问题背景

人类的运动能力涵盖了广泛的范围 [1]，能够适应各种任务所需的不同步态，例如站立、行走、快步走、小跑、奔跑和跳跃 [2]。这种适应性甚至包括蹲姿行走，使人类能够在广泛的动作范围中表现出卓越的灵活性和机动性。幸运的是，大多数人类运动可以被视为周期性或准周期性 [3] [4]，这为人形机器人运动控制器的设计提供了极大的便利。

在过去几十年中，已经开发了许多机器人及相应的控制算法，以实现周期性步态 [5] [6] [7]。尽管取得了显著进展，人形机器人在运动范围方面仍无法与人类相提并论。因此，尤其是像跑步这样的高动态步态的实现，已引起越来越多的关注。

通常有两种方法可以生成稳定的周期性步态，其中之一是基于模型的方法 [8]。这种方法通过建立简化或精确的模型来分析运动学或动力学，从而生成理想的参考轨迹，并通过基于模型计算的反馈控制来跟踪该轨迹，以确保机器人的行走模式稳定 [9] [10]。线性倒立摆模型（LIPM）是人形和四足机器人中应用最广泛的简化模型之一 [11] [12]。Kajita等人 [13] 将机器人建模为固定高度的倒立摆，将复杂的动力学简化为一个质心点质量，并使用预览控制方法结合零力矩点（ZMP）控制。基于LIPM，Takenaka等人 [14] 将其扩展为可变高度倒立摆模型（VHIP），并分析了运动的发散分量（DCM），以生成周期性跑步步态。Han等人 [15] 结合3D弹簧加载倒立摆（3D-SLIP）模型与任务空间非线性最小二乘优化器，实现了6.5 m/s的稳定跑步。单刚体模型（SRBM）在基于优化的方法中被广泛用于质心运动规划 [16] [17] [18]。通过手动指定接触顺序，并参考SRBM在线求解优化问题，Bellicoso等人 [19] 在四足机器人上实现了多种步态及平滑过渡。García等人 [20] 提出了结合SRBM和方向变量的MPC运动控制方法。通过模仿生物体的神经机制，中央模式发生器（CPG）使用耦合振荡器模型生成可调节频率和幅度的步态周期，以响应感官反馈和环境变化 [21] [22] [23] [24]。

模型方法的一个主要缺点是，这些简化模型通常假设机器人某些关节是被动的或锁定的 [24] [25]，只允许一种或有限的步态模式。这导致运动空间仅代表了机器人全部潜力的一小部分，限制了对所有驱动关节所提供的完整运动范围和灵活性的探索和利用。

另一种实现人形机器人稳定周期性步态的方法是强化学习。在过去十年中，强化学习越来越多地应用于人形机器人，实现了如爬台阶 [25] 和后空翻 [26] 等复杂动作。这种方法不需要精确的机器人建模，而是通过智能体与环境的交互逐步学习运动控制策略 [27] [28]，从而缩短了开发周期 [29]。通过在GPU上进行大规模并行计算，平地行走的训练时间已缩短至仅5分钟 [29]。强化学习更高效地生成了周期性步态。Wu等人 [30] 表明，通过使用简单的正弦信号来获得预期的摆动-支撑时序并将其作为观测的一部分，并设计奖赏函数以跟踪该时序，机器人可以实现固定频率的行走。Li等人 [31] 提出了傅里叶潜在动力学（FLD）方法，从人类运动中提取周期特征，通过傅里叶变换将动作捕捉的关节角度转化为频域信息，并利用解码器重建这些特征，以模仿原始人类动作。即使直接在奖赏函数中引入周期性的脚部提升和落地，强化学习也能够在不依赖周期性参考信号的情况下生成自适应步态模式。尽管RL目前仅限于特定步态，覆盖了运动能力的一小部分，但其简单性和自主学习特性为探索人形机器人的运动能力和扩展步态空间提供了潜力。

研究目标

在人形机器人领域，多地形步态规划是一个复杂且关键的问题。这种人形机器人需要能够在不同的地形上稳定行走，包括但不限于平坦地面、楼梯、斜坡、不平坦的地面等。步态规划的目标是确保机器人在这些多变环境中保持平衡，同时高效地移动。

为了实现这一目标，我预期开发一套基于强化学习的步态规划算法，该算法能够自主学习并优化机器人的步态，以适应不同的地形条件。这套算法将基于机器学习的原理，通过与环境的交互来学习如何调整步态参数，如步长、步速、足部位置和姿态调整等，以实现最佳的行走效率和稳定性。

具体预期成果包括：

1. **算法开发**：开发一种基于强化学习的控制算法，能够根据机器人的传感器输入（如关节角度、姿态信息、平衡状态等）来调整步态参数。
2. **仿真验证**：在仿真环境中训练并测试这些算法，以验证它们在不同地形上的有效性。仿真环境将模拟真实世界的复杂性，包括各种地形和可能的干扰因素。
3. **实验验证**：如果条件允许，在实体机器人上实施算法，并在控制的实验环境中进行测试，以评估算法在现实世界条件下的表现。
4. **性能优化**：根据仿真和实验结果，对算法进行调整和优化，以提高机器人的行走效率和稳定性。
5. **泛化能力**：确保算法不仅在特定地形上表现良好，而且能够泛化到未见过的地形上。。

工作计划

设计研究方法和实验方案

目标：选择合适的强化学习算法，搭建步态规划的模拟环境。

时间表：第4-6周完成。

任务清单：

选择或设计适合多地形步态规划的强化学习算法。

搭建仿真环境，包括不同地形的模拟。

设计初步的实验方案。

实施研究计划

目标：进行算法开发和实验验证。

时间表：第7-14周完成。

任务清单：

开发强化学习算法。

在仿真环境中测试算法。

收集实验数据。

分析实验结果

目标：总结实验数据，撰写论文草稿。

时间表：第15-18周完成。

任务清单：

分析实验数据，提炼关键发现。

撰写开题报告。

准备论文草稿。

根据反馈进行调整

目标：优化研究方法和实验方案。

时间表：第19-22周完成。

任务清单：

根据导师的反馈调整研究方案。

优化算法和实验设计。

形成最终论文和结题报告。

预期成果

开题报告：详细阐述研究背景、方法、预期成果和实验设计。

论文草稿：包含研究的所有关键部分，包括引言、文献综述、方法、结果和讨论。

风险评估

技术风险：算法开发可能遇到技术难题，需要提前规划备选方案。

时间风险：研究进度可能因各种因素延迟，需要定期检查进度并调整计划。

资源风险：仿真环境和实验设备可能不足，需要提前申请和准备。