

基于深度学习的机器人运动控制研究

杜亮

长春电子科技学院

DOI:10.32629/acair.v3i4.18119

[摘要] 深度学习方法为自动设计机器人运动控制提供了新的思路,不仅有助于简化设计流程,而且可以提高机器人适应高度复杂地形的能力。因此,文章提出了基于深度学习的机器人运动控制框架,聚焦高度复杂地形下机器人多步态灵活运动,论述基于深度学习的机器人运动控制方法实践要点,并通过仿真环境构建,研究基于深度学习的机器人运动控制训练及效果,以供参考。

[关键词] 深度学习; 机器人; 运动控制

中图分类号: TP242 **文献标识码:** A

Research on Robot Motion Control Based on Deep Learning

Liang Du

Changchun College Of Electronic Technology

[Abstract] Deep learning methods offer a novel approach to the automated design of robot motion control, which not only helps streamline the design process but also enhances a robot's ability to adapt to highly complex terrains. Accordingly, this paper proposes a deep learning-based framework for robot motion control, focusing on multi-gait flexible locomotion over highly complex terrains. It discusses the practical key points of deep learning-based robot motion control methods and investigates the training and effectiveness of such control through the construction of a simulation environment, for reference purposes.

[Key words] deep learning; robot; motion control

前言

近几年,随着人口老龄化加剧、劳动力成本上升,社会对机器人的需求急剧增长,包括服务机器人、工业机器人等。但是,由于应用场景多样、应用环境复杂,机器人在不同地形中实现稳定运动仍然面临严峻挑战。现今时期,国内外关于机器人运动控制的研究仅停留在理想化环境、预设场景内,无法有效应对真实世界。因此,研究基于深度学习的机器人运动控制具有非常突出的现实意义。

1 基于深度学习的机器人运动控制框架设计

深度学习是通过神经网络模型学习数据复杂表示的过程,核心目标是从原始数据中自动提取复杂特征并用于分类、模式识别、回归等任务,高度契合机器人运动控制训练需求^[1]。本次研究选择由教师大型预训练模型、学生轻量级模型构成的教师-学生网络,指导机器人运动控制迁移训练。基于教师-学生网络的机器人运动控制框架见图1。

图1中,Estimator MLP(Multilayer Perceptron,多层感知器)是用于构建、训练多层感知器的神经网络高阶接口,允许快速部署学习模型;Encoder MLP属于前馈神经网络,用于对输入特征进行非线性变换,增强模型表达;Adaptation MLP是自适应

多层感知器,用于动态调整网络结构与学习率适应多数据分布;Adaptation前向网络是神经网络的一种结构,信息由输入层单向传递至输出层;Actor前向网络是用于生成动作的策略网络,输出给定状态下最优的机器人运动策略选择;PD控制器即比例-微分控制器,通过计算系统误差及变化率精准控制输入;Critic网络是通过深度神经网络估算状态-动作对应价值函数,调整动作选择概率,优化整体策略。由图1可知,基于教师-学生网络的机器人运动控制框架中,教师网络的输入是状态估计器估计的机器人线速度,教师网络指导学生网络学习复现相关动作表征,并分别输入到Actor前向网络、Critic网络,Critic网络输出关节位置,经PD控制器输入控制器,驱使机器人做出相应动作。在教师-学生网络中,根据平地与上下斜坡等地形特点,持续分区训练,迭代更新机器人运动策略,直到机器人运动控制策略收敛。

2 基于深度学习的机器人运动控制方法实践

2.1 总体实践思路

周期信号是机器人运动控制的关键信号之一,可引导机器人运动控制策略生成与显式步态启发轨迹控制^[2]。基于周期信号的显式步态启发轨迹控制的实现需要在深度学习框架内嵌入

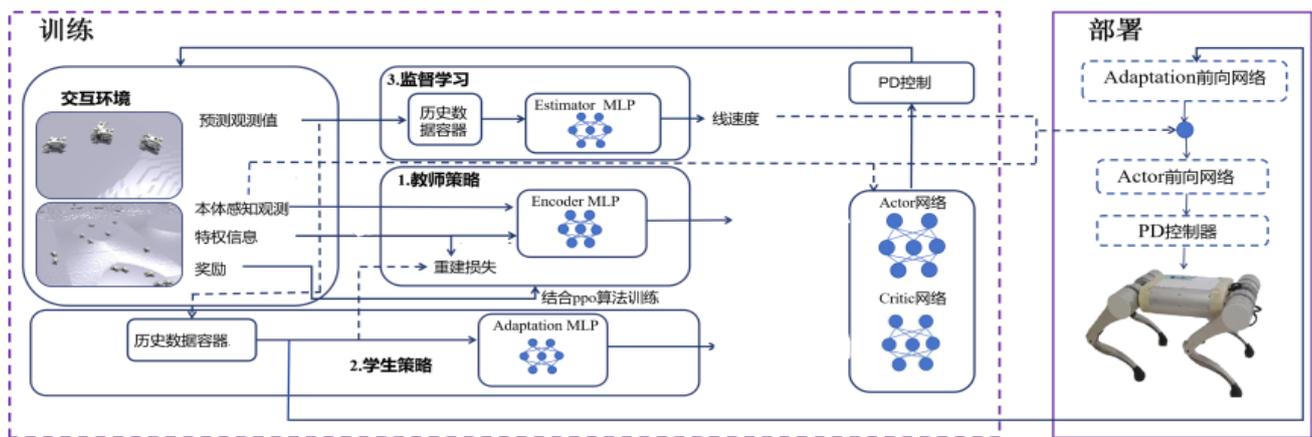


图1 基于教师—学生网络的机器人运动控制框架

1个可引导机器人步态、周期运动的生成器与策略网络,生成器包括上层步态规划器与下层相位发生器,上层步态规划器可直接接收策略网络输出变量,下层相位发生器负责提供机器人周期性运动信号,并引导机器人学习运动控制的热启动训练;策略网络定义为教师—学生网络迭代决策过程,整合概率转移函数、奖励函数,分析观测空间、动作空间、状态空间,寻找一个最大化期望累计奖励的运动策略,输出调节机器人运动轨迹参数与关节电机残差向量。

2.2 基于周期信号的显式步态启发轨迹控制

基于周期信号的显式步态启发轨迹控制需要从上层步态规划、下层相位发生两个方面进行分析。

(1) 上层步态规划器的输出是下层相位发生器的输入,可通过运动频率信号向量转化调节机器人运动中的步态参数,促使机器人按照预定步态动作,如小跑、行走等。在上层步态规划器中,机器人步态模式受运动通用频率、步态类型系数比例、步态占空系数比例等变量的影响^[3]。比如,将左前腿作为整个机器人运动步态参考腿,设定小跑、行走步态对应的期望左前腿频率、相位,可引导全部腿的运动步态。

(2) 下层相位发生器输出引导机器人周期性运动的信号,即将机器人运动模块接收频率信号转化为关节电机旋转角度信号。以四足机器人为例,因其运动模块有4个,需要逐一一定义对应的相位发生器模块,每一相位发生器接收1个运动频率并输出2个周期信号,包括某一时刻机器人抬腿频率向量元素、一号电机角度、二号电机角度等^[4]。在一个完整的运动周期内,机器人的一个运动模块需要经历腾空相、着地相动作,某个运动模块的电机位置周期信号向量与机器人单个运动模块逆运动学解析、本地坐标系内足部位置向量、本地坐标系某个运动模块足端距地面高度映射、最大允许抬腿高度参数等有关。比如,在机器人左前腿动作时,先接收1个腿频率信号,再通过周期信号区分腾空相、着地相,紧接着,获得机器人左前腿的足端距地面高度,最终经并联腿逆运动学解析获得左前腿一号电机、二号电机位置分量,引导左前腿按规定路程运动。

2.3 基于深度学习的机器人运动控制策略网络

基于深度学习的机器人运动控制策略网络由概率转移函数、奖励函数与初始状态分布、观测空间、动作空间、状态空间几个元组构成。

在深度学习框架下,观测空间由机器人坐标系内与重力向量旋转角度有关的侧倾角度的余弦函数、横滚角度的正弦函数、电机编码器自测电机旋转角度、显式步态启发轨迹生成器的行为信息向量、每一个运动模块频率向量、前进速度指令、步态相关调节信息等构成^[5]。而动作空间则由上层步态规划器输入步态周期调节信号、步态调节信号、电机旋转角度残差信号向量等构成。状态空间在融合机器人运动惯性测量单元与关节编码器测量值的基础上,增设源于状态估计器网络的机器人运动线速度,最终形成由12维关节位置误差、12维机体角速度、惯性测量单位坐标系下的重力单位向量、12维当前动作、3维关节速度、线速度、3维期望速度控制指令、12维历史时间步对应关键位置误差与速度误差等构成的状态空间。

在机器人运动观测空间与运动空间确定后,以机器人真实运动速度为依据,设置机器人奖励函数,奖励机器人精准执行前进速度指令、步态准确,持续缩小机器人实际运动速度与给定速度的偏差。

3 基于深度学习的机器人运动控制仿真训练及效果验证

3.1 仿真环境构建

为验证基于深度学习的机器人运动控制效果,利用Isaac Gym仿真软件,设置100个大小均为8m×8m的地形,相邻地形之间为平地,宽2m。仿真训练场景涵盖斜坡、台阶、平地等,同一列地形类型相同,地形难度随着行数的增加而增加。其中,斜坡地形包括上行斜坡、下行斜坡,坡度在0°~40°之间变化,每一个斜坡地形坡度变化前设置2m×2m平台;台阶地形为金字塔状,包括上行台阶、下行台阶,台阶高度在5cm~23cm之间,每一个台阶高度变化前设置3m×3m平台。

在仿真训练地形确定后,设定仿真步长为0.005,机器人运

动频率为50Hz, 每一个时刻机器人状态张量由3维本体角速度、3维速度指令、3维姿态向量、3维本体线速度、12维关节角速度、12维关节角、12维上一时刻网络输出等构成^[6]。基于深度学习的机器人运动控制策略网络处理方式批正则化, 激活函数为指数式线性单元, 奖励函数为线速度Z轴惩罚与基座高度惩罚, 训练方法为域随机化。

3.2 运动仿真训练

在Isaac Gym仿真环境下, 同时训练400个机器人, 将机器人平均分配到不同地形类型中, 初始地形难度随机分配。分配后, 每一个地形中心生成1个机器人, 以机器人在所处地形完成1个周期运动为标准判定机器人适应当前难度, 随后, 机器人被重新分配到更高级难度的地形。若机器人无法完成1个周期运动, 则表明机器人无法适应当前难度, 分配到更低一级难度地形接受训练。持续训练, 直到机器人适应全部难度地形。

在训练期间, 随机采样1个目标朝向, 根据目标朝向与仿真器内实际朝向之间的差值, 输出角速度遥控指令, 每间隔10s更新一次, 确保机器人按照控制器给出期望速度运动。

3.3 仿真结果分析

经过1500次迭代, 基于深度学习的机器人运动控制策略网络得到收敛, 在[0, 1]速度区间内, 不同地形上机器人运动控制成功率见表1。

表1 不同地形上机器人运动控制成功率

迭代次数	平地	斜坡	台阶
100	90%	89%	76%
500	93%	90%	86%
1500	96%	94%	93%

由表1可知, 经过1500次迭代训练, 基于深度学习的机器人运动控制成功率较高。其中, 机器人在地面上的运动控制成功率最高, 达到96%, 台阶地面上机器人运动控制成功率最低, 为93%, 均高于90%。表明基于深度学习的机器人运动控制效果较好, 推广应用价值较高。

4 总结

综上所述, 从机器人电机旋转角度指令划分着手, 聚焦周期信号策划机器人运动控制训练方案, 持续提升高度复杂环境下机器人运动控制效果。经仿真验证得出, 基于深度学习的机器人运动控制具有灵活的多步态运动能力, 可稳定通过复杂地形。下一步, 将继续深耕更多随机障碍地形条件下的机器人运动控制方案, 为机器人应用推广提供支持。

[参考文献]

- [1]黄欢, 邱涛, 甄庆凯, 等. 基于深度学习的运动分析数字孪生系统[J]. 中国体育科技, 2025, 61(03): 44-54.
- [2]郑泛舟. 基于深度强化学习的机器人运动状态估计研究[J]. 湖南文理学院学报(自然科学版), 2023, 35(01): 34-39.
- [3]刘春晖, 王思长. 基于深度学习的室内导航机器人避障规划算法[J]. 吉林大学学报(工学版), 2023, 53(12): 3558-3564.
- [4]宋士国, 陈二军. 基于深度学习的配电高压操作机器人运动控制技术[J]. 自动化与仪器仪表, 2024, 44(02): 209-212.
- [5]秦建军, 孟圆. 基于强化学习的四足机器人牵引运动控制研究[J]. 重庆理工大学学报(自然科学), 2022, 36(12): 80-90.
- [6]江雨霏, 朱其新. 深度强化学习下的管道气动软体机器人控制[J]. 西安工程大学学报, 2025, 39(02): 65-74.

作者简介:

杜亮(1991-), 男, 汉族, 吉林白城人, 博士, 助教, 长春电子科技学院, 主要研究方向: 机器人智能控制、人工智能。