

基于生物智能算法的智能控制研究与 实践应用

谢文光, 阎芳, 汪克念

1. 民航航空器适航审定技术重点实验室, 天津 300300
2. 中国民航大学适航学院, 天津 300300

摘要 传统的自动控制方法因为其固定参数等弊端极大限制了控制效果, 生物智能算法因其环境自适应与自学习机制的特性, 为突破传统控制方法的瓶颈提供了一种新的思路, 并且随着强化学习等机器学习理论与方法的不断完善与发展, 生物智能算法的性能也得到了极大的提高。总结了在智能控制中常用的7种生物智能算法, 分析了经典的自动控制方法与生物智能算法, 尤其是强化学习、深度学习等新型智能算法的结合的应用实例。结合近年来兴起的深度学习, 强化学习及类脑智能科学对智能控制的发展现状, 以及未来的发展趋势进行展望。强调一种智能辅助控制方法, 将智能算法与传统控制方法相结合, 为智能控制的研究提供新的思路与实用范例。

关键词 智能控制; 生物智能; 自动控制; 强化学习

智能控制^[1]是指在极少或是几乎没有人为干预的情况下, 计算机能够独立驱动机器实现目标的控制过程。把人工智能的理论和方法与传统的自动控制技术相结合就得到了智能控制, 智能控制是一种研制智能控制系统的方法和技术。智能控制系统^[2]是一种具有自学习、自组织、自适应等功能的智能系统。智能控制方法是一种面向过程的控制方法。智能控制与经典控制的本质区别是其控制过程具有学习与决策2个特点。智能控制的学习特

点是该过程不仅具有从外界获取并存储信息的能力, 还能够不断积累经验, 通过不断学习获得更优的控制策略, 进而增强对环境的应变能力。智能控制的决策特点是智能控制可以根据已经收集过的样本信息及所感知到的外界环境的改变, 及时调整控制策略, 使得控制过程朝着改善系统性能的方向发展。

经典的智能控制方法包括模块控制、神经网络控制^[3]、专家控制^[4]等, 其主要应用领域包括智能机

收稿日期: 2019-01-03; 修回日期: 2019-05-10

基金项目: 国家自然科学基金青年科学基金项目(61601468)

作者简介: 谢文光, 助理研究员, 研究方向为嵌入式计算机软件、计算机控制、机载软件适航审定, 电子信箱: caucxwg@foxmail.com

引用格式: 谢文光, 阎芳, 汪克念. 基于生物智能算法的智能控制研究与实践应用[J]. 科技导报, 2019, 37(17): 65-72; doi: 10.3981/j.issn.1000-7857.2019.17.012

器人系统、物联网系统、计算机集成制造系统、视频游戏等。2006年深度学习的提出为解决优化神经网络相关结构的难题带来曙光,之后深层自动编码器结构被提出^[5,6]。深度学习为解决大规模特征训练问题提供了方法,提高了系统的学习能力与自适应性,因而被广泛应用于视觉图像处理、无人系统研究等领域。强化学习^[7]是机器学习的一个重要分支,在智能控制中有着广泛的应用。强化学习除了应用在线性控制、博弈论、机器人智能控制^[8]等方向外,在人机交互、无人驾驶、机器翻译、文本序列预测等方向也有很多应用。

生物智能算法是通过利用自然界的一般规律,以及根据生物生存与进化的原理,模仿求解问题的算法。生物智能算法的内容有很多,如神经网络、遗传算法、模拟退火算法和群集智能技术等。经典的生物智能算法可以作为辅助控制器为自动控制提供智能辅助控制^[9],也可以与传统的控制方法结合形成智能控制方法甚至可以做为单独的控制器对控制系统进行分析、预测、故障诊断和解耦控制等。

类脑智能^[10](brain-inspired intelligence, BII)是当今人工智能技术研究的热点与前沿,当下许多重要成果都反映了这种重要的技术发展趋势。即使是对脑科学理论的部分借鉴,都可以有效提升现有人工智能系统的智能水平。

1 智能控制中的生物智能算法模型

1.1 人工神经网络

神经网络(neural network, NN)是连接主义智能实现的范例,其采用了有效的学习机制与极为广泛的级联结构来模拟大脑信息处理的过程^[11]。人工神经网络是一种重要的人工智能方法,同时也是当下类脑智能研究过程中的一种有效工具。神经网络可以模仿人类的神经级联结构,它通过一种对大量人工神经元的大规模互联的机制,进而构造人工神经网络系统,模拟生物神经系统。单个神经元之间的联结模式构成了人工神经网络的互联结构,同时它也是构建神经网络的基础。从网络互联结构的角度看,神经网络分为前向型神经网络与反馈

型神经网络2种不同的类型。在现有的神经网络系统中,最常用的神经网络包括传统感知器模型、具有误差反向传播功能的BP(back propagation)神经网络模型。神经网络一般可以作为单独的控制器实现控制系统的故障诊断、模式识别^[12]等,也可以作为辅助控制器^[13],甚至可以结合传统的控制方法形成智能控制方法。

神经网络有多种分类方法,一般按照神经网络模型的类型,可将神经网络分为感知器神经网络、BP神经网络、径向基神经网络以及近些年来兴起的深度神经网络等。

1.1.1 感知器神经网络

感知器学习^[14]使用了一种迭代的思想调整神经网络的连接权值和连接阈值,这个过程一直持续到满足约束条件为止,是一种基于纠错学习的规则的学习方法。感知器模型是一种可以自学习与自调节的神经网络模型,同时它是第一个机器学习模型。它把人们对神经网络的研究从纯理论探讨引向工程实践。

1.1.2 BP神经网络

经典的BP神经网络是由输入层、隐含层、输出层3个部分组成,每一层可以为一层或是多层^[15]。BP神经网络使用反向传播算法调节网络的权重。反向传播算法由2部分组成:信息的正向传播与误差的反向传播。在正向传播的过程中,输入信号从输入层经过隐含层逐层传递直到输出层,每一层神经元的状态只受上一层神经元状态的影响。如果在神经网络的输出层没有得到期望的输出,则计算输出层的实际输出值与期望值之间的误差,然后使用反向传播的方式,通过将误差信号沿原来的连接通路反向传递、更新各层神经元的权值直至达到期望值^[16]。

1.1.3 径向基神经网络

径向基神经网络是一种前向神经网络。径向基神经网络可以以任意的精度逼近任意的非线性函数,并且具有全局逼近能力,因而从根本上解决了BP网络的局部最优问题。而且其拓扑结构紧凑,结构化参数可实现分离学习,快速收敛的功能。径向基神经网络通常由隐含层、输出层2个神

神经元层组成,是一种多层前向传递型神经网络。径向基神经元通常采用的是径向基传递函数,网络结构通常包括2层神经元,第1层为隐含的径向基层,第2层为输出线性层。

1.1.4 模糊神经网络

模糊神经网络就是将模糊系统与神经网络相结合,即将神经网络模糊化,是将经典的神经网络赋予模糊输入信号与模糊化的权值。模糊神经网络的第1层与经典的神经网络相同,为输入层,作用是存储输入信号。第2层为模糊化层,模糊化输入信号。第3层为模糊规则层,设定模糊逻辑规则。第4层是模糊决策层,分类满足一定条件的量并去模糊化模糊量。第5层为输出层,输出最终的计算结果。

1.1.5 深度神经网络

深度神经网络^[17]具有提取数据深层特征的能力,它有效地解决了传统多层神经网络难以进行训练的问题。目前已经出现了很多种深度神经网络,这些神经网络的目标函数是不同的,如深度受限玻尔兹曼机、Autoencoder、Convolutional Neural Networks(CNN)。其中,Autoencoder因其具有较快的训练速度,以及较强的原有数据恢复能力等特点,常用于实际的模式识别与智能控制场景与任务中^[18]。

人工神经网络由于其独有的仿生模型结构以及高度的自适应和泛化特性等突出特征,广泛应用于智能控制领域中,已取得了许多重要的成果。该应用几乎覆盖了智能控制理论研究中的绝大多数问题,应用形式主要有机器学习、自然语言理解、计算机视觉、图像分割、系统建模和辨别等。

本文给出了使用BP神经网络进行人脸不同角度检测的实例。神经网络的输入层是8维的向量,输出层是3维的向量,有2个隐藏层,分别有10个、3个神经元。经典的BP神经网络使用的激活函数是logsig,传递函数为traingd,经典的BP神经网络与基于动量加速的BP神经网络进行对比实验。设定学习率为0.01,实验设置最大循环次数为10000,期望误差最小值err_goal设为0.001。图1与图2展示了分别使用加速BP神经网络与经典的BP神经

网络进行训练得到的收敛曲线。

在图1、图2所示的实验结果中,经典的BP神经网络在8219回合时收敛,改进BP神经网络在1888回合时收敛。改进的BP神经网络的收敛时间短于经典BP神经网络,说明了改进BP神经网络的收敛效果优于经典的BP神经网络。

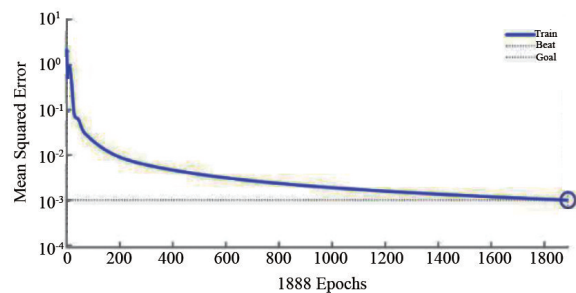


图1 加速BP神经网络的收敛曲线

Fig. 1 Convergence curve for the accelerated BP neural network

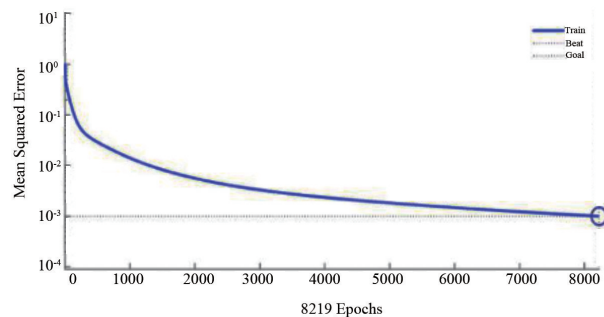


图2 经典BP神经网络的收敛曲线

Fig. 2 Convergence curve for the classical BP neural network

1.2 遗传算法

遗传算法(genetic algorithm, GA)是在1969年由Michigan大学Holland提出的,经DeJong、Goldberg等归纳总结形成一类模拟进化算法^[19]。遗传算法是使用自然界生物生存与进化的原理与机制求解极值问题的一类自适应的人工智能方法。遗传算法具有深厚的生物学进化理论基础,是一种模拟生物进化机制与过程而形成的一种过程搜索最优解的算法。遗传算法具有十分鲜明的认知学意

义,通过一种智能生成过程观点对生物智能进行模拟。遗传算法可以实现有效的并行计算,适合于任何类有表达或是无表达的非线性函数。

到目前为止,遗传算法广泛应用于智能控制、机器人自动控制等领域,被用于解决一类复杂的非线性、多维空间最优解问题。遗传算法在智能控制中的应用主要包括以下方面:基于遗传算法的智能系统设计,基于遗传算法的机器人智能控制方法设计,包括机器人路径规划、步态规划等,以及基于遗传算法的智能辅助控制器设计,基于遗传算法的多智能体控制问题。

1.3 人工免疫算法

人工免疫算法(immune algorithm, IA)一直以来是人工智能研究领域的一个重要内容^[20]。优化问题的可行解与目标函数分别对应于人工免疫算法中的抗体与抗原。人工免疫算法是通过抗体和抗原之间的亲和力描述目标函数与可行解之间的匹配程度:对于人工免疫算法,抗体之间的亲和力可以保证多样的可行解。抗体期望生存的概率可以促进抗体的遗传和变异,并且记忆细胞可以保存存留的可行解、抑制相似可行解的生成并引导算法加速搜索到全局最优解。同时,再次出现相似的问题需要求解时,人工免疫算法能够较快得到针对该问题的较优解甚至可能是最优解。

2 智能控制中群生物智能算法模型

2.1 粒子群算法

粒子群算法(particle swarm optimization, PSO)属于群智能算法的一种,是通过模拟鸟群捕食行为设计的^[21]。假设某一区域内存在一块食物,这个食物就是通常所说的最优解,而鸟群的任务就是寻找这个食物源。在整个食物搜索过程中,鸟群中的个体通过互相传递各自携带的信息,这样其他的鸟就会知道自己的位置,通过这样的协作方式,每只鸟可以判断出自己找到的是否是最优的解。这种方式同时将最优解的信息传递给鸟群中的其他个体,最后整个鸟群都能聚集在食物源周围,即求解到了最优解,从而问题得到收敛。粒子群算法通常用于

移动机器人路径规划、控制系统优化、调节控制器参数以及与其他智能算法相结合实现智能控制。

2.2 细菌觅食算法

细菌觅食优化算法(bacteria foraging optimization algorithm, BFOA)是一种全局范围的随机性的求解搜索算法,是进化算法的一员。研究人员通过模拟大肠杆菌在人体肠道内觅食时所表现出来的行为而提出了该算法^[22]。该算法主要通过3种操作,即趋向性操作、复制操作和迁徙操作迭代计算来解决并优化问题。

趋向操作是该算法的关键,它指引着细菌向食物丰富的区域移动,具有局部寻优的特点。趋向操作包括翻转和前进2种基本行为,细菌通过翻转朝着任意的方向移动一段距离,再计算翻转后的适应度函数值;然后根据适应度函数值是否得到改善,来决定细菌是否继续按当前方向前进。

复制操作体现了生物界优胜劣汰的生存机制。在细菌的觅食过程中,该算法为了优化细菌的觅食能力,会根据细菌在所有趋向周期内的适应度函数值,淘汰掉觅食能力差、复制觅食能力强的细菌,并维持种群规模不变。

3 强化学习与神经科学

3.1 强化学习

强化学习(reinforcement learning, RL)是一种智能体通过不断与环境进行交互而获取经验的自学习算法。强化学习模仿人或动物的学习过程,一直以来在智能控制领域有着广泛的应用^[23]。

马尔可夫决策过程是一个状态连续变换的过程,当前状态与下一状态之间没有关联。马尔可夫决策过程可以用状态、动作、奖励、状态转移概率和折扣因子五元组表示。为了获得在马尔可夫过程下的累计奖励最大,就有了强化学习,累积奖励可以使用状态-动作的值函数表示,于是有描述状态-动作的贝尔曼方程

$$V_{\pi}(s, a) = E_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t R_{t+p+1} | s_t = s, a_t = a \right] \quad (1)$$

其中, $V_{\pi}(s, a)$ 是强化学习中的状态与动作值函数,

γ 是折扣因子, 为常数; E_{π} 是累积奖励的期望值; s_t 是 t 时刻智能体的状态, a_t 是智能体采取的动作, R_{t+p+1} 是获得的奖励。

最优状态-动作值函数 $V^*(s, a)$ 为所在所有策略中值最大的状态-行为值函数。

$$\begin{cases} V^*(s, a) = \max_{\pi} V_{\pi}(s, a) \\ V^*(s, a) = \sum_{s' \in S} P(s, a, s') [R + \gamma \max_{a' \in A} V^*(s', a')] \end{cases} \quad (2)$$

其中, $P(s, a, s')$ 为状态转移概率, R 为获得的即时奖励, γ 为折扣因子, s 与 a 分别为智能体的当前状态与所采取的动作, s' 为下一时刻的动作。

时间差分方法是一种无模型的强化学习方法, 被用于解决系列状态决策问题。因为缺少模型, 因此该方法无法获得当前状态的全部后继状态, 智能体通过采样的方法获得当前状态的下一时刻的状态。使用时间差分的思想将状态-动作值函数改写为

$$V(s_t, a_t) = V(s_t, a_t) + \alpha [R + \gamma \max_{a_{t+1} \in A} V(s_{t+1}, a_{t+1}) - V(s_t, a_t)] \quad (3)$$

其中, α 为学习率, R 为获得的即时奖励, s_t 与 a_t 分别为智能体在 t 时刻时的状态与所采取的动作。下一时刻的状态与动作分别用 s_{t+1} 与 a_{t+1} 表示。

如果该状态-动作值函数满足采用贪心策略选择动作, 就构成了 Q 学习算法。

Q 学习算法是 Watkins 提出的一种模型无关的强化学习算法, 它具有简单、有效、无模型的特点。Q 学习是一种离线策略的强化学习, 它简单易用, 收敛速度快, 是一种十分常用的强化学习算法^[24]。定义 Q 值为

$$Q(s, a) = E[r | (s, a)] + \gamma \sum_{s'} T_a^{s'} \max_{a'} Q(s', a') \quad (4)$$

使用时间差分估计状态-动作值函数 Q , 得到

$$Q_{t+1}(s, a) = (1 - \alpha) Q_t(s, a) + \alpha [r_t + \gamma \max_{a'} Q_t(s_{t+1}, a')] \quad (5)$$

其中, α 为学习率, 取值范围是 $(0, 1)$, 学习率反映了学习的效率。

当智能体到达目标状态时, 一次迭代学习的过

程结束, 智能体回退到初始状态, 开始下一次循环直到整个学习过程结束, 即可以获得最优的 Q 值。

随着强化学习在理论和应用等方面研究的不断深入, 强化学习算法在实际的过程优化和智能控制中得到广泛应用。目前, 强化学习方法已经在智能控制、非线性控制、组合优化和调度、通信和数字信号处理、多智能体协同与对抗、模式识别、交通信号控制等智能控制相关领域取得了成功的应用。如何使得机器人在复杂场景中有效地学习是近年机器人学研究热点之一。而随着强化学习越来越多地应用在智能控制中, 基于强化学习的机器人智能控制与决策算法成为研究热点^[25, 26]。

对于经典的强化学习算法, 随着模型的状态与动作空间的不断扩大, 传统的强化学习遇到了维数灾难的困境。深度强化学习方法依靠深度学习强大的计算力为解决传统强化学习的维数灾难的困境提供了一种新的方式。深度强化学习结合了强化学习的学习与决策能力与深度学习的感知能力, 能够通过一种端到端的方式实现从感知的输入到决策输出的控制。在机器人无地图导航控制过程中, 传统的控制方法例如 PID (proportional integral derivative) 控制与视觉伺服控制需要通过复杂的数学推导建立机器人的导航控制模型, 而且为了提高模型的精度往往需要花费大量的时间修正模型。一种基于深度强化学习的端到端控制为实现无地图导航提供了一种新的方法, 并且该方法不需要复杂的动力学建模。通过移动机器人在环境中不断探索与试错, 最终自主学习到最佳的策略达到预定的目标位置。

3.2 类脑智能科学

类脑智能具有比传统人工智能更为广泛的应用领域, 因此是实现通用人工智能的重要途径。目前, 类脑智能在智能控制上的应用主要体现在智能机器人的自学习与自主决策上。由于目前几乎所有的智能机器人在运动机制上都不具备仿人的外围神经系统, 因此机器人的自适应性和灵活性与人类运动系统还具有较大差距。

4 应用分析

经典的控制方法是通过视觉等传感器感知外界环境信息,然后对信息进行处理,将处理后得到的讯息传入控制系统,控制系统控制机器人产生动作作用于环境中(图3)。经典的控制方法无法通过学习去适应未知的环境,因此效率不高。生物智能算法为经典控制方法性能的提高提供了一种全新的思路,而生物智能算法主要在感知信息处理阶段或控制阶段与经典控制方法结合。综上所述,生物智能算法与经典控制方法相结合主要有2个研究方向:图像处理与智能辅助控制方法。

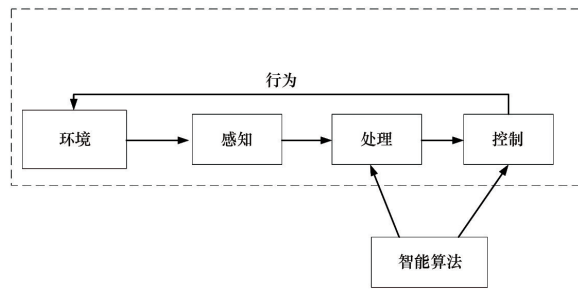


图3 智能辅助控制的基本框架
Fig. 3 Basic framework for the intelligent aided control

4.1 基于强化学习调节增益的视觉伺服模型

基于图像的视觉伺服(image-based visual servo, IBVS)^[27]是一种使用当前特征点与目标特征点之间误差的闭环反馈控制系统。该方法首先通过机器人感知外界图像信息,然后基于当前位置与目标位置的误差反馈控制机器人的运动。但是经典的视觉伺服控制的增益是固定的,因此效率不高,可以用强化学习的方法训练增益值,使得旋翼无人机可以在不同的环境中自适应选择增益值。以旋翼无人机为研究对象,图4展示了基于强化学习调节增益的视觉伺服模型的方法框架。在图4中,无人机通过视觉感知图像信息,然后提取图像的特征点,视觉伺服控制系统输出线速度与角速度进而控制旋翼无人机的运动。由于强化学习的学习率是固定的,因此使用模糊控制调节学习率。

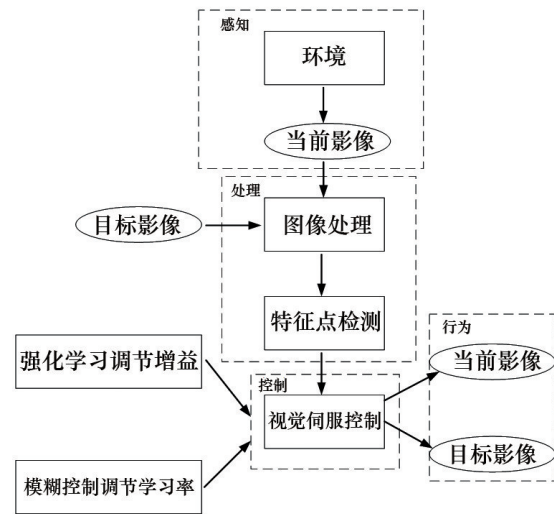


图4 基于强化学习的视觉伺服控制
Fig. 4 Visual servo control based on reinforcement learning

4.2 基于智能算法的PID控制器参数调节模型

使用神经网络调节经典PID控制器增益^[28]的方法:环境给定一个输入信息,输入信息与输出信息的差值通过PID控制器输出得到输出信息和机器人的行为,同时将输入信息与输出信息输入神经网络中训练神经网络。神经网络调节PID控制器的增益值,设计了一组实验展示这个过程。

二阶延时控制系统为

$$G(s) = \frac{2.0}{s^2 + 1.6s + 2.0} e^{-0.3s} \quad (6)$$

其中, $G(s)$ 为二阶延迟系统, s 为输入。

控制系统的系统输入为: $\text{rin}(x)=1$ 。使用PID控制器对该控制系统进行系统逼近,当输出等于输入的时候,系统逼近完成。实验对分别使用不同的方法对PID控制器参数进行整定及探索。分别使用人群优化算法(SOA)、粒子群算法、遗传算法、改进的神经网络算法(GBP)对PID控制器进行参数整定,实验结果如图5与图6所示。图5展示了分别使用这几种方法的信号输入与输出曲线图,GBP方法在1.5 s时收敛,PSO方法在3 s时收敛,SOA方法在2 s以后收敛,GA方法在3 s以后收敛,由此可以看出GBP方法具有相对较快的收敛速度。图6展示了误差变化,误差由使用输入减去输出得到,如果误差为0,则方法收敛,系统逼近完成。

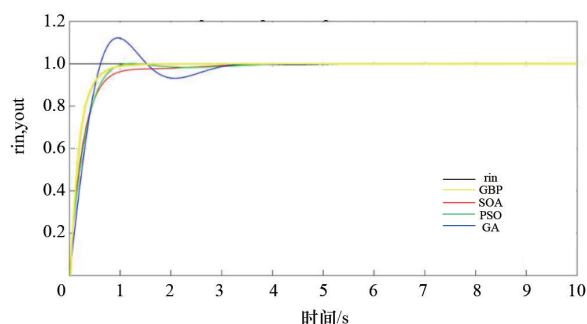


图5 4种方法的输出信号曲线

Fig. 5 Output signal curves for the four methods

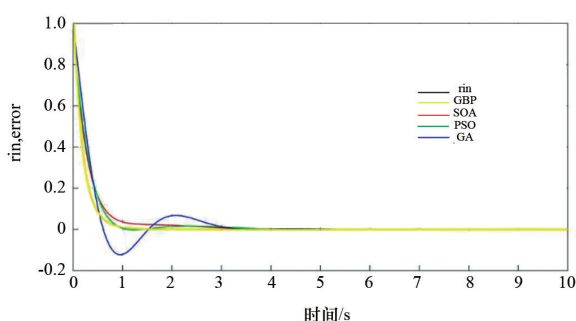


图6 4种方法的误差变化曲线

Fig. 6 Curve of error for the four methods

5 结论

总结了智能控制领域常用的7种生物智能算法,分析了生物智能算法尤其是强化学习与传统的控制方法相结合的智能控制方法。生物智能算法与传统控制方法的结合克服了传统控制方法固定增益等问题,极大提高了传统方法的控制性能。随着人工智能的不断发展,特别是强化学习和脑科学的不断发展,强化学习和脑科学为智能控制带来了新的研究方法和思路。目前,强化学习和脑科学在智能控制领域有许多应用案例,但仍存在许多机遇和挑战。

参考文献(References)

[1] Åström K J, McAvoy T J. Intelligent control[J]. Journal of Process Control, 1992, 2(3): 115-127.

- [2] Das S K. Modeling intelligent decision-making command and control agents: An application to air defense[J]. IEEE Intelligent Systems, 2014, 29(5): 22-29.
- [3] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [4] Huang Y N, Zhang A J, Hu M Y. Single Neuron PID control based on expert experiences for temperature difference control system of a digester[J]. Paper and Biomaterials, 2017, 2(1): 52-58.
- [5] Le T T, Huff T B, Cheng J X. Coherent anti-stokes raman scattering imaging of lipids in cancer metastasis[J]. BMC Cancer, 2009, 9(1): 42-42.
- [6] Ma J, Ni S, Xie W, et al. Deep auto-encoder observer multiple-model fast aircraft actuator fault diagnosis algorithm[J]. International Journal of Control Automation & Systems, 2017, 15(11): 1-10.
- [7] Liang X, Li Y, Dai C, et al. Reinforcement learning-based NOMA power allocation in the presence of smart jamming[J]. IEEE Transactions on Vehicular Technology, 2018, 67(4): 3377-3389.
- [8] Shi H, Li X, Hwang K S, et al. Decoupled visual servoing with fuzzy Q-Learning[J]. IEEE Transactions on Industrial Informatics, 2018, 14(1): 241-252.
- [9] Peng J I, Song A, Changcheng W U, et al. Human-robot-robot-cooperation based uncalibrated visual servoing control for mobile robotic manipulators without joint-state feedback[J]. Robot, 2017, 39(2): 197-204.
- [10] Zhao F, Yi Z, Wang G, et al. A brain-inspired decision making model based on top-down biasing of prefrontal cortex to basal ganglia and its application in autonomous UAV explorations[J]. Cognitive Computation, 2017, 10(6): 1-11.
- [11] He W, Chen Y, Yin Z. Adaptive neural network control of an uncertain robot with full-state constraints[J]. IEEE Transactions on Cybernetics, 2017, 46(3): 620-629.
- [12] Long W, Li X, Liang G, et al. A new convolutional neural network-based data-driven fault diagnosis method [J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5990-5998.
- [13] Merayo N, Juárez D, Aguado J C, et al. PID controller based on a self-adaptive neural network to ensure QoS bandwidth requirements in passive optical networks[J]. IEEE/OSA Journal of Optical Communications & Networking, 2017, 9(5): 433-445.
- [14] Block H D. The perceptron: A model for brain functioning[J]. Reviews of Modern Physics, 1962, 35(1): 123-134.
- [15] Fan W, Li Z S. Prediction of the creep of piezoelectric ceramic based on BP neural network optimized by genet-

- ic algorithm[J]. *Acta Metrologica Sinica*, 2017, 38(4): 429–434.
- [16] Kuang Y, Singh R, Singh S, et al. A novel macroeconomic forecasting model based on revised multimedia assisted BP neural network model and ant colony algorithm [J]. *Multimedia Tools & Applications*, 2017, 76(18): 1–22.
- [17] Qian Y, Bi M, Tian T, et al. Very deep convolutional neural networks for noise robust speech recognition[J]. *IEEE/ACM Transactions on Audio Speech & Language Processing*, 2017, 24(12): 2263–2276.
- [18] Wang Y B, You Z H, Li X, et al. Predicting protein–protein interactions from protein sequences by a stacked sparse autoencoder deep neural network[J]. *Molecular Biosystems*, 2017, 13(7): 1336–1344.
- [19] Friedrich T, Kötzing T, Krejca M, et al. The compact genetic algorithm is efficient under extreme Gaussian noise [J]. *IEEE Transactions on Evolutionary Computation*, 2017, 21(3): 477–490.
- [20] Gosciniak I. Discussion on semi-immune algorithm behaviour based on fractal analysis[J]. *Soft Computing*, 2017, 21(14): 3945–3956.
- [21] Bonyadi M R, Michalewicz Z. Particle swarm optimization for single objective continuous space problems: A review[J]. *Evolutionary Computation*, 2017, 25(1): 1–54.
- [22] Majumder A, Laha D. Bacteria foraging optimization algorithm for robotic cell scheduling problem[J]. *Materials Today Proceedings*, 2017, 4(2): 2129–2136.
- [23] Liang X, Li Y, Dai C, et al. Reinforcement learning-based NOMA power allocation in the presence of smart jamming[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(4): 3377–3389.
- [24] Deng Y, Bao F, Kong Y, et al. Deep direct reinforcement learning for financial signal representation and trading[J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2017, 28(3): 653–664.
- [25] Qiang Z, Yuan Q, Chao Z, et al. Missing data reconstruction in remote sensing image with a unified spatial-temporal-spectral deep convolutional neural network[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2018, 56(8): 1–15.
- [26] Zhao D, Liu D, Lewis F L, et al. Special issue on deep reinforcement learning and adaptive dynamic programming[J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2018, 29(6): 2038–2041.
- [27] Serra P, Cunha R, Hamel T, et al. Landing of a quadrotor on a moving target using dynamic image-based visual servo control[J]. *IEEE Transactions on Robotics*, 2017, 32(6): 1524–1535.
- [28] Cong S, Liang Y. PID-like neural network nonlinear adaptive control for uncertain multivariable motion control systems[J]. *IEEE Transactions on Industrial Electronics*, 2009, 56(10): 3872–3879.

An overview and the practical application of the biological intelligence algorithms in intelligence control

XIE Wenguang, YAN Fang, WANG Kenian

1. Key Laboratory of Civil Aircraft Airworthiness Technology, Civil Aviation Administration of China, Tianjin 300300, China
2. College of Airworthiness, Civil Aviation University of China, Tianjin 300300, China

Abstract The classical control methods have shortcoming, such as its fixed parameters, with a limited control effect, and the Biological Intelligence Algorithm provides a new way to break the bottleneck of the classical control methods because of its adaptive and learning mechanism. With the improvement of the theory of the reinforcement learning and the deep learning, the performance of the Biological Intelligence Algorithm is greatly improved. This paper reviews seven kinds of intelligent algorithms commonly used in intelligence control, and the application examples of combining the classical automatic control methods and the intelligent algorithms, especially, the reinforcement learning. The development status and the future development trend of the intelligent control based on the reinforcement learning, the deep learning and the Brain-inspired Intelligence Technology in recent years are discussed. The purpose of this paper is to emphasize a new idea of combining the intelligent algorithms with the classical control methods, and provide some new ideas and practical examples for the intelligent control.

Keywords intelligent control; biological intelligence; automatic control; reinforcement learning ●



(责任编辑 徐丽娇)