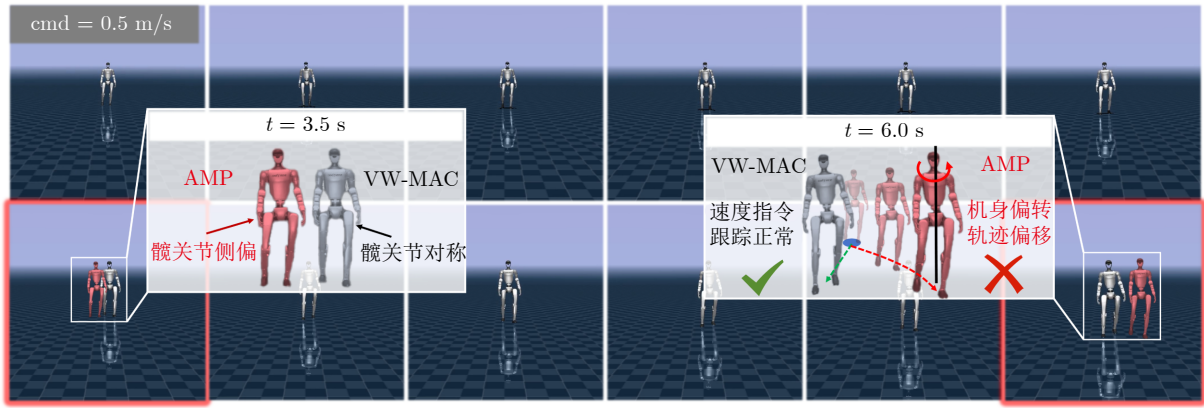


图 9 速度指令 0.5 m/s 下人形机器人行走关键帧切片

Fig.9 Keyframe slices of humanoid robot walking at a 0.5 m/s velocity command

表 1 AMP^[30] 与 VW-MAC 在人形机器人行走动作中的左右腿对称性误差对比 (rad)

Table 1 Comparison of left-right leg symmetry error between AMP^[30] and VW-MAC during humanoid robot walking (rad)

有无可变融合机制	运动速度指令	算法	髋关节俯仰↓	髋关节滚转↓	髋关节偏航↓	膝关节滚转↓	踝关节俯仰↓	踝关节滚转↓
有可变融合机制	cmd = 0.5 m/s	AMP	0.029 943	0.038 185	0.038 442	0.167 267	0.049 829	0.064 260
		VW-MAC	0.008 752	0.025 621	0.025 128	0.028 924	0.016 321	0.019 843
	cmd = 1.0 m/s	AMP	0.006 264	0.015 836	0.024 013	0.232 104	0.248 462	0.052 555
		VW-MAC	0.005 491	0.014 215	0.011 845	0.119 843	0.135 621	0.048 926
无可变融合机制	cmd = 1.5 m/s	AMP	0.021 222	0.029 997	0.034 885	0.241 722	0.186 154	0.067 396
		VW-MAC	0.019 843	0.027 621	0.032 128	0.128 924	0.175 321	0.062 843
	cmd = 0.5 m/s	AMP	0.013 942	0.045 621	0.048 166	0.188 946	0.065 191	0.072 661
		VW-MAC	0.011 667	0.041 237	0.042 465	0.178 993	0.058 221	0.068 186
无可变融合机制	cmd = 1.0 m/s	AMP	0.009 331	0.022 136	0.034 272	0.262 222	0.279 343	0.074 722
		VW-MAC	0.008 541	0.026 443	0.031 635	0.249 911	0.266 621	0.068 837
	cmd = 1.5 m/s	AMP	0.029 396	0.039 741	0.052 112	0.263 180	0.211 401	0.092 015
		VW-MAC	0.026 962	0.035 516	0.047 535	0.257 881	0.203 851	0.082 095

控制. 主要贡献包括: 1) 构建虚拟腰部关节耦合的多关节主动补偿机制, 有效缓解了腰部物理自由度缺失引起的上下肢运动失配; 2) 在 AMP 框架下引

入虚拟关节的对抗式动作先验风格奖励与任务奖励可变融合机制, 提升仿人运动的自然度与稳定性. 实验结果表明, 所提方法可在多变速度指令下保持

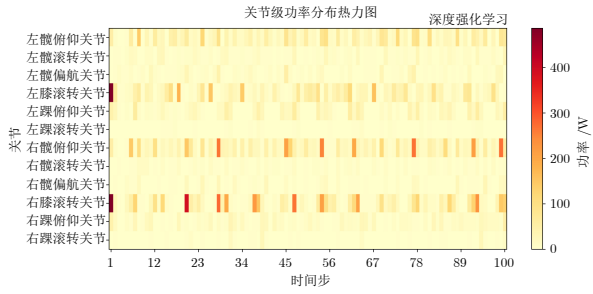


图 10 基于深度强化学习的关节能耗分布

Fig.10 Joint power consumption distribution based on deep reinforcement learning

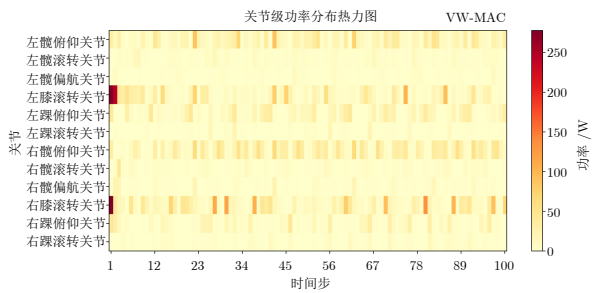


图 11 基于 VW-MAC 的关节能耗分布

Fig.11 Joint power consumption distribution based on VW-MAC

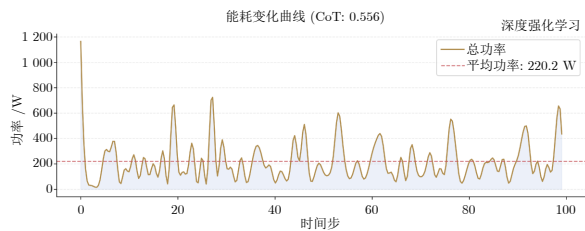


图 12 基于深度强化学习的总能耗

Fig.12 Total power consumption based on deep reinforcement learning

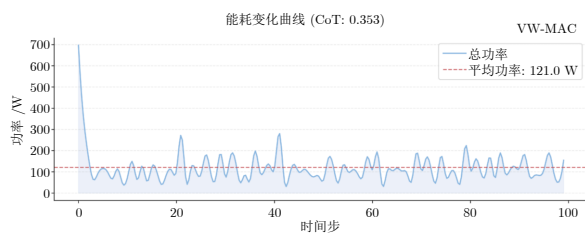


图 13 基于 VW-MAC 的总能耗

Fig.13 Total power consumption based on VW-MAC

质心轨迹稳定跟踪, 显著缓解腰部物理自由度缺失条件下的人形运动畸变. 未来工作将重点研究虚拟关节在复杂地形下的全身协调控制效果, 以提升方法的泛化性与鲁棒性.

参考文献

- 1 Agravante D J, Cherubini A, Sherikov A, Wieber P-B, Kheddar A. Human-humanoid collaborative carrying. *IEEE Transactions on Robotics*, 2019, **35**(4): 833–846
- 2 Ding J, Santina C D, Lam T L, Pang J, Xiao X, Tsagarakis N G, et al. Robust humanoid locomotion via sequential stepping and angular momentum optimization. *IEEE Transactions on Industrial Electronics*, 2025, **72**(3): 2779–2789
- 3 Han Lian-Qiang, Chen Xue-Chao, Yu Zhang-Guo, Gao Zhi-Fa, Huang Yan, Huang Qiang. Balance control of underactuated biped robot for discrete terrain. *Acta Automatica Sinica*, 2022, **48**(9): 2164–2174
(韩连强, 陈学超, 余张国, 高志发, 黄岩, 黄强. 面向离散地形的欠驱动双足机器人平衡控制方法. 自动化学报, 2022, **48**(9): 2164–2174)
- 4 Dallard A, Benallegue M, Scianca N, Kanehiro F, Kheddar A. Robust bipedal walking with closed-loop MPC: Adios stabilizers. *IEEE Transactions on Robotics*, 2025, **41**: 4679–4698
- 5 Kim M-J, Lim D, Park G, Lee K, Park J. A model predictive capture point control framework for robust humanoid balancing via ankle, hip, and stepping strategies. *IEEE Transactions on Robotics*, 2025, **41**: 3297–3316
- 6 Xin Ya-Xian, Li Yi-Bin, Chai Hui, Rong Xue-Wen, Li Bin. Research on jumping method of two wheeled-leg robot based on whole-body torque control. *Acta Automatica Sinica*, 2023, **49**(8): 1635–1644
(辛亚先, 李贻斌, 柴汇, 荣学文, 李彬. 基于全身力矩控制的双腿轮机器人跳跃方法研究. 自动化学报, 2023, **49**(8): 1635–1644)
- 7 Li G, Ijspeert A, Hayashibe M. AI-CPG: Adaptive imitated central pattern generators for bipedal locomotion learned through reinforced reflex neural networks. *IEEE Robotics and Automation Letters*, 2024, **9**(6): 5190–5197
- 8 Tong Y, Liu H, Zhang Z. Advancements in humanoid robots: A comprehensive review and future prospects. *IEEE/CAA Journal of Automatica Sinica*, 2024, **11**(2): 301–328
- 9 Zhang C, Bai J, Chen Z, Gao J, Qiao H. Hierarchical control for robust standing stability and fall recovery of task-performing humanoid robots. *IEEE Transactions on Automation Science and Engineering*, 2025, **22**: 20920–20933
- 10 Zhang Nan-Jie, Chen Yu-Quan, Ji Mao-Qin, Sun Yun-Kang, Wang Bing. Adaptive control method for quadruped robot facing floors of different roughness. *Acta Automatica Sinica*, 2025, **51**(7): 1585–1598
(张楠杰, 陈玉全, 季茂沁, 孙运康, 王冰. 面向不同粗糙程度地面的四足机器人自适应控制方法. 自动化学报, 2025, **51**(7): 1585–1598)
- 11 Wu Xiao-Guang, Liu Shao-Wei, Yang Lei, Deng Wen-Qiang, Jia Zhe-Heng. A gait control method for biped robot on slope based on deep reinforcement learning. *Acta Automatica Sinica*, 2021, **47**(8): 1976–1987
(吴晓光, 刘绍维, 杨磊, 邓文强, 贾哲恒. 基于深度强化学习的双足机器人斜坡步态控制方法. 自动化学报, 2021, **47**(8): 1976–1987)
- 12 Wu F, Nal X, Jang J, Zhu W, Gu Z, Wu A, et al. Learn to teach: Sample-efficient privileged learning for humanoid locomotion over real-world uneven terrain. *IEEE Robotics and Automation Letters*, 2025, **10**(9): 9048–9055
- 13 Rozlivek J, Roncone A, Pattacini U, Hoffmann M. Harmonious: Human-like reactive motion control and multimodal perception for humanoid robots. *IEEE Transactions on Robotics*, 2025, **41**: 378–393
- 14 Gu Z, Li J, Shen W, Yu W, Xie Z, McCrory S, et al. Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and learning [Online], available: <https://arxiv.org/abs/2501.02116>, April 19, 2025
- 15 Viceconte P M, Camoriano R, Romualdi G, Ferigo D, Daffarra S, Traversaro S, et al. Adherent: Learning human-like trajectory generators for whole-body control of humanoid robots. *IEEE Ro-*

- botics and Automation Letters*, 2022, **7**(2): 2779–2786
- 16 Oh J, Sim O, Cho B, Lee K, Oh J-H. Online delayed reference generation for a humanoid imitating human walking motion. *IEEE/ASME Transactions on Mechatronics*, 2021, **26**(1): 102–112
 - 17 Ji M, Peng X, Liu F, Li J, Yang G, Cheng X, et al. ExBody2: Advanced expressive humanoid whole-body control. In: Proceedings of the RSS 2025 Workshop on Whole-body Control and Bimanual Manipulation: Applications in Humanoids and Beyond. Los Angeles, USA: OpenReview.net, 2025.
 - 18 Huang P, Hu Y, Nechyporenko N, Kim D, Talbott W, Zhang J. EMOTION: Expressive motion sequence generation for humanoid robots with in-context learning. *IEEE Robotics and Automation Letters*, 2025, **10**(8): 7699–7706
 - 19 Wang Y, Li W, Togo S, Yokoi H, Jiang Y. Design of a 3-DOF coupled tendon-driven waist joint. In: Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA). Xi'an, China: IEEE, 2021. 7046–7051
 - 20 Tian Y, Su X, Shen C, Ma X. Exponentially extended dissativity-based filtering of switched neural networks. *Automatica*, 2024, **161**: Article No. 111465
 - 21 Tian Y, Su X, Kwong S K, Shen C. A hybrid output feedback control scheme of Markovian jump systems via partly transition rates/probabilities design. *Science China Information Sciences*, 2026, **69**(1): Article No. 112206
 - 22 Li Q, Meng F, Yu Z, Chen X, Huang Q. Dynamic torso compliance control for standing and walking balance of position-controlled humanoid robots. *IEEE/ASME Transactions on Mechatronics*, 2021, **26**(2): 679–688
 - 23 Prasad V, Heitlinger L, Koert D, Stock-Homburg R, Peters J, Chahvatzaki G. Learning multimodal latent dynamics for human-robot interaction. *IEEE Transactions on Robotics*, 2025, **41**: 4418–4438
 - 24 Tang A, Hiraoka T, Hiraoka N, Shi F, Kawaharazuka K, Kojima K, et al. HumanMimic: Learning natural locomotion and transitions for humanoid robot via Wasserstein adversarial imitation. In: Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA). Yokohama, Japan: IEEE, 2024. 13107–13114
 - 25 Tazaki Y. Key pose-based dynamic humanoid motion generation using parallel multi-fidelity model predictive control. *IEEE Robotics and Automation Letters*, 2025, **10**(12): 12979–12986
 - 26 Dong R, Chang Q, Er M J, Zhong J, Ikuno S. Motion capture-based robotic imitation: A keyframeless implementation method using multivariate empirical mode decomposition. *IEEE/ASME Transactions on Mechatronics*, 2025, **30**(3): 1947–1958
 - 27 Ju F, Jin H, Yin H, Liu J, Zhang X, Zhao M, et al. A joint-space humanoid robot torso motion planning method based on spiking neural networks. *IEEE/ASME Transactions on Mechatronics*, 2026, **31**(2): 1377–1387
 - 28 Zambella G, Caporale D, Grioli G, Pallottino L, Bicchi A. Composite whole-body control of two-wheeled robots. *IEEE Transactions on Robotics*, 2025, **41**: 2301–2321
 - 29 Beranek R, Karimi M, Ahmadi M. A behavior-based reinforcement learning approach to control walking bipedal robots under unknown disturbances. *IEEE/ASME Transactions on Mechatronics*, 2022, **27**(5): 2710–2720

- 30 Peng X B, Ma Z, Abbeel P, Levine S, Kanazawa A. AMP: Adversarial motion priors for stylized physics-based character control. *ACM Transactions on Graphics*, 2021, **40**(4): 1–20



刘煜程 重庆大学自动化学院博士研究生。主要研究方向为人形机器人全身协调运动控制。

E-mail: yucheng_l@cqu.edu.cn

(**LIU Yu-Cheng** Ph.D. candidate at the School of Automation, Chongqing University. His main research interest is whole-body coordinated motion control of humanoid robots.)



田羽锋 重庆大学自动化学院副教授。主要研究方向为具身智能控制及机器人应用。

E-mail: tyf@cqu.edu.cn

(**TIAN Yu-Feng** Associate professor at the School of Automation, Chongqing University. His main research interest is embodied intelligence control and robotic applications.)



杨玥 西安建筑科技大学信息与控制工程学院副教授。主要研究方向为智能无人系统的安全控制。

E-mail: yangyue@xauat.edu.cn

(**YANG Yue** Associate professor at the College of Information and Control Engineering, Xi'an University of Architecture and Technology. Her main research interest is the security control of intelligent unmanned systems.)



苏晓杰 重庆大学自动化学院教授。主要研究方向为具身智能及无人系统应用。本文通信作者。

E-mail: sxj@cqu.edu.cn

(**SU Xiao-Jie** Professor at the School of Automation, Chongqing University. His research interests include embodied intelligence and unmanned systems applications. Corresponding author of this paper.)