

文章编号: 1674-8190(2023)06-081-10

航空发动机虚拟自学习控制方法研究

董建华¹, 朱建铭¹, 黎瀚涛², 刘文烁¹, 唐炜¹

(1. 西北工业大学 自动化学院, 西安 710129)

(2. 南昌航空大学 测试与光电工程学院, 南昌 330063)

摘要: 随着人工智能技术的发展, 智能航空发动机逐渐成为当今航空领域研究的热点。传统的航空发动机控制对发动机模型的依赖性过强, 而基于发动机气热动力学公式的机理建模会引入较大的建模误差, 给控制器设计带来困难。对此, 提出一种基于强化学习的航空发动机控制虚拟自学习方法, 首先利用航空发动机的试验数据通过LSTM神经网络建立虚拟学习环境, 然后采用深度强化学习TD3算法, 在虚拟环境中训练智能控制器, 最后采用JT9D发动机模型验证智能控制器的性能。结果表明: 相比于传统PID控制, 智能控制器产生的超调量更小, 调节时间更短。

关键词: 航空发动机; 智能控制; 强化学习; LSTM神经网络; TD3算法

中图分类号: V233.7

文献标识码: A

DOI: 10.16615/j.cnki.1674-8190.2023.06.09

Research on virtual self-learning control method for aero-engine

DONG Jianhua¹, ZHU Jianming¹, LI Hantao², LIU Wenshuo¹, TANG Wei¹

(1. School of Automation, Northwestern Polytechnical University, Xi'an 710129, China)

(2. School of Measuring and Optical Engineering, Nanchang Hangkong University, Nanchang 330063, China)

Abstract: With the development of artificial intelligence technology, intelligent aircraft engines have gradually become a hot spot in the field of aviation today. Traditional aero-engine control heavily relies on the engine model, and the theoretical modeling approach based on aerothermodynamic formula introduces modeling error that may degrade the performance of controller. In this paper a virtual self-learning control method for aero-engine intelligent controller design is proposed. Firstly, a virtual environment is established from the testing data of the aero-engine via LSTM neural network. Secondly, the reinforcement learning algorithm based on TD3 is employed for intelligent controller training in the virtual environment. Finally, the JT9D aero-engine model is utilized for controller performance evaluation. The simulation comparisons between intelligent controller and traditional PID control show that the intelligent controller has remarkable performance due to the less overshoot and shorter setting time.

Key words: aero-engine; intelligent control; reinforcement learning; LSTM (long short-term memory) neural network; TD3 algorithm

收稿日期: 2022-11-04; 修回日期: 2023-02-14

基金项目: 先进航空动力创新工作站(依托中国航空发动机研究院设立)资助(HKCX2020-02-019)

通信作者: 唐炜, tangwei@nwpu.edu.cn

引用格式: 董建华, 朱建铭, 黎瀚涛, 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.

DONG Jianhua, ZHU Jianming, LI Hantao, et al. Research on virtual self-learning control method for aero-engine[J]. Advances in Aeronautical Science and Engineering, 2023, 14(6): 81-90. (in Chinese)

0 引言

航空发动机被誉为是现代工业“皇冠上的明珠”,其发展水平是国家科技、军事实力和综合国力的重要象征。随着航空技术的不断发展,对飞机性能的需求越来越高,航空发动机结构越来越复杂,传统的发动机控制方案已经无法满足其日益增长的控制性能需求,与人工智能相结合的智能控制技术已经成为热门研究领域。

传统的PID控制^[1-2]是基于某一特定飞行状态下的发动机线性模型进行控制器设计。在实际应用中需要对飞行包线进行划分,并对每个区域单独设计控制器,最终通过控制器切换的方式实现对发动机全包线范围的控制,增加了控制系统设计的复杂程度,当发动机部件经过长时间的运行而性能下降时,传统的PID控制器不能很好地适应新的控制环境。为此,国内外提出了很多改进方法,包括调度增益控制^[3-4]和自适应控制^[5-6]。但在实际应用中发现,调度增益控制不适用于系统参数快速变化的非线性发动机系统,局部调试最优的线性性能不能确保全局性能最优;自适应控制系统很难处理非线性、变结构或大延迟的控制问题,无法保证模型的收敛性和系统稳定性。

21世纪提出的新概念智能航空发动机,将人工智能技术与发动机控制相结合。陈勇等^[7]设计了基于模糊误差判断的PID控制,验证了模糊控制在发动机控制中应用的可能性;薛红阳等^[8]成功将神经网络应用于发动机自适应控制,所设计的控制器对外界环境的变化具有极好的适应性,对参数摄动也具有很好的鲁棒性。然而这两种控制方法本身也具有一定的局限性,模糊控制中启发式规则的局部限制,缺乏自适应性;神经网络无法充分利用专家知识,存在样本过拟合、参数优化算法容易陷入局部极值等问题。

随着深度强化学习理论的发展与完善,越来越多的人将其与发动机智能控制相结合,并在实际应用中取得了一些进展。Gao Wenbo等^[9]通过利用Actor-Critic网络进行控制器设计,显著加快了航空发动机的加速过程;G. Waxenegger-Wilfing等^[10]利用深度强化学习技术对液体火箭发动机进行控制,仿真结果表明,学习到的策略能够在不同的稳态工作点上适应系统的参数变化并进行控制,与传统的PID控制器相比,具有更好的性能。

但是上述控制器设计方法都离不开航空发动机模型,模型的精确程度决定了最终的实际控制效果。传统的发动机建模过程^[11-12]不仅需要求解由大量热力学公式构成的共同工作方程,而且为了解决发动机部件级特性数据产生的误差,通常要进行大量的试凑来确定修正系数,整个过程复杂繁琐且耗费周期长。基于模型的深度强化学习训练过程耗时长,对模型依赖程度高,对此,Zheng Qiang等^[13]提出了一种无模型、可在线执行的Q学习智能算法,在满足发动机所有极限的前提下,与传统PID控制相比,新的控制方法加速时间缩短了1.525 s。但Q学习方法只能处理离散的动作问题,不适用于连续动作空间探索的强化学习,在此基础上改进提出的DDPG算法虽然弥补了这一缺陷,却存在Q值过估计和对值空间中的窄峰估计过拟合的问题。

本文为应对航空发动机传统控制的不足以及智能控制在应用中面临的挑战,提出一种基于深度强化学习的航空发动机控制虚拟自学习方法。该方法采用DDPG算法的进阶版本——双延迟深度确定性策略梯度(TD3)算法,并结合虚拟环境开展交互式训练,实现航空发动机智能控制。

1 研究方案

强化学习(Reinforcement Learning)研究的问题是一个智能体如何在一个复杂不确定的环境中学习到最优策略。在智能航空发动机控制过程中,智能控制器可以看作智能体,航空发动机的状态可以看作环境,控制器根据发动机的状态实时调整自己的输出,使发动机在较短的时间内迅速、稳定地达到给定状态,学习到最优控制策略,从而实现发动机的智能控制。强化学习是一种试错学习,智能控制器只需要在与发动机不断交互的过程中便能学习到最优的控制策略,而无需对发动机的模型进行参数辨识或数学建模,有效避免了复杂的建模过程并消除了发动机模型精度对控制方案带来的影响。但在实际应用过程中,由于发动机结构精密,价格昂贵,不适宜用来进行成千上万次的试错练习,为了节约成本、减少安全隐患和提高研发效率,本文提出通过建立虚拟仿真环境来完成智能控制器虚拟自学习的方法。

本文的研究内容分三部分进行(如图1所示):

一是虚拟仿真环境的搭建,将采集得到的真实航空发动机试车数据,通过LSTM循环时间神经网络进行训练学习,用来搭建深度强化学习的虚拟学习环境;二是智能控制器设计,通过设计深度强化学习TD3算法,在虚拟仿真环境中完成智能控制器的交互训练,在较短的时间内学习到最优控制策略;三是实验结果测试,结合JT9D双轴转子涡扇航空发动机模型验证智能控制器性能。采集JT9D发动机的燃油流量和高压转子转速的仿真数据作为发动机试车数据,希望设计智能控制器实现发动机转速对给定参考指令的跟踪控制。

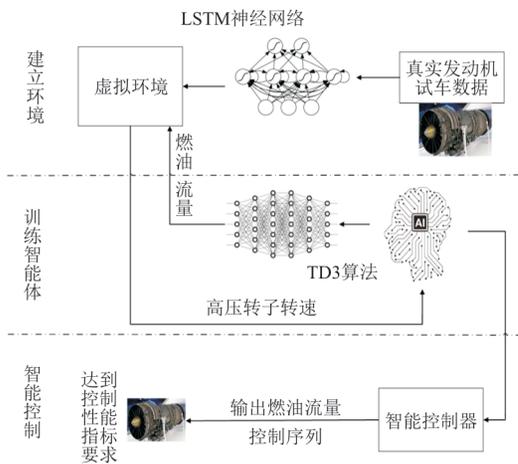


图1 虚拟自学习方法基本框架
Fig. 1 The basic framework of virtual self-learning method

2 虚拟环境的建立

本节主要介绍航空发动机虚拟自学习方法中虚拟环境的搭建,根据给定的发动机试车数据,为智能控制器训练搭建适合的交互环境。神经网络具有能逼近高度复杂的非线性、并行高速处理、高度容错、联想记忆等优点,常用于发动机的建模过程。周俊^[14]与李杰等^[15]分别通过建立数据驱动模型实现对发动机剩余使用寿命的预测。本文基于发动机转速控制特点,提出一种基于长短期记忆时间(LSTM)循环神经网络建立虚拟环境的方法。

2.1 LSTM 算法原理

LSTM神经网络可以对 $t-1$ 时刻的数据和 t 时刻数据保持相关性,能很好地用来进行长时序数据的训练。LSTM网络结构的核心是细胞状态,如图2所示,当前的LSTM网络接收来自上一

时刻的细胞状态 C_{t-1} 和隐含信息 H_{t-1} ,并与当前接收的信号 X_t 共同作用产生当前细胞状态 C_t ,再将新的细胞状态 C_t 和隐含信息 H_t 传递到下一时刻的细胞体中。LSTM算法采取“门”信号来修改细胞状态。

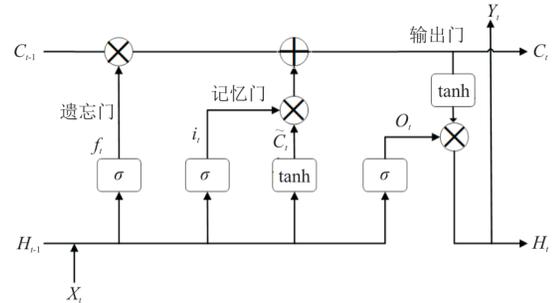


图2 LSTM基本结构示意图
Fig. 2 Schematic diagram of the basic structure of LSTM

1) 遗忘门:有选择地使上一时刻细胞状态 C_{t-1} 的部分信息遗忘。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

式中: σ 为 sigmoid 激活函数; W_f 和 b_f 分别为遗忘门神经网络的权重和偏置矩阵。

2) 记忆门:有选择地使上一时刻细胞状态 C_{t-1} 的部分信息被传入下一时刻的细胞状态中。

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

式中: W_i 和 b_i 分别为 sigmoid 层的权重和偏置矩阵; W_c 和 b_c 分别为 tanh 层的权重和偏置矩阵。

3) 输出门:将记忆门和遗忘门输出的信息加和后得到新的细胞状态 C_t 和隐含信息 H_t 并传给下一时刻的细胞。

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (4)$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$H_t = O_t \cdot \tanh(C_t) \quad (6)$$

式中: W_o 和 b_o 分别为输出神经网络神经门的权重和偏置矩阵。

新的细胞状态经 tanh 函数激活后与 O_t 相乘得到新的隐含信息 H_t 。

2.2 LSTM 神经网络搭建

本文选用双轴转子涡扇发动机 JT9D 模型来代替真实航空发动机进行实验。JT9D 模型的内部结构示意图如图3所示,其工作原理为:从进气

道进入发动机的空气首先经过压气机的等熵压缩后速度降低,压力增加;而后高压气体流向燃烧室,与燃料混合后进行等压燃烧,获得大量热能;燃气经过涡轮时通过等熵膨胀将热能转换为机械能,推动涡轮转动,涡轮同时带动压气机转动;燃气最终在尾喷管进一步膨胀,将剩余的部分热能转换为动能并高速喷出,为飞机提供飞行前进所需动力^[16]。

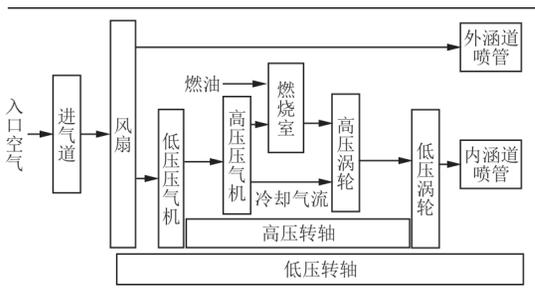


图3 JT9D涡扇发动机内部结构

Fig. 3 Internal structure of JT9D turbofan engine

航空发动机 JT9D 模型的工作参数如表 1 所示,限定其高压转子转速 n_H 范围在 6 009~8 881 r/min,对应的燃油流量 W_f 选取范围为 0.211 9~2.703 0 kg/s。

表1 JT9D航空发动机主要工作参数

Table 1 Main working parameters of JT9D aero engine

参数	数值
风扇进口气流流量 $W/(\text{kg}\cdot\text{s}^{-1})$	305.82
风扇进口流量总焓 $H_f/[\text{kJ}\cdot(\text{kg}\cdot\text{K})^{-1}]$	543.40
风扇进口流量温度 T_f/K	833.72
风扇进口压力 P_{amb}/Pa	25 000.39
风扇出口压力 P_f/Pa	38 114.22

获取 100 组不同给定燃油流量下、100 s 内发动机高压转子转速输出,采样时间间隔为 0.04 s,每组数据序列长为 2 500。由于获取到的高压转速数量级为 10^3 ,与燃油流量的 10^{-1} 数量级相差太大,不能直接用来进行神经网络训练,需要进行数据归一化处理。本文采用最小-最大规范化方法将输入输出参数归一化至 $[0,1]$ 之间:

$$y_{\min-\max} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (7)$$

式中: x 为原始样本数据; $\min(x)$ 、 $\max(x)$ 分别为原始样本数据最小值和最大值。

将归一化后的数据用来进行神经网络训练,如图 4 所示,选取其中的 90% 作为训练集用来对

神经网络进行训练,剩余 10% 作为测试集对已训练好的神经网络进行测试。将输入参数燃油流量及输出参数高压转速数据序列通过神经网络进行迭代预测,将理论输出与模型输出进行比较计算损失误差,进而更新神经网络。

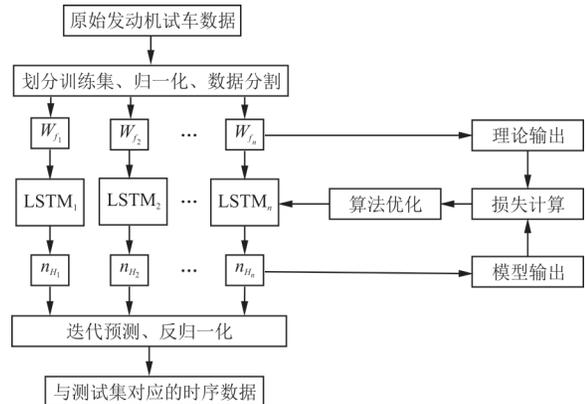


图4 LSTM神经网络训练流程图

Fig. 4 Training flow chart of LSTM neural network

设置神经网络参数(如表 2 所示),输入层特征维数为 1;LSTM 层包含 128 个隐含神经单元;抛弃层默认值为 0,不丢弃任何数据;全连接层输出特征维数为 1,最后进行回归分析输出。训练 250 个回合,采取分段学习率的学习策略,初始学习率为 0.005,125 个回合后学习率开始下降,下降因子为 0.2,每次迭代使用的数据量为 256。

表2 LSTM神经网络训练参数设置

Table 2 Training parameter settings of LSTM neural network

参数	范围/数值
输入层维数	1
输出层维数	1
LSTM 神经元数目	128
数据丢弃率	0
初始学习率	0.005
学习率下降因子	0.2
迭代数据量	256

3 智能控制器设计

本节主要介绍航空发动机虚拟自学习方法中智能控制器的设计,智能控制器如何在与仿真环境的交互过程中不断逼近控制目标,在最少的迭代次数中学习最优策略。由于 Q-Learning 算法不适用于对连续动作空间的强化学习,在此基础上的 DDPG 算法虽然弥补了 Q-Learning 算法的缺

陷,但在实际应用中存在 Q 值过估计和对于值空间中的窄峰估计过拟合的问题,因此本文选用 DDPG 的优化版本 TD3 算法。

3.1 TD3 算法

双延迟深度确定性策略梯度(TD3)算法采用 Actor-Critic 的网络结构,在智能控制器与虚拟环境交互的过程中,Actor 网络产生确定性策略 π_ϕ , Critic 网络评判策略 π_ϕ 的优劣,然后进行迭代更新策略,优化智能控制器的参数,使得智能控制器获得累计奖励的最大值。具体流程如图 5 所示: t 时刻智能控制器接收从环境观测到的状态信息 S_t , 根

据 Actor 网络产生实际动作 $\mu(S_t)$,与高斯噪声产生的探索加和后输出执行动作 A_t 给虚拟环境,虚拟环境在动作执行后反馈一个新的状态信息 S_{t+1} 和即时奖励 R_{t+1} 给智能控制器,并将这一过程存储在经验池中。智能控制器在学习生成最优控制策略的过程中,需要从经验池中采样获取历史信息对 Actor 目标网络参数 θ^{*1} 、 θ^{*2} 和 Critic 目标网络参数 θ^{Q1} 、 θ^{Q2} 进行更新,利用更新后的目标网络对策略网络和评价网络进行缓慢的参数优化,使得智能控制器生成的策略获得最大的累计奖励值,同时评价行为价值的 Q 值最大化,得到最优控制策略 π^* 。

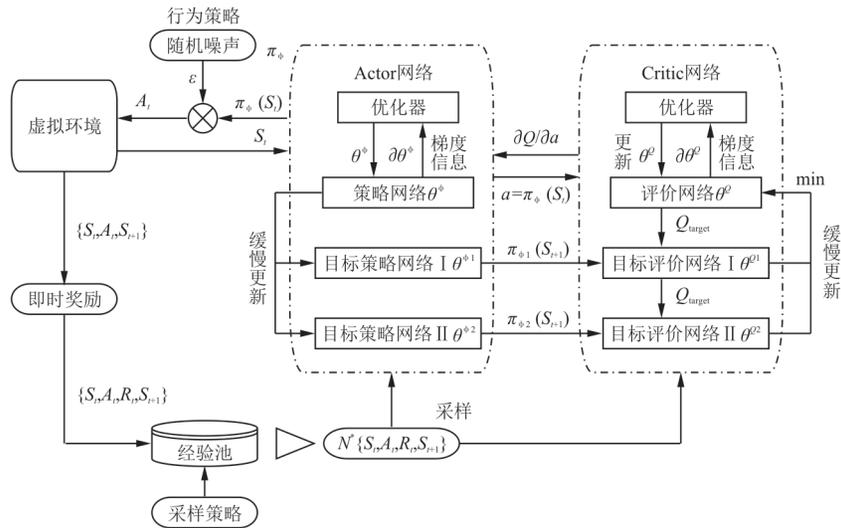


图 5 TD3 算法流程图

Fig. 5 Flow chart of TD3 algorithm

TD3 算法在 DDPG 算法的基础上进行了以下 3 点改进:

1) 采用双重网络,通过学习两个 Critic 目标网络的 Q 值函数,选取其中较小的值更新 Critic 网络:

$$y = r + \gamma \min_{i=1,2} Q'_i(s', a' | \theta^{Qi}) \quad (i = 1, 2) \quad (8)$$

式中: y 为 Critic 网络的目标值; Q'_i 为 Critic 目标网络的目标值; θ^{Qi} 为两个 Critic 目标网络参数; s' 为预测得到的 $t+1$ 时刻的状态; a' 为根据预测状态产生的加了正态分布噪声之后的动作。

选取两个目标值中较小的 Q'_i 值用来更新 Critic 网络,避免了 Q 值估计过高的问题。

2) 目标策略平滑正则化,通过在每个动作加入截断的正态分布噪声:

$$y = r + \gamma Q_\sigma(s', \pi_{\phi_i}(s') + \epsilon) \quad (i = 1, 2) \quad (9)$$

式中: $\pi_{\phi_i}(s')$ 为依据 Actor 目标网络输出的动作; $\epsilon \sim \text{clip}(N(0, \sigma), -c, c)$ 为截断后的正态分布噪声,可以作为正则化来平滑 Q 值的计算。

3) 对 Actor 网络采取延迟更新,只有在 Critic 网络更新 d 次后才进行更新:

$$\nabla_{\phi} J(\phi) = N^{-1} \sum \nabla_a Q_{\theta_i}(S_t, A_t) |_{A_t = \pi_{\phi}(S_t)} \nabla_{\phi} \pi_{\phi}(S_t) \quad (10)$$

$$\hat{\theta}_i \leftarrow \rho \theta_i + (1 - \rho) \hat{\theta}_i \quad (11)$$

$$\hat{\phi} \leftarrow \rho \phi + (1 - \rho) \hat{\phi} \quad (12)$$

ρ 很小,使得参数更新缓慢,确保 Q 值函数的估计具有更小的方差,从而得到更优化的策略。

3.2 TD3 算法神经网络设计

在 TD3 算法中, Actor 网络和 Critic 网络均使用深度神经网络来建立近似函数, Actor 网络可以生成确定的动作, 并由 Critic 网络评价该状态下执行此动作的优劣程度。Actor 网络的结构如图 6 所示, 包含: 一个输入层, 三个隐含层和一个输出层。输入层拥有 3 个神经元, 对应智能控制器从环境中获取的三维状态空间信息; 隐含层采用全连接层神经网络, 每层拥有 64 个神经元, 选取 ReLU 函数作为激活函数; 输出层拥有 1 个神经元, 对应智能控制器的一维动作空间, 使用 Sigmoid 激活函数。

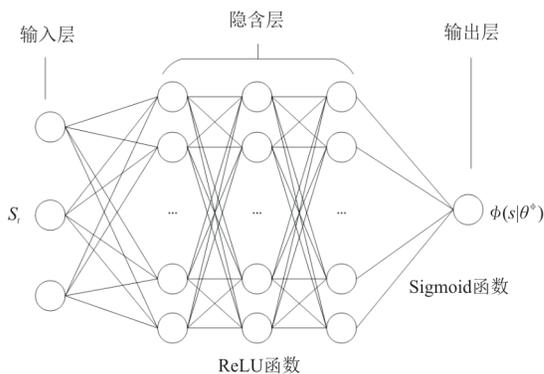


图 6 Actor 神经网络结构示意图

Fig. 6 Structure diagram of Actor neural network

Critic 神经网络结构与 Actor 神经网络类似, 如图 7 所示: 同样包含一个输入层, 三个隐含层和一个输出层。Critic 神经网络的输入层有 4 个神经元, 分别代表三维状态空间和一维动作空间; 隐含层为采用激活函数 Tanh 的全连接层; 输出层有 1 个神经元, 代表所处时刻下状态行为对的价值 $Q(s, a)$ 。

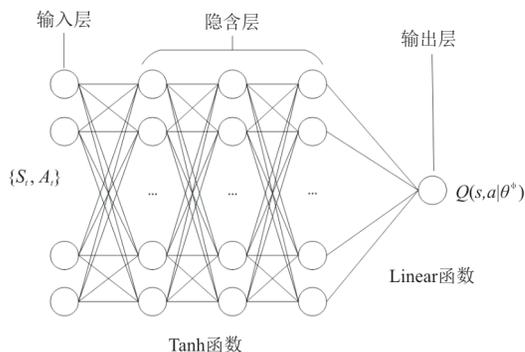


图 7 Critic 神经网络结构示意图

Fig. 7 Structure diagram of Critic neural network

在利用 TD3 算法对智能控制器进行训练的过程中, 神经网络参数设置如表 3 所示, 共训练 500 回合, 经验池的容量设置为 100 000, 每次采样批次获取 256 组数据。Actor 网络的学习率为 0.005; Critic 网络的学习率要低于 Actor 网络, 设为 0.002, 确保在 Actor 网络更新前最小化价值更新的误差; 神经网络的缓慢更新速率为 0.005; 折扣因子 γ 设为 0.99。

表 3 TD3 算法神经网络参数设置
Table 3 Parameter settings of TD3 algorithm neural network

参数	范围/数值
输入状态维数	3
输出状态维数	1
Actor 网络隐含神经元数目	[3, 64×3, 1]
Critic 网络隐含神经元数目	[4, 64×3, 1]
Actor 网络学习率	0.005
Critic 网络学习率	0.002
Soft 更新率	0.005
Mini-Batch 大小	256
经验回放池大小	1 000 000
折扣因子	0.99
训练回合数	500

3.3 状态空间和动作空间设计

航空发动机智能控制需要实现发动机对给定参考信号的跟踪控制, 智能控制器可以从虚拟环境中获取发动机的高压转子转速 n_H , 将高压转速信号 n_H 和高压转速与参考转速之间的误差信号 $e(t)$ 作为状态空间观测值对智能控制器进行训练。但在实际训练中发现, 智能控制器学到的策略并不能稳定跟踪参考信号, 而在状态空间中加入误差信号的积分项可以有效减小系统稳态误差并提高系统的稳定性。因此, 虚拟自学习方法中的状态空间可以表示为

$$S_t = \{n_H(t), e(t), \int e(t) dt\} \quad (13)$$

智能控制器通过直接输出燃油流量 W_f , 实现对发动机高压转子转速的控制, 动作空间可以表示为

$$A_t = \{W_f\} \quad (14)$$

3.4 奖励函数设计

深度强化学习的训练过程中奖励函数的优劣

直接影响了神经网络的收敛速度和网络最终表现的好坏,进而决定了智能控制器表现的好坏。智能发动机转速控制中,要求发动机可以稳、准、快地跟踪给定参考信号,因此设置的奖励函数包含三个部分,分别为 $r_1(t)$ 、 $r_2(t)$ 和 $r_3(t)$ 。

1) 奖励函数 $r_1(t)$ 与误差信号 $e(t)$ 有关,输出转速信号与给定参考信号的偏差越小,智能控制器获得的奖励值越大,当偏差为0时,奖励值达到最大值1,修改奖励项的设定可以使智能控制器准确地学习到最优控制策略。

$$r_1(t) = e^{-0.02|e(t)|} - 0.002e^2(t) \quad (15)$$

2) 奖励函数 $r_2(t)$ 与智能控制器的输出大小有关,为了使智能控制器更有效地对动作空间进行探索,在探索过程中引入惩罚项,当动作值输出 $W_f > 3$ 或 $W_f < 0$ 时,给予智能控制器 -0.2 的惩罚。

$$r_2(t) = \begin{cases} 0 & (0 < W_f < 3) \\ -0.2 & (\text{其他}) \end{cases} \quad (16)$$

3) 奖励函数 $r_3(t)$,在实际训练过程中发现,如果仅以 $r_1(t)$ 、 $r_2(t)$ 为奖励项构建奖励函数,训练出的智能控制器容易出现超调现象。为了抑制超调,引入发动机转速误差信号的微分项作为惩罚,微分项越大,惩罚越大。

$$r_3(t) = -0.01 \left| \frac{de(t)}{dt} \right| \quad (17)$$

最终奖励函数 $r(t)$ 由 $r_1(t)$ 、 $r_2(t)$ 和 $r_3(t)$ 共同组成,即:

$$r(t) = r_1(t) + r_2(t) + r_3(t) \quad (18)$$

4 仿真结果

4.1 虚拟环境的仿真结果分析

对使用LSTM神经网络搭建的航空发动机虚拟环境进行测试,把2.2节中的测试集输入数据送入LSTM神经网络,将得到的虚拟环境输出与测试集理论输出相比较,如图8所示,查看训练后的LSTM神经网络对测试集数据的拟合近似效果。输入范围在0.211 9~2.703 0 kg/s的燃油流量,得到范围在6 009~8 881 r/min的高压转子转速。

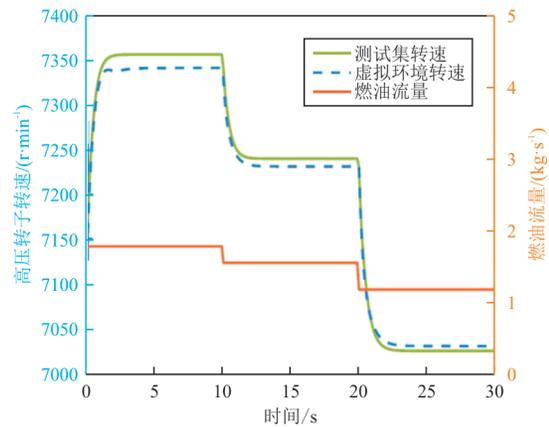


图8 虚拟环境与测试集输出拟合比较

Fig. 8 Comparison of virtual environment and test set outputs

从图8可以看出:在30 s的测试过程中,前10 s内燃油流量为1.77 kg/s,后降为1.54 kg/s维持10 s,最后10 s燃油流量稳定在1.17 kg/s;测试集高压转速的变化趋势与燃油流量变化趋势相同,前2 s内高压转速开始上升并稳定在7 356.79 r/min,然后在10 s处开始下降,在1.9 s内降至7 240.08 r/min并保持稳定,最后在20 s处发生二次下降,且2.2 s内下降并保持稳定在7 024.97 r/min;虚拟环境高压转速的变化趋势与测试集高压转速曲线变化趋势相同,在前20 s内高压转速略低于测试集输出转速,在2.2 s内高压转速迅速上升并保持稳定在7 341.5 r/min,与测试集输出转速的误差约为15 r/min,在10 s处发生转速下降,2.3 s内降至7 231.5 r/min并保持稳定,与测试集输出转速相差约为9 r/min,最后10 s内虚拟环境的输出略高于测试集输出转速,在20 s处转速发生二次下降,2.6 s内高压转子转速下降并稳定在7 030.4 r/min,与测试集输出的转速误差约为5 r/min。

选取不同测试集在各种燃油流量下进行多次测试,实验结果如表4所示。

表4 不同测试集下的虚拟环境与理论输出比较
Table 4 Comparison of virtual environment and theoretical outputs under different test sets

燃油流量/(kg·s ⁻¹)	高压转速/(r·min ⁻¹)	
	理论输出	虚拟环境输出
0.781 8	6 742.53	6 750.11
0.893 6	6 831.74	6 840.92
1.208 4	7 052.00	7 047.63
1.543 7	7 231.40	7 240.08
2.387 8	7 616.50	7 621.13

从表4可以看出:由LSTM神经网络搭建的虚拟环境能够很好地与测试集输出拟合,经计算虚拟环境与测试集输出的误差仅为9.459 5 r/min。

为了验证模型的泛化能力,本文在更多输入形式下对比虚拟学习环境的建模误差。选用JT9D双轴转子涡扇发动机模型代替真实发动机,输入相同的燃油流量,比较虚拟环境的输出与JT9D模型的输出,检测虚拟环境对真实航空发动机的近似程度。

1) 当给定输入为斜坡信号时,虚拟环境与JT9D模型的高压转速输出如图9所示。

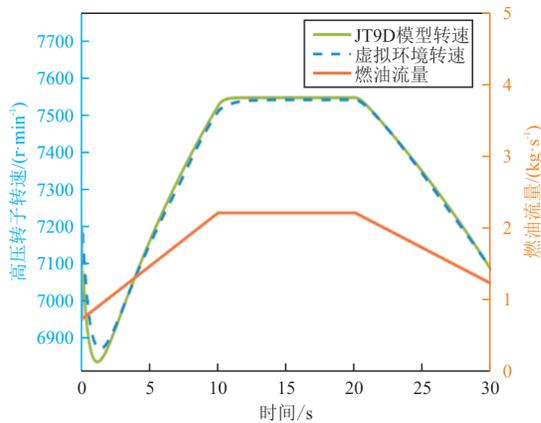


图9 斜坡信号下虚拟环境与JT9D输出拟合比较
Fig. 9 Comparison of virtual environment and JT9D outputs under ramp signal

从图9可以看出:燃油流量从0.7 kg/s开始,在10 s内上升至2.2 kg/s,稳定输出10 s后,在10 s内又降低至1.2 kg/s;JT9D模型的高压转速输出在初始阶段会有一段下降的过程,在1.3 s内降至6 835 r/min,然后开始上升且变化趋势与燃油流量变化趋势相同,在10 s内上升至7 547.8 r/min并稳定维持转速不变,在20 s处发生转速下降,最终降至7 078.9 r/min;虚拟环境高压转速输出曲线变化趋势和JT9D模型相同,在开始阶段同样有转速下降的过程,1.4 s内下降至6 870.8 r/min,与JT9D模型输出的转速误差约为35 r/min,但在上升阶段和下降阶段均与JT9D模型的输出转速拟合效果很好,转速误差仅在3 r/min以内,在10~20 s的稳态过程中,虚拟环境输出高压转速为7 542.8 r/min,与JT9D模型的高压转速输出误差仅为5 r/min。综上,在输入燃油流量为斜坡信号下,虚拟环境输出的高压转速与JT9D模型输出的高压转速虽然在开始阶段转速下降时有较大的误差,但在后续过程都有很好的拟合效果。

2) 当给定信号为方波信号时,虚拟环境与JT9D模型的高压转速输出如图10所示。

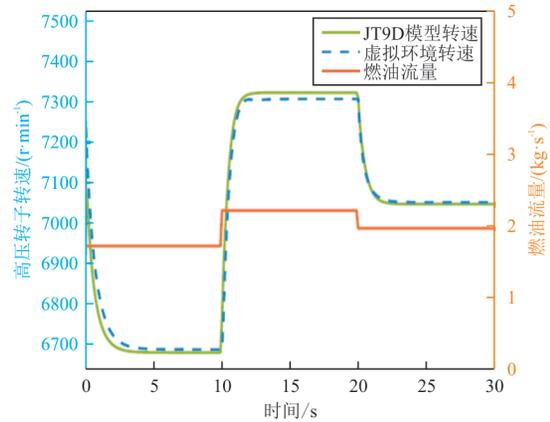


图10 方波信号下虚拟环境与JT9D输出拟合比较
Fig. 10 Comparison of virtual environment and JT9D outputs under square wave signal

从图10可以看出:燃油流量在前10 s内维持在1.70 kg/s,10 s处发生阶跃上升至2.20 kg/s,维持10 s后发生二次阶跃下降,最后10 s内稳定在1.95 kg/s;JT9D模型输出的高压转速在前3.5 s有转速下降的过程,降至6 680.1 r/min并稳定一段时间,在10 s处发生阶跃上升,2 s内转速上升并稳定在7 323.5 r/min,在20 s处发生阶跃下降,2.2 s内下降并稳定在7 047.64 r/min;虚拟环境输出的高压转速变化趋势与JT9D模型输出的高压转速相同,3 s内转速降至6 687.8 r/min,略高于JT9D模型,转速误差约为6 r/min,随后在10 s处发生阶跃上升,2.3 s内快速升至7 308.51 r/min并稳定维持至20 s,转速略低于JT9D模型的输出转速,相差约为15 r/min,最后在20 s处发生阶跃,2.4 s内下降并稳定维持在7 051.9 r/min,转速略高于JT9D模型的输出,转速误差约为4 r/min。综上,输入燃油流量为方波信号时,虚拟环境的输出在开始和结束阶段略高于高压JT9D模型的输出,转速误差在5 r/min以内,中间阶段虚拟环境的输出略低于JT9D模型的输出,转速误差在15 r/min之内。通过比较可以发现,虚拟环境对JT9D模型的近似程度较高,符合控制系统要求,可以用来对智能控制器进行交互式训练。

4.2 智能控制的仿真结果分析

将设计好的智能控制器与建立的虚拟仿真环境进行交互仿真,通过不断试错的方法,学习最优控制策略。智能控制器共训练500回合,每回合与

虚拟环境交互 750 次,训练后得到的回合累计奖励函数曲线如图 11 所示。

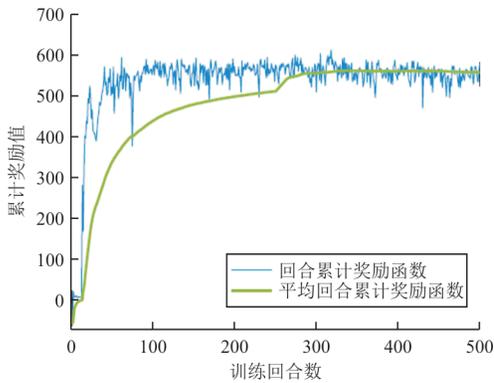


图 11 TD3 算法回合累计奖励函数

Fig. 11 Round cumulative reward function of TD3 algorithm

从图 11 可以看出:智能控制器经过 50 个回合的试错学习后可以获得较好的控制策略,回合累计奖励值较高;平均回合累计奖励在 250 回合处发生了一次阶跃;最终回合累计奖励值收敛至 590,说明智能控制器在学到最优控制策略后,每一步都可以获得较好的即时奖励。

将训练好的智能控制器用于航空发动机高压转速的控制,并与传统 PID 控制器进行比较。

1) 参考信号为斜坡信号时,TD3 智能控制与传统 PID 控制性能比较如图 12 所示。

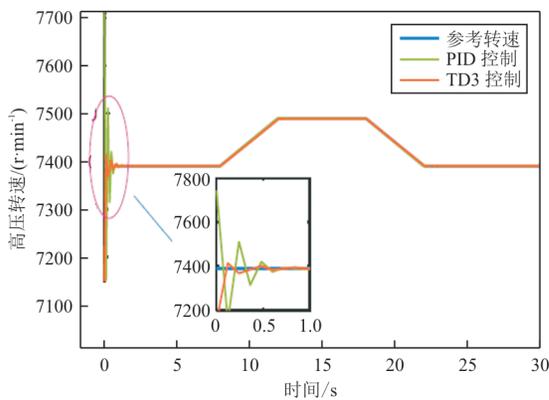


图 12 斜坡信号下的 TD3 智能控制器与传统 PID 控制比较

Fig. 12 Comparison between TD3 intelligent control and traditional PID control under ramp signal

从图 12 可以看出:给定参考转速在前 8 s 内稳定在 7 390 r/min,后在 4 s 内上升至 7 490 r/min,稳定维持 6 s 后,在 4 s 内下降至 7 390 r/min,最后稳定维持 8 s;传统 PID 控制调节参数 $K_p=0.009$ 、 $K_i=0.04$ 、 $K_d=0$,在开始阶段,高压转速有剧烈的

振荡,最高达 7 510 r/min,最低为 7 314 r/min,超调量为 1.48%,0.5 s 后高压转速可以稳定维持在 7 390 r/min,在上升阶段、下降阶段和稳态阶段对参考信号的跟踪效果较好;TD3 智能控制下的高压转速在开始阶段虽然也有小范围的振荡,但转速误差在 15 r/min 以内,超调量仅为 0.2%,且在 0.2 s 内就迅速收敛至 7 390 r/min,在上升阶段、下降阶段和稳态阶段,TD3 智能控制同样可以较好地跟踪给定参考信号。综上,在斜坡信号下,智能控制器产生的振荡远小于传统 PID 控制器产生的振荡,且对给定参考信号的控制效果要更好。

2) 参考信号为方波信号时,TD3 智能控制与传统 PID 控制性能比较如图 13 所示。

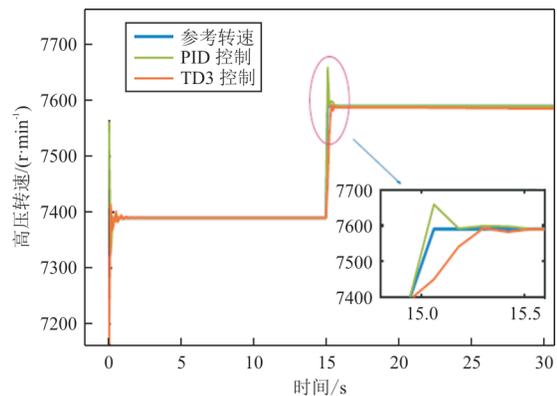


图 13 阶跃信号下的 TD3 智能控制器与传统 PID 控制比较

Fig. 13 Comparison between TD3 intelligent controller and traditional PID control under step signal

从图 13 可以看出:给定参考转速在前 15 s 内为 7 490 r/min,在 15 s 处发生阶跃,上升至 7 590 r/min;PID 调节参数 $K_p=0.006$ 、 $K_i=0.03$ 、 $K_d=0$,在阶跃过程中,TD3 智能控制的动态性能指标要远高于传统 PID 控制,如表 5 所示,传统 PID 控制器下的超调量为 0.909 0%,TD3 智能控制器下几乎没有超调,而且 TD3 智能控制器的上升速度虽然略慢于传统 PID 控制,但调节时间仅为传统 PID 控制器的一半,具有较好的动态性能。

表 5 PID 传统控制与 TD3 智能控制性能指标比较
Table 5 Comparison of performance indicators between PID traditional control and TD3 intelligent control

控制性能指标	上升时间/s	超调量/%	调节时间/s
PID 传统控制	0.2	0.909 0	0.90
TD3 智能控制	0.3	0	0.45

5 结 论

本文提出的航空发动机虚拟自学习方法,为航空发动机的智能控制进行了有益探索。它通过建立虚拟学习环境,为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。仿真结果表明,本文建立的航空发动机虚拟仿真环境能够满足后续的控制要求,同时证明了基于该方法设计的智能控制器在燃油流量—高压转子转速控制任务上的有效性。后续将立足现有理论成果,开展实验验证研究。

参 考 文 献

- [1] 刘毅男,李述清,张胜修,等. 航空发动机闭环增益成形PID控制器设计研究[J]. 电光与控制, 2011, 18(11): 74-76.
LIU Yinan, LI Shuqing, ZHANG Shengxiu, et al. Design of a PID controller for aero-engine based on closed-loop gain shaping[J]. Electronics Options & Control, 2011, 18(11): 74-76. (in Chinese)
- [2] RICHTER H. Advanced control of turbofan engines[M]. Berlin, Germany: Springer Science & Business Media, 2011.
- [3] 吴斌,黄金泉. 航空发动机全包线鲁棒变增益LPV控制律设计[J]. 南京航空航天大学学报, 2014, 46(2): 252-258.
WU Bin, HUANG Jinquan. Robust gain scheduling LPV control design for aeroengine in full envelope[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2014, 46(2): 252-258. (in Chinese)
- [4] WOLODKIN G, BALAS G J, GARRARD W L. Application of parameter-dependent robust control synthesis to turbofan engines[J]. Journal of Guidance, Control and Dynamics, 1999, 22(6): 833-838.
- [5] 任立通,谢寿生,王磊,等. 不确定航空发动机分布式控制系统自适应滑模控制[J]. 航空动力学报, 2017, 32(8): 2032-2040.
REN Litong, XIE Shousheng, WANG Lei, et al. Adaptive sliding mode control for uncertain aero-engine distributed control system[J]. Journal of Aerospace Power, 2017, 32(8): 2032-2040. (in Chinese)
- [6] FAKHARI V, OHADI A, TALEBI H A. A robust adaptive control scheme for an active mount using a dynamic engine model[J]. Journal of Vibration and Control, 2015, 21(11): 2223-2245.
- [7] 陈勇,蔡开龙. 基于模糊误差判断算法的航空发动机PID控制[J]. 自动化仪表, 2021, 42(4): 73-76.
CHEN Yong, CAI Kailong. Aero-engine PID control based on fuzzy error judgment algorithm[J]. Process Automation Instrumentation, 2021, 42(4): 73-76. (in Chinese)
- [8] 薛红阳,蔡开龙,李黄琪,等. 基于RBF神经网络补偿的航空发动机 H_{∞} 自适应控制研究[J]. 航空工程进展, 2023, 14(1): 128-134.
XUE Hongyang, CAI Kailong, LI Huangqi, et al. Research on H_{∞} adaptive control of aero-engine based on RBF neural network compensation[J]. Advances in Aeronautical Science and Engineering, 2023, 14(1): 128-134. (in Chinese)
- [9] GAO Wenbo, PAN Muxuan, ZHOU Wenxiang, et al. Acceleration control strategy for aero-engines based on model-free deep reinforcement learning method[J]. Aerospace Science and Technology, 2022, 120: 112-118.
- [10] WAXENEGGER-WILFING G, DRESIA K, DEEKEN J C, et al. A reinforcement learning approach for transient control of liquid rocket engines[J]. IEEE Transactions on Aerospace and Electronic Systems, 2021, 57(5): 2938-2952.
- [11] 孙传宗. 航空发动机双转子系统高精度动力学建模与碰撞响应研究[D]. 哈尔滨: 哈尔滨工业大学, 2017.
SUN Chuanzong. Accurate dynamical modelling for dual-rotor aero-engine system and its responses induced by rub impact[D]. Harbin: Harbin Institute of Technology, 2017. (in Chinese)
- [12] 石岩. 航空双轴涡扇发动机的建模及切换控制设计[D]. 沈阳: 东北大学, 2019.
SHI Yan. Modeling and switching control design for two-spool turbofan aero-engines[D]. Shenyang: Northeastern University, 2019. (in Chinese)
- [13] ZHENG Qiangang, JIN Chongwen, HU Zhongzhi, et al. A study of aero-engine control method based on deep reinforcement learning[J]. IEEE Access, 2019, 7: 123-129.
- [14] 周俊. 数据驱动的航空发动机剩余使用寿命预测方法研究[D]. 南京: 南京航空航天大学, 2017.
ZHOU Jun. Research on data-driven prediction methods for remaining useful life of aero-engine[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017. (in Chinese)
- [15] 李杰,贾渊杰,张志新,等. 基于融合神经网络的航空发动机剩余寿命预测[J]. 推进技术, 2021, 42(8): 1725-1734.
LI Jie, JIA Yuanjie, ZHANG Zhixin, et al. Remaining useful life prediction of aeroengine based on fusion neural network[J]. Journal of Propulsion Technology, 2021, 42(8): 1725-1734. (in Chinese)
- [16] 罗茂春. 基于FMI的航空发动机控制系统多学科联合仿真[D]. 南京: 南京航空航天大学, 2019.
LUO Maochun. FMI-based multi-domain simulation for aero-engine control systems[J]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2019. (in Chinese)

作者简介:

董建华(1999—),男,硕士研究生。主要研究方向:航空发动机智能控制。

朱建铭(1999—),男,硕士研究生。主要研究方向:航空发动机智能控制。

黎瀚涛(1996—),男,硕士研究生。主要研究方向:航空发动机智能控制。

刘文烁(2001—),男,硕士研究生。主要研究方向:航空发动机智能控制。

唐 炜(1977—),男,博士,副教授。主要研究方向:航空发动机智能控制,机器人与智能系统,振动控制。

(编辑:马文静)