

深度强化学习为智能航空发动机控制赋能

Deep Reinforcement Learning Empowers Intelligent Aero Engine Control

■ 唐炜 李鹏源 董建华 朱建铭 / 西北工业大学

深度强化学习可以通过交互式训练实现控制策略的智能优化，使航空发动机更加智能地适应不同飞行环境和任务需求，为提升航空发动机性能、降低耗油率、增强飞机安全性等方面带来了新的机遇，有望成为实现智能航空发动机的有效途径之一。

航空发动机是一个高度复杂的系统，涉及时变性、非线性、耦合性和多变量等问题，这些特点使得其控制成为一项极具挑战的任务，主要包括控制策略优化、燃油效率提升、故障检测与诊断、环境自适应等方面。传统控制方法主要是在特定飞行状态下基于线性模型设计控制器，但这种方法对于在广泛飞行包线内运行的发动机效果不佳。此外，由于发动机部件长期运行导致的性能下降，传统控制器将难以适应新的控制环境。随着人工智能技术的发展，特别是深度强化学习的进步，国内外研究人员开始将其与航空发动机控制结合。深度强化学习旨在探索智能体如何在复杂和不确定的环境中学习到最优策略。通过与发动机不断交互，智能控制器能够学习并优化控制策略，适应发动机工作状态的变化，从而表现出优越的环境适应性。

航空发动机控制的发展

航空发动机的构型及控制系统自20世纪以来经历了多次变革，如图1所示。诞生于20世纪中期的燃气涡轮航空发动机构型较为简单，采用液压-机械式控制系统，这要求飞

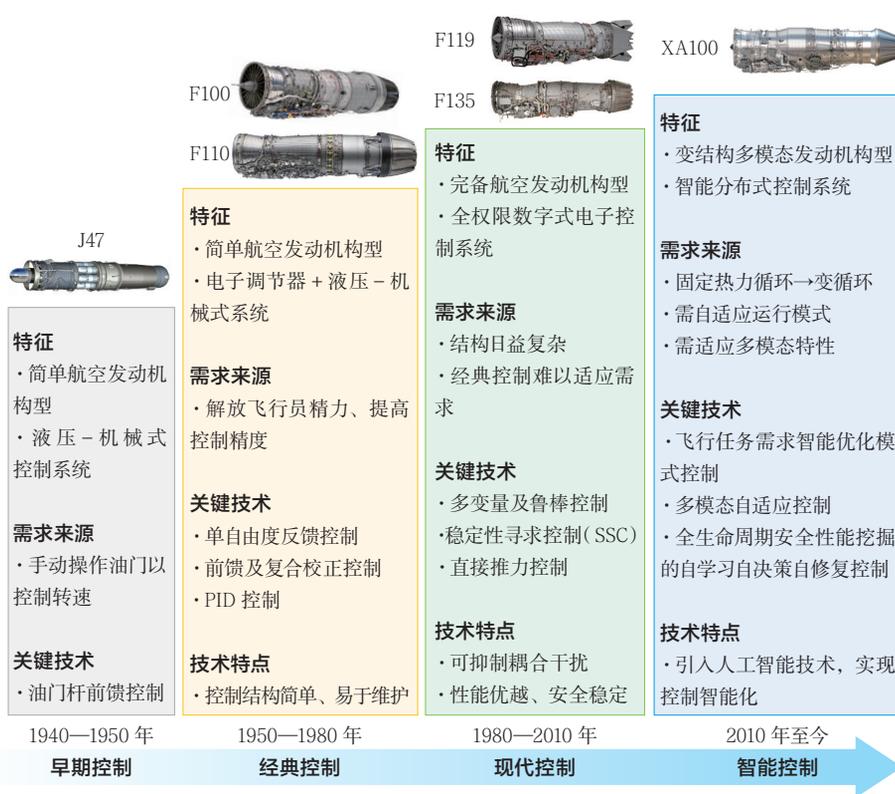


图1 发动机控制系统的发展历程

行员根据飞行高度和速度手动操作油门以调整燃油流量，将发动机转速控制在一定范围内保持不变。随后为解放飞行员的精力并提高发动机控制精度，1950—1980年，以比例-积分-微分(PID)为代表的单变量反馈控制器被应用且代替了早期的油门杆前馈控制，形成了结合电子调节器的液压机械式控制系统。但随着飞机性能需求的提升，发动

机的结构逐渐复杂化，可调部件从早期的1个增加至10个以上，系统的耦合性也逐渐变强，这也导致面向线性系统的经典控制方法已难以适应发动机高性能的控制要求，从而促进了现代控制理论技术的应用，线性二次型调节器(LQR)、线性二次型高斯/回路传输恢复(LQG/LTR)及多变量鲁棒控制等的各种控制方法在1980—2010年间被广泛研

究和应用。美国针对航空发动机多变量控制开展过许多研究，其中包括利用多变量控制解决F-35战斗机短距起飞垂直降落的特殊飞行任务、为确保控制精度和发动机的最佳性能对F135发动机进行多变量控制等。尽管传统控制方法能在特定飞行状态下稳定控制航空发动机，但面对发动机来流条件的变化，其控制系统性能可能会下降，无法实现环境自适应，因此智能航空发动机控制的概念被也逐渐被提出^[1]。

深度强化学习基本原理及无模型控制

深度强化学习融合了深度学习的高效特征提取能力和强化学习的先进策略学习机制，提供了一种强大的自主学习框架，能够在无需人工干预的情况下，通过与环境的动态交互，自我学习以完成特定任务。这一技术的显著特点是其知识构建和学习过程完全依赖原始输入信号，不需要预先的人工编码。因此，深度强化学习形成了一种端到端的感知控制系统，这种系统具有广泛的

应用潜力。

在多种领域中，深度强化学习已经展现出强大的适用性和灵活性，其训练过程可以概述为以下几个步骤：智能体感知目前所处状态并根据当前策略选择动作；在环境中执行该动作从而转移到新的状态并获得相应奖励；将当前状态、动作、奖励和新状态作为经验元组存入经验回放池中；从经验回放池抽取一批经验用于学习最优策略，使智能体更倾向于选择奖励较高的动作。循环以上过程，如图2所示。通过以上的训练过程，智能体会逐步优化其决策策略最终确定出实现既定目标的最优策略。

自动控制领域当前面临的主要挑战源于被控系统的复杂性，尤其是当系统表现出显著的非线性、时变性和不确定性时，这些特性使得建立高精度数学模型异常困难。PID控制算法虽不依赖精确的数学模型，但在处理复杂的非线性或时变系统时，其性能有限。而像LQR、模型预测控制（MPC）和自适应控制等算法，则通常在高度依赖精确系统

模型的基础上设计，面对模型不准确或不完整的实际情况时，控制效果可能显著下降。与传统控制算法不同，深度强化学习可以不依赖系统的数学模型，通过从系统中收集的数据来逐步学习和优化控制策略，实现最优控制。显然深度强化学习不仅给控制领域注入了新的活力，还为复杂系统控制策略的探索开辟了新的途径。

深度强化学习的应用探索

航空发动机系统常表现出噪声、不确定性和非线性，而深度强化学习因其面对复杂系统时的显著优势而被推向了航空发动机智能控制的前沿。

在21世纪初，英国学者马克·豪威尔（Mark Howell）率先开展了将强化学习技术融入发动机控制系统的研究。其工作聚焦于PID控制算法，该算法以结构简单、适用范围广而著称，但其参数调整过程往往繁琐且依赖于人工操作。因此他提出了一种基于连续动作空间的强化学习自动机，用以在线自动调整发动机怠速转速控制的PID参数，如图3所示。此方法相较于传统的人工调参而言大幅降低了人力成本，同时显著提高了调参效率及发动机的动态性能。这不仅为深度强化学习在实际控制系统中的应用奠定了实证基础，而且对发动机控制技术的发展做出了实质性贡献。

自2018年以来，南京航空航天大学也开展了一系列相关研究，郑前钢等率先使用深度强化学习中的无模型Q学习对航空发动机控制进行了仿真研究。该项研究中将燃油流量作为控制动作，将当前和过去时刻的燃油流量、风机和压气机的转子转速和喘振裕度、高压涡轮进

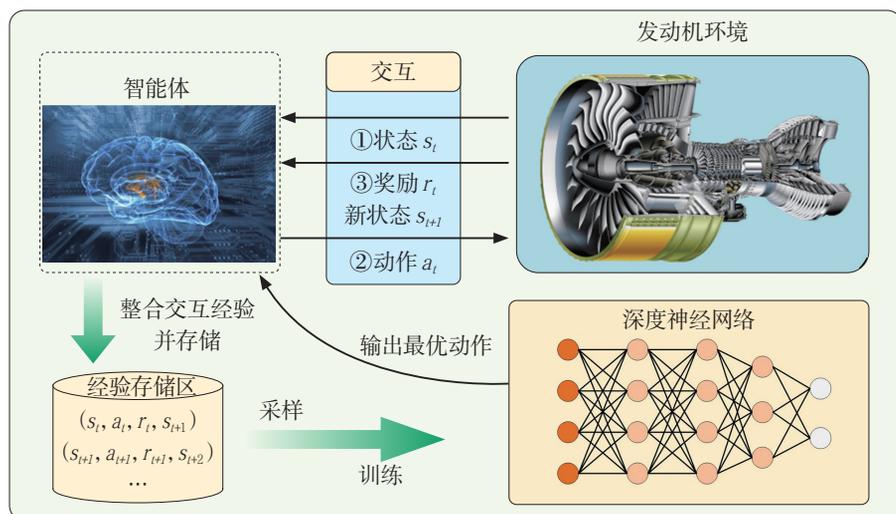


图2 深度强化学习的训练过程示意

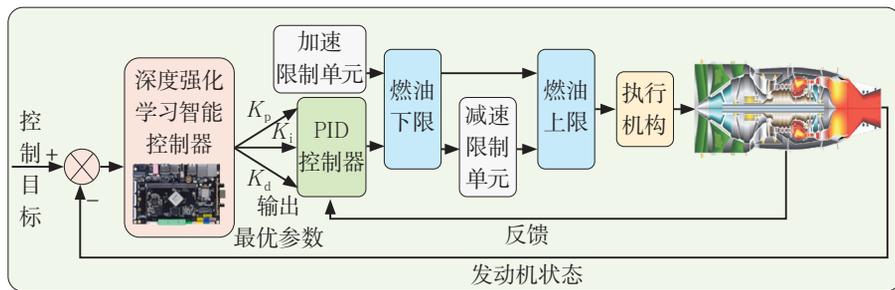


图3 基于深度强化学习的PID参数自整定

口温度等一系列状态作为深度神经网络的输入，从而映射为动作价值函数 $Q(s, a)$ ，在经过大量的训练后，该控制器便能够智能地选择燃油流量的最优值，完成对发动机转子转速的控制。此仿真研究的结果显示，深度Q学习控制算法在动态响应能力上相比传统的PID控制具有显著优势。

但深度Q学习算法通常用来解决离散动作空间的问题，因此随着深度强化学习技术的进步，面向连续域问题的深度确定性策略梯度(DDPG)算法应运而生。然而深度神经网络中普遍存在的近似误差问题会导致控制时的稳态误差难以收敛至零。针对这一问题，黄金泉等在使用DDPG设计控制器时提出了一种新颖的解决方案^[2]。他们引入了

互补积分器来校正深度神经网络的输出，有效减少稳态误差，从而提高了控制器的渐进性能。此外，还加入了动量项来限制系统输入，以抑制早期学习和探索过程中的超限现象。仿真结果表明，这种方法相比于鲁棒控制能显著提升加速过程的效率，如图4所示。

在此之后，郑前钢等又使用深度强化学习进行了全飞行包线下的加速度控制器设计，以解决航空发动机加速过程的多维约束优化问题^[3]。由于不同飞行高度和马赫数条件下，发动机进气流的总温度和压强等特性的变化，发动机工作状态存在显著差异。仅考虑地面条件下的加速性能无法确保发动机在全飞行包线范围内的正常运行。单个神经网络难以捕捉全飞行包线下的发动机状态，

而传统的相似转换法又无法保证控制的高精度。因此该团队提出了一种新思路：通过聚类方法将飞行包线划分为若干区域，每个区域内保持较高的状态相似性，然后在每个聚类区域内应用双延迟深度确定性策略梯度(TD3)算法进行控制律的训练。通过线性插值法，将控制器扩展至整个飞行包线。最终仿真结果也表明该方法相比于PID控制加速时间显著缩短，最大加速时间减少了48.33%。此外，由于采用了深度强化学习中改进后的TD3算法，有效避免了改进前DDPG算法可能引起的Q值过高估计问题，从而加快了训练过程的稳定性和神经网络的收敛速度。

北京航空航天大学朱美印等也曾在2022年开展过基于深度强化学习的涡扇发动机控制器研究^[4]。虽然该研究的大体框架仍是PID参数值的智能调节，但该项研究考虑到了发动机的性能退化。发动机部件老化后常会引起响应性能的下降。该团队通过降低压气机效率来模拟发动机的退化，并使控制系统能够智能地适应这种退化，从而保持住较好的控制性能，显示出比增益调度

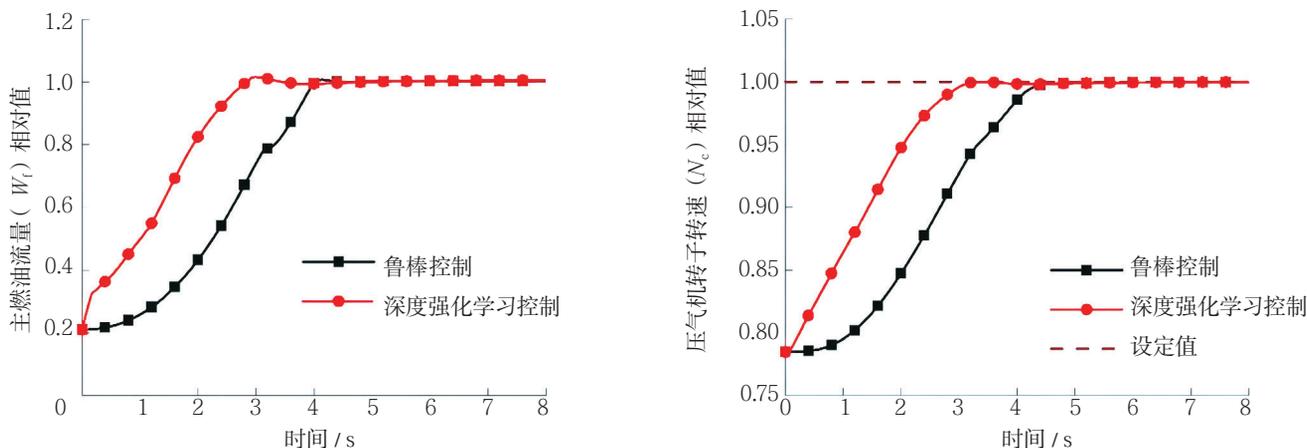


图4 深度强化学习和鲁棒控制的控制效果对比^[2]

控制（GSC）更强的适应能力。

未来展望

目前，业界虽已结合深度强化学习在航空发动机控制方向展开了许多研究，相比传统控制也有了显著的性能提升，有望为未来智能航空发动机的实现提供有效途径，但大部分研究还停留在仿真阶段，应用于真实发动机的案例较少，目前限制相关技术应用的原因有以下几点。

第一个问题是深度强化学习的过拟合天性。过拟合是所有类型的机器学习模型都可能遇到的问题，主要表现为对训练环境的过度优化：模型可能在特定的训练环境下表现出色，但当环境发生变化时，其性能会显著下降。这意味着模型过度适应了训练环境的特定特征，而没有学习到更加普适的策略。过拟合的模型在训练数据上表现可能非常好，但在新的、未见过的数据或全新的环境中表现通常会差很多，因为它没有很好地泛化到新的情况。反映到航空发动机控制上即为：在不进行在线学习的情况下，基于深度强化学习的智能控制器可能只在刚出厂阶段对发动机进行最优控制，随着发动机的老化或故障及维修以后，真实发动机环境可能相比最初的训练环境产生了一定变化，此时控制器已无法进行最优控制而是变为次优控制，最极端的情况下控制器可能因为发动机环境的改变导致面对全新的环境时无所适从，引起控制发散。

但是过拟合的问题并非无法解决，因训练数据单一化导致过度适应特定环境的特定特征的过拟合，往往可以通过使用更多数量和更多

样化的训练样本来克服，但是也由此引出深度强化学习的第二个问题：样本效率低下。深度强化学习本就需要大量的训练样本以学习有效策略，其中样本源自与发动机环境的交互，若通过用更多数量以及不同性能状态下的样本模拟发动机环境在后续使用过程中可能经历的变化，进而解决过拟合问题，则又会增加获取样本的难度、时间和成本等。

以上是先离线训练再进行部署时会遇到的问题，能否通过在线训练以克服智能控制器对发动机环境变化后产生的不适应？这样又暴露出深度强化学习的第三个问题：无法预知训练过程是否发生大的波动及何时收敛，并且波动程度和收敛时长还与超参数选择有关。对于在线训练而言，样本效率低、训练过程的波动性和收敛速度未知等因素将有可能影响发动机运行稳定性和安全性，这也是目前非仿真的在线训练研究案例只停留在为传统控制算法进行参数寻优的原因，不完全将控制任务交给深度强化学习能在

一定程度上保证安全性。

因此，针对上述过拟合、样本效率低、交互成本高等问题，未来应该将交互低成本化、便捷化视作研究重点之一。若能得到一种便于发动机与深度强化学习控制器交互的方式，从而使训练样本的获得更为容易，则大概率能推进该项技术的实际应用，而在此之前王潘、刘永泉等已分析过大数据、数字孪生等技术对于航空发动机建模、预测等的方向的助力^[5-6]。因此，针对以上问题，笔者团队尝试使用航空发动机虚拟学习环境技术以克服与真实发动机交互带来的问题^[7-9]。深度强化学习的本质在于从交互训练中试错，为了避免产生高昂的试错成本，可以凭借深度神经网络较强的非线性拟合能力，基于真实的试验数据建立起航空发动机的虚拟学习环境，从而让深度强化学习控制器只需与虚拟学习环境进行交互（见图5），这也为深度强化学习应用于发动机智能控制提供了一条低风险、高效率的实施路径。

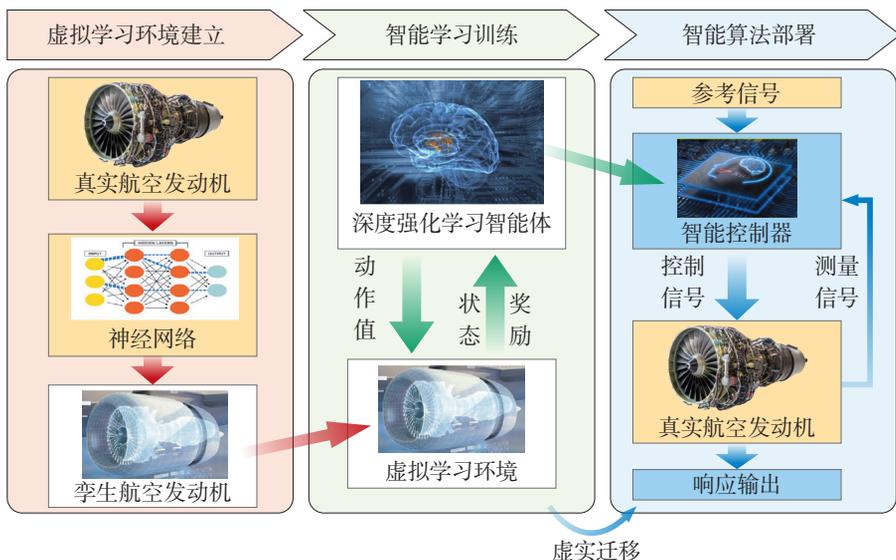


图5 航空发动机虚拟自学习控制示意

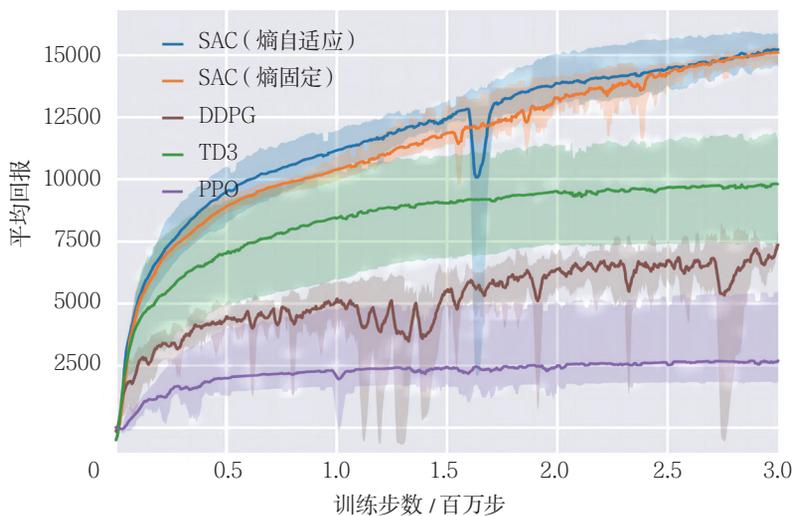


图6 不同深度强化学习算法的学习效率对比^[10]

另外，对于训练过程波动性、收敛时间较长、受超参数影响等问题，首先可以明确的一点是采用离线训练时无须过于关注训练过程的波动性和收敛速度，只需将最终学习到的最优策略进行部署即可；若未来有在线训练的需求，同样可以从超参数自整定、算法择优、算法创新等研究角度来克服，因为不同算法对于不同环境的学习效率也不尽相同，有研究表明，随着深度强化学习算法的不断革新，其学习效率也在上升，如SAC、TD3等算法相比于提出较早的DDPG等算法学习效率逐步升高（见图6）。特别是SAC算法中提到的平衡探索与利用的全新机制，不再需要在算法中主动添加噪声来探索环境，极大地减少了需要调整的超参数，且自身算法对超参数的变化也不敏感，相比于其他算法表现出显著的优势。可见算法的创新也将为在线训练的部署奠定基础。

结束语

控制技术的革新能推动航空发动机的构型升级及性能提升，因此智能

控制成为了智能航空发动机研究的一个重要分支。而深度强化学习凭借其强大的非线性拟合、特征提取、策略学习等能力逐渐展现出了优异的控制性能，尽管目前基于安全性的考虑，影响了其直接的应用，但是短期内完全可以将其融入常规的控制方法中，用于各种控制的参数自寻优。随着其算法的不断发展和虚拟学习环境的逐渐成熟，未来有望成为推动智能航空发动机发展的关键技术。

航空动力

（唐炜，西北工业大学，副教授，主要从事航空发动机智能控制等方向研究）

参考文献

[1] LV C, CHANG J, BAO W, et al. Recent research progress on airbreathing aero-engine control algorithm[J]. Propulsion and Power Research, 2022, 11(1):1-57.
 [2] GAO W, ZHOU X, PAN M, et al. Acceleration control strategy for aero-engines based on model-

free deep reinforcement learning method[J]. Aerospace Science and Technology, 2022, 120:107248.
 [3] FANG J, ZHENG Q, CAI C, et al. Deep reinforcement learning method for turbofan engine acceleration optimization problem within full flight envelope[J]. Aerospace Science and Technology, 2023, 136:108228.
 [4] MIAO K, WANG X, ZHU M, et al. Transient controller design based on reinforcement learning for a turbofan engine with actuator dynamics[J]. Symmetry, 2022, 14(4):684.
 [5] 王潘, 刘魁. 大数据技术在航空发动机中的应用[J]. 航空动力, 2018(1):48-51.
 [6] 刘永泉, 黎旭, 任文成, 等. 数字孪生助力航空发动机跨越发展[J]. 航空动力, 2021(2):24-29.
 [7] ZHU J, TANG W, DONG J. Design of intelligent controller for aero-engine based on TD3 algorithm[J]. Information Technology and Control, 2023, 52(4):1010-1024.
 [8] ZHU J, TANG W, DONG J, et al. A virtual reinforcement learning method for aero-engine intelligent control[C]. 2023 8th International Conference on Automation, Control and Robotics Engineering (CACRE). IEEE, 2023: 138-143.
 [9] 董建华, 朱建铭, 唐炜, 等. 航空发动机虚拟自学习控制方法研究[J]. 航空工程进展, 2023, 14(6):81-90.
 [10] HAARNOJA T, ZHOU A, HARTIK K, et al. Soft actor-critic algorithms and applications[J]. arXiv preprint, arXiv:1812.05905, 2018.