

摘要

近年来,深度强化学习已在众多控制任务上超越人类水平。然而,对于机器人等移动设备,其主板能量资源受限,难以满足深度强化学习算法的计算需求。由于脉冲神经网络具有低能耗、高生物合理性与复杂时空信息处理能力等优势,研究人员将其应用于强化学习领域,提出了脉冲强化学习算法,为连续控制任务提供了一种极具前景的低能耗解决方案。然而,由于网络结构的差异,目前主流的强化学习算法不能直接迁移到脉冲神经网络中。**如何根据脉冲神经网络设计出兼顾高性能与低能耗的强化学习算法**,是本文的核心研究问题。本文从脉冲强化学习中的信号处理、网络结构和环境探索三个层面,创新性地解决了以下科学问题:**“脉冲神经网络的输出与强化学习算法不匹配”、“脉冲神经网络的表征能力不足”以及“脉冲神经网络的环境探索效率不足”**。其主要创新点如下:

- (1) **针对脉冲神经网络的输出与强化学习算法不匹配的问题,提出膜电压编码的深度脉冲 Q 网络算法,实现从高维感知输入到鲁棒策略的端到端脉冲强化学习。**现有工作不直接采用强化学习算法训练脉冲神经网络,或者引入与神经形态芯片不兼容的浮点矩阵运算。本文受昆虫内部的非发放中间神经元启发,提出膜电压编码的深度脉冲 Q 网络算法,将脉冲序列转化为对应的 Q 值,即动作价值函数的值,有效解决输出不匹配导致的训练与部署问题。实验结果表明,本方法在性能、稳定性与鲁棒性上都优于深度 Q 网络算法。
- (2) **针对脉冲神经网络的表征能力不足的问题,提出层内连接增强的全脉冲行动器网络方法,实现兼顾高性能与低能耗的强化学习算法。**现有工作重点关注脉冲编码方法与脉冲神经元,缺乏对网络内部复杂时空动态的深入挖掘,限制网络的表征能力。本文仿照生物体内的层内连接机制,提出层内连接增强的全脉冲行动器网络方法,将输出群体中的神经元在时域和空域上进行连接。实验结果表明,本方法有效提高网络的动作表征能力,在多个连续控制任务的性能上优于对应深度网络,并将能效提高 71 倍以上。
- (3) **针对脉冲神经网络的环境探索效率不足的问题,提出面向环境探索的噪声脉冲行动器网络方法,实现脉冲神经网络在多样化环境下的高效探索。**现有工作在不同的环境中采用相同的环境探索策略,难以满足多样化的环境对不同探索策略的需求。本文研究噪声驱动的环境探索特性,提出面向环境探索的噪声脉冲行动器网络方法,将时序相关的噪声引入脉冲神经元。实验结果表明,本方法

显著提高智能体在不同环境中的探索效率，在多个连续控制任务的性能上相比当前最先进的同类型算法提升 9.41%。

为了减轻脉冲强化学习算法实现的工作量，本文搭建了一个**开源脉冲强化学习函数库**，将上述脉冲强化学习算法进行了模块化集成，并通过演示实验证明了其高效性与灵活性，便于后续脉冲强化学习的研究。

综上所述，本文研究兼顾高性能与低能耗的脉冲强化学习算法，从输出表示方法、网络表征能力和环境探索方法出发，挖掘了脉冲强化学习算法的巨大潜力，为脉冲强化学习领域的后续研究奠定了基础。

关键词：强化学习，脉冲神经网络，类脑智能，神经形态计算