

# 研究报告

题目：足式机器人强化学习控制调研

研 究 生 彭道杰

## 摘 要

足式机器人相对于常见的轮式和履带式机器人有着更加灵活，环境适应能力强的优势。与启发它们的人类或者动物一样，这类机器人拥有着更强的越野能力，能够适应复杂的地形环境，到达普通轮式或履带式机器人无法到达的地方。然而，这类机器人往往具有更多的自由度，这使得对它们的控制变得具有挑战性。如何让这类足式机器人像它们的动物模仿对象一样能够灵活、优雅地运动成为一个重要的研究课题。在这其中，模仿四足动物的四足机器人（机械狗）是一类被广泛研究的对象。

机械狗是一个多自由度运动系统，其典型自由度数量为 12 个。每条腿有三个自由度分别实现小腿伸缩、大腿伸缩和整腿伸缩。由于它的求解空间巨大而可用的解空间相对较小，因此机械狗的运动控制常常是启发式的——预先定义一些运动，以缩小求解时的搜索空间。经典的基于运动动力学建模和优化问题的控制方式主要有轨迹优化 (*Trajectory Optimization, TO*) 和模型预测控制 (*Model Predictive Control, MPC*)。由于求解需要涉及大量的非线性优化问题，这类控制方法求解算力开销较大往往不能够在线完成，需要上位机的辅助。相比较而言近年来发展起来的基于深度学习的强化学习控制方法在运行期间的算力开销很小。强化学习控制是一种通过构建神经网络并依靠大量训练优化神经网络参数以得到预期的控制效果的控制方法，可以完全不需要动力学建模。其训练过程往往在仿真环境中完成，因此需要对机械狗的物理模型进行建模，这些建模包括：机械狗的躯干及四肢的刚体模型、电机驱动关节的响应模型。

不过由于神经网络的黑盒性，单纯的强化学习控制往往只能针对特定的训练目标而很难与其它模型融合，缺乏跨模式泛化能力。因此经典控制和强化学习控制的融合控制变得十分重要。

**关键词：**机器人控制；轨迹优化；深度学习；强化学习控制；Isaac 仿真

# 目 录

摘 要	I
第 1 章 基于经典方式的机械狗控制	1
1.1 基于线性倒立摆模型的控制	1
1.1.1 机械狗的运动模型	1
1.2 基于质心动力学模型的控制	1
1.3 基于直接刚体动力学模型的控制	1
1.4 基于 MPC 的控制	1
第 2 章 基于强化学习的机械狗控制	2
2.1 单层神经网络直接驱动关节方式	2
2.2 双层神经网络本体感知方式	2
2.3 双层网络本体和外部感知融合方式	2
2.4 电机驱动关节点模型	2
第 3 章 强化学习部署实践和关键点	3
3.1 强化学习平台 Isaac	3
3.2 强化学习案例 legged_gym 库	3
3.3 PyTorch 的使用	3
第 4 章 几种构想的机器人类型草图	4
4.1 弹跳轮足 + 机械手机器人	4
4.2 轮鞋足式 + 机械手机器人	5
4.3 弹跳轮足 + 飞行器机器人	5
结 论	7
参考文献	8

## 第 1 章 基于经典方式的机械狗控制

足式机器人的运动规划是比较困难的，不仅因为它的自由度较多，更因为它的机体运动不能被直接得出，而是要通过四肢状态及四肢与环境的接触产生。

经典控制的基础是对机器人的运动学和动力学建模。经典控制主要方式有轨迹优化 (*Trajectory Optimization, TO*) 和模型预测控制 (*Model Predictive Control, MPC*) 两类。常用的轨迹优化机械狗动力模型有：1. 基于有线性倒立摆的模型；2. 基于直接刚体动力学的模型；3. 基于质心动力学的模型。基于这些模型的优化控制都被描述为非线性优化问题，这些问题的求解算力开销较大往往不能够在线完成，需要上位机的辅助。对于 MPC 还不太熟悉，等看完了补充下 MPC 描述...

### 1.1 基于线性倒立摆模型的 TO 控制<sup>[1]</sup>p2-5

基于参考文献阐述基于 ZMP 模型的控制方法...

#### 1.1.1 机械狗的运动模型

基于参考文献阐述机械狗的一般运动学表述...

### 1.2 基于质心动力学模型的 TO 控制<sup>[2]</sup>p2-6

基于参考文献阐述基于接触力模型的控制方法...

### 1.3 基于直接刚体动力学模型的 TO 控制<sup>[3]</sup>p2-6

基于参考文献阐述基于刚体动力学模型的控制方法...

### 1.4 基于 MPC 的控制<sup>[4]</sup>p2-4

基于参考文献阐述基于 MPC 的控制方法...

## 第 2 章 基于强化学习的机械狗控制

### 2.1 单层神经网络直接驱动关节方式<sup>[5]</sup>p3

基于参考文献阐述单层神经网络直接驱动关节方式控制方法...

### 2.2 双层神经网络本体感知方式<sup>[6]</sup>p8

基于参考文献阐述双层神经网络本体感知方式控制方法...

### 2.3 双层网络本体和外部感知融合方式<sup>[7]</sup>p7

基于参考文献阐述双层网络本体和外部感知融合方式控制方法...

### 2.4 电机驱动关节点模型<sup>[8]</sup>p4

基于参考文献阐述电机驱动关节点模型描述...

## 第 3 章 强化学习部署实践和关键点

**Notes 3.0.1.** 两个参考的 RL 深度学习案例库：

*isaacgym* 库

*legged\_gym* 库

### 3.1 关于强化学习平台 Isaac<sup>[9]</sup>p6-10

这部分参考文献介绍对 Isaac 平台的基本理解，主要是它的并行工作流程和提供的读取和设置 API...

### 3.2 强化学习案例 legged\_gym 库<sup>[10]</sup>p1

这部分展示 legged\_gym 库关于 ANYmal-c 的训练代码，主要关注整个训练流程、奖励函数的设计、对 Isaac-API 的使用、地形设计、训练课程设计等与论文中提到的概念对应...

### 3.3 PyTorch 的使用

这部分结合 Pytorch 的官方文档和 legged\_gym 库提取其中关于 .pth 模型的构建和使用流程和实例...

## 第4章 几种构想的机器人类型草图

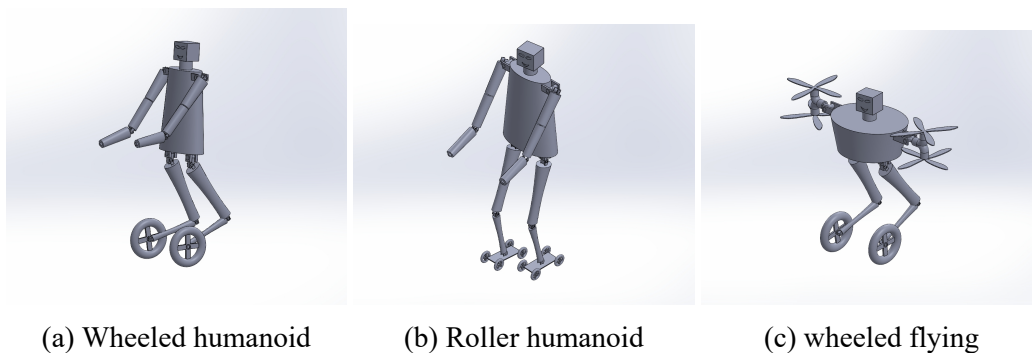


图 4-1 三种构想机器人类型的示意图汇总

如图4-1所示，这里给出几种近来构想的机器人结构草图和简单介绍，可以作为随后研究的机器人备选平台参考。

### 4.1 弹跳轮足 + 机械手机器人

如图4-2所示，该类型的机器人由类似于 Ascento<sup>[1]p1</sup> 类型机器人的双轮足，在此基础上添加两个机械臂构成类人形机器人。该类型的机器人可以轮式地行走和跳跃，同时可以用机械臂模仿双手进行各种操作。

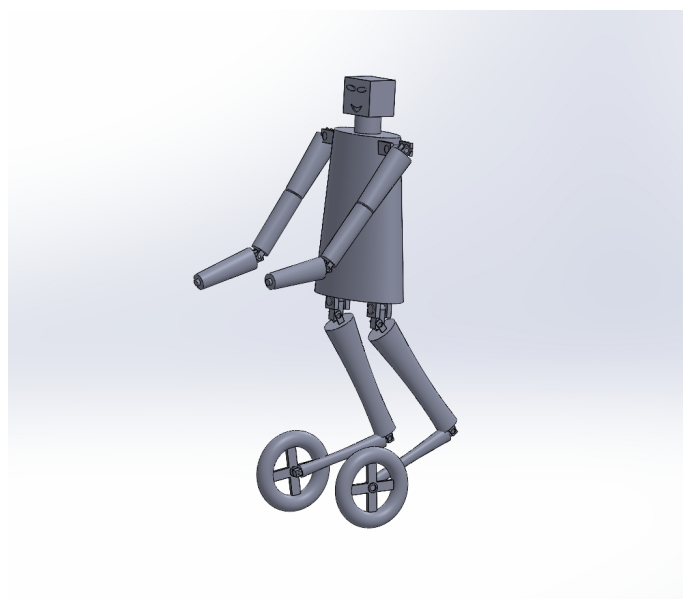


图 4-2 Schematic view of wheeled-humanoid robot

该类型机器人的控制基础是倒立摆模型和机械臂运动学，它相对于现有双轮

机器人，如 Ascento<sup>[1]p1</sup> 等，新的控制问题难点是：

- (1) 在机械手运动的同时保持双轮足的稳定，比如从空手到搬起重物的过程；
- (2) 如何让手臂配合轮足实现更加灵巧和优雅的动作；

## 4.2 轮鞋足式 + 机械手机器人

如图4-3所示，该类型的机器人由类似一般人形机器人的基本结构，在此基础上为每个脚添加四个轮子

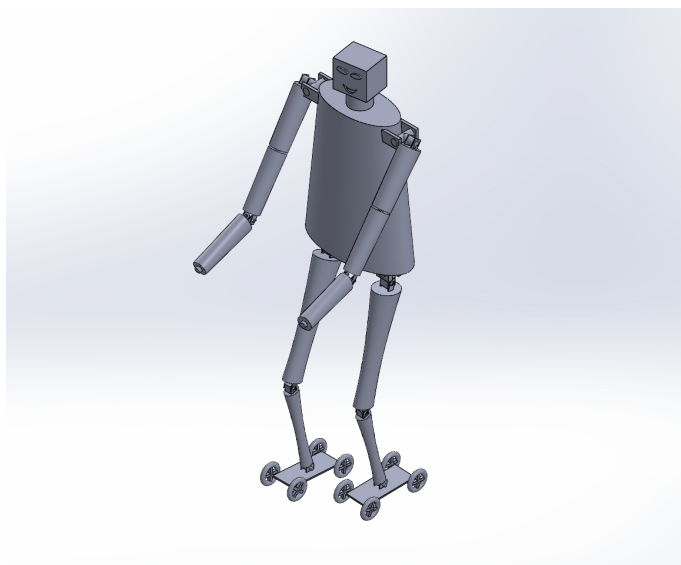


图 4-3 Schematic view of roller humanoid.

该类型机器人的控制基础是 ZMP 约束和机械臂运动学，它相对于现有人形机器人，如 Xxx 等，新的控制问题难点是：

- (1) 如何实现步态和轮式之间的运动切换；
- (2) 如何实现步态和轮式的联合运动平衡；
- (3) 在上面两点的基础上，如何结合机械臂实现更加灵巧、优雅、节能的动作；

## 4.3 弹跳轮足 + 飞行器机器人

如图4-4所示，该类型的机器人由类似于 Ascento<sup>[1]p1</sup> 类型机器人的双轮足，在此基础上添加一对旋翼使其具备飞行能力。

该类型机器人的控制基础是倒立摆模型和飞行器控制，它相对于现有双轮机器人，如 Ascento<sup>[1]p1</sup> 等，新的控制问题难点是：

- (1) 飞行器的矢量控制；



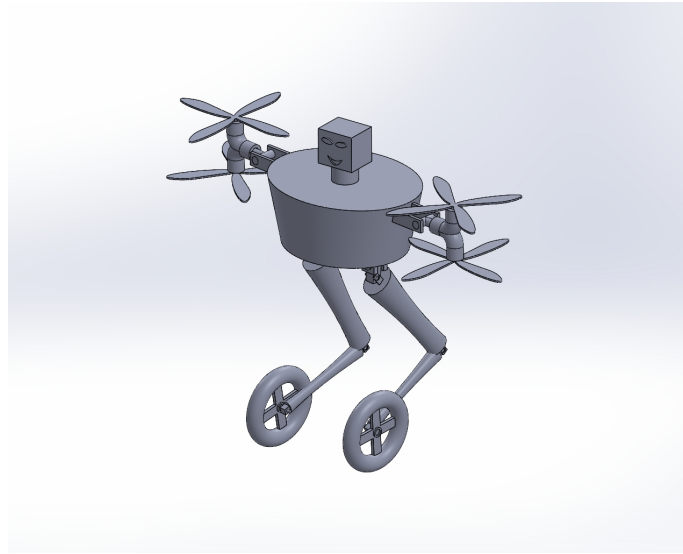


图 4-4 Schematic view of wheeled-flying robot.

- (2) 如何优化设计使得旋翼提供的升力可以满足飞行控制
- (3) 续航安全问题；
- (4) 如何让旋翼配合轮足实现更加灵巧、优雅、节能、稳定的动作；

## 结 论

对于经典控制的实现,涉及到复杂的运动模型和动力学建模,难点在于选择优化目标构建优化策略和优化问题的求解。经典控制的优点在于控制过程清晰,可以方便地进行迭代优化和泛化部署;缺点在于建模和实际机器调试复杂,运行时算力开销大。对于强化学习控制的实现,涉及到机器人的仿真模型建模和神经网络设计,难点在于设计合适的优化目标和训练策略以使得训练过程能较快收敛且结果达到预期目标。强化学习控制的优点在于模型设计和调试应用简单迅速,运行时算力开销小;缺点在于控制过程不清晰,导致模型的泛化能力较弱。

经典控制与深度学习控制各有优缺点,在实践应用中两者的有机结合有助于实现更好的控制。在面对一种机器人设计控制策略时,如何分配经典控制和深度学习控制是一个十分重要的问题。两者的融合也是未来的重要发展方向之一。

对于接下来的研究,就当前的情况来看,实现强化学习到实际机器人 *ANYmal* 上的部署还欠缺的是:

- (1) 一套 *ANYmal* 机器人硬件;
- (2) 可用的例程代码;
- (3) 电机、控制手柄、各类传感器的驱动和软件接口。
  - 在有直接可用例程的情况下,可以采用例程的架构和电机接口直接部署如2.1节所述的神经网络策略。
  - 在没有直接可用例程有类似参考例程的情况下,可以花一些时间阅读相关代码,逐步将电机、控制手柄、各类传感器等驱动重构到 *ANYmal* 上,然后在此基础上设计控制结构实现策略部署。
  - 在没有任何软件接口的情况下至少应该有各种器件相应的手册。会消耗比较长的时间来完成驱动和信息格式的整理和构建。完成后设计控制结构进行策略部署。
- (4) 对 *Isaac* 仿真环境的深入学习;
- (5) 对 *PyTorch* 框架的深入学习;

对于后续的研究,在上面的基础上需要进一步学习和尝试将经典策略应用到机器人上,以更加深入的理解经典控制方法。与此同时加强对强化学习理论的理解和仿真环境的使用能力,在接下来的研究中构思经典-强化学习融合的控制策略。在此基础上,考虑设计和实现新结构的机器人。

## 参考文献

- [1] BELLICOSO C, JENELTEN F, FANKHAUSER P, et al. Dynamic locomotion and whole-body control for quadrupedal robots[C/OL]//2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2017. <http://dx.doi.org/10.1109/iros.2017.8206174>.
- [2] WINKLER A W, BELLICOSO C D, HUTTER M, et al. Gait and Trajectory Optimization for Legged Systems Through Phase-Based End-Effector Parameterization[J/OL]. IEEE Robotics and Automation Letters, 2018: 1560–1567. <http://dx.doi.org/10.1109/lra.2018.2798285>.
- [3] PARDO D, NEUNERT M, WINKLER A, et al. Hybrid direct collocation and control in the constraint-consistent subspace for dynamic legged robot locomotion[C/OL]//Robotics: Science and Systems XIII. 2017. <http://dx.doi.org/10.15607/rss.2017.xiii.042>.
- [4] NEUNERT M, STAUBLE M, GIFTTHALER M, et al. Whole-Body Nonlinear Model Predictive Control Through Contacts for Quadrupeds[J/OL]. IEEE Robotics and Automation Letters, 2018: 1458–1465. <http://dx.doi.org/10.1109/lra.2018.2800124>.
- [5] HWANGBO J, LEE J, DOSOVITSKIY A, et al. Learning agile and dynamic motor skills for legged robots[J/OL]. Science Robotics, 2019. <http://dx.doi.org/10.1126/scirobotics.aau5872>.
- [6] LEE J, HWANGBO J, WELLHAUSEN L, et al. Learning Quadrupedal Locomotion over Challenging Terrain[J/OL]. Science Robotics, 2020. <http://dx.doi.org/10.1126/scirobotics.abc5986>.
- [7] MIKI T, LEE J, HWANGBO J, et al. Learning robust perceptive locomotion for quadrupedal robots in the wild[J/OL]. Science Robotics, 2022(VOL. 7, NO. 62). DOI: DOI:, [10.1126/scirobotics.abk2822](https://doi.org/10.1126/scirobotics.abk2822).
- [8] GEHRING C, COROS S, HUTTER M, et al. An Optimization-based Approach to Controlling Agile Motions for a Quadruped Robot[J/OL]. IEEE Robotics and Automation, 2016. DOI: [10.3929/ethz-a-010644954](https://doi.org/10.3929/ethz-a-010644954).
- [9] MAKOVICHUK V, WAWRZYNIAK L, GUO Y, et al. Isaac Gym: High Performance GPU-Based Physics Simulation For Robot Learning[A]. 2021.
- [10] RUDIN N, HOELLER D, REIST P, et al. Learning to Walk in Minutes Using Massively Parallel Deep Reinforcement Learning[A]. 2021.
- [11] KLEMM V, MORRA A, SALZMANN C, et al. Ascento: A Two-Wheeled Jumping Robot [C/OL]//2019 International Conference on Robotics and Automation (ICRA). 2019. <http://dx.doi.org/10.1109/icra.2019.8793792>.