

# 机械与动力工程学院

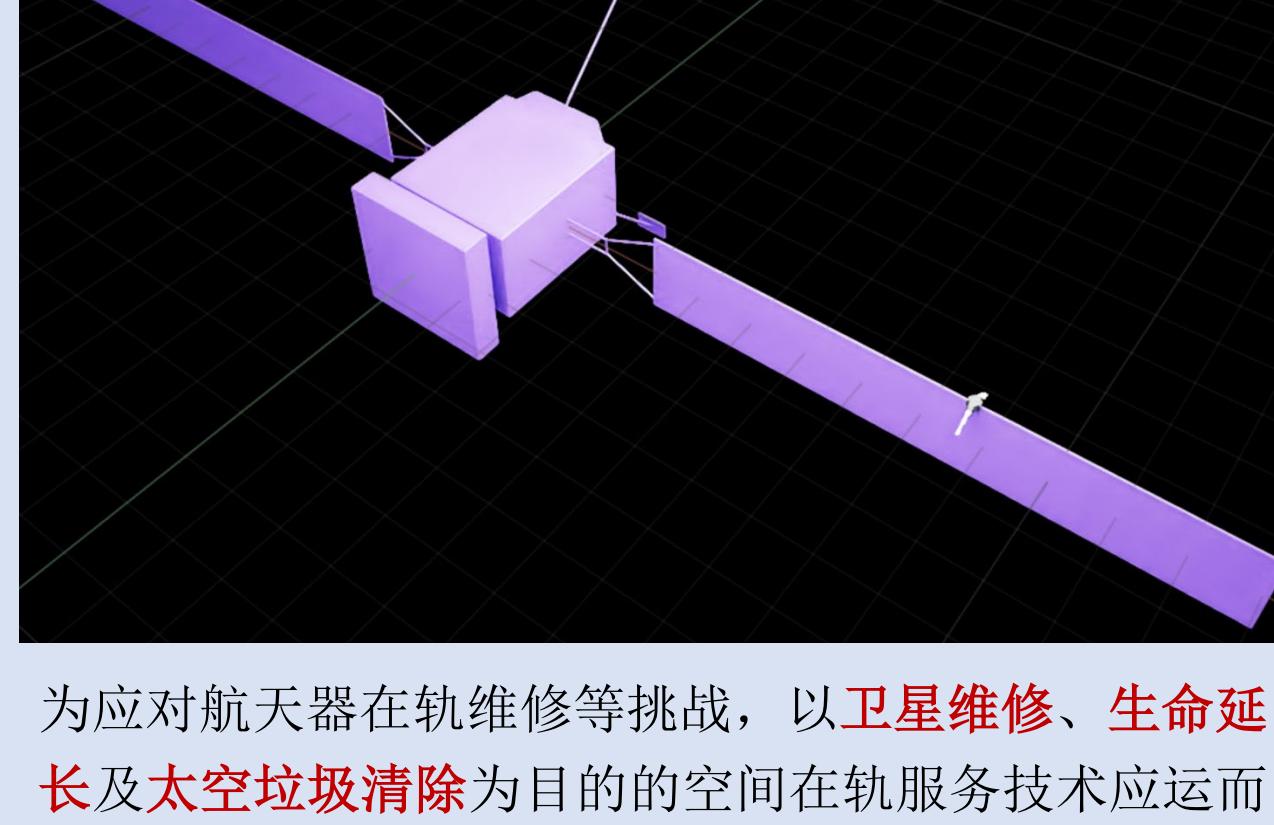
## 2025届本科优秀毕业设计/毕业论文展

2025年6月

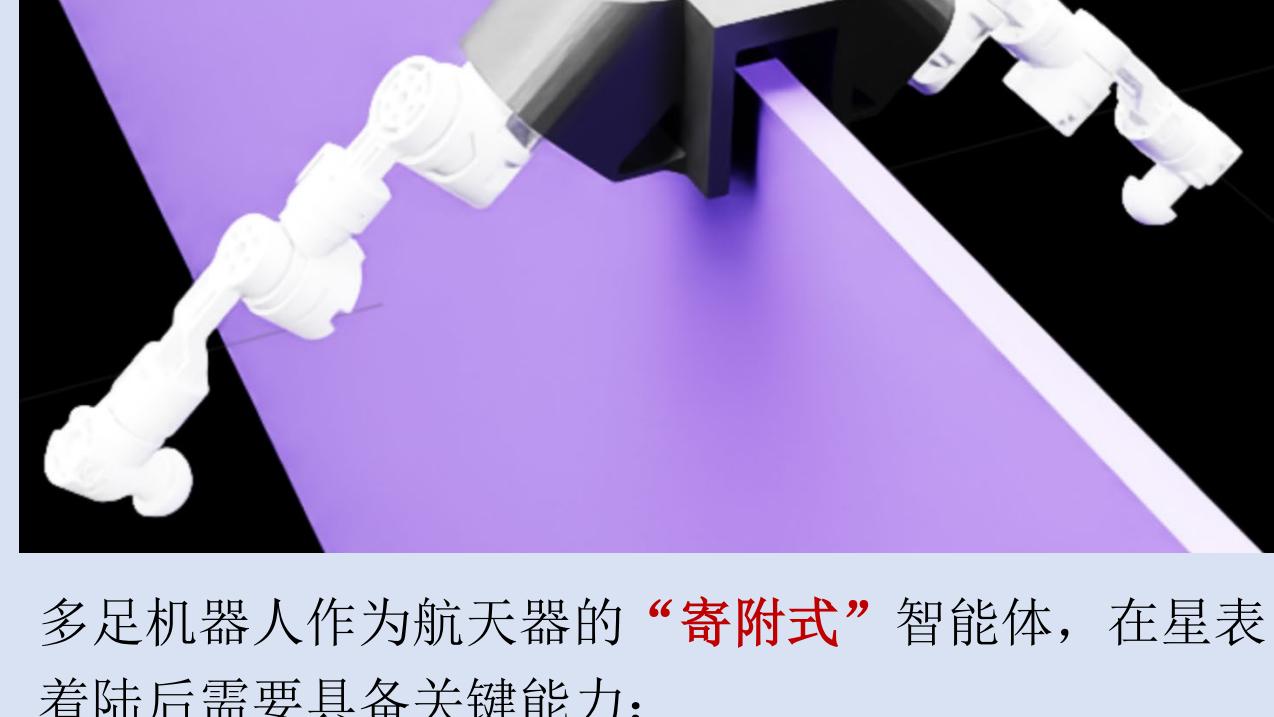
### 基于多模态的具身机器人感知与交互系统的开发与研究

学生姓名：应一源 学号：21011404 班级：机设213 指导教师：易建军教授

#### 课题背景



为应对航天器在轨维修等挑战，以卫星维修、生命延长及太空垃圾清除为目的的空间在轨服务技术应运而生，空间机器人成为关键研究方向。



多足机器人作为航天器的“寄附式”智能体，在星表着陆后需要具备关键能力：

- 在未知环境中实现自主感知与决策；
- 构建任务导向的智能闭环系统。

机器人技术要求：

- 建立“感知—决策—规划—控制”架构；
- 具备多模态运动切换能力；
- 适应复杂星表地形构型。

#### 技术体系

##### 基于多约束强化学习的星表行进方法

融合示教数据的模仿学习与并行强化学习的机器人智能行进控制技术

模仿学习数据采集 太空环境约束构建 多智能体并行训练

采集示教轨迹，提供运动参考

大量机器人并行训练，提升学习效率



##### 星表价值驱动的端到端运动方法

“感知—运动”一体化的行进运动控制技术

星表价值与原始感知融合

“感知—运动”控制联合训练

渐进式迭代训练

##### 星表机器人感知-行进地面训练验证方法

星表机器人本体设计

高保真仿真训练平台搭建

多自由度构型、灵巧末端设计

生成式星表场景仿真

多模态传感器部署

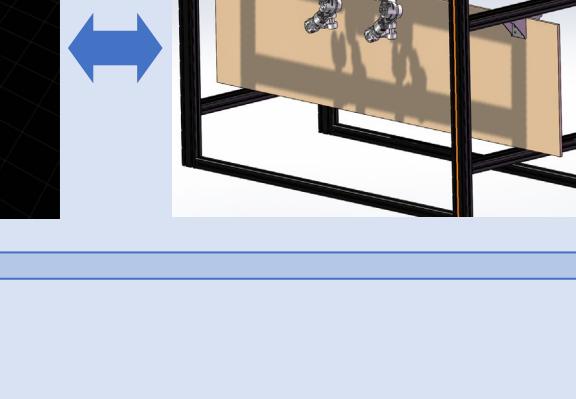
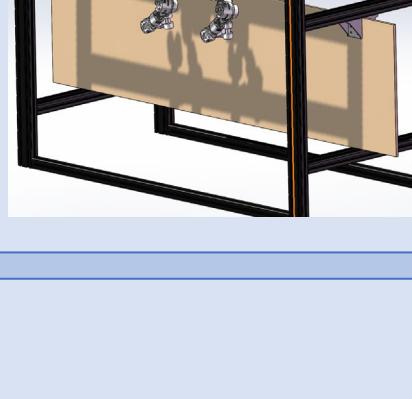
机器人端到端行为训练

高性能边缘计算平台部署

仿真-真实环境(Sim2Real)迁移学习

无重力或重力环境下机器人行进模拟

配合动捕系统的实时机器人行进模拟

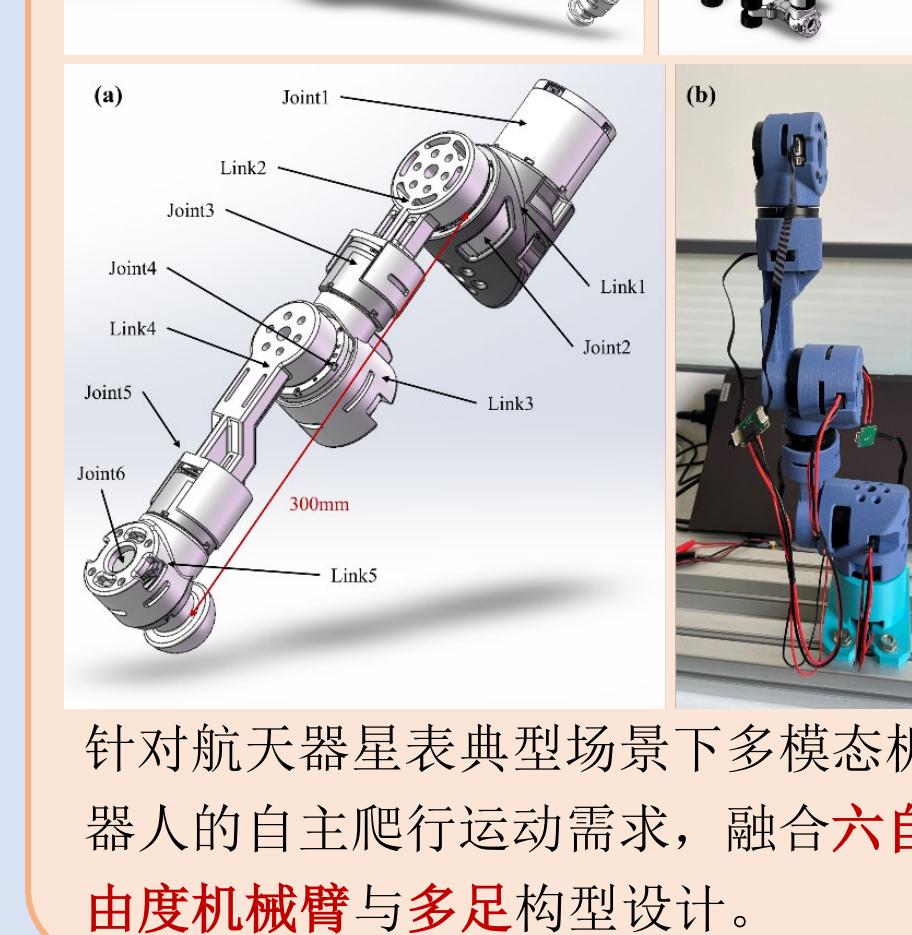


动作孪生映射

#### 本文架构

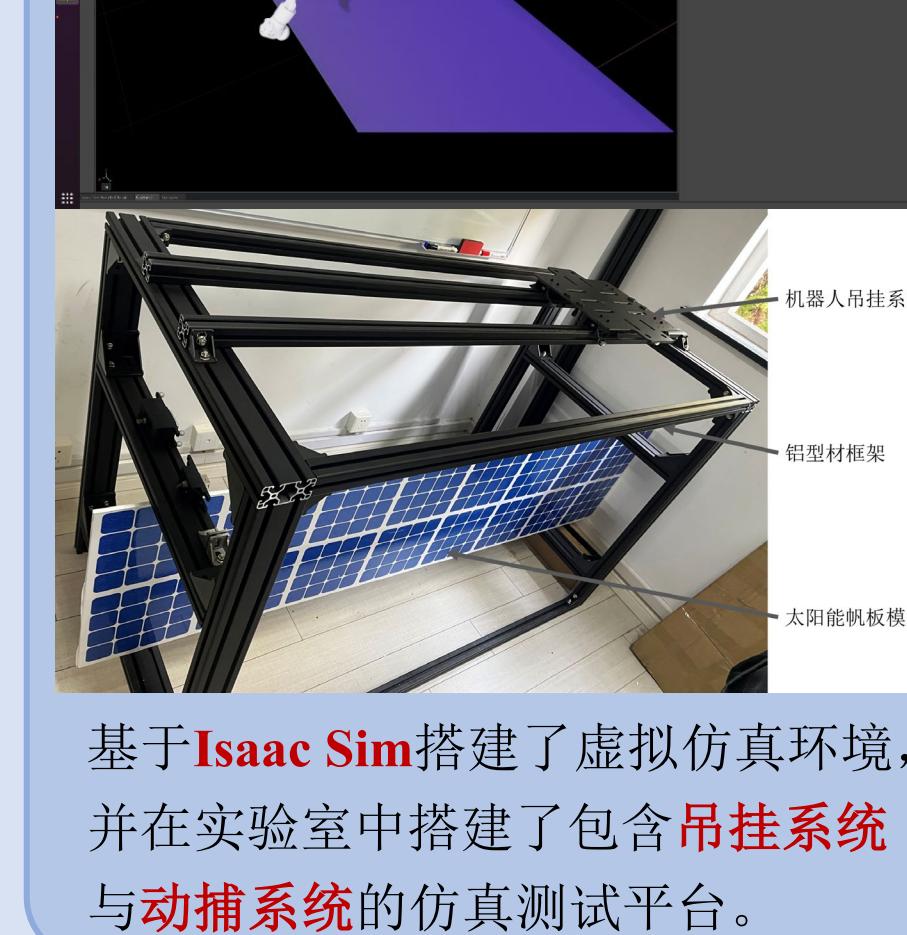
### 基于多模态的具身机器人感知与交互系统的开发与研究

#### 星表爬行机器人结构设计



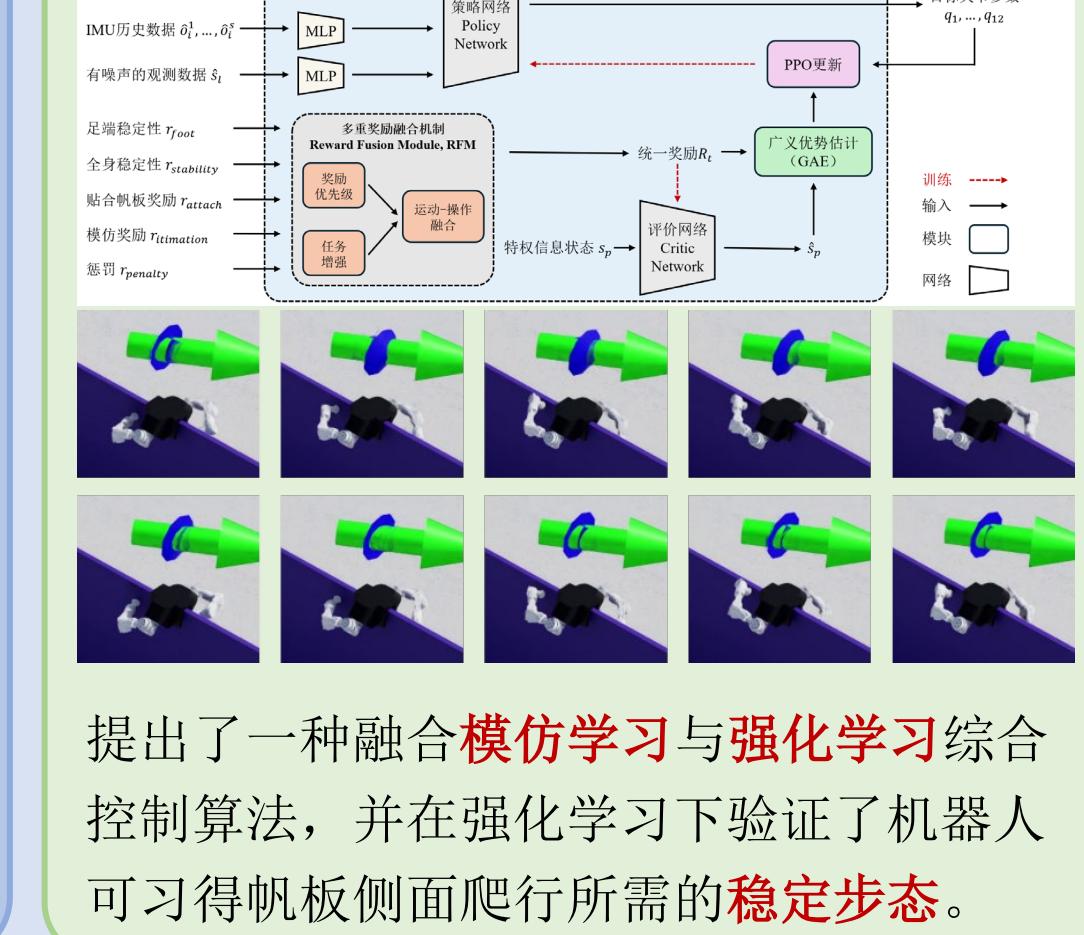
针对航天器星表典型场景下多模态机器人的自主爬行运动需求，融合六自由度机械臂与多足构型设计。

#### 仿真与测试平台搭建



基于Isaac Sim搭建了虚拟仿真环境，并在实验室中搭建了包含吊挂系统与动捕系统的仿真测试平台。

#### 强化学习爬行控制方法



提出了一种融合模仿学习与强化学习综合控制算法，并在强化学习下验证了机器人可习得帆板侧面爬行所需的稳定步态。

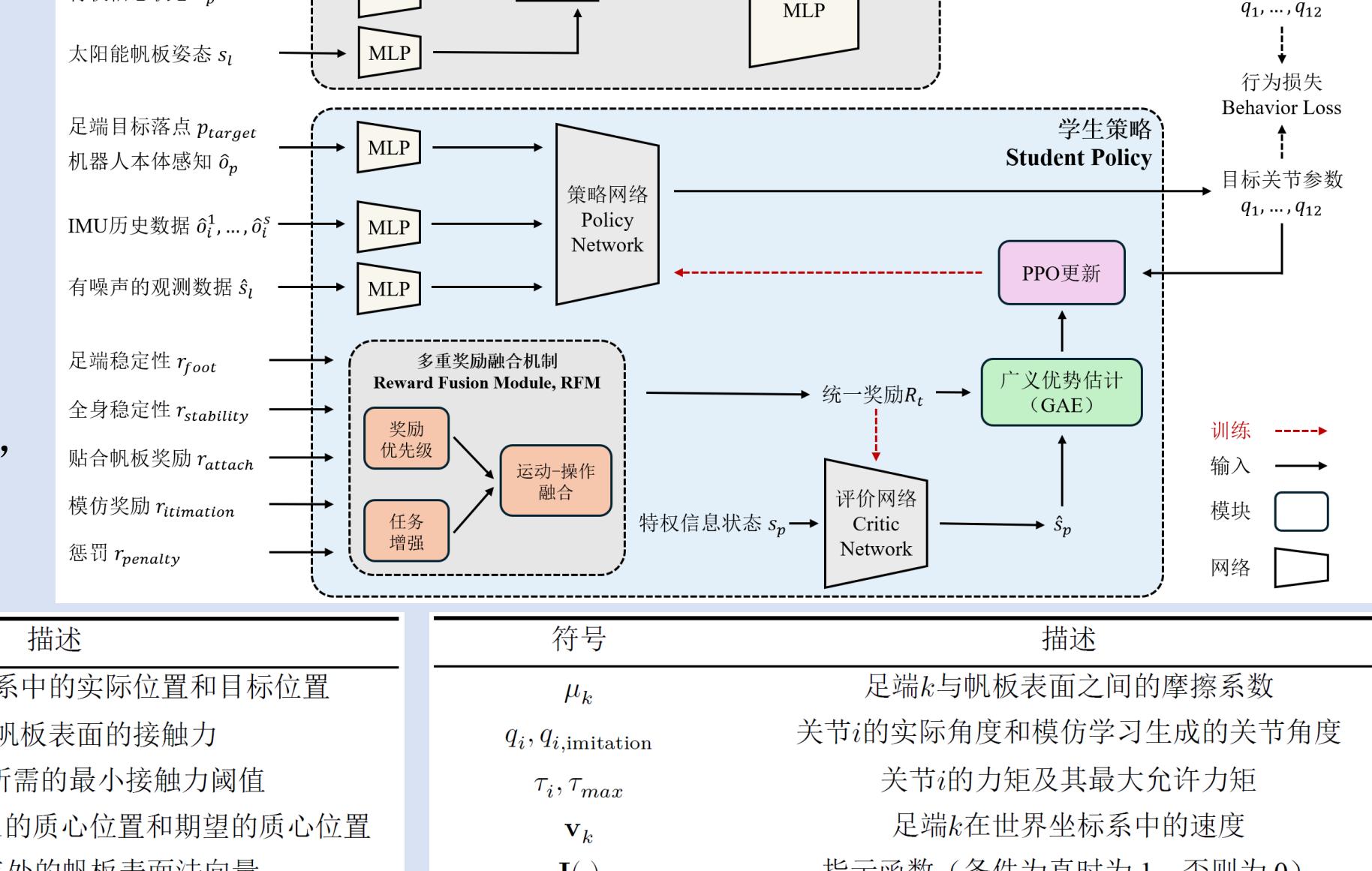
#### 机器人智能控制算法设计

结合强化学习与模仿学习，采用“教师-学生”训练模式。

教师策略：通过在无噪声环境

中采集示教数据，构建高质量轨迹，获得表现最优的动作序列，用以指导后续策略学习。

学生策略：通过与教师轨迹对齐并结合实际带噪声环境观测，学习出具有迁移性和泛化能力的策略模型。



总奖励函数  $r$ ：主要由多个子奖励项加权组合而成

$$r = w_1 r_{foot} + w_2 r_{stability} + w_3 r_{attach} + w_4 r_{imitation} + w_5 r_{penalty}$$

足端稳定性  $r_{foot}$ ：足端能够到达目标落点并保持稳定

$$r_{foot} = \mu_1 \left( - \sum_{k \in \text{feet}} \| \mathbf{p}_k - \mathbf{p}_k.\text{goal} \|^2 \right) + \mu_2 I(\| \mathbf{f}_k \| > f_{min})$$

全身稳定性  $r_{stability}$ ：确保机器人重心保持稳定，不偏移

$$r_{stability} = \lambda_1 \exp(-\lambda_3 \| \mathbf{c}_t - \mathbf{c}_{t-1} \|^2) - \lambda_2 \| \mathbf{c}_t - \mathbf{c}_{desired} \|^2$$

贴合帆板奖励  $r_{attach}$ ：确保足端与帆板侧面保持接触，避免打滑脱离

$$r_{attach} = v_1 \sum_{k \in \text{feet}} I(\mathbf{f}_k \cdot \mathbf{n}_k > \mu_k \| \mathbf{f}_k \|) - v_2 \sum_{k \in \text{feet}} I(\mathbf{f}_k \cdot \mathbf{n}_k < 0)$$

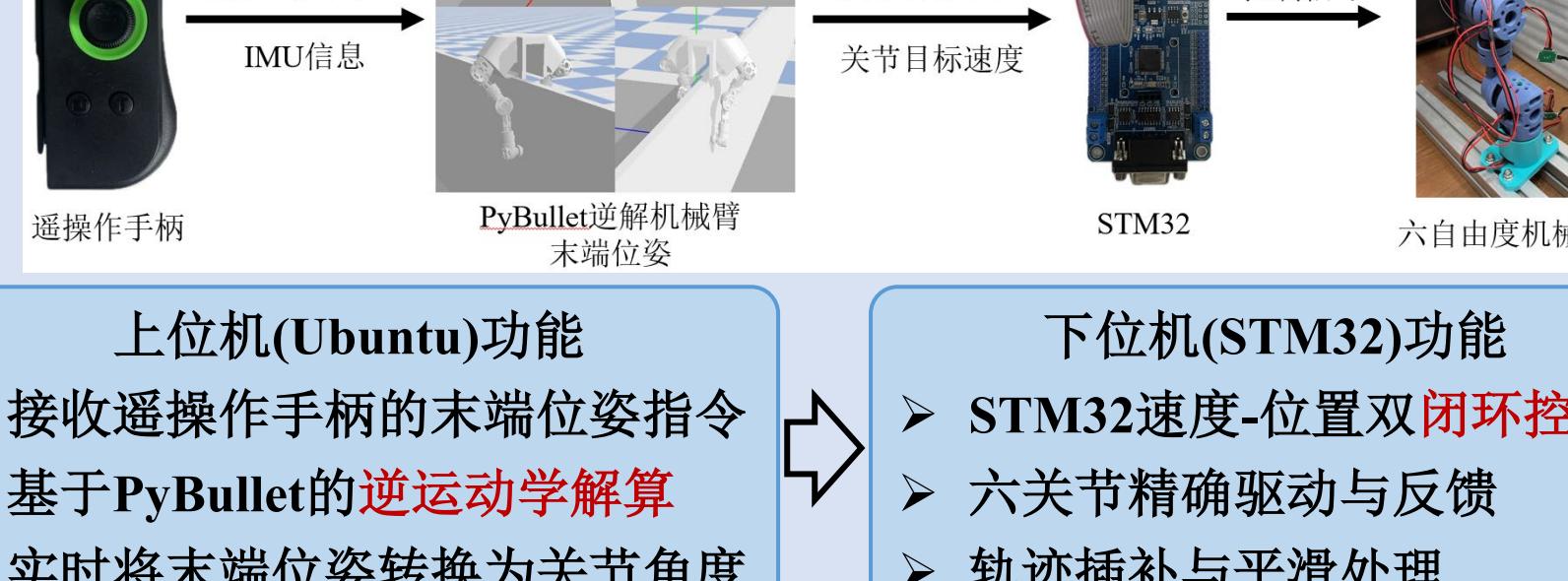
模仿奖励  $r_{imitation}$ ：确保生成的足端轨迹接近模仿学习的结果

$$r_{imitation} = \xi_1 \exp(-\xi_2 \sum_{t=1}^{12} \| q_i - q_{i,imitation} \|^2)$$

惩罚  $r_{penalty}$ ：惩罚不合理的动作（例如关节力矩超限、足端打滑等）

$$r_{penalty} = -\eta_1 \sum_{i=1}^{12} I(\tau_i > \tau_{max}) - \eta_2 \sum_{k \in \text{feet}} I(\mathbf{v}_k \cdot \mathbf{n}_k \neq 0)$$

#### 机器人遥操作控制实验



上位机(Ubuntu)功能

- 接收遥操作手柄的末端位姿指令
- 基于PyBullet的逆运动学解算
- 实时将末端位姿转换为关节角度

串口发送目标关节位置指令序列

下位机(STM32)功能

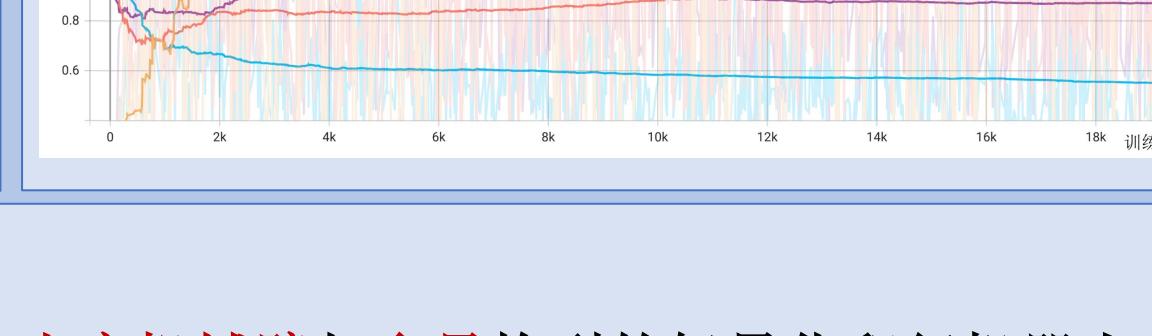
- STM32速度-位置双闭环控制
- 六关节精确驱动与反馈
- 轨迹插补与平滑处理

#### 机器人强化学习训练过程与结果

##### 四组训练实验中关键奖励项系数对比

	训练1	训练2	训练3	训练4
XY平面线速度奖励	2.0	1.0	5.0	3.0
足端接触奖励	5.0	8.0	5.0	7.0
足端无指令接触惩罚	-0.05	-0.1	-0.1	-0.2
足端绊倒惩罚	-0.001	-0.001	-0.01	-0.02
腿部关节姿态惩罚	-1.0	-1.0	-1.5	-1.5

##### 四次不同参数训练下的价值函数损失



训练4中的奖励函数项系数配置有利于机器人在帆板侧面这一典型环境下，通过自主探索与主学习获得较稳定的运动步态。

#### 研究总结

- 针对星表环境下多模态机器人的自主运动控制需求，设计了融合六自由度机械臂与多足构型的轻量化爬行机器人；
- 搭建基于NVIDIA Isaac Sim的仿真环境与实验室仿真测试平台，配合NVIDIA Isaac Lab支持高效并行策略训练与Sim2Real迁移；
- 提出了一种融合模仿学习与强化学习的控制框架，基于PPO与SAC算法，通过“教师-学生”策略与多重奖励融合机制，实现了稳定步态与高效运动控制。

#### 后续研究

- 基于已经验证的遥操作控制实验，未来将遥操作采集的示教数据作为模仿学习的教师策略数据样本引入整体算法框架，以让机器人获得期望的更为稳定的运动步态；
- 优化奖励函数设计，确保各奖励项在数值空间上的统一性与合理性，同时考虑引入宏观评论家网络，以增强对长时序和任务整体表现的评估能力，提升策略效果；
- 通过优化奖励函数并微调经模仿学习与强化学习训练的模型，实现Sim2Real迁移学习，提升策略在真实机器人中的表现，为未来空间机器人在轨作业提供有力支撑。

#### 活动组织

#### 动力工程系

#### 机械工程系

#### 智能制造工程系

#### 机械与动力工程学院本科教务办公室

#### 企业支持

上海良信电器股份有限公司

上海恩井汽车科技有限公司

上海森永工程设备股份有限公司

上海维科精密模塑股份有限公司

中国航发商用航空发动机有限责任公司