



清华大学
Tsinghua University

清华大学自动化系综合论文训练 选题汇报

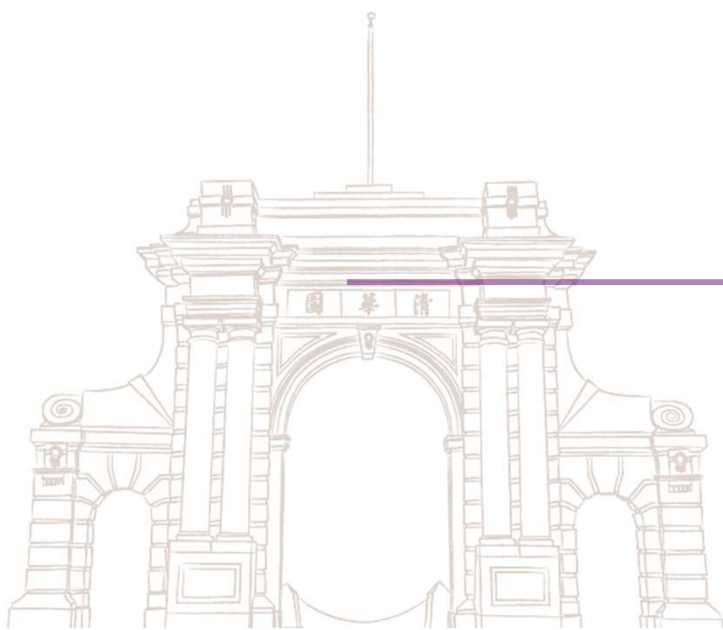
——基于人形机器人本体感觉的多地形
强化学习步态生成方法



指导教师：张涛

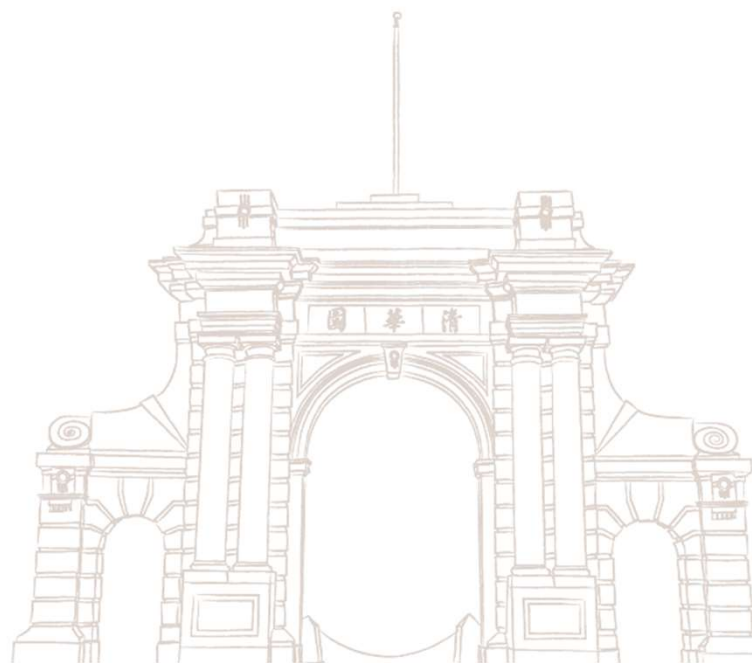
报告人：李昭阳

2024.11.28



CONTENTS

- 01 研究背景和意义
- 02 文献综述
- 03 研究内容
- 04 方案设计
- 05 日程计划与安排



01

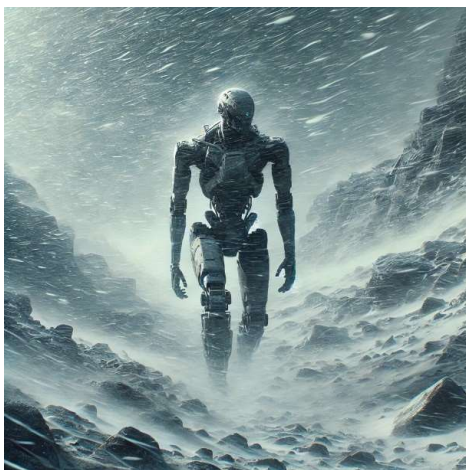
研究背景和意义





研究背景

- 人类的运动能力范围广、步态类型多
- 人类生活的环境丰富多样，而且主要为双足行走而设计
- 在人形机器人的计划使用场景中存在大量低外界感知的场景（雨雪、灾区、夜间等）
- 传统的强化学习算法追求泛化性，希望以同一个网络解决所有地形的步态规划



研究意义

- 仅依靠本体感觉来识别和适应各种地形，增强人形机器人在极端环境中的自主性和适应性
- 根据识别结果显式选取合适步态，确保人形机器人在多变环境中保持平衡并更高效地移动



02

研究现状





目前已有的人形机器人

- Boston Dynamic的Atlas机器人
- Tesla的Optimus机器人
- 宇树科技的人形机器人H1、G1
- 北京人形机器人创新中心的“天工”机器人
- 上海人形机器人创新中心的“青龙”机器人

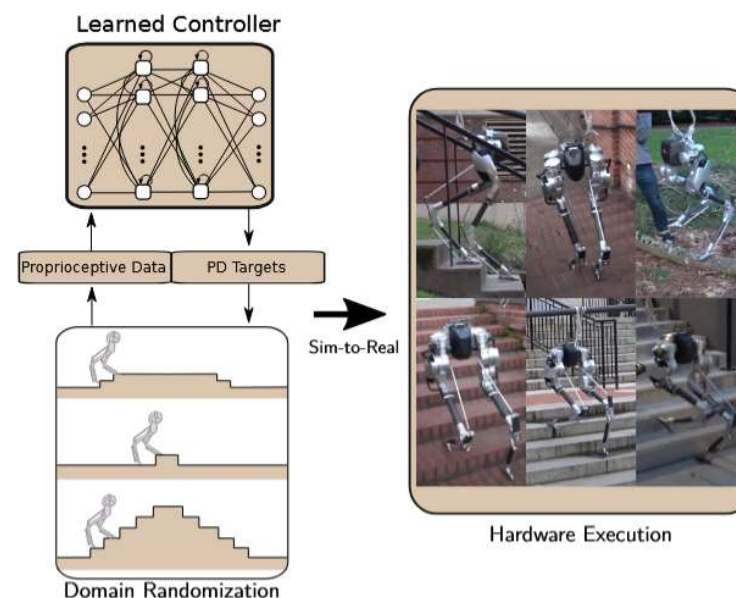




人形机器人步态规划方法

——Blind Bipedal Stair Traversal via Sim-to-Real Reinforcement Learning

- 修改现有的平地训练RL框架，加入楼梯状地形的随机化，无需改变奖励函数
- 在训练中使用镜像损失项以确保控制策略不学习不对称的步态
- 首次展示双足机器人仅使用本体感知穿越真实世界的楼梯

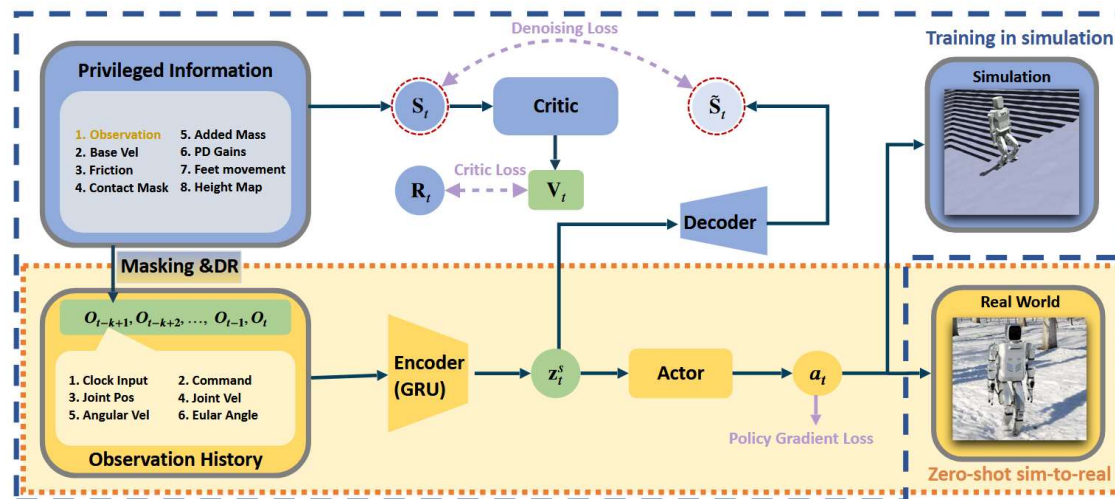




人形机器人步态规划方法

——Advancing Humanoid Locomotion: Mastering Challenging Terrains with Denoising World Model Learning

- 第一个通过端到端强化学习掌握真实世界挑战性地形的人形机器人
- 提出了一个新颖的强化学习框架，以弥合模拟到现实的差距，并实现鲁棒的泛化性能





基于本体感知的地形预测方法

文章索引	机器人类型	使用非本体感觉	具体方法
[42]	四轮月球车	x	主成分分析
[43]	四轮车	x	快速傅里叶变换
[44]	扫地机器人	√	触觉探针
[45]	四轮车	√	多传感器融合
[46]	三轮车	x	机器学习
[47]	异形月球车	x	机器学习



03

研究内容





研究内容

- 通过强化学习预训练一个基础“骨架网络”，设计专家强化学习算法，进一步训练步态库
- 通过深度学习等方法，仅利用关节角度、姿态信息和平衡状态等本体感觉数据（不依赖视觉和雷达信息），预测当前地形
- 结合PID控制实现自适应时变步态顺滑切换

研究创新点

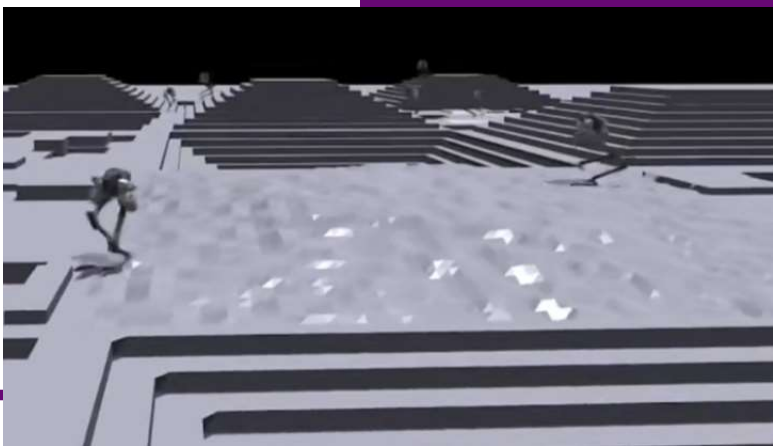
- 专注地形的专家强化学习算法
- 基于深度学习的本体感觉的地形预测算法
- 自适应时变步态顺滑切换



04

方案设计





Isaac Gym仿真环境搭建

- Isaac Gym是由NVIDIA推出的一个高效且可靠的机器人仿真器
- 模拟复杂的多地形环境，包括平坦地面、斜坡和不平坦地面等关键地形条件
- 我将在Isaac Gym中先建造一个平面地形，以训练整体算法的骨架部分
- 再建造多种不同的地形，训练不同的网络参数





PPO强化学习算法部署与数据收集

- PPO算法被用来优化机器人在不同地形的步态
- 反“泛化性”，设置“专家”神经网络，为每种地形情况采用不同的神经网络参数，形成“步态库”
- 收集最终神经网络模型中机器人的关节角度、力矩、速度、姿态变化等数据，为地形预测分类积累数据集





分类神经网络搭建

- 对上述收集的数据集进行**预处理**，将收集到的关节数据、力矩数据和姿态数据进行标签化处理
- 使用**支持向量机或深度学习**方法，训练地形分类神经网络
- 采用交叉验证方法，确保网络的表现不会受到**过拟合或欠拟合**的影响
- 最终目标是构建一个高效、可靠的地形分类预测器，以指导多输出头强化学习神经网络选择正确输出头





测试实验设计

- 仿真实验
 - 评估地形分类器的准确率、平均精确度、平均召回率
 - 记录机器人步长、步速及姿态等关键参数
 - 对照传统强化学习方法的步态效果
- 实机实验
 - 在完成上述实验的情况下，考虑对神经网络进行sim2real优化
 - 在条件允许的情况下，将经过充分验证的训练算法应用在实体机器人上进行测试
 - 仿真结果进行对比，评估算法的泛化能力与适应性



05

日程计划与安排





日程计划与安排

内容	时间安排
搭建步态规划的模拟环境，继续调研强化学习算法	2024.12.01-2024.12.31
通过PPO强化学习算法训练步态库	2024.01.01-2025.02.15
设计并训练基于本体感觉的地形分类预测算法	2025.02.16-2025.03.15
对未预期的异常进行处理	2025.03.16-2025.03.31
仿真和实机实验	2025.04.01-2025.05.15
论文撰写	2025.05.16-2025.05.30





清华大学
Tsinghua University

Thank you for watching



李昭阳 2024.11.28

