

# 基于神经网络的机器人学习与控制：回顾与展望

谢正泰<sup>1, 2</sup>, 樊佳亮<sup>1, 2</sup>, 刘梅<sup>1, 2</sup>, 金龙<sup>1, 2</sup>

1. 兰州大学信息科学与工程学院, 甘肃 兰州 730000;

2. 中国科学院重庆绿色智能技术研究院大数据与智能计算重庆市重点实验室, 重庆 400714

基金项目: 国家自然科学基金(62176109); 腾讯 Robotics X 犀牛鸟专项研究计划(2021-01); 甘肃省自然科学基金项目(21JR7RA531, 22JR5RA427, 22JR5RA487); 中央高校基本科研业务费(lzujbky-2021-65, lzujbky-2022-it02, lzujbky-2022-23); 兰州大学超算中心

通信作者: 金龙, jinlongsysu@foxmail.com 收稿/录用/修回: 2022-09-16/2022-11-09/2023-01-09

## 摘要

机器人因其高效的感知、决策和执行能力, 在人工智能、信息技术和智能制造等领域中具有巨大的应用价值。目前, 机器人学习与控制已成为机器人研究领域的重要前沿技术之一。各种基于神经网络的智能算法被设计, 从而为机器人系统提供同步学习与控制的规划框架。首先从神经动力学(ND)算法、前馈神经网络(FNNs)、递归神经网络(RNNs)和强化学习(RL)四个方面介绍了基于神经网络的机器人学习与控制的研究现状, 回顾了近30年来面向机器人学习与控制的智能算法和相关应用技术。最后展望了该领域存在的问题和发展趋势, 以期促进机器人学习与控制理论的推广及应用场景的拓展。

## Learning and Control of Robots Based on Neural Networks: Review and Outlook

XIE Zhengtai<sup>1, 2</sup>, FAN Jialiang<sup>1, 2</sup>, LIU Mei<sup>1, 2</sup>, JIN Long<sup>1, 2</sup>

1. School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China;

2. Chongqing Key Laboratory of Big Data and Intelligent Computing, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China

## Abstract

Robots have great application value in the fields of artificial intelligence, information technology, and intelligent manufacturing due to their efficient perception, decision-making, and execution capabilities. At present, robot learning and control have become one of the most critical frontier technologies in the field of robotics. Meanwhile, different intelligent algorithms based on neural networks have been designed to provide a planning framework for synchronous learning and control of robot systems. Specifically, the research status of neural-network-based robot learning and control is reviewed from four aspects: neural dynamics (ND) algorithms, feedforward neural networks (FNNs), recurrent neural networks (RNNs), and reinforcement learning (RL). The intelligent algorithms and related application technologies utilized for robot learning and control in the past three decades are reviewed in detail. Finally, the remaining challenges and development trends in this field are provided to promote the development of robot learning and control theory and the extension of application scenarios.

## 关键词

机器人学习与控制

神经动力学方法

前馈神经网络

递归神经网络

强化学习

中图法分类号: TP242

文献标识码: A

## Keywords

robot learning and control;

neural dynamics algorithm;

feedforward neural network;

recurrent neural network;

reinforcement learning

## 0 引言

机器人被誉为 20 世纪人类最伟大的发明之一。它具有感知能力、决策能力和执行能力，能够协助甚至替代人类完成各种复杂、繁重和危险的任务。任何涉及重复性工作的工业系统都可以通过机器人实现自动化作业，进而在降低人力成本的同时显著地提高生产效率和产品质量<sup>[1]</sup>。这也预示着现代智能工厂能够根据需求扩大或缩小生产规模。典型的机器人设备包括冗余度机器人<sup>[2]</sup>、柔性机器人<sup>[3]</sup>和移动机器人<sup>[4]</sup>等。近年来，机器人在军事战争、太空探索、医疗手术和其他重要行业中得到了广泛的应用<sup>[5-7]</sup>。

在第四次工业革命的推动下，机器人系统，特别是机械臂，受到世界各国的广泛关注<sup>[8]</sup>。机器人技术与多学科技术相结合，逐步推动了人类社会生产和生活方式的转变。以机器人视觉和触觉为代表的机器人感知技术加速了机器人学习与控制概念的产生和发展<sup>[9]</sup>。因此，机器人系统能够克服源于自身或环境的不确定性因素，从而提高任务执行的效率和稳定性。具体而言，机器人需要解决由磨损、老化、故障等因素引起的结构不确定性问题和由环境、任务等因素决定的决策问题<sup>[10]</sup>。因此，亟需开发其自我学习和实时校准功能。近年来，相关的学习技术提高了具有不确定性的机器人系统的容错能力和决策能力<sup>[10-12]</sup>。另一方面，在任务执行过程中，机器人期望保持稳定且高效的工作性能。因此，它需要基于特定的控制方案和高效的网络算法以实时获得合适的控制信号。这要求机器人系统具备优越的容错能力和可靠的控制精度，并在此基础上以人工智能的模式实现复杂多变的任务<sup>[13-14]</sup>。近年来，脑机接口技术<sup>[15]</sup>、手术机器人<sup>[16]</sup>、康复机器人<sup>[17-18]</sup>、多机器人集群<sup>[19]</sup>等新兴概念的发展对机器人控制技术与应用提出了实时性、精确性、稳定性、容错性、鲁棒性等多方面实际要求。随后出现了一系列先进的控制技术，如远中心控制<sup>[20]</sup>、视觉伺服控制<sup>[21]</sup>、模型预测控制<sup>[22]</sup>、阻抗控制<sup>[23]</sup>。因此，探索机器人学习与运动控制相结合的新模式是近年来国内外机器人研究的重点。

基于神经网络的机器人学习与控制技术已成为前沿科技之一，也是智能制造、智慧医疗和无人驾驶领域中的重要瓶颈技术之一。近 30 年来，相关的研究成果显著地提高了机器人系统的功能性、精确性和智能性<sup>[24]</sup>。机器人学习技术利用数据驱动

方法分析、预测和估计带有不确定性的模型或策略<sup>[25]</sup>；机器人控制技术结合学习信息与控制算法实现功能性运动规划<sup>[26]</sup>。在多源传感器的协助下，机器人系统可以收集、处理和开发数据，进而集成数据库，甚至搭建数据云<sup>[27]</sup>。随后，基于神经网络的机器人算法根据测量数据实现优化的学习与控制过程。

神经网络具备强大的拟合能力和并行处理能力，可以被理解为机器人的“大脑”。面向机器人学习与控制的主流神经网络方法包括神经动力学(neural dynamics, ND)方法<sup>[28-32]</sup>、前馈神经网络(feedforward neural network, FNN)方法<sup>[33-34]</sup>、递归神经网络(recurrent neural network, RNN)方法<sup>[35-36]</sup>和强化学习(reinforcement learning, RL)方法<sup>[11, 37]</sup>。ND 方法通过构建常微分方程(ordinary differential equation, ODE)形式或对应的离散变体系统以求解机器人学习与控制问题，并实现实时参数辨识<sup>[28-30]</sup>和机器人控制<sup>[31-32]</sup>。FNN 方法属于一种机器学习算法，能够通过更新权值逼近机器人系统的特定函数，从而开发相应的数据驱动技术<sup>[33-34]</sup>。与 FNN 方法不同，RNN 方法的网络节点之间可以构建循环连接，允许一些节点的输出影响其后续输入，在机器人学习领域有显著的应用价值<sup>[35-36]</sup>。值得注意的是，ND 方法和 RNN 方法在结构上存在交集。当 ODE 形式的计算网络的输入神经节点数大于或等于 2 时，该系统即属于 RNN 方法，也可被称为 ND 方法<sup>[38]</sup>。为了便于阐述，本文分别对 ND 方法与 RNN 方法进行单独介绍。RL 方法能够解决机器人与环境交互过程中的决策问题，并通过试错的方式实现机器人特定技能的学习<sup>[11, 37]</sup>。以上 4 种机器人学习算法均属于人工智能算法的范畴<sup>[39]</sup>。其主要任务是开发智能信息处理应用，从而设计与人类智能相似的机器人系统。在真实场景中，机器人的交互过程产生的学习与控制问题通常需要多种技术协同解决<sup>[40]</sup>。这要求机器人系统通过传感器获取大量的测量数据，进而驱动学习和控制过程。因此，机器人学习与控制系统的发展必然伴随着传感器技术、神经网络方法、数学理论等重要技术的进步<sup>[41]</sup>。

如图 1 所示，本文回顾了用于解决机器人学习与控制问题的神经网络方法以及相关应用，其中，机器人学习包括机器人模型学习和机器人策略学习两个方面。相关的神经网络方法被归纳为 ND 方法、FNN 方法、RNN 方法和 RL 方法。

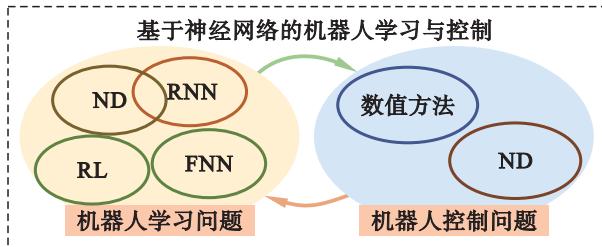


图1 机器人学习与控制的技术框架

Fig.1 Technical framework for robot learning and control

## 1 机器人模型与问题描述

本节介绍了机器人运动学模型、动力学模型和决策模型，并说明了机器人系统的学习与控制问题。

### 1.1 机器人模型

常用的机器人模型包括机器人运动学模型、动力学模型和决策模型。机器人运动学模型在机器人控制中扮演着至关重要的角色，是机器人轨迹规划、运动控制、视觉伺服等重要功能的实现基础<sup>[13-14]</sup>。机器人运动学包括机器人正运动学和机器人逆运动学。对于一个  $n$  自由度的机器人而言，其正运动学模型为  $\mathbf{T}(\boldsymbol{\theta}) = \mathbf{p}$ ，式中， $\boldsymbol{\theta} \in \mathbb{R}^n$  表示机器人的关节角度向量； $\mathbf{p} \in \mathbb{R}^m$  表示机器人的末端执行器位姿向量； $\mathbf{T}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  表示非线性映射函数。此外，速度层运动学模型为  $\mathbf{J}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}_k) \dot{\boldsymbol{\theta}}(t) = \dot{\mathbf{p}}$ ，式中， $\dot{\boldsymbol{\theta}}$  表示机器人的关节速度向量； $\dot{\mathbf{p}}(t)$  表示末端执行器的笛卡儿速度向量； $\mathbf{J}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}_k) = \partial \mathbf{T}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \in \mathbb{R}^{m \times n}$  表示机器人的雅可比矩阵； $\dot{\boldsymbol{\theta}}_k \in \mathbb{R}^h$  表示由运动学参数组成的常数向量。

与机器人运动学不同，机器人动力学模型不仅与关节速度、关节加速度等运动学变量有关，还涉及机器人的质量、惯性、外部受力等因素<sup>[42]</sup>。机器人动力学阐述了机器人驱动力/力矩与机器人运动状态之间的关系。对于一个  $n$  自由度的机器人而言，其关节空间的动力学模型为  $\mathbf{M}(\boldsymbol{\theta}) \ddot{\boldsymbol{\theta}} + \mathbf{C}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) \dot{\boldsymbol{\theta}} + \mathbf{g}(\boldsymbol{\theta}) = \boldsymbol{\tau} + \mathbf{J}^T \mathbf{f}_e$ ，式中， $\mathbf{M}(\boldsymbol{\theta}) \in \mathbb{R}^{n \times n}$  被定义为惯性矩阵； $\ddot{\boldsymbol{\theta}} \in \mathbb{R}^n$  表示机器人的关节加速度向量； $\mathbf{C}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) \in \mathbb{R}^{n \times n}$  表示科氏和离心矩阵； $\mathbf{g}(\boldsymbol{\theta}) \in \mathbb{R}^n$  表示重力向量； $\boldsymbol{\tau} \in \mathbb{R}^n$  表示机器人的关节扭矩向量； $\mathbf{f}_e \in \mathbb{R}^m$  表示施加在末端执行器的力。

机器人的决策模型可以用马尔可夫决策过程来表述，它由  $(S, A, P, R, \delta)$  五个部分组成<sup>[43]</sup>，其中， $S$  表示环境状态空间/集合， $s \in S$  表示具体状态； $A$  为机器人的动作空间/集合， $a \in A$  为具体动

作； $P(s'|s, a)$  表示机器人在执行动作  $a$  的状态下到达下一个状态  $s'$  的转移概率分布函数； $R$  表示奖励函数， $r(s'|s, a) \in R$  表示机器人在状态  $s$  下执行动作  $a$  以到达状态  $s'$  时获得的即时奖励； $\delta \in [0, 1]$  表示折扣因子，能够在计算累积奖励时削弱未来奖励的影响。在执行策略阶段，机器人在当前状态  $s_t$  ( $t$  表示当前时刻) 依据策略  $\pi$  执行动作  $a_t$  得到奖励  $r_{t+1}$ ，并根据状态概率分布函数  $P(s_{t+1}|s_t, a_t)$  到达新状态  $s_{t+1}$ ，直到机器人到达终止状态。

### 1.2 学习与控制问题

基于模型的运动学控制方法(如伪逆法<sup>[28]</sup>、模型预测控制<sup>[22]</sup>等)都是在运动学模型框架上设计的。在机器人系统中，运动学模型可以由机器人的 Denavit-Hartenberg 参数确定，这取决于机器人的出厂参数设置。然而，由于诸如磨损、老化、故障等因素，机器人结构参数(例如雅可比矩阵)可能发生变化<sup>[26]</sup>。因此，机器人运动学中不可避免地存在不确定因素。与此同时，机器人的硬件设计与开发过程也会增加模型的不确定性。基于模型的运动学控制方案难以精确控制此类具有不确定性的机器人，这可能显著地降低机器人系统的任务执行性能。对于机器人的动力学模型而言，多个动力学参数、外部负载和其他因素(如难以辨识的摩擦力)会增加其不确定性。此外，对于多自由度机器人手臂，各关节的运动会引起机器人的质量分布、惯性矩大小、力臂长度等参数的改变，进而显著影响其动力学结构<sup>[13]</sup>。其次，机器人系统存在时间延时特性，例如，因周期性采样导致的单位时延，以及因系统机电部分的惯性导致的时间滞后。再者，传感器测量误差、建模误差和噪声影响给机器人系统的动态控制带来了很大的困难。上述不确定性情况均被归纳为机器人模型学习问题。

机器人学习的另一个重要研究方向是机器人策略学习。不同于机器人模型学习，机器人策略学习的目标是求解马尔科夫决策过程的最优策略  $\pi^*(als)$  以指导机器人运动，并实现最大化折扣累积奖励  $U_t = \sum_{k=t}^{\infty} \delta^{k-t} r_k$ 。这种建模方式给机器人策略学习提供了一个具有通用性和可移植性的系统框架<sup>[44]</sup>。在此框架中，机器人策略学习问题涉及定义状态空间、环境模型、运动策略模型等。机器人强化学习算法在给定状态空间的前提下学习运动策略模型，而环境模型通常是未知的。强化学习会通过与环境交互进行试错学习，并以此为基础进行策略优化。机器人策略学习能够克服机器人与环境交互过程中

的不确定因素，从而提高系统功能性和智能性<sup>[45]</sup>。

探究基于学习技术的机器人控制器与方案求解器，继而建立具有通用求解性能的智能控制系统，是机器人技术开发中的另一个难题。首先，针对具备不确定性的机器人系统，需要对不确定性参数进行有效学习，并且在学习的同时将所学知识传递给控制系统。输出的控制信号在控制机器人的同时反馈于学习系统进行同步参数学习。如此学习与控制结合的迭代循环需要在给定的采样间隔内完成，以满足机器人系统的实时性要求。此外，随着系统的功能性需求提升，机器人控制算法需实现诸如稳定性优化<sup>[23]</sup>等功能。再者，智能机器人系统可能存在诸如雅可比矩阵奇异、所构建的控制模型非凸、噪声干扰等棘手问题。上述3个方面被归纳为机器人控制问题，严重制约了系统的高效控制。

面向机器人系统的学习与控制问题，下文将重点介绍以ND方法、FNN方法、RNN方法和RL方法为代表的神经网络技术以及相关的应用。

## 2 基于ND方法的机器人学习

ND方法是一种基于数据驱动技术且无需训练的实时求解算法。它通过开发具有全局收敛特征的

ODE系统或对应的离散变体系统以不断减小系统误差。当ND方法的输入变量大于或等于2时，ND方法属于RNN方法的特殊形式，能够解决机器人学习与控制<sup>[42]</sup>、矩阵求逆<sup>[46]</sup>等问题。单变量ND算法的应用包括时变非线性方程求解<sup>[47]</sup>。

以雅可比矩阵学习为例，ND算法的学习目标为实现 $\hat{\mathbf{J}}\dot{\boldsymbol{\theta}} = \hat{\mathbf{p}} \rightarrow \dot{\mathbf{p}}$ 或 $\hat{\mathbf{J}} \rightarrow \mathbf{J}$ ，其中 $\hat{\mathbf{J}}$ 表示估计的雅可比矩阵； $\hat{\mathbf{p}}$ 表示末端执行器的估计速度。这个参数学习问题可从优化的角度转化为最小化问题，并设计误差函数 $\epsilon = \hat{\mathbf{p}} - \dot{\mathbf{p}} = \hat{\mathbf{J}}\dot{\boldsymbol{\theta}} - \dot{\mathbf{p}}$ 。ND算法通过构造收敛系统以求解该最小化方案。其它基于ND算法的机器人学习技术也类似以上过程。为了便于理解，图2提供了基于ND方法的机器人系统结构图。如图所示，传感器系统将机器人状态数据传递给由神经动力学算法和机器人控制系统组成的计算网络。测量数据难免会受到噪声和误差等因素影响，网络的输出经过数据滤波系统后，作为控制信号输入用于执行任务的机器人系统。此后，机器人状态数据被传感器系统再次测量，使得整个系统进入一次新的循环。这样持续的循环过程实现了机器人系统的学习与控制。

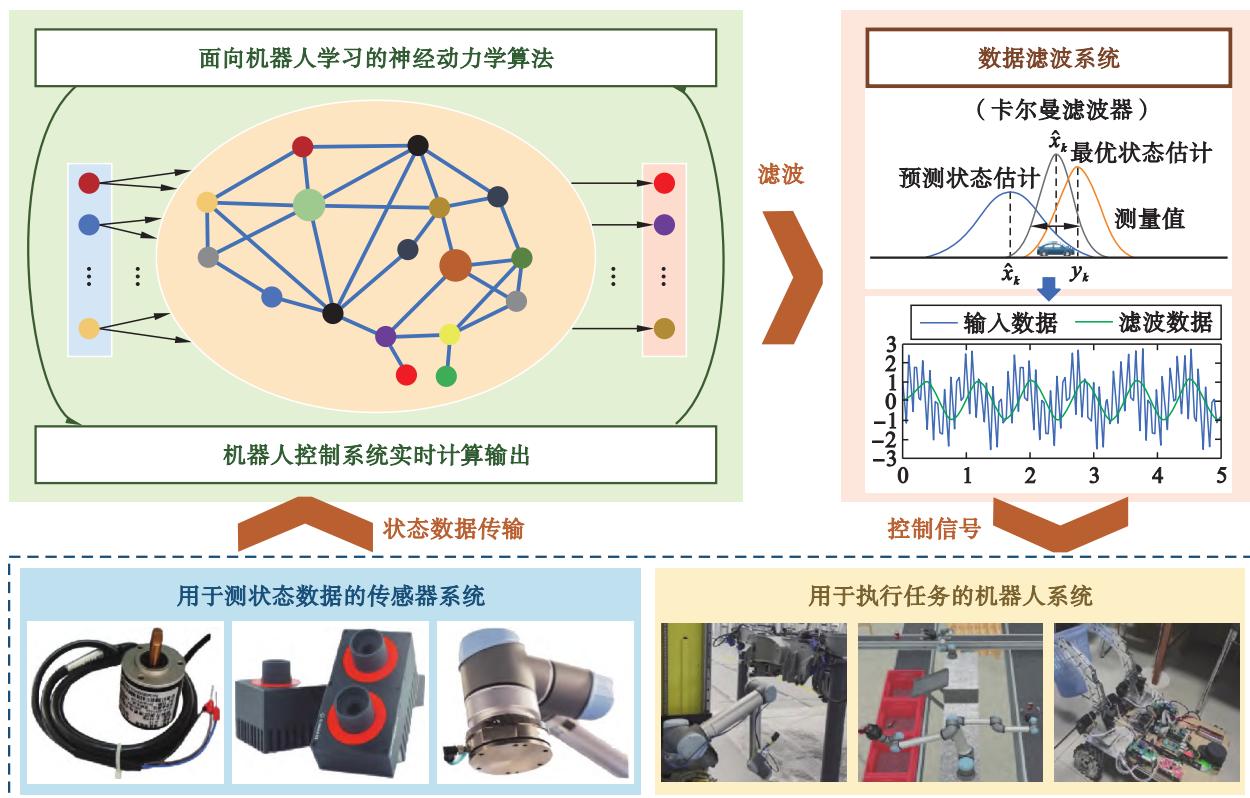


图2 基于ND方法的机器人系统结构图

Fig.2 Structure diagram of ND-based robot systems

## 2.1 Broyden 算法

Broyden 算法是一种有效求解非线性系统的拟牛顿法<sup>[48]</sup>, 已广泛应用于机器人学习中。Broyden 算法仅在第 1 次迭代时计算雅可比矩阵, 在后续迭代中进行参数更新。基于 Broyden 算法, 文[49–50]设计了面向机器人的动态和静态离散时间学习公式:

$$\hat{\mathbf{J}}_k = \hat{\mathbf{J}}_{k-1} + \frac{(\Delta\mathbf{p} - \hat{\mathbf{J}}_{k-1}\mathbf{h}_\theta - \dot{\mathbf{p}}_k(t_k - t_{k-1}))\mathbf{h}_\theta^\top}{\mathbf{h}_\theta^\top \mathbf{h}_\theta} \quad (1)$$

$$\hat{\mathbf{J}}_k = \hat{\mathbf{J}}_{k-1} + \frac{(\Delta\mathbf{p} - \hat{\mathbf{J}}_{k-1}\mathbf{h}_\theta)\mathbf{h}_\theta^\top}{\mathbf{h}_\theta^\top \mathbf{h}_\theta} \quad (2)$$

式中,  $\Delta\mathbf{p} = \mathbf{p}_k - \mathbf{p}_{k-1}$ ,  $\mathbf{h}_\theta = \boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1}$ , 下标  $k$  和  $k-1$  表示第  $k$  和  $k-1$  时刻。基于 Broyden 方法的静态离散时间学习公式(2)相对于动态式(1)去除了速度项  $\dot{\mathbf{p}}_k(t_k - t_{k-1})\mathbf{h}_\theta^\top$ 。在跟踪静态目标或极小采样间隔的情况下, 式(1)和式(2)的学习性能是类似的。然而, 对于动态目标, 式(1)能够获得更小的学习误差并取得更好的收敛性能。此外, 面向机器人模型不确定的情况, 该方法需要传感器获得真实末端执行器的笛卡儿速度  $\dot{\mathbf{p}}_k$  来促进雅可比矩阵的学习。基于 Broyden 方法的机器人学习公式(1)和学习公式(2)多次被应用于处理未标定机器人视觉伺服问题。文[51]基于 Broyden 方法和测量数据在线估计了视觉雅可比矩阵, 进而实现了无模型视觉伺服的任务模式。文[52]针对机器人与未知弹性物体相互作用的问题, 提出了一种基于视觉的变形控制器, 利用视觉反馈实时估计变形后的雅可比矩阵, 避免了繁琐的传统模型辨识步骤。值得注意的是, 上述学习公式需要与控制算法进行结合以实现机器人的同步学习与控制。虽然 Broyden 方法在机器人参数估计中是有效的, 但是其收敛速度主要取决于机器人的状态且无法调整。

## 2.2 ZND 算法

近年来, 零化神经动力学(zeroing neural dynamic, ZND)算法被设计以收敛误差函数  $\epsilon$ , 进而解决机器人系统的无模型控制问题<sup>[53]</sup>。该算法最初开发神

经动力学模型  $\dot{\epsilon} = -\mu\epsilon$  (其中  $\mu > 0$  为收敛参数) 构建 ODE 形式的收敛系统。例如, 文[28]基于 ZND 算法和伪逆法构造了机器人雅可比矩阵的学习算法:

$$\dot{\hat{\mathbf{J}}} = (\ddot{\mathbf{p}} - \hat{\mathbf{J}}\dot{\boldsymbol{\theta}} + \mu(\dot{\mathbf{p}} - \hat{\mathbf{J}}\dot{\boldsymbol{\theta}}))\dot{\boldsymbol{\theta}}^\dagger \quad (3)$$

式中,  $\dot{\hat{\mathbf{J}}}$  为  $\hat{\mathbf{J}}$  的时间导数;  $\ddot{\mathbf{p}}$  表示末端执行器的笛卡儿加速度; 上标  $\dagger$  表示向量或矩阵的伪逆且  $\dot{\boldsymbol{\theta}}^\dagger = \dot{\boldsymbol{\theta}}^\top / (\dot{\boldsymbol{\theta}}^\top \dot{\boldsymbol{\theta}})$ 。当关节速度等于零向量时,  $\dot{\boldsymbol{\theta}}^\dagger$  会陷入奇异状态。为了避免该问题, 奇异状态下的关节速度的伪逆形式被其转置形式所取代。此外, 式(3)要求从外部传感器获取末端执行器的笛卡儿速度和加速度的实时信息。文[28]中的理论分析表明, 由式(3)生成的误差函数  $\epsilon$  满足  $\epsilon(t) = \epsilon(0) \exp(-\mu t)$ 。换言之, 误差函数  $\epsilon$  能通过对  $\hat{\mathbf{J}}$  的实时更新全局收敛于零向量。后续的拓展研究基于式(3)开发了意图识别应用<sup>[54]</sup>和离散学习公式研究<sup>[53]</sup>, 进一步验证了该学习算法的可行性。此后, 文[55]提出了一个附带积分项和激活函数的机器人学习公式:

$$\dot{\epsilon} = -\gamma(t)s(\epsilon) - \int_0^t \beta(\tau)l(\epsilon(\tau))d\tau \quad (4)$$

式中,  $\gamma(t) > 0$  和  $\beta(t) > 0$  表示时变收敛参数;  $s(\cdot)$  和  $l(\cdot)$  表示激活函数。文[55]借助额外的视觉传感器实时获取末端执行器状态数据, 并给出了详细的机器人实验证。此外, 文[55]的理论证明显示, 式(4)的误差  $\epsilon$  能在噪声的干扰下实现有限时间收敛。

然而, 上述 ZND 研究预先假设了未知雅可比矩阵满足某些结构特性。这些先验信息在实际应用中通常很难获得。其次, ZND 方法估计的雅可比矩阵  $\hat{\mathbf{J}}$  与实际雅可比矩阵  $\mathbf{J}$  是有区别的。其原理在于引导末端执行器的估计速度接近实际速度。换言之, ZND 方法实现了误差函数  $\epsilon = \hat{\mathbf{p}} - \mathbf{p}$  趋近于零向量, 而不是雅可比矩阵估计误差  $\hat{\mathbf{J}} - \mathbf{J}$  趋近于零矩阵。此外, 不同 ND 方法的比较结果被总结在表 1 中。

表 1 不同 ND 方法的特点比较  
Tab.1 Comparisons of different ND methods

方法	优势	劣势	主要应用
Broyden 方法	不涉及矩阵求逆或求伪逆运算	收敛速度取决于机器人的状态且难以调整	视觉伺服控制
ZND 方法	系统收敛速度可调整, 拓展研究多	计算关节速度的伪逆, 需要关节加速度的信息	轨迹跟踪控制

## 2.3 基于噪声驱动的方法

根据近几年的研究发现, 在机器人学习算法中注入必要的随机噪声  $\mathbf{n}$  能够促进估计的雅可比矩阵  $\hat{\mathbf{J}}$

的学习过程。随机噪声  $\mathbf{n}$  可能源于真实场景中的计算误差、测量误差、环境噪声等方面。LI 等基于梯度下降法提出了噪声驱动的学习算法<sup>[56]</sup>:

$$\alpha \dot{\hat{J}} = -(\hat{J}(\dot{\theta} + n) - \dot{p})(\dot{\theta} + n)^T \quad (5)$$

式中,  $\alpha > 0$  表示收敛参数。附加噪声  $n$  能够丰富学习算法的输入多样性, 可以被视为一种数据增强形式。对此, 文[56]在理论上证明了雅可比矩阵学习误差  $\hat{J} - J$  能够全局收敛于零矩阵。相比之下, 不添加噪声而仅基于梯度下降法的学习算法无法学习准确的雅可比矩阵<sup>[32]</sup>。由此可见, 基于噪声驱动的方法能够有效地提高 ND 算法的学习能力。在后续的研究中, 这种噪声驱动的学习模式被广泛扩展到不同的机器人应用中, 包括运动学参数估计<sup>[57]</sup>、循环运动控制<sup>[58]</sup>和运动与力混合控制<sup>[59]</sup>等。

基于噪声驱动的方法能够与 ND 算法结合, 进而增强机器人系统的学习能力。然而, 附加噪声  $n$  会直接影响雅可比矩阵估计结果和控制信号, 不利于机器人系统的稳定控制。

#### 2.4 其他相关技术

为了减轻噪声对机器人系统的负面影响, 各种抗噪方法和滤波技术被研究并集成于机器人的学习与控制系统中<sup>[60]</sup>。例如, 卡尔曼滤波算法具有预测和估计能力, 能够处理机器人的测量噪声和过程噪声, 从而提高系统的精确性和稳定性<sup>[61-62]</sup>。在文[63]中, 一个自适应卡尔曼滤波器被设计以实现雅可比矩阵的估计过程。该方法提出了一个递归的线性系统状态估计方程, 并通过测量关节角速度和末端执行器的数据对雅可比矩阵进行最优估计。其优势在于只需当前时刻的测量值和前一个采样周期的估计值即可进行雅可比矩阵估计, 不需要大量的存储空间, 非常适合计算机处理。因此, 卡尔曼滤波算法与 ND 算法结合能够有效提高机器人学习与控制的稳定性和鲁棒性。

滑模控制是一种广为研究的非线性控制方法<sup>[64-65]</sup>。其非线性表现为控制的不连续性, 可以在机器人运动过程中根据系统的当前状态不断变化, 迫使收敛系统按照预定的状态轨迹运动。滑模控制的优点是可以克服机器人系统的不确定性, 对机器人学习问题具有鲁棒性。文[42]设计了多个自适应雅可比矩阵估计器来实现机器人雅可比矩阵学习和动力学与运动学参数估计。其核心原理在于构建自适应滑动矢量  $s = \hat{J}\dot{\theta} - \dot{p}$ , 并根据线性化公式和参数控制原理, 设计多个全局收敛于滑动矢量的控制系统。在该系统下, 机器人可以通过更新估计参数来跟踪期望轨迹。滑模控制的原理和 ND 算法的原理是类似的, 因此滑模控制及其相关理论可

以为 ND 方法提供可借鉴的改进方向。例如, 文[66]借鉴滑模控制的设计公式改进了 ZND 算法, 从而引入了有限时间收敛机制。

#### 2.5 传感器系统

传感器系统是基于 ND 算法的机器人学习中不可缺少的一部分, 能够提供机器人系统必要的数据以促进学习过程。传感器系统的精确程度直接影响基于 ND 算法的机器人学习性能。基于 ND 算法的机器人学习的文[26, 55-66]均介绍了数据来源和相关传感器设备。作为最常见的传感器系统, 末端深度相机能够获取物体、末端执行器等位置信息, 进而实现机器人定位、姿态识别、运动检测等功能<sup>[67]</sup>。一套精密而复杂的相机能够搭建动作捕捉系统<sup>[68]</sup>。在该系统中, 多个高速相机从不同角度对目标特征点进行跟踪, 进而显著提高定位能力。此外, 力(扭矩)传感器可以对机器人的受力情况进行监控<sup>[69]</sup>, 并根据数据分析, 对机器人行为作出指导。编码器负责测量机器人的各个关节状态, 并及时反馈机器人的运动状态<sup>[70]</sup>; 距离传感器能够基于光波或声波进行测距, 有助于移动智能机器人的避障功能开发<sup>[71]</sup>; 声音传感器用于感受和捕捉在气体、液体或固体中的声波, 进而实现语言识别<sup>[72]</sup>。速度或加速度层的数据信息可以通过差分算法、滤波器等手段从位置层数据近似获得或通过特定的传感器进行测量<sup>[28, 55]</sup>。例如, 物体的加速度可以通过惯性测量单元进行获取<sup>[73]</sup>。

本节系统性地总结了 ND 算法及其相关技术。ND 算法需要基于数据驱动技术实现参数学习, 进而实现机器人系统的无模型控制。噪声驱动方法、卡尔曼滤波算法、滑模控制、传感器系统可以与 ND 方法结合进一步提升机器人系统的性能。与机器学习算法相比, 其优势在于所需数据量小且不需要预训练。然而, 实际机器人系统的实时性特征也对数据处理、噪声抑制、算法设计等多方面提出了更高的要求。

### 3 基于 FNN 的机器人学习

FNN 方法是一种函数拟合方法。它能够通过简单非线性函数的多次复合, 获得输入空间到输出空间的复杂映射关系。针对机器人系统的学习与控制问题, FNN 方法被设计以实现机器人参数学习, 进而弥补系统的不确定性。例如, 反向传播(back propagation, BP)神经网络和径向基函数(radial basis function, RBF)神经网络在一定条件下能够拟合

任意连续曲线<sup>[74]</sup>。它们能通过测量的机器人状态数据集, 训练并更新权重参数, 实现对机器人系统的参数估计。机器人逆运动学模型、正运动学模型、动力学模型和相关物理参数都可以成为 FNN 方法的拟合对象。如图 3 所示, FNN 的结构包括输入层、隐藏层和输出层。其输出可以被表述为

$$y_i = \sum_{j=1}^{N_h} \left( w_{ij} F \left( \sum_{k=1}^{N_i} v_{jk} x_k + \zeta_j \right) + \omega_i \right), \quad i = 1, 2, \dots, N_o \quad (6)$$

式中,  $x_k$  和  $y_i$  分别表示神经网络的输入和输出;  $N_i$ 、 $N_h$  和  $N_o$  分别表示输入层、隐藏层和输出层的神经元数量;  $v_{jk}$  和  $w_{ij}$  分别表示隐藏层和输出层的权重;  $\zeta_j$  和  $\omega_i$  分别表示隐藏层和输出层的偏置项;  $F(\cdot)$  表示隐藏层的激活函数, 可以选择为 sigmoid 激活函数  $F(\chi) = 1/(1 + \exp(-\chi))$ 。进而, 式(6)

可以被简化为  $\mathbf{y} = \bar{\mathbf{F}}(\bar{\mathbf{x}})$ , 式中,  $\mathbf{A} \in \mathbb{R}^{N_o \times (N_h+1)}$  和  $\mathbf{P} \in \mathbb{R}^{N_h \times (N_i+1)}$  表示两个增广权重矩阵;  $\bar{\mathbf{x}} = [1, x_1, x_2, \dots, x_{N_i}]^T \in \mathbb{R}^{N_i+1}$ ;  $\mathbf{y} = [y_1, y_2, \dots, y_{N_o}]^T \in \mathbb{R}^{N_o}$ ;

$$\bar{\mathbf{F}}(\bar{\mathbf{x}}) = [1, \mathbf{F}(\mathbf{P}^r \bar{\mathbf{x}}), \dots, \mathbf{F}(\mathbf{P}_{N_h}^r \bar{\mathbf{x}})]^T \in \mathbb{R}^{N_h+1}, \mathbf{P}_i^r$$

是矩阵  $\mathbf{P}$  的第  $i$  行。此外,  $\bar{\mathbf{F}}(\cdot)$  可以看成从  $\mathbb{R}^{N_h}$  到  $\mathbb{R}^{N_h+1}$  的映射。 $\zeta_j$  和  $\omega_i$  分别构成矩阵  $\mathbf{P}$  和  $\mathbf{A}$  的第 1 列。基于上述结构, 文[75]开发了一种基于 FNN 的学习方法以估计机器人动力学和执行器模型中的动力学参数并补偿系统的不确定性, 实现了良好的无模型控制性能。文[76]在柔性机器人上测试了基于 FNN 方法和基于雅可比矩阵方法的两种逆运动学控制问题。其结果显示 FNN 方法不但能解决系统中的学习问题, 而且具备响应速度快、控制精度高的特点。文[77]通过构建 FNN 结构以估计 Baxter 机器人的动力学参数, 进而实现了同步避障和抓取任务。随着神经网络技术的发展, BP 神经网络<sup>[78]</sup>、RBF 神经网络<sup>[79]</sup>、极限学习机<sup>[33]</sup>等多种 FNN 方法显著提升了机器人系统的学习能力。

### 3.1 BP 神经网络

BP 神经网络能够学习并存储丰富的输入和输出映射关系, 而不需要预先获得映射对应的数学方程。其学习原理采用最速下降法, 通过反向传播不断调整网络的权值和偏差, 实现网络误差最小化。其主要特征在于神经元数据是正向传播的, 而误差是反向传播的。文[80]基于 BP 神经网络提出了一种简单且灵活的并联机器人标定方法, 能够拟合校准参数与测量参数之间的关系, 并提高了 6 自由度并联机器人的定位精度。文[81]提出了一种改进

的 BP 神经网络以学习机器人的正运动学公式, 进而解决逆运动学控制问题。

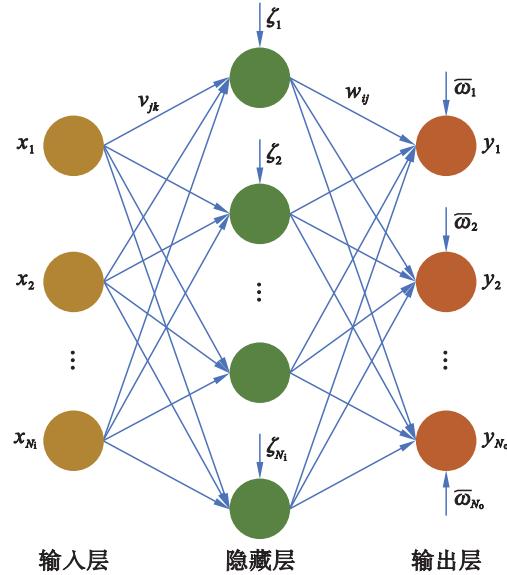


图 3 FNN 的网络结构图

Fig.3 Structure diagram of FNN

### 3.2 RBF 神经网络

与 BP 神经网络相比, RBF 神经网络在泛化能力、逼近精度和收敛速度方面具有优势。因此, RBF 神经网络在机器人学习中能取得更高效的性能<sup>[74]</sup>。具体而言, RBF 神经网络中的径向基函数可以表示为

$$W_i(\mathbf{x}) = \exp \left( -\frac{1}{2\sigma_i^2} \|\mathbf{x} - \mathbf{c}_i\| \right), \quad i = 1, 2, \dots, N_h \quad (7)$$

式中,  $\mathbf{x}$  为输入样本,  $\mathbf{c}_i$  是第  $i$  中心点,  $\sigma_i$  表示基函数第  $i$  个标准差。式(7)构成了 FNN 的隐藏层空间。因此, 其输入向量可以直接映射到隐藏层空间且无需权值连接。RBF 神经网络的输出表示为

$$y_j = \sum_{i=1}^{N_h} w_{ij} \exp \left( -\frac{1}{2\sigma_i^2} \|\mathbf{x} - \mathbf{c}_i\| \right), \quad j = 1, 2, \dots, N_o \quad (8)$$

通过上述结构, 从隐藏层空间到输出层空间的映射关系变为线性关系。换言之, RBF 神经网络的输出是隐藏层元素的线性加权和。基于此, 低维空间线性不可分的情况可以通过 RBF 转换成高维空间中线性可分的情况。该转换过程能够显著提升学习过程的收敛速度, 同时避免局部极值问题。RBF 神经网络在运动学和动力学系统的参数估计和系统辨识中有着重要的作用。在文[82]中, RBF 神经网络被开发用于估计机器人的逆运动学关系, 进而提高了机器人在无模型条件下的容错控制能力。该

研究验证了 RBF 神经网络在实验中的性能优于 BP 神经网络。文[83]提出了一种分布式神经网络用于协同控制多机器人系统，并利用 RBF 神经网络估计逼近误差以补偿在通信带宽受限和时延存在下的系统不确定性。

### 3.3 其他 FNN 方法

其他 FNN 方法和相关的拓展研究给机器人学习与控制提供了有效的可行方案。极限学习机采用最小二乘法更新网络权值，在保证学习精度的前提下，其计算速度比传统学习算法更快。文[84]提出的基于极限学习机的方法能够利用数据进行实时学

习以补偿机器人的动力学不确定性，进而实现末端执行器的触觉识别。此外，文[85]基于宽度学习思想，研究了无深度结构的增量学习系统。该方法能够通过增加神经节点以激发神经网络的收敛性与适应性，进而有效应用于微型机器人的伺服控制<sup>[86]</sup>。

FNN 方法已在机器人学习方面取得了一定的成功。作为总结，不同的 FNN 方法的比较见表 2。为了获得满意的训练结果，一些结构复杂的 FNN 方法对数据和设备的要求很高。此外，FNN 方法的学习过程可被理解为缺乏可解释性的黑箱，在一定程度上影响了学习结果的可信度。

表 2 不同的 FNN 方法的特点比较  
Tab.2 Comparison of different FNN methods

方法	优势	劣势
BP 神经网络	结构简单、自适应学习能力	训练速度慢、容易陷入局部极值、数据需求大
RBF 神经网络	全局逼近能力、自适应学习	数据需求大
极限学习机	训练速度快、自适应学习能力	数据需求大、难以适应高维大数据
宽度学习	训练速度快、增量学习、自适应学习能力	数据需求大、难以适应高维大数据

## 4 基于 RNN 的机器人学习

RNN 是一类具有短期记忆能力，能够用于处理和预测序列数据的神经网络。其神经元可以接受其他神经元的信息，也可以接受自身信息，进而形成具有环路的网络结构<sup>[35]</sup>。如图 4 所示，RNN 的隐藏层  $h_t$  不仅取决于当前时刻的输入  $x_t$ ，还取决于上一时刻的隐藏层状态  $h_{t-1}$ 。这使得 RNN 方法可以有效地处理手写识别、语音识别等任务<sup>[35]</sup>。RNN 方法包括简单 RNN、长短时间记忆网络、门循环单元、连续时间 RNN 和回声状态网络等<sup>[35]</sup>。此外，RNN 的方法在机器人控制方面有很多应用，例如文[87–89]，但很少涉及机器人学习。近年来，回声状态网络被应用于机器人学习中，并取得了良好的学习与控制效果。因此，在机器人学习方面，下文重点介绍回声状态网络。

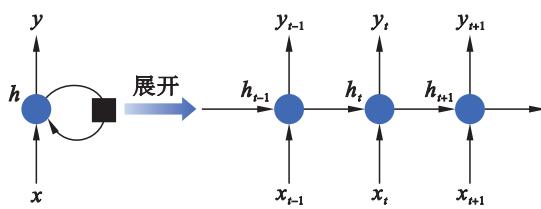


图 4 RNN 的网络结构时间展开图

Fig.4 Time expansion diagram of the network structure of RNN

回声状态网络是一种特殊的 RNN，具有一个稀

疏连接的随机隐含层。该网络克服了基于梯度下降的学习算法所难以避免的收敛速度慢和易陷入局部极值等问题<sup>[90]</sup>。在机器人学习与控制方面，文[91]利用回声状态网络建立了受小脑启发的预测模型，能够准确预测机器人在时序控制信号作用下的运动结果，并以此对原始控制信号进行准确调制。文[92]开发了基于回声状态网络的小脑启发的预测模型。该方法具体将任务空间误差转化为关节空间误差，并通过误差序列训练小脑预测模型，从而获得关节角度增量以实现机器人的高效控制。文[93]利用回声状态网络对机器人视觉系统的不确定性进行学习，进而实现机器人的视觉伺服控制。文[94]将回声状态网络应用于柔性机器人的物体跟踪任务中。该方法对时序因素的可控性强。

RNN 方法能够在计算过程中体现过程动态特性，因此比 FNN 方法具有更强的动态计算能力。目前，结合 RNN 和机器人学习与控制技术的研究仍然不成熟，有待进一步开发。

## 5 基于神经网络的强化学习

基于神经网络的 RL 方法在机器人策略学习中发挥着至关重要的作用，能够解决机器人码垛、机械臂抓取、移动导航、自动驾驶等重要问题<sup>[44]</sup>。图 5 展示了机器人强化学习的原理图。在机器人执行动作  $a_t$  后，环境转换到新的状态  $s_{t+1}$ ，并提交奖励信

号  $r_{t+1}$  作为反馈。随后, 机器人根据反馈执行新的动作  $a_{t+1}$ 。RL 方法利用这种试错机制与环境进行交互, 从而不断改善机器人策略。神经网络在强化学习系统中的作用在于拟合机器人交互过程中的动作分布<sup>[44]</sup>。根据不同的原理, RL 方法可以被划分为基于价值的 RL、基于策略的 RL 和基于 Actor-Critic 框架的 RL。

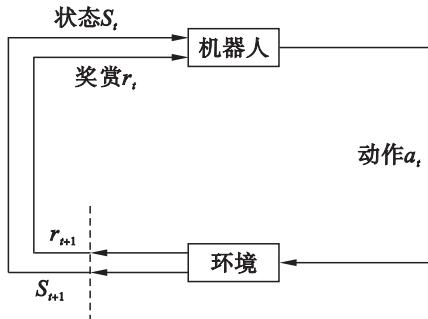


图 5 RL 方法的原理图

Fig.5 Schematic of RL methods

### 5.1 基于价值的强化学习

为了有效地评估策略  $\pi$ , 动作价值函数被设计为  $Q_\pi(s_t, a_t) = E(U_t | s_t, a_t)$ , 式中  $E(\cdot)$  表示求均值操作。基于价值的 RL 的目标为获取最优动作值函数  $Q^*(s_t, a_t) = \max_a Q_\pi(s_t, a_t)$ 。而后, 通过最大化最优动作值函数得到最优的动作策略  $a^* = \operatorname{argmax}_a Q^*(s_t, a_t)$  以指导机器人系统进行决策。时序差分方法是常用的基于价值的 RL 方法, 包括 Sarsa 方法和 Q-learning 方法<sup>[95]</sup>。该方法能考虑少量状态和动作以更新动作值函数。多步时序差分方法是时序差分方法的改进算法, 能同时处理多个奖励, 从而获得更好的性能<sup>[96]</sup>。当涉及大量或连续的状态或动作时, 深度 Q 网络(deep Q-network, DQN)被设计为  $Q(s_t, a_t; g)$  以估计最优动作值函数  $Q^*(s_t, a_t)$ , 其中  $g$  表示深度神经网络的参数<sup>[97]</sup>。基于 DQN, 机器人被设定尝试各种与环境进行交互的动作, 从而不断获得奖励。DQN 的输入是状态  $s_t$  和动作  $a_t$ , 而 DQN 的输出是评估动作的值。此外, 该网络利用奖励  $r_t$  训练参数  $g$ , 同时选择 DQN 的最大值所对应的动作为下一时刻的动作。经过足够的训练, DQN 能够学习最优动作值函数  $Q^*(s_t, a_t)$ , 间接获得最优策略。近年来, 许多改进的框架和算法被开发出来, 用以提高基于价值的 RL 方法的性能。美国人工智能企业 DeepMind 总结了 DQN 的 6 种主要改进方法<sup>[98]</sup>, 包括双 DQN 和基于竞争构架的 DQN 等, 并提供了丰富的测试和比较结果。此后,

文[99]提出了这些改进算法的深入分析和不同见解。在机器人技能学习方面, 文[100]提出了基于双 DQN 网络的编队策略, 能够控制多机器人系统接近目标并避免碰撞。文[101]提出了一种改进的 DQN 方法, 能够实现多机器人系统的自主轨迹规划功能。文[102]将 DQN 开发为机器人逆运动学求解器, 能够基于给定轨迹任务生成高自由度机械臂的关节空间轨迹。

### 5.2 基于策略的强化学习

基于策略的 RL 方法需要学习策略函数, 进而利用学习的策略逼近状态价值函数  $V_\pi(s) = E_A(Q_\pi(s, A))$ 。该方法需要将策略参数化, 并通过策略网络  $\pi(a|s; \iota)$  替代策略  $\pi(a|s)$ , 其中  $\iota$  表示神经网络模型的参数<sup>[95]</sup>。因此, 状态值函数可以近似写成  $V(s; \iota) = \sum_a \pi(a|s; \iota) Q_\pi(s, a)$ 。 $V(s; \iota)$  能够评价状态  $s$  和策略网络  $\pi(a|s; \iota)$ 。在给定状态  $s$  的情况下, 状态值函数  $V(s; \iota)$  的值越大, 策略网络  $\pi(a|s; \iota)$  越好。更新参数  $\iota$  以提高状态值函数  $V(s; \iota)$  是一种可行的策略优化方法。为了消除状态  $s$  的随机性影响, 目标函数被设计为  $O(\iota) = E_s(V(S; \iota))$ 。由此, 基于策略的 RL 的目标是通过更新参数  $\iota$  以获得最大的  $O(\iota)$ 。策略梯度法是一种主流的参数更新方法, 能够有效解决连续空间决策问题<sup>[103]</sup>。该方法通过引入学习速率  $\zeta > 0$  和策略梯度以更新参数  $\iota$ , 具体更新公式为  $\iota \rightarrow \iota + \zeta \cdot \partial V(s; \iota) / \partial \iota$ , 式中  $\partial V(s; \iota) / \partial \iota = \sum_a \partial \pi(a|s; \iota) / \partial \iota \cdot Q_\pi(s, a)$ 。然而, 策略梯度法的缺点是策略估计结果存在较大方差。此外, 该方法还存在样本利用率低、收敛速度慢的缺点<sup>[104]</sup>。基于策略的 RL 方法在机器人学习与控制应用中获得了丰富的研究成果<sup>[105]</sup>。文[106]开发一种改进的策略梯度算法, 能够直接学习机器人的深度神经网络控制策略, 不需要人工分析机器人控制模型。文[107]提出一种样本利用率高的 RL 算法, 通过学习机器人的最佳运动参数以赋予机器人准确回击乒乓球的技能。

### 5.3 基于 Actor-Critic 框架的强化学习

基于 Actor-Critic 框架的 RL 方法是一种同时学习动作价值函数和策略函数的前沿算法<sup>[108]</sup>。该方法结合了价值网络  $Q(s_t, a_t; g)$  和策略网络  $\pi(a|s; \iota)$  以估计状态价值函数  $V(s; \iota, g) = \sum_a \pi(a|s; \iota) Q(s, a; g)$ 。因此, 参数  $\iota$  和参数  $g$  的训练过程需要同步执行以提高系统动作评价的准确性并获得最优策略<sup>[109]</sup>。

Actor-Critic 框架能够通过时序差分方法更新参数  $g$ , 并通过策略梯度方法更新参数  $\tau^{[104]}$ 。在训练过程中, 策略网络  $\pi(a|s; \tau)$  负责控制机器人运动, 价值网络  $Q(s_t, a_t; g)$  负责评估动作<sup>[109]</sup>。基于 Actor-Critic 框架的 RL 方法的优势在于能够进行单步更新, 且更新速度比基于策略的 RL 方法更快。然而, 其缺点在于同时结合两个神经网络导致系统难以收敛。基于 Actor-Critic 框架的 RL 方法可应用于机器人在各种复杂环境中的最优阻抗控制, 并保持一定的安全性<sup>[110]</sup>。文[111]提出了一种基于 Actor-Critic 的 RL 方法并应用于机器人膝关节阻抗参数的自动调整, 能够学习完整的膝关节运动学模型。文[112]提出了一种 Actor-Critic 高斯混合方法, 该方法能够基于动态系统学习机器人技能, 并通过与环境的相互作用将学习的技能调整于机器人的轨迹分布空间中。

#### 5.4 基于 Sim2Real 的强化学习

基于 RL 方法的大部分研究都是在非真实物理环境中进行的。在真实机器人系统中, 基于神经网络的 RL 方法仍然存在着难以解决的瓶颈问题。具体而言, RL 方法需要机器人与现实环境之间丰富的交互数据。这要求机器人长时间执行各种交互动作, 从而导致了昂贵的现实试错成本。为了解决这个问题, Sim2Real 能够帮助现实机器人在模拟环境中使用数据进行学习, 属于一种迁移学习方法<sup>[113]</sup>。该技术在模拟环境中完成机器人控制策略的数据采集和网络训练过程, 然后将获得的控制策略应用于实际的机器人<sup>[114]</sup>。基于 Sim2Real 的 RL 技术被广泛应用于移动机器人、四足机器人、类人机器人等多种重要机器人的策略开发中<sup>[115]</sup>。然而, 模拟环境和现实环境之间存在不可避免的差异, 会在一定程度上影响策略的执行效率<sup>[115]</sup>。对此, 神经增强仿真法<sup>[116]</sup>和动态随机化处理法<sup>[117]</sup>被提出以减小模拟环境和现实环境的差距, 从而提高学习的可靠性。

综上所述, 基于神经网络的 RL 方法是一种以目标为导向的学习方法。该方法通过不断与环境交互试错以获得足够的知识, 从而使机器人能够做出合理的行为决策。其优势在于能够学习多变的环境模型, 并解决无模型动态交互问题。然而, 其缺点在于深度学习所需的昂贵数据成本和设备成本。此外, RL 方法通常严重依赖于仿真, 难以直接应用于实际机器人系统。

### 6 基于学习的机器人控制

机器人学习与控制系统的正常运行依靠两个功

能相异的内外循环。外循环管理系统的控制功能, 而内循环主要负责系统的参数学习功能。从系统控制的角度来看, 内部循环的收敛速度必须比外部循环更快, 这对网络的参数设定给出特殊的限制。换言之, 只有优越的学习能力才能保证机器人系统控制的精确性<sup>[26]</sup>。这强调了面向不同功能性提升的机器人控制方案的重要性。为解决机器人的控制问题, 传统方法通常会设计特定的优化方案来约束相关的性能指标, 从而规划机器人的运动轨迹。常用的优化指标包括可操作性<sup>[118]</sup>、关节漂移<sup>[119]</sup>和动能消耗<sup>[120]</sup>。基于模型学习的机器人控制方法将机器人优化方案表述为时变线性方程的数学求解问题<sup>[119-120]</sup>。求解这些机器人优化方案的常用算法包括伪逆法、数值算法、神经网络法等<sup>[121-122]</sup>。通常而言, 机器人优化方案可以通过基于二次规划的统一框架整合目标函数和等式与不等式约束条件<sup>[123-124]</sup>, 从而得到泛化形式:

$$\begin{aligned} & \min \frac{1}{2} s^T A s + v^T s \\ & \text{s.t. } As = b \\ & \quad s \in \omega \end{aligned} \quad (9)$$

式中,  $s \in \mathbb{R}^n$  表示机器人优化方案的决策变量, 可以根据任务需求选择为关节速度、关节加速度、关节突加度、关节扭矩等;  $A \in \mathbb{R}^{n \times n}$  表示一个正定的矩阵;  $v \in \mathbb{R}^n$  表示待优化的指标, 如复运动指标<sup>[125]</sup>、可操作度优化指标<sup>[126]</sup>和位置误差指标<sup>[121]</sup>;  $As = b$  表示根据任务要求设定的等式约束, 可以选择为动力学约束等式<sup>[127]</sup>、运动学约束等式<sup>[125]</sup>、视觉约束等式<sup>[21]</sup>等, 其中  $A \in \mathbb{R}^{c \times n}$ ,  $b \in \mathbb{R}^c$ ;  $\omega = \{s \in \mathbb{R}^n, s^- \leq s \leq s^+\}$  表示决策变量的不等式约束集, 其中  $s^-$  和  $s^+$  分别表示决策变量的下限和上限。二次规划方案能够以统一的框架描述基于运动学模型、动力学模型、混合模型等的机器人控制方案。近年来相关的研究成果有效解决了避障规划<sup>[120]</sup>、运动规划<sup>[42]</sup>、可操作度优化<sup>[112]</sup>、位置与力混合控制<sup>[59]</sup>、视觉伺服<sup>[21]</sup>等重要的机器人控制问题。基于模型的机器人控制方案中的学习问题可以通过结合 ND 方法和 FNN 方法的学习功能解决, 从而构建机器人系统的同步学习与控制的统一框架<sup>[74]</sup>。不同于工业机器人中使用的比例-积分-微分控制, 方案更加注重于待优化的指标和不等式约束, 更加具备功能性。此外, 方案可以基于模型预测控制思想来构建, 利用已有的模型、机器人系统当前的状态和未来的控制量以预测系统未来的输出。由于未来的控

制量是未知的, 模型预测控制需要设计误差函数进行辅助求解以得到有效的未来控制量序列。在每次控制周期结束后, 机器人系统根据当前实际状态重新预测未来输出。ZHANG 等提出了一种基于模型预测控制的二次规划方案, 能够基于机器人运动学模型预测未来系统的输出行为, 并通过求解最优控制问题将未来的跟踪误差最小化<sup>[121]</sup>。文[128]将模型预测控制的思想融入机器人动力学控制中, 有效考虑了关节扭矩的约束。不同于基于模型的机器人控制方案, 基于学习的机器人控制方案需要实现参数学习并基于学习结果计算控制信号。此外, 方案可以通过二次规划求解器直接进行求解, 从而得到方案的最优控制信号。需要说明的是, 二次规划的求解器有很多, 商业的求解器包括 Gurobi、Cplex、Xpress 和 Mosek<sup>[129]</sup>; 免费求解器包括 SCIP、CBC、GLPK 和 Matlab 中的 quadprog 求解器等。与此同时, 方案可以等价地转化为非线性方程组求解问题。因此, 常用的二次规划求解器也包括处理非线性方程组求解问题的数值方法, 例如 94LVI(94 linear variational inequality, LVI)算法、E47 算法和 M4 法<sup>[130]</sup>。

基于方案开发的机器人控制器通常拥有较高的求解速度以满足实时性要求。机器人控制问题的传统方法是伪逆法<sup>[131]</sup>。基于该方法的控制器在驱动机器人实现给定任务的同时, 能够优化单个额外目标。对于方案而言, 基于伪逆法的控制器可以表述为

$$\mathbf{s} = \mathbf{A}^\dagger \mathbf{b} - (\mathbf{I} - \mathbf{A}^\dagger \mathbf{A}) \mathbf{v} \quad (10)$$

式中,  $\mathbf{I} \in \mathbb{R}^{n \times n}$  表示单位阵。虽然控制器结构简单且广泛应用于工业机器人控制中, 但是它只能求解标准的单约束控制方案, 无法处理涉及多个等式约束和不等式约束的复杂情况。这在一定程度上限制了伪逆法的应用范围。例如, 以最小化关节速度无穷范数为目标的机器人控制方案就涉及多个不等式约束<sup>[132]</sup>, 无法通过基于伪逆法的控制器实现有效求解。

不同于基于伪逆法的控制器, 基于 ND 方法的控制器凭借其强大的并行处理能力和计算能力在机器人学习与控制研究中取得了卓有成效的理论与实践成果。就其本质而言, ND 方法属于一种 ODE 形式的实时求解算法, 能够凭借全局指数收敛特性实现系统误差控制<sup>[123-125]</sup>。因此, 基于学习的机器人控制问题等价于求解 ODE 系统的平衡解问题。值得注意的是, 所设计 ODE 系统可以直接基于任务要求与机器人模型构建动力学控制器, 也可以基于方案的最优解问题构建线性方程组。后者可以通过 Karush-Kuhn-Tucker 条件获得, 并满足线性方程组:

$$\mathbf{H}\mathbf{u} = \mathbf{r} \quad (11)$$

式中,  $\mathbf{H} = \begin{bmatrix} \mathbf{A} & \mathbf{A}^T \\ -\mathbf{A} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{(c+n) \times (c+n)}$ ;  $\mathbf{u} = \begin{bmatrix} \mathbf{s} \\ \boldsymbol{\lambda} \end{bmatrix} \in \mathbb{R}^{c+n}$ ,  $\boldsymbol{\lambda} \in \mathbb{R}^c$ , 为等式约束的拉格朗日乘子;  $\mathbf{r} = \begin{bmatrix} -\mathbf{v} \\ -\mathbf{b} \end{bmatrix} \in \mathbb{R}^{c+n}$ ;  $\mathbf{u} \in \Omega$ ,  $\Omega = \{ \mathbf{u} \in \mathbb{R}^{c+n}, \mathbf{u}^- \leq \mathbf{u} \leq \mathbf{u}^+ \}$ , 其中,  $\mathbf{u}^+ = \begin{bmatrix} \mathbf{s}^+ \\ \mathbf{h} \end{bmatrix} \in \mathbb{R}^{c+n}$ ,  $\mathbf{u}^- = \begin{bmatrix} \mathbf{s}^- \\ -\mathbf{h} \end{bmatrix} \in \mathbb{R}^{c+n}$ ,  $\mathbf{h} \in \mathbb{R}^{c+n}$  表示由极大的正数组成的向量。

本节的后续内容将基于 ND 算法的控制器归纳为 ZND 控制器、梯度神经网络(gradient neural network, GNN)控制器和动态神经网络(dynamic neural network, DNN)控制器, 并在下文中进行了详细介绍。以 GNN 控制器为例, 其在机器人系统中的工作流程如图 6 所示。基于该网络结构, 机器人系统的状态数据被测量作为反馈信息, 进而驱动机器人实现给定任务。

## 6.1 ZND 控制器

ZND 控制器能够将机器人时变最优化问题转化为时变等式求解问题, 进一步归纳为零化问题进行求解<sup>[133]</sup>。2002 年, ZHANG 等首次提出了一种基于神经动力学模型  $\dot{\boldsymbol{\epsilon}} = -\boldsymbol{\mu}\boldsymbol{\epsilon}$  的 ZND 控制器以在线求解线性方程组<sup>[134]</sup>。以式(11)为例, ZND 控制器将零化误差函数定义为  $\boldsymbol{\epsilon} = \mathbf{H}\mathbf{u} - \mathbf{r}$ , 并得到:

$$\dot{\mathbf{u}} = \mathbf{H}^{-1}(\dot{\mathbf{r}} - \dot{\mathbf{H}}\mathbf{u} - \boldsymbol{\mu}(\mathbf{H}\mathbf{u} - \mathbf{r})) \quad (12)$$

式中,  $\boldsymbol{\mu} > 0$  为收敛参数; 参数  $\dot{\mathbf{u}}$ 、 $\dot{\mathbf{r}}$ 、 $\dot{\mathbf{H}}$  分别为  $\mathbf{u}$ 、 $\mathbf{r}$ 、 $\mathbf{H}$  的时间导数;  $\mathbf{H}^{-1}$  表示矩阵  $\mathbf{H}$  的逆矩阵。式(12)的零化误差函数满足  $\boldsymbol{\epsilon}(t) = \boldsymbol{\epsilon}(0)\exp(-\boldsymbol{\mu}t)$ , 能够指数收敛于 0, 其收敛速度与  $\boldsymbol{\mu}$  成正相关。然而, 式(12)无法处理关于决策变量的不等式约束条件。对此, 文[132]提出了一种将不等式约束转化为等式约束的方式, 进而赋予 ZND 控制器解决不等式约束的能力, 在一定程度上拓展了 ZND 控制器在机器人学习与控制方面的求解能力。此外, 后续改进的 ZND 控制器(如抗噪 ZND<sup>[135]</sup>、有限时间收敛 ZND<sup>[136]</sup>等)提高了系统的鲁棒性、收敛速度和容错能力。基于欧拉差分方法或其他多步差分法的离散 ZND 控制器能够顺利地嵌入机器人系统中<sup>[137]</sup>。近年来, ZND 控制器在机器人控制方面的拓展性研究包括重复运动控制、关节速度无穷范数最小化控制、多机器人协同控制等<sup>[138-141]</sup>。此外, 文[28]将 ZND 控制器和机器人学习算法相结合, 实现了机器人系统无模型条件下的有效控制。然而, ZND 控制器中涉及参数的导数项, 可能由于外界噪声干扰而发生控制性能衰退的情况, 这不利于机器人系统的长期运行。

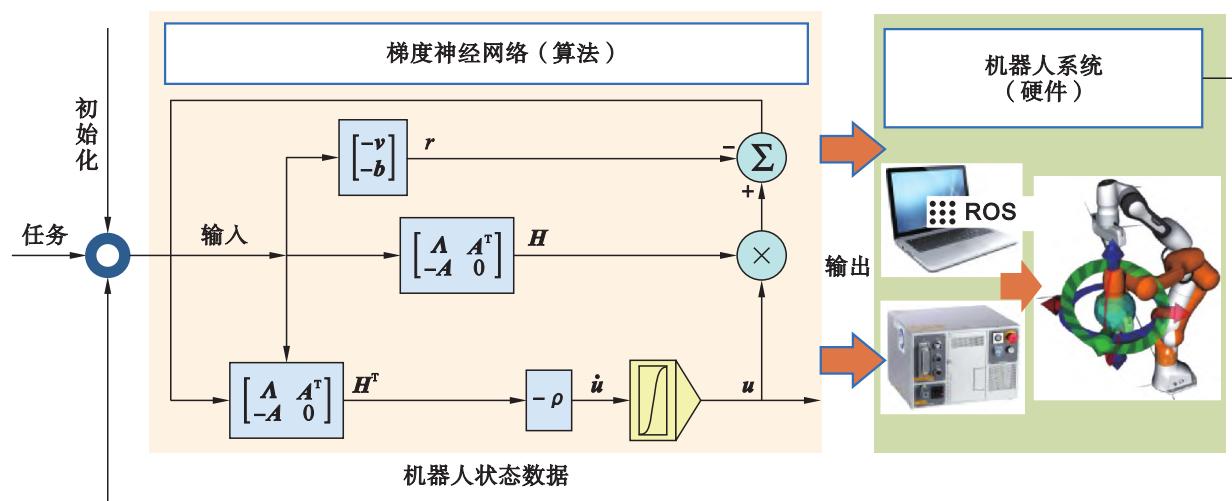


图 6 基于 GNN 控制器的机器人系统

Fig.6 GNN controller based robotic system

## 6.2 GNN 控制器

GNN 控制器属于一种基于梯度下降法的实时求解算法，能够利用梯度信息寻找目标函数最小化的演化方向，并通过不断迭代调整机器人系统状态以寻找最优解。对于式(11)而言，GNN 控制器的输出会驱使系统沿负梯度方向进行衍化，并收敛至方案的最优解<sup>[46]</sup>。具体地，面向机器人控制问题的 GNN 控制器可以被推导为

$$\dot{u} = -\rho \frac{\partial \|Hu - r\|^2}{\partial u} = -\rho H^T (Hu - r) \quad (13)$$

式中， $\rho > 0$  表示收敛系数，决定了机器人系统收敛速度和稳定状态的精度； $\|\cdot\|^2$  表示向量 2 范数的平方。GNN 控制器的输出  $\dot{u}$  决定了决策变量  $u$  的变化方向，且  $u$  的前  $n$  个元素表示机器人的控制信号。文[142]证明了该控制器的收敛性，并验证了求解矩阵逆问题的可行性。然而，GNN 控制器没有考虑动态参数的速度补偿，适合于求解静态问题。根据文[143]证明，对于动态问题而言，GNN 控制器的误差函数  $\epsilon = Hu - r$  将收敛于非零的边界，进而增大了机器人系统的任务误差。对此，文[144]提出了一种自适应 GNN 控制器，通过开发自适应收敛系数提高计算精度和收敛速度，并在 PUMA 560 机器人上验证了该控制器的可行性。此后，文[56]对 GNN 控制器设计了速度补偿项，从理论上消除了机器人系统的滞后误差并通过投影函数施加了关节约束以确保机器人系统的安全性。文[145]中的研究工作将该控制器拓展至机器人的关节加速度控制中。此外，WEI 等利用非线性互补函数将机器人系统中的不等式约束转变为等式约束，并引入误差积

分项以有效提高机器人系统的抗噪能力<sup>[143]</sup>。此后，文[146]提出了基于正交投影矩阵的 GNN 控制器，成功地解耦了机器人关节空间误差和位置误差，从而实现了更高的控制精度。上述研究表明，GNN 控制器及其拓展方法的优势在于不涉及矩阵求逆计算且网络结构简单。

## 6.3 DNN 控制器

ZND 控制器和 GNN 控制器在处理机器人不等式约束时，需要通过一定的转化技巧将不等式约束转化为等式约束并重新构造线性方程组进行求解。这个转化过程不但增加了控制器的搭建难度，而且增加了机器人系统的计算开销。针对这个问题，DNN 控制器能够通过引入投影函数直接处理不等式约束，并且具备网络结构简单和计算复杂度低的优势。具体而言，一种基于线性变分不等式的原对偶神经网络(linear variational inequality-based primal-dual neural network, LVI-PDNN)在文[37]中被设计以求解线性方程组。作为一种特殊的 DNN 求解器，该方法提出了双桥理论，并将二次规划问题转化成了分段线性投影方程求解问题<sup>[147]</sup>。具体地，面向机器人控制问题的 LVI-PDNN 被设计为

$$\dot{u} = \eta(I + H^T)(P_\alpha(u - (Hu - r)) - u) \quad (14)$$

式中， $\eta > 0$  表示收敛系数；投影函数  $P_\alpha(\cdot)$  是神经网络的有界激活函数，并被表述为

$$P_\alpha(u) = \begin{cases} u^-, & u \leq u^- \\ u, & u^- < u < u^+ \\ u^+, & u \geq u^+ \end{cases}$$

LVI-PDNN 能够全局收敛于线性方程组的解，即机器人控制方案的最优解。类似于 LVI-PDNN 方法

的离散数值算法包括 94LVI 算法和 E47 算法<sup>[130]</sup>。这类方法被开发用于求解移动机器人视觉伺服、动力学系统控制等问题, 并在仿真和实验中达到了毫米级的控制精度<sup>[148-149]</sup>。

在后续的研究中, 简化的 DNN 求解器被设计为

$$\dot{\boldsymbol{u}} = \eta(P_{\Omega}(\boldsymbol{u} - (\boldsymbol{H}\boldsymbol{u} - \boldsymbol{r})) - \boldsymbol{u}) \quad (15)$$

其控制框图如图 7 所示。文[150]证明了当矩阵  $\boldsymbol{H}$  连续可导且单调时, DNN(15) 的解将全局收敛于  $\boldsymbol{H}\boldsymbol{u} = \boldsymbol{r}$  的最优解。文[56]将简化的 DNN 控制器与基于梯度下降的学习算法结合在同一个方案中, 从而实现了雅可比矩阵学习误差和机器人跟踪误差的同步收敛。基于类似的框架, 一种用于求解多层变量约束方案的 DNN 控制器在文[151] 中被研究。该求解器能够在缺乏机器人模型先验知识的条件下约束更深层决策变量。简言之, DNN 控制器的优势在于能够直接处理不等式约束。此外, 上述 3 种

ND 控制器的比较如表 3 所示。

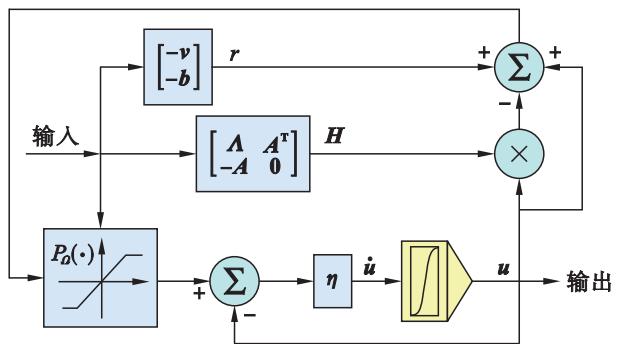


图 7 简化的 LVI-PDNN 的框图

Fig.7 Block diagram of simplified LVI-PDNN

综上所述, 基于 ND 方法的控制器有效提升了机器人系统的高效性、安全性和稳定性。目前, 关于机器人控制器的抗噪性研究和交互研究仍然是机器人系统在实际应用中的重要问题<sup>[152]</sup>。

表 3 不同 ND 控制器的特点比较

Tab.3 Comparisons of characteristic of different ND controllers

控制器	优势	劣势
ZND	全局收敛速度快, 理论误差为 0	需要时间导数信息, 需要矩阵求逆运算, 难以考虑关节约束
GNN	考虑关节约束, 只需计算矩阵转置	存在滞后误差
DNN	直接处理不等式约束, 不需要矩阵求逆	抗噪能力弱

## 7 挑战与展望

ND 方法、FNN 方法、RNN 方法和 RL 方法为解决机器人系统内部的学习与控制问题做出了巨大贡献, 并取得了丰硕的理论与实践成果。然而, 机器人还无法普及于人类社会中的各个领域中。其发展仍然存在许多问题与挑战。因此, 本节将介绍机器人学习与控制领域的 5 个问题。

1) 数据问题: 机器人学习算法依赖数据驱动技术进行学习或估计过程。这种学习过程要求获得机器人系统的状态数据, 例如, Broyden 算法需要获取末端执行器的实时速度信息<sup>[49-50]</sup>。有限的传感设备条件会极大地限制机器人学习算法的可行性。另一方面, 机器人系统的学习与控制精度也会受到传感器测量误差的影响。其次, 传感器系统故障所导致的错误数据可能引发致命的事故。例如, 2019 年 3 月 10 日, 埃塞俄比亚航空一架波音 737 Max 8 航班由于迎角探测器的错误数据在起飞 6 min 后坠毁<sup>[153]</sup>。以上问题对高性能传感器、数据可信度、故障应对和数据处理方法提出了更高的要求。

2) 噪声问题: 在实际机器人系统中, 噪声是不

可避免的, 可能来源于测量、计算、环境等因素。现有的 ND 算法甚至会引入较小的附加噪声来促进学习过程<sup>[26, 56, 59]</sup>。附加噪声能够增加系统输入的多样性, 在一定程度上避免鞍点问题<sup>[154]</sup>, 进而提高收敛能力。然而, 这一现象是缺乏可解释性的。此外, 相对较小的附加噪声有利于学习过程的进行, 但噪声本身也会影响机器人系统的控制精度。相关研究并没有分析和消除附加噪声对机器人系统的潜在影响。因此, 提高基于噪声驱动的学习方法的可解释性和解决噪声对学习与控制系统的不利影响具备一定的研究意义。

3) 安全问题: 机器人学习与控制系统的安全性体现在两个方面。一方面, 用于控制机器人系统的算法需要是严格稳定的。在各种特殊情况下(如障碍物、时滞), 其应保持稳定的全局收敛特性, 这是保证机器人作业安全的前提。另一方面, 基于神经网络的学习算法需要对外界环境的噪声干扰具有鲁棒性。机器人系统的实际执行环境是复杂多变且非结构化的, 这会影响网络的输入数据。现有的神经网络技术鲁棒性还有待提高。当输入数据被污染后, 其学习准确率会急剧下降。这必然会使机器人系统

失控，并产生较大的控制误差，甚至损坏机器人。

4) 成本问题：作为机器人学习与控制技术的核心算法，神经网络在各种机器人应用场景中得到了广泛的应用。不同的实际场景也对神经网络的计算能力提出了多样化的要求。然而，复杂的深度神经网络拥有多隐藏层，在增强模型表达能力的同时也提高了其结构复杂度和数据需求量。因此，部分神经网络需要高性能的多核高速CPU和GPU进行复杂的单精度运算<sup>[155]</sup>。对此，成本问题的重点是在不影响性能的前提下减少神经网络的计算量、数据需求和存储需求<sup>[156]</sup>。因此，未来的研究方向需要开发更加高效、简洁的神经网络计算架构，进而降低硬件成本，提高数据利用效率。

5) 泛化问题：当遇到同类型的任务时，机器人需要有效利用所学习的知识来完成新任务，从而提高机器人系统的学习效率和灵活性。换言之，机器人需要根据原有的知识和技能适应新的环境和任务。然而，这方面的相关研究相对缺乏。因此，亟需提出增强机器人泛化性能的可行方案，进而提高机器人系统的智能化程度。

如前所述，当前机器人学习与控制研究仍面临着巨大挑战。这也表明机器人技术潜在的研究趋势。在未来，随着传感器技术、神经网络方法和数

学理论的进步，机器人学习与控制系统将会变得更加智能、更加通用。同时，随着近年来软体机器人技术<sup>[76]</sup>、脑机接口技术<sup>[15]</sup>、自动驾驶技术<sup>[157]</sup>、多机器人云服务技术<sup>[27]</sup>等一系列新兴概念的成熟，机器人系统的服务性能也将趋于智能化和人性化。

## 8 结论

近30年来，机器人系统借助ND方法、FNN方法、RNN方法和RL方法解决了学习与控制问题，并取得了丰富的理论与实践成果。然而，多方面尚存的挑战和问题仍然阻碍着机器人技术的进一步发展。在这样的研究背景下，本文对机器人学习与控制的研究主题进行了详细的回顾和讨论，综述了4种主要的神经网络技术。针对不确定性学习问题，重点介绍了ND方法、FNN方法、RNN方法和RL方法。这些方法能够完成模型学习任务或者动态决策任务。面向机器人控制问题，基于ND的控制器和数值方法凭借其全局指数收敛特性负责机器人系统的高效控制。最后，分析了基于神经网络的机器人学习与控制技术的5大挑战。笔者相信随着新兴技术的发展与成熟，未来机器人会变得更加高效、智能且易于沟通，更具安全性、稳定性和通用性。

## 参考文献

- [1] CERIANI N M, ZANCHETTIN A M, ROCCO P, et al. Reactive task adaptation based on hierarchical constraints classification for safe industrial robots[J]. IEEE/ASME Transactions on Mechatronics, 2015, 20(6): 2935–2949.
- [2] FICUCIELLO F, VILLANI L, SICILIANO B. Variable impedance control of redundant manipulators for intuitive human-robot physical interaction[J]. IEEE Transactions on Robotics, 2015, 31(4): 850–863.
- [3] KIM J, MATHELIN M DE, IKUTA K, et al. Advancement of flexible robot technologies for endoluminal surgeries[J]. Proceedings of the IEEE, 2022, 110(7): 909–931.
- [4] 徐翔斌, 马中强. 基于移动机器人的拣货系统研究进展[J]. 自动化学报, 2022, 48(1): 1–20.  
XU X B, MA Z Q. Robotic mobile fulfillment systems: State-of-the-art and prospects[J]. Acta Automatica Sinica, 2022, 48(1): 1–20.
- [5] 谭民, 王硕. 机器人技术研究进展[J]. 自动化学报, 2013, 39(7): 963–972.  
TAN M, WANG S. Research progress on robotics[J]. Acta Automatica Sinica, 2013, 39(7): 963–972.
- [6] 张宇, 刁亚楠, 梁升云, 等. 认知-运动康复医疗机器人应用设计[J]. 信息与控制, 2021, 50(6): 740–747, 760.  
ZHANG Y, DIAO Y N, LIAN S Y, et al. Cognitive-motion rehabilitation medical robot application design[J]. Information and Control, 2021, 50(6): 740–747, 760.
- [7] HE Y, DENG Z, ZHANG J. Design and voice-based control of a nasal endoscopic surgical robot[J]. CAAI Transactions on Intelligence Technology, 2021, 6(1): 123–131.
- [8] KAELBLING L P. The foundation of efficient robot learning[J]. Science, 2020, 369(6506): 915–916.
- [9] LEE M A, ZHU Y K, ZACHARES P. Making sense of vision and touch: Learning multimodal representations for contact-rich tasks[J]. IEEE Transactions on Robotics, 2020, 36(3): 582–596.
- [10] SCHAAL S, ATKESON C G. Learning control in robotics[J]. IEEE Robotics & Automation Magazine, 2010, 17(2): 20–29.

- [11] 刘乃军, 鲁涛, 蔡莹皓, 等. 机器人操作技能学习方法综述[J]. 自动化学报, 2019, 45(3): 458–470.  
LIU N J, LU T, CAI Y H, et al. A review of robot manipulation skills learning methods[J]. Acta Automatica Sinica, 2019, 45(3): 458–470.
- [12] RATLIFF N, BAGNELL J A, SRINIVASA S S. Imitation learning for locomotion and manipulation[C]//27th IEEE-RAS International Conference on Humanoid Robots. Piscataway, USA: IEEE, 2007: 392–397.
- [13] YANG C G, ZENG C, CONG Y, et al. A learning framework of adaptive manipulative skills from human to robot[J]. IEEE Transactions on Industrial Informatics, 2019, 15(2): 1153–1161.
- [14] XIE Z T, JIN L, LUO X. Kinematics-based motion-force control for redundant manipulators with quaternion control[J/OL]. IEEE Transactions on Automation Science and Engineering, 2022 [2022-08-04]. <http://ieeexplore.ieee.org/document/9817106>. DOI:10.1109/TASE.2022.3186668.
- [15] 伏云发, 王越超, 李洪谊, 等. 直接脑控机器人接口技术[J]. 自动化学报, 2012, 38(8): 1229–1246.  
FU Y F, WANG C Y, LI H Y, et al. Direct brain-controlled robot interface technology[J]. Acta Automatica Sinica, 2012, 38(8): 1229–1246.
- [16] 李江, 苏鹏, 冯冬冬, 等. 基于准确度评估的7自由度手术机器人术前摆位优化算法[J]. 信息与控制, 2022, 51(3): 280–287, 297.  
LI J, SU P, FENG D D, et al. Preoperative positioning optimization algorithm based on accuracy evaluation for 7-DOF surgery robot[J]. Information and Control, 2022, 51(3): 280–287, 297.
- [17] NICHOLSON-SMITH C, MEHRABI V, ATASHZAR S F, et al. A multi-functional lower- and upper-limb stroke rehabilitation robot[J]. IEEE Transactions on Medical Robotics and Bionics, 2020, 2(4): 549–552.
- [18] 徐克峰, 蒋近, 季心宇, 等. 腰部外骨骼机器人线性自抗扰控制器参数优化[J]. 信息与控制, 2022, 51(3): 288–297.  
XU K F, JIANG J, JI X Y, et al. Parameters optimization of the linear active disturbance rejection controller for lumbar exoskeleton robot[J]. Information and Control, 2022, 51(3): 288–297.
- [19] SHAHRIARI M, BIGLARBEGIAN M. A new conflict resolution method for multiple mobile robots in cluttered environments with motion-liveness[J]. IEEE Transactions on Cybernetics, 2018, 48(1): 300–311.
- [20] SU H, QI W, CHEN J H, et al. Fuzzy approximation-based task-space control of robot manipulators with remote center of motion constraint[J]. IEEE Transactions on Fuzzy Systems, 2022, 30(6): 1564–1573.
- [21] LI W B, CHIU P W Y, LI Z. An accelerated finite-time convergent neural network for visual servoing of a flexible surgical endoscope with physical and RCM constraints[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(12): 5272–5284.
- [22] GRIMM G, MESSINA M J, TUNA S E, et al. Nominally robust model predictive control with state constraints[J]. IEEE Transactions on Automatic Control, 2007, 52(10): 1856–1870.
- [23] SUN T R, PENG L, CHENG L, et al. Stability-guaranteed variable impedance control of robots based on approximate dynamic inversion[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(7): 4193–4200.
- [24] JIN L, LI S, YU J, et al. Robot manipulator control using neural networks: A survey[J]. Neurocomputing, 2018, 285(1): 23–34.
- [25] LYU S K, CHEAH C C. Data-driven learning for robot control with unknown Jacobian[J/OL]. Automatica, 2020, 120(1) [2022-01-12]. <https://www.sciencedirect.com/science/article/pii/S0005109820303186>. DOI:10.1016/j.automatica.2020.109120.
- [26] XIE Z T, JIN L, LUO X, et al. An acceleration-level data-driven repetitive motion planning scheme for kinematic control of robots with unknown structure[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(9): 5679–5691.
- [27] KEHOE B, PATIL S, ABBEEL P, et al. A survey of research on cloud robotics and automation[J]. IEEE Transactions on Automation Science and Engineering, 2015, 12(2): 398–409.
- [28] CHEN D C, ZHANG Y N, LI S. Tracking control of robot manipulators with unknown models: A Jacobian-matrix-adaption method[J]. IEEE Transactions on Industrial Informatics, 2018, 14(7): 3044–3053.
- [29] PIEPMEIER J A, MCMURRAY G V, LIPKIN H. Uncalibrated dynamic visual servoing[J]. IEEE Transactions on Robotics and Automation, 2004, 20(1): 143–147.

- [30] 常振强, 金龙, 李帅, 等. 微型扑翼飞行器的自适应位置跟踪控制器[J]. 信息与控制, 2021, 50(1): 88–94.  
CHANG Z Q, JIN L, LI S, et al. Adaptive position tracking controller for the flapping wing micro aerial vehicle[J]. Information and Control, 2021, 50(1): 88–94.
- [31] KHAN A H, LI S, LUO X. Obstacle avoidance and tracking control of redundant robotic manipulator: An RNN-based metaheuristic approach[J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4670–4680.
- [32] TAN N, YU P, LIAO S, et al. Recurrent neural networks as kinematics estimator and controller for redundant manipulators subject to physical constraints[J]. Neural Networks, 2022, 153(1): 64–75.
- [33] 徐睿, 梁循, 齐金山, 等. 极限学习机前沿进展与趋势[J]. 计算机学报, 2019, 42(7): 1640–1670.  
XU R, LI X, QI J S, et al. Advances and trends in extreme learning machine[J]. Chinese Journal of Computers, 2019, 42(7): 1640–1670.
- [34] YANG C G, WANG X J, LI Z J, et al. Teleoperation control based on combination of wave variable and neural networks[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(8): 2125–2136.
- [35] 杨丽, 吴雨茜, 王俊丽, 等. 循环神经网络研究综述[J]. 计算机应用, 2018, 38(S2): 1–6.  
YANG L, WU Y X, WANG J L, et al. Research on recurrent neural network[J]. Journal of Computer Applications, 2018, 38(S2): 1–6.
- [36] CHEN X F, LUO X, JIN L, et al. Growing echo state network with an inverse-free weight update strategy[J/OL]. IEEE Transactions on Cybernetics, 2022[2022-08-14]. <https://ieeexplore.ieee.org/document/9740037/>. DOI: 10.1109/TCYB.2022.3155901.
- [37] ZHANG T, ZHANG K F, LIN J T, et al. Sim2real learning of obstacle avoidance for robotic manipulators in uncertain environments[J]. IEEE Robotics and Automation Letters, 2022, 7(1): 65–72.
- [38] XIA Y S, FENG G, WANG J. A recurrent neural network with exponential convergence for solving convex quadratic program and related linear piecewise equations[J]. Neural Networks, 2004, 17(7): 1003–1015.
- [39] RUSSELL S J, NORVIG P. Artificial intelligence: A modern approach[M]. 4th ed. Hoboken, US: Pearson, 2021.
- [40] MEDINA J R, HIRCHE S. Considering uncertainty in optimal robot control through high-order cost statistics[J]. IEEE Transactions on Robotics, 2018, 34(4): 1068–1081.
- [41] 王飞跃. 智能控制五十年回顾与展望: 傅京孙的初心与萨里迪斯的雄心[J]. 自动化学报, 2021, 47(10): 2301–2320.  
WANG F Y. Intelligent control in 50 years: From K. S. Fu's vision to G. N. Saridis' inspiration[J]. Acta Automatica Sinica, 2021, 47(10): 2301–2320.
- [42] CHEAH C C, LIU C, SLOTINE J J E. Adaptive Jacobian tracking control of robots with uncertainties in kinematic, dynamic and actuator models[J]. IEEE Transactions on Automatic Control, 2006, 51(6): 1024–1029.
- [43] 李满宏, 张建华, 张小俊, 等. 基于马尔可夫决策过程的六足机器人自由步态规划[J]. 机器人, 2015, 37(5): 529–537.  
LI M H, ZHANG J H, ZHANG X J, et al. Free gait planning for a hexapod robot based on Markov decision process[J]. Robot, 2015, 37(5): 529–537.
- [44] 陈佳盼, 郑敏华. 基于深度强化学习的机器人操作行为研究综述[J]. 机器人, 2022, 44(2): 236–256.  
CHEN J P, ZHENG M H. A survey of robot manipulation behavior research based on deep reinforcement learning[J]. Robot, 2022, 44(2): 236–256.
- [45] WANG L R, XIANG Y, YANG W. Goal-auxiliary actor-critic for 6D robotic grasping with point clouds[C]//5th Conference on Robot Learning. Cambridge, USA: JMLR, 2022: 70–80.
- [46] ZHANG Y N. Revisit the analog computer and gradient-based neural system for matrix inversion[C]//IEEE International Symposium on Intelligent Control. Piscataway, USA: IEEE, 2005: 1411–1416.
- [47] ZHANG Y N, YANG W Y. Real-time solutions of time-varying separable integral equations (TVSIE) using Zhang equivalency [C]//2021 China Automation Congress. Piscataway, USA: IEEE, 2021: 3937–3943.
- [48] BROYDEN C G. On the discovery of the “good Broyden” method[J]. Mathematical Programming, 2000, 87(1): 209–213.
- [49] PIEPMEIER J A, McMURRAY G V, LIPKIN H. Tracking a moving target with model independent visual servoing: A predictive estimation approach[C]//IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 1998:

- 2652–2657.
- [50] PIEPMEIER J A. A dynamic quasi-Newton method for model independent visual servoing[D]. Atlanta, USA: Georgia Institute of Technology, 1999.
- [51] PIEPMEIER J A, LIPKIN H. Uncalibrated eye-in-hand visual servoing[J]. The International Journal of Robotics Research, 2003, 22(10/11): 805–819.
- [52] NAVARRO-ALARCON D, LIU Y H, ROMERO J G, et al. Model-free visually servoed deformation control of elastic objects by robot manipulators[J]. IEEE Transactions on Robotics, 2013, 29(6): 1457–1468.
- [53] YANG M, ZHANG Y N, ZHANG Z J, et al. Adaptive discrete ZND models for tracking control of redundant manipulator[J]. IEEE Transactions on Industrial Informatics, 2020, 16(12): 7360–7368.
- [54] JIN L, LI J C, SUN Z B, et al. Neural dynamics for computing perturbed nonlinear equations applied to ACP-Based lower limb motion intention recognition[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(8): 5105–5113.
- [55] TAN N, YU P, NI F L. New varying-parameter recursive neural networks for model-free kinematic control of redundant manipulators with limited measurements[J/OL]. IEEE Transactions on Instrumentation and Measurement, 2022 [2022-05-20]. <https://ieeexplore.ieee.org/document/9740118>. DOI:10.1109/TIM.2022.3161713.
- [56] LI S, SHAO Z, GUAN Y. A dynamic neural network approach for efficient control of manipulators[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2019, 49(5): 932–941.
- [57] ZHANG Y Y, CHEN S Y, LI S, et al. Adaptive projection neural network for kinematic control of redundant manipulators with unknown physical parameters[J]. IEEE Transactions on Industrial Electronics, 2018, 65(6): 4909–4920.
- [58] XIE Z T, JIN L, LUO X, et al. A data-driven cyclic-motion generation scheme for kinematic control of redundant manipulators [J]. IEEE Transactions on Control Systems Technology, 2022, 49(1): 53–63.
- [59] FAN J L, JIN L, XIE Z T, et al. Data-driven motion-force control scheme for redundant manipulators: A kinematic perspective [J]. IEEE Transactions on Industrial Informatics, 2022, 18(8): 5338–5347.
- [60] 吕涛, 杨芳艳, 李清都, 等. 基于时延估计的扰动卡尔曼滤波器外力估计[J]. 信息与控制, 2022, 51(2): 157–167.  
LYU T, YANG Y F, LI Q D, et al. External force estimation of perturbed Kalman filter based on time delay estimation[J]. Information and Control, 2022, 51(2): 157–167.
- [61] 袁健, 唐功友. 基于扩展卡尔曼滤波的移动机器人变结构线性化复合跟踪控制[J]. 信息与控制, 2010, 39(6): 749–754.  
YUAN J, TANG G Y. Compound tracking control with variable structure linearization for mobile robots based on extended Kalman Filter[J]. Information and Control, 2010, 39(6): 749–754.
- [62] AGARWAL V, PARTHASARATHY H. Disturbance estimator as a state observer with extended Kalman filter for robotic manipulator[J]. Nonlinear Dynamic, 2016, 85(1): 2809–2825.
- [63] LI M H, KANG R J, BRANSON D T, et al. Model-free control for continuum robots based on an adaptive Kalman filter[J]. IEEE/ASME Transactions on Mechatronics, 2018, 23(1): 286–297.
- [64] 刘金琨, 孙富春. 滑模变结构控制理论及其算法研究与进展[J]. 控制理论与应用, 2007, 24(3): 407–418.  
LIU J K, SUN F C. Research and development on theory and algorithms of sliding mode control[J]. Control Theory & Applications, 2007, 24(3): 407–418.
- [65] ROY S, BALDI S, FRIDMAN L M. On adaptive sliding mode control without a priori bounded uncertainty[J/OL]. Automatica, 2020, 111(1)[2022-07-04]. <https://www.sciencedirect.com/science/article/pii/S0005109819305114>. DOI:10.1016/j.automatica.2019.108650.
- [66] CHEN D C, LI S, LIN F J, et al. New super-twisting zeroing neural-dynamics model for tracking control of parallel robots: A finite-time and robust solution[J]. IEEE Transactions on Cybernetics, 2020, 50(6): 2651–2660.
- [67] 葛运建, 张建军, 戈瑜, 等. 无所不在的传感与机器人感知[J]. 自动化学报, 2002(S1): 125–133.  
HE Y J, ZHANG J J, GE Y, et al. Ubiquitous sensing and robot perception[J]. Acta Automatica Sinica, 2002(S1): 125–133.
- [68] MOSTASHIRI N, DHUPIA J S, VERL A W, et al. A novel spatial mandibular motion-capture system based on planar fiducial markers[J]. IEEE Sensors Journal, 2018, 18(24): 10096–10104.

- [69] 张立建, 胡瑞钦, 易旺民. 基于六维力传感器的工业机器人末端负载受力感知研究[J]. 自动化学报, 2017, 43(3): 439–447.
- ZHANG L J, HU R Q, YI W M. Research on force sensing for the end-load of industrial robot based on a 6-axis force/torque sensor[J]. Acta Automatica Sinica, 2017, 43(3): 439–447.
- [70] 王兴, 贾晓虎, 郝春丽. 基于增量式光电编码器位移传感器研究[J]. 电子设计工程, 2012, 20(5): 155–158.
- WANG X, JIA X H, HE C L. Research of the sensor about displacement based on the increment photoelectricity encoder[J]. Electronic Design Engineering, 2012, 20(5): 155–158.
- [71] LU X J, WOO H, FARAGASSO A, et al. Socially aware robot navigation in crowds via deep reinforcement learning with resiliency reward functions[J]. Advanced Robotics, 2022, 36(8): 388–403.
- [72] LYU X L, ZHANG M L, LI H. Robot control based on voice command[C]//2008 IEEE International Conference on Automation and Logistics. Piscataway, USA: IEEE, 2008: 2490–2494.
- [73] HÖFLINGER F, MÜLLER J, ZHANG R L, et al. A wireless micro inertial measurement unit (IMU)[J]. IEEE Transactions on Instrumentation and Measurement, 2013, 62(9): 2583–2595.
- [74] 王旭东, 邵惠鹤. RBF 神经网络理论及其在控制中的应用[J]. 信息与控制, 1997, 26(4): 32–44.
- WANG X D, SHAO H H. The theory of RBF neural network and its application in control[J]. Information and Control, 1997, 26(4): 32–44.
- [75] CHENG L, HOU Z G, TAN M. Adaptive neural network tracking control for manipulators with uncertain kinematics, dynamics and actuator model[J]. Automatica, 2009, 45(10): 2312–2318.
- [76] GIORELLI M, RENDA F, CALISTI M, et al. Neural network and Jacobian method for solving the inverse statics of a cable-driven soft arm with nonconstant curvature[J]. IEEE Transactions on Robotics, 2015, 31(4): 823–834.
- [77] GUO K L, SU H, YANG C G. A small opening workspace control strategy for redundant manipulator based on RCM method [J]. IEEE Transactions on Control Systems Technology, 2022, 30(6): 2717–2725.
- [78] GUTIERREZ-MARTINEZ V J, CAÑIZARES C A, FUERTE-ESQUIVEL C R, et al. Neural-network security-boundary constrained optimal power flow[J]. IEEE Transactions on Power Systems, 2011, 26(1): 63–72.
- [79] ZOU Y, ZHENG Z W. A robust adaptive RBFNN augmenting backstepping control approach for a model-scaled helicopter[J]. IEEE Transactions on Control Systems Technology, 2015, 23(6): 2344–2352.
- [80] ZHANG D, ZHANG G, LI L. Calibration of a six-axis parallel manipulator based on BP neural network[J]. Industrial Robot, 2019, 46(5): 692–698.
- [81] BAI Y H, LUO M Z, PANG F L. An algorithm for solving robot inverse kinematics based on FOA optimized BP neural network [J]. Applied Science, 2021, 11(15): 1–12.
- [82] ZHANG P Y, LÜ T S, SONG L B. RBF networks-based inverse kinematics of 6R manipulator[J]. The International Journal of Advanced Manufacturing Technology, 2005, 26(1): 144–147.
- [83] LI Y L, WANG L P, LIU K, et al. Distributed neural-network-based cooperation control for teleoperation of multiple mobile manipulators under round-robin protocol[J/OL]. IEEE Transactions on Neural Networks and Learning Systems, 2021 [2022–06–05]. <http://ieeexplore.ieee.org/document/9612720>. DOI:10.1109/TNNLS.2021.3125004.
- [84] YANG C G, HUANG K X, CHENG H, et al. Haptic identification by ELM-controlled uncertain manipulator[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(8): 2398–2409.
- [85] CHEN C L P, LIU Z L. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(1): 10–24.
- [86] XU S, LIU J, YANG C G, et al. A learning-based stable servo control strategy using broad learning system applied for microrobotic control[J/OL]. IEEE Transactions on Cybernetics [2022–04–05]. <https://ieeexplore.ieee.org/document/9612720>. DOI:10.1109/TCYB.2021.3121080.
- [87] KHAN A H, LI S, LUO X. Obstacle avoidance and tracking control of redundant robotic manipulator: An RNN-based metaheuristic approach[J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4670–4680.
- [88] LI W B, SONG C Z, LI Z. An accelerated recurrent neural network for visual servo control of a robotic flexible endoscope with joint limit constraint[J]. IEEE Transactions on Industrial Electronics, 2020, 67(12): 10787–10797.

- [89] LI Y M, LI S, HANNAFORD B. A model-based recurrent neural network with randomness for efficient control with applications [J]. IEEE Transactions on Industrial Informatics, 2019, 15(4): 2054–2063.
- [90] 罗熊, 黎江, 孙增圻. 回声状态网络的研究进展[J]. 北京科技大学学报, 2012, 34(2): 217–222.
- XIONG L, LI J, SUN Z Q. Review on echo state networks[J]. Journal of University of Science and Technology Beijing, 2012, 34(2): 217–222.
- [91] ZHANG J H, CHEN J H, WU W, et al. A cerebellum-inspired prediction and correction model for motion control of a musculo-skeletal robot[J/OL]. IEEE Transactions on Cognitive and Developmental Systems[2022-06-09]. <https://ieeexplore.ieee.org/document/9612720/>. DOI:10.1109/TCDS.2022.3200839.
- [92] TAN N, YU P, NI F L. A cerebellum-inspired network model and learning approaches for solving kinematic tracking control of redundant manipulators[J/OL]. IEEE Transactions on Cognitive and Developmental Systems[2022-07-07]. <https://ieeexplore.ieee.org/document/9706313/>. DOI:10.1109/TCDS.2022.3149622.
- [93] 李国友, 苏侠飞, 秦晓磊, 等. 基于回声状态网络的机械臂的视觉伺服控制[J]. 系统仿真学报, 2017, 29(8): 1845–1850.
- LI G Y, SU X F, QIN X L, et al. Robotic manipulator's visual servo control based on echo state networks[J]. Journal of System Simulation, 2017, 29(8): 1845–1850.
- [94] KUWABARA J, NAKAJIMA K, KANG R J, et al. Timing-based control via echo state network for soft robotic arm[C/OL]//2012 International Joint Conference on Neural Networks. Piscataway, USA: IEEE, 2012[2022-06-07]. <http://ieeexplore.ieee.org/document/6252774/>. DOI:10.1109/IJCNN.2012.6252774.
- [95] 刘建伟, 高峰, 罗雄麟. 基于值函数和策略梯度的深度强化学习综述[J]. 计算机学报, 2019, 42(6): 1406–1438.
- LIU J W, GAO F, LUO X L. Survey of deep reinforcement learning based on value function and policy gradient[J]. Chinese Journal of Computers, 2019, 42(6): 1406–1438.
- [96] 陈学松, 杨宜民. 基于递推最小二乘法的多步时序差分学习算法[J]. 计算机工程与应用, 2010, 46(8): 52–55.
- CHEN X S, YANG Y M. Multi-step temporal difference learning algorithm based on recursive least-squares method[J]. Computer Engineering and Applications, 2010, 46(8): 52–55.
- [97] GAO N, QIN Z J, JING X J, et al. Anti-intelligent UAV jamming strategy via deep Q-networks[J]. IEEE Transactions on Communications, 2020, 68(1): 569–581.
- [98] HESSEL M, MODAYIL J, HASSELT H V, et al. Rainbow: Combining improvements in deep reinforcement learning[C]//32nd AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2018: 3215–3222.
- [99] CERON J O, CASTRO P S. Revisiting rainbow: Promoting more insightful and inclusive deep reinforcement learning research[C]//38th International Conference on Machine Learning. New York, USA: ACM, 2021: 1373–1383.
- [100] BAI C C, YAN P, PAN W, et al. Learning-based multi-robot formation control with obstacle avoidance[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(8): 11811–11822.
- [101] YANG Y, LI J T, PENG L L. Multi-robot path planning based on a deep reinforcement learning DQN algorithm[J]. CAAI Transactions on Intelligence Technolog, 2020, 5(3): 177–183.
- [102] MALIK A, LISCHUK Y, HENDERSON T, et al. A deep reinforcement-learning approach for inverse kinematics solution of a high degree of freedom robotic manipulator[J]. Robotics, 2022, 11(2): 1–17.
- [103] 俞胜平, 韩忻辰, 袁志明, 等. 基于策略梯度强化学习的高铁列车动态调度方法[J]. 控制与决策, 2022, 37(9): 2407–2417.
- YU S P, HAN X C, YUAN Z M, et al. A policy gradient reinforcement learning algorithm for high-speed railway dynamic scheduling[J]. Control and Decision, 2022, 37(9): 2407–2417.
- [104] 朱斐, 朱海军, 刘全, 等. 一种解决连续空间问题的真实在线自然梯度 AC 算法[J]. 软件学报, 2018, 29(2): 267–282.
- ZHU F, ZHU H J, LIU Q, et al. True online natural actor-critic algorithm for the continuous space problem[J]. Journal of Software, 2018, 29(2): 267–282.
- [105] PETERS J, SCHAALE S. Policy gradient methods for robotics[C]//2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, USA: IEEE, 2006: 2219–2225.

- [106] LI X, SHANG W W, CONG S. Model-based reinforcement learning for robot control [C]//5th International Conference on Advanced Robotics and Mechatronics. Piscataway, USA: IEEE, 2020: 300–305.
- [107] TEBBE J, KRAUCH L, GAO Y P. Sample-efficient reinforcement learning in robotic table tennis [C]//2021 IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2021: 4171–4178.
- [108] 孙歧峰,任辉,段友祥. 基于异步优势执行器评价器学习的自适应PID控制设计[J]. 信息与控制, 2019, 48(3): 323–328, 338.  
SUN Q F, REN H, DUAN Y X. The adaptive PID control design based on asynchronous advantage actor-critic learning [J]. Information and Control, 2019, 48(3): 323–328, 338.
- [109] KONDA V R, TSITSIKLIS J N. Actor-critic algorithms [C]//1999 Advances in Neural Information Processing Systems. Cambridge, USA: MIT Press, 2000: 1008–1014.
- [110] ZHAO X W, HAN S B, TAO B, et al. Model-based actor-critic learning of robotic impedance control in complex interactive environment [J]. IEEE Transactions on Industrial Electronics, 2022, 69(12): 13225–13235.
- [111] WU R F, YAO Z K, SI J N, et al. Robotic knee tracking control to mimic the intact human knee profile based on actor-critic reinforcement learning [J]. IEEE/CAA Journal of Automatica Sinica, 2022, 9(1): 19–30.
- [112] NEMATOLLAHI I, ROSETE-BEAS E, RÖFER A, et al. Robot skill adaptation via soft actor-critic Gaussian mixture models [C]//2022 International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2022: 8651–8657.
- [113] NARANG Y, SUNDARALINGAM B, MACKLIN M, et al. Sim-to-real for robotic tactile sensing via physics-based simulation and learned latent projections [C]//2021 IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2021: 6444–6451.
- [114] 《机器人产业》编辑部. DeepMind“提出从模拟到真实”的新型迁移学习方案[J]. 机器人产业, 2019(6): 16–20.  
Editorial Board of Robot industry. DeepMind proposes a new transfer learning scheme called “simulation to reality” [J]. Robot industry, 2019(6): 16–20.
- [115] ZHU W, GUO X, OWAKI D, et al. A survey of sim-to-real transfer techniques applied to reinforcement learning for bioinspired robots [J/OL]. IEEE Transactions on Neural Networks and Learning Systems [2022–05–09]. <https://ieeexplore.ieee.org/document/9552429>. DOI:10.1109/TNNLS.2021.3112718.
- [116] GOLEMO F, TAIGA A A, COURVILLE A C, et al. Sim-to-real transfer with neural-augmented robot simulation [C]//2nd Conference on Robot Learning. Cambridge, USA: JMLR, 2018: 817–828.
- [117] PENG X B, ANDRYCHOWICZ M, ZAREMBA W, et al. Sim-to-real transfer of robotic control with dynamics randomization [C]//2018 IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2018: 3803–3810.
- [118] KENNEL-MAUSHART F, PORANNE R, COROS S. Manipulability optimization for multi-arm teleoperation [C]//2021 IEEE International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2021: 3956–3962.
- [119] XIAO L, ZHANG Y N. A new performance index for the repetitive motion of mobile manipulators [J]. IEEE Transactions on Cybernetics, 2014, 44(2): 465–474.
- [120] ZHANG Z J, YANG S, CHEN S Y, et al. A vector-based constrained obstacle avoidance scheme for wheeled mobile redundant robot manipulator [J]. IEEE Transactions on Cognitive and Developmental Systems, 2021, 13(3): 465–474.
- [121] ZHANG F, JIN L, LUO X. Error-summation enhanced newton algorithm for model predictive control of redundant manipulators [J/OL]. IEEE Transactions on Industrial Electronics [2022–06–03]. <https://ieeexplore.ieee.org/document/9756224/>. DOI:10.1109/TIE.2022.3165277.
- [122] 邱志成. 柔性机械臂的振动测量和控制研究进展综述[J]. 信息与控制, 2021, 50(2): 141–161.  
QIU Z C. Review on research progress in vibration measurement and control of flexible manipulators [J]. Information and Control, 2021, 50(2): 141–161.
- [123] HASSAN A A, EL-HABROUK M, DEGHEDIE S. Inverse kinematics of redundant manipulators formulated as quadratic programming optimization problem solved using recurrent neural networks: A review [J]. Robotica, 2019, 38(8): 1495–1512.
- [124] ZHANG X Y, CHEN L M, LI S, et al. Design and analysis of recurrent neural network models with non-linear activation functions for solving time-varying quadratic programming problems [J]. CAAI Transactions on Intelligence Technology, 2021, 6(4): 394–404.

- [125] XIE Z T, JIN L, DU X J, et al. On generalized RMP scheme for redundant robot manipulators aided with dynamic neural networks and nonconvex bound constraints[J]. IEEE Transactions on Industrial Informatics, 2019, 15(9): 5172–5181.
- [126] ALAMBEIGI F, WANG Z, HEGEMAN R. Autonomous data-driven manipulation of unknown anisotropic deformable tissues using unmodelled continuum manipulators[J]. IEEE Robotics and Automation Letters, 2019, 4(2): 254–261.
- [127] ZHANG Y N, WANG J. A dual neural network for constrained joint torque optimization of kinematically redundant manipulators [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2002, 32(5): 654–662.
- [128] NICOLIS D, ALLEVI F, ROCCO P. Operational space model predictive sliding mode control for redundant manipulators [J]. IEEE Transactions on Robotics, 2020, 36(4): 1348–1355.
- [129] SOU K C, SANDBERG H, JOHANSSON K H. Computing criticalk-tuples in power networks[J]. IEEE Transactions on Power Systems, 2012, 27(3): 1511–1520.
- [130] ZHANG Z J, ZHENG L, YU J M, et al. Three recurrent neural networks and three numerical methods for solving a repetitive motion planning scheme of redundant robot manipulators [J]. IEEE/ASME Transactions on Mechatronics, 2017, 22(3): 1423–1434.
- [131] DING H, TSO S K. Redundancy resolution of robotic manipulators with neural computation[J]. IEEE Transactions on Industrial Electronics, 1999, 46(1): 230–233.
- [132] LI W B, CHIU P W Y, LI Z. A novel neural approach to infinity-norm joint-velocity minimization of kinematically redundant robots under joint limits[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(1): 409–420.
- [133] 金龙, 李嘉昌, 常振强, 等. 基于ACP理论的微型扑翼飞行器的姿态控制[J/OL]. 自动化学报, 2021 [2022-04-05]. <http://www.aas.net.cn/article/doi/10.16383/j.aas.c210646>. DOI:10.16383/j.aas.c210646.
- JIN L, LI J C, CHANG Z Q, et al. Attitude control for flapping wing micro aerial vehicle based on ACP theory[J/OL]. Acta Automatica Sinica, 2021 [2022-04-05]. <http://www.aas.net.cn/article/doi/10.16383/j.aas.c210646>. DOI:10.16383/j.aas.c210646.
- [134] ZHANG Y N, JIANG D C, WANG J. A recurrent neural network for solving Sylvester equation with time-varying coefficients [J]. IEEE Transactions on Neural Networks, 2002, 13(5): 1053–1063.
- [135] JIN L, ZHANG Y N, LI S, et al. Noise-tolerant ZNN models for solving time-varying zero-finding problems: A control-theoretic approach[J]. IEEE Transactions on Automatic Control, 2017, 62(2): 992–997.
- [136] LI W B. Predefined-time convergent neural solution to cyclical motion planning of redundant robots under physical constraints [J]. IEEE Transactions on Industrial Electronics, 2020, 67(12): 10732–10743.
- [137] LI J, ZHANG Y N, MAO M Z. Continuous and discrete zeroing neural network for different-level dynamic linear system with robot manipulator control[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 50(11): 4633–4642.
- [138] STANIMIROVIĆ P S, KATSIKIS V N, LI S. Higher-order ZNN dynamics[J]. Neural Process Letter, 2020, 51(1): 697–721.
- [139] JIN L, LI S, LIAO B L, et al. Zeroing neural networks: A survey[J]. Neurocomputing, 2017, 267(1): 597–604.
- [140] KATSIKIS V N, STANIMIROVIĆ P S, MOURTAS S D, et al. Zeroing neural network with fuzzy parameter for computing pseudoinverse of arbitrary matrix[J]. IEEE Transactions on Fuzzy Systems, 2022, 30(9): 3426–3435.
- [141] ZHANG J Z, JIN L, WANG Y. Collaborative control for multi-manipulator systems with fuzzy neural networks[J/OL]. IEEE Transactions on Fuzzy Systems [2022-06-08]. <https://ieeexplore.ieee.org/document/9857613/>. DOI: 10.1109/TFUZZ.2022.3198855.
- [142] ZHANG Y N, CHEN K, TAN H Z. Performance analysis of gradient neural network exploited for online time-varying matrix inversion[J]. IEEE Transactions on Automatic Control, 2009, 54(8): 1940–1945.
- [143] JIN L, WEI L, LI S. Gradient-based differential neural-solution to time-dependent nonlinear optimization[J/OL]. IEEE Transactions on Automatic Control [2022-04-23]. <https://ieeexplore.ieee.org/document/9684931/>. DOI: 10.1109/TAC.2022.3144135.
- [144] LIAO S, LIU J Y, QI Y M, et al. An adaptive gradient neural network to solve dynamic linear matrix equations[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(9): 5913–5924.
- [145] JIN L, XIE Z T, LIU M, et al. Novel joint-drift-free scheme at acceleration level for robotic redundancy resolution with tracking

- error theoretically eliminated[J]. IEEE/ASME Transactions on Mechatronics, 2021, 26(1): 90–101.
- [146] XIE Z T, JIN L, LUO X, et al. RNN for repetitive motion generation of redundant robot manipulators: An orthogonal projection-based scheme[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(2): 615–628.
- [147] KATSIKIS V N, MOURTASA S D, STANIMIROVIĆ P S. Time-varying mean – variance portfolio selection problem solving via LVI-PDNN[J/OL]. Computers & Operations Research, 2022, 138(1)[2022-05-05]. <https://www.sciencedirect.com/science/article/abs/pii/S0305054821003099>. DOI:10.1016/j.cor.2021.105582.
- [148] LI Z J, DENG J, LU R Q, et al. Trajectory-tracking control of mobile robot systems incorporating neural-dynamic optimized model predictive approach[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2016, 46(6): 740–749.
- [149] KE F, LI Z J, YANG C G. Robust tube-based predictive control for visual servoing of constrained differential-drive mobile robots[J]. IEEE Transactions on Industrial Electronics, 2018, 65(4): 3437–3446.
- [150] GAO X B. Exponential stability of globally projected dynamic systems[J]. IEEE Transactions on Neural Networks, 2003, 14(2): 426–431.
- [151] LIU M, FAN J L, ZHENG Y, et al. A simultaneous learning and control scheme for redundant manipulators with physical constraints on decision variable and its derivative[J]. IEEE Transactions on Industrial Electronics, 2022, 69(10): 10301–10310.
- [152] 李军强, 吕瑞武, 杨冬, 等. 基于柔性人机接口的人机协调运动控制方法[J]. 信息与控制, 2022, 51(2): 237–246, 256.
- LI J Q, LYU R W, YANG D, et al. Control method of human-robot coordinated motion based on flexible human-robot interface [J]. Information and Control, 2022, 51(2): 237–246, 256.
- [153] JAMIESON G A, SKRAANING G, JOE J. The B737 MAX 8 accidents as operational experiences with automation transparency [J]. IEEE Transactions on Human-Machine Systems, 2022, 52(4): 794–797.
- [154] GE R, HUANG F R, JIN C, et al. Escaping from saddle points – Online stochastic gradient for tensor decomposition[C]// 28th Conference on Learning Theory. Cambridge, USA: JMLR, 2015: 797–842.
- [155] SZE V, CHEN Y H, EMER J, et al. Hardware for machine learning: Challenges and opportunities[C/OL]// IEEE Custom Integrated Circuits Conference. Piscataway, USA: IEEE, 2017. <https://ieeexplore.ieee.org/document/7993626>. DOI:10.1109/CICC.2017.7993626.
- [156] 张政馗, 庞为光, 谢文静, 等. 面向实时应用的深度学习研究综述[J]. 软件学报, 2020, 31(9): 2654–2677.
- ZHANG Z K, PANG W G, XIE W J, et al. Deep learning for real-time applications: A survey[J]. Journal of Software, 2020, 31(9): 2654–2677.
- [157] CLAUSSMANN L, REVILLOUD M, GRUYER D, et al. A review of motion planning for highway autonomous driving[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(5): 1826–1848.

## 作者简介

谢正泰(1996-), 男, 博士生。研究领域为智能机器人, 神经网络算法, 优化理论。

樊佳亮(1997-), 男, 硕士生。研究领域为智能机器人, 神经网络算法。

刘 梅(1988-), 女, 硕士, 工程师。研究领域为智能机器人, 智能计算。