

基于强化学习的燃气轮机转速控制策略研究

张启航, 李学民, 祝存玺

(哈尔滨工程大学 动力与能源工程学院, 黑龙江 哈尔滨 150001)

摘要:以某三轴燃气轮机作为研究对象,针对燃气轮机各工况下转速失稳、不同海况下的动力响应需求、外部干扰等因素对燃气轮机工作状态的影响问题,设计了基于Q-learning强化学习理论的控制策略,以智能算法在线调试代替人工手动调试过程,实现人工智能理论到工程实际的应用在进行软件层智能算法验证后,通过强化学习算法控制参数的自动优化完成了基于强化学习的转速控制策略硬件在环试验。研究表明:该设计算法可以在燃气轮机运行过程中判断触发、自我训练、自我调整控制参数,保证燃气轮机在各种突发情况下的运行稳定;在三轴燃气轮机转速失稳时,在较短时间内即可完成转速的回稳,并可将转速误差控制在2 r/min内,从而实现了转速失稳时的自救。

关键词:燃气轮机;强化学习算法;转速稳定控制;硬件在环

中图分类号:TH113 文献标识码:A DOI:10.16146/j.cnki.rndlgc.2022.12.019

[引用本文格式]张启航,李学民,祝存玺.基于强化学习的燃气轮机转速控制策略研究[J].热能动力工程,2022,37(12):148-156. ZHANG Qi-hang, LI Xue-min, ZHU Cun-xi. Research on gas turbine speed control strategy based on reinforcement learning[J]. Journal of Engineering for Thermal Energy and Power, 2022, 37(12):148-156.

Research on Gas Turbine Speed Control Strategy based on Reinforcement Learning

ZHANG Qi-hang, LI Xue-min, ZHU Cun-xi

(College of Power and Energy Engineering, Harbin Engineering University, Harbin, China, Post Code:150001)

Abstract: This paper took a three-shaft gas turbine as the research object, under various operating conditions for gas turbine speed instability problem, dynamic response requirements of different sea conditions and the influence of external disturbance factors on the working state of gas turbine, designed the control strategy of the reinforcement learning theory based on the Q-learning, and realized the application of artificial intelligence theory to engineering practice by using the intelligent algorithm online debugging to replace the manual debugging process. After verifying the intelligent algorithm of software layer, through the automatic optimization of the control parameters of the reinforcement learning algorithm, the hardware-in-the-loop test of speed control strategy based on reinforcement learning was completed. The research indicates that the algorithm designed in this paper can judge the trigger, train itself, and self-adjust the control parameters in the gas turbine operation process to ensure the stable operation of the gas turbine in various emergencies. When the three-shaft gas turbine speed is unstable, the speed stabilization can be completed in a short time, and the speed error can be controlled within 2 r/min, so as to realize the self-rescue when the speed is unstable.

收稿日期:2022-04-14; 修订日期:2022-06-30

基金项目:国家科技重大专项(2017-V-0005-0055, 2017-V-0015-0067)

Fund-supported Project: National Science and Technology Major Project(2017-V-0005-0055, 2017-V-0015-0067)

作者简介:张启航(1998-),男,黑龙江大庆人,哈尔滨工程大学硕士研究生。

通讯作者:李学民(1971-),男,吉林长春人,哈尔滨工程大学研究员。

Key words: gas turbine, reinforcement learning algorithm, speed stability control, hardware-in-the-loop

引 言

燃气轮机是涉及热能工程、气动力学、机械工程、控制工程及材料科学等多学科、多物理场耦合的非线性动力机械,同时燃气轮机属于结构紧凑的高温旋转叶轮机机械,许多性能和状态参数难以直接测量或准确测试,因此有必要开展非线性、自适应控制理论应用研究^[1-3]。

基于强化学习的控制方法是一种新兴的理论,已经成为国内外学者广泛重视的研究方向,在控制理论领域占有较高的关注度。结合人工智能控制数据库的建立,根据工况的变化实现控制策略的调整,是对燃气轮机控制系统进行的新探索与实践,对传统 PID 控制方法的缺点与不足进行改善和补充。Sanusi 等人^[4]在 2019 年提出了一种基于强化学习和自适应动态规划的燃气轮机状态控制框架,使用表现出的控制效果作为奖励信号来学习和调整系统效率,这是在不了解系统变化或退化实际动力学 的情况下实现的,提供一种贯穿生命周期的适应策略,从而提高了系统性能。通过采用基于强化学习的先进控制理论,使得所设计的燃气轮机控制系统能够适应燃气轮机运行条件的变化、燃气轮机装置的特性变化,在燃气轮机的整个运行周期内始终保持其高效、快速以及安全、低污染运行,满足燃气轮机控制需求,发挥最佳控制性能^[5]。

针对燃气轮机由于工况变化导致转速失稳的问题,本文提出一种基于强化学习的转速控制算法,利用人工智能对数据的分析学习能力进行燃气轮机运行过程中控制策略的迭代,实现对 PI 控制器参数调节的功能。算法可以模拟人工调整 PI 参数试凑法的过程,通过燃气轮机模型与算法的信息交互以及 Q-learning 中价值计算公式迭代的训练方法得到最终用于控制转速的策略。借助快速控制原型机与 MATLAB/Simulink 自动代码控制箱,得到可以写入嵌入式开发系统的代码,通过硬件在环系统(Hard in-the-loop, HIL)验证了整个策略的控制效果。

1 基于强化学习的燃气轮机控制器设计

1.1 强化学习原理

强化学习是机器学习领域的一类算法,目的是让一个由算法定义的智能体去学习如何在一个环境中采取行动,其中的反馈由标量形式的奖励信号以及转移后的环境状态组成,其原理如图 1 所示。其中 S 代表状态, R 代表奖励, A 代表动作, t 和 $t+1$ 代表当前循环以及下一循环。

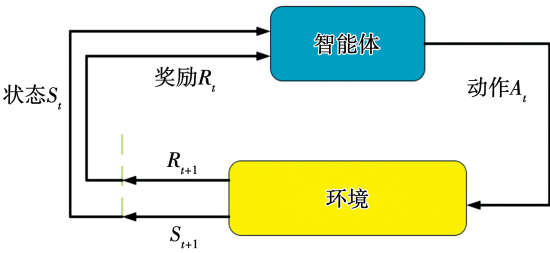


图 1 马尔可夫决策过程原理框图

Fig. 1 Functional block diagram of Markov decision process

1.2 强化学习问题建模

对控制对象——三轴燃气轮机的转速控制建模,搭建了 Simulink 模型。人工调节 PID 参数试凑法过程如图 2 所示。

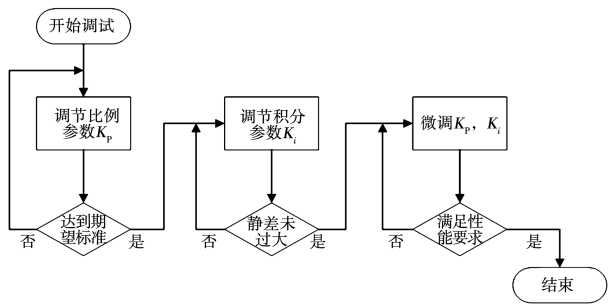


图 2 试凑法调 PID 参数流程图

Fig. 2 PID parameter adjustment flow chart by trial and error method

1.2.1 模型动态规划(MDP)状态定义

根据建模目的明确了包含在状态集合中的参数量有 3 个:当前时间下的比例系数 K_p 、积分系数 K_i

与目标转速 y_r 。

当前时间下的比例系数与积分系数是智能体做出下一步参数调整的基准,目标转速是判断控制效果并给出奖励值反馈的重要参数,将状态集合中第 j 个状态记为 $S^j = [K_p, K_i, y_r]$, 状态集合 S 由 K_p, K_i 和 y_r 3 个元素做笛卡尔积构成,即 $S = K_p \otimes K_i \otimes y_r$ 。将比例系数 K_p 的集合定义为 $K_p = \{K_p^1, K_p^2, \dots, K_p^M\}$, 积分系数 K_i 的集合定义为 $K_i = \{K_i^1, K_i^2, \dots, K_i^M\}$, 目标转速的集合定义为 $y_r = \{y_r^1, y_r^2, \dots, y_r^M\}$ 。将 3 个集合的所有元素排列组合可以得到最终状态的集合中元素个数为 $m = M(K_p)M(K_i)M(y_r)$ 。燃气轮机的 K_p, K_i 和 y_r 参数经过离散化后的集合并不能包含其在实际运行中的所有参数值,但是取大小合适的领域半径后会使得算法具有易收敛、计算量减少等优点,且不会影响控制效果的精度。

MDP 状态空间离散化过程如图 3 所示,假设状态参数的值域为 N , 则其邻域半径 $\sigma = N/M$ 。另外, σ 也可以理解为状态参数的离散步长,用来描述状态参数的离散程度。

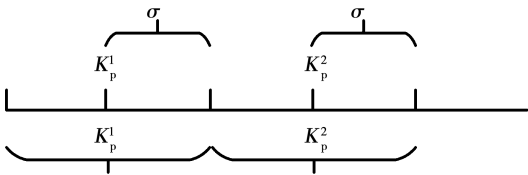


图 3 MDP 状态空间离散化过程

Fig.3 MDP state space discretization process

1.2.2 MDP 动作定义

根据人工调整 PID 参数试凑法进行算法设计,人工调参试凑法的动作可以分为对比例系数 K_p 和积分系数 K_i 的增加、减少和不变 3 种动作的排列组合,每次对 K_p 和 K_i 做出的改变可以称为对这两个参数施加的增量: ΔK_p 和 ΔK_i 。此调整步长也是一个超参数,通过定义每次调整参数增量的步长可以改变智能体对参数调整的细腻程度,根据燃气轮机的控制要求和工作条件选择合适的步长可以使控制效果在满足精确度的同时减少计算量,易于收敛寻优。表 1 为定义动作的向量库。

表 1 MDP 动作向量库

Tab.1 MDP action vector library

A		ΔK_i		
		$+I$	0	$-I$
ΔK_p	$+P$	$[+P, +I]$	$[+P, 0]$	$[+P, -I]$
	0	$[0, +I]$	$[0, 0]$	$[0, -I]$
	$-P$	$[-P, +I]$	$[-P, 0]$	$[-P, -I]$

智能体选择动作后的参数变化与动作增量之间的关系可以由式(1)计算:

$$K_{p, \text{final}} = K_{p, \text{init}} + \sum_{t=0}^n \Delta K_p(t) \tag{1}$$

式中: $K_{p, \text{final}}$ —调整后的比例系数值; $K_{p, \text{init}}$ —比例系数初始值; $\Delta K_p(t)$ —智能体在第 t 步时执行的动作。

同理可计算积分系数调整值。

1.2.3 奖励函数定义

奖励函数是 MDP 过程最重要的一环,反映了整个强化学习过程的最终目的。在设计奖励函数时,对能使转速趋向稳定的动作赋予更高的奖励,对使燃气轮机转速偏离目标值的动作赋予更低的奖励值。

将选择当前转速与目标转速差值的绝对值 $e(t)$ 作为奖励值的主要判定标准,首要任务是消除外界环境变化如风浪或海浪等因素以及负载量变化引起的转速波动与误差。同时,应在奖励函数中增加控制参数变化量 $\Delta u(t)$ 的绝对值积分,避免因控制器参数改变导致的机械结构受损或寿命减少。根据上述因素建立性能指标函数:

$$J = \int_0^T (\rho |e(t)| + (1 - \rho) |\Delta u(t)|) dt \tag{2}$$

式中: T —Q-Learning 算法的一个采样周期; $e(t)$ —目标值与实际值的偏差; ρ —权重系数。

值得注意的是性能指标函数在实际计算时一般会采用离散积分形式。

根据性能指标函数设计奖励函数:

$$R(t) = \begin{cases} -100, y > \max \parallel y < \min \\ 10, J < \xi/4 \\ 0, \xi/4 \leq J < \xi \\ -10, J \geq \xi \end{cases} \tag{3}$$

目标转速与实际转速误差超出预设范围的情况

为转速控制最差情况,则赋予这种状态 -100 的最小奖励值, ξ 为性能指标阈值,根据目标的闭环回路性能估计得到。

1.2.4 探索策略设计

在设计探索策略时应考虑智能体在探索过程中陷入局部最优的情况,这种情况通常是因为在完全贪婪策略下智能体每次在选择动作时总是选择上一次策略迭代后最大价值的动作,未能考虑到后期策略迭代的过程会使动作的价值改变,从而导致不能对整个动作库进行探索,错失全局最优解。为避免这种情况发生,且在智能体探索后期,策略迭代至接近成熟时采取近似完全贪婪策略,需要引入一个随着训练回合数逐渐减少的贪婪因子。设置一个参数 $\varepsilon \in [0,1]$,智能体在学习过程中,有 ε 的概率选择随机的动作执行,最优价值动作的选择概率与其他非最优价值动作一致。有 $1 - \varepsilon$ 的概率选择最优价值动作,综合后可以得到选择最优价值动作的概率和选择其他动作的概率为:

$$\pi(a | s) = \begin{cases} 1 - \varepsilon + \varepsilon/n, a^* = \arg \max Q(s, a) \\ \varepsilon/n, a^* \neq \arg \max Q(s, a) \end{cases}$$

(4)

其中 n 代表动作向量的个数,设计如下的贪婪因子计算式:

$$\varepsilon = \varepsilon_0 + \varepsilon_1 e^{-0.05h}$$

(5)

式中: ε_1 一初始设置的贪婪因子取值; ε_0 一设置贪婪因子可以衰减到的最小值; h 一训练回合数。

为规范智能体探索学习时的边界,还应当设定一个判断训练处于何时何种状态终止的标志,该标志可以包含被控对象的输出状态值、训练进行的时间或训练过程中出现的任何信号标志。添加此信号可在任何希望训练终止的情况出现时停止训练的功能。

基于 Q-Learning 方法的 PI 控制器参数自整定算法伪代码过程如下:

- (1) 定义训练参数
- (2) 建立状态行为策略库表格并初始化
- (3) for 每一训练回合至最大回合数
- (4) 随机初始化低速燃气轮机转速闭环回路初始状态

- (5) for 每一步至每回合最大探索步数
- (6) 随机数 $p = rand(0,1)$
- (7) If $p \geq \varepsilon$
- (8) 智能体随机选择执行动作
- (9) else
- (10) 智能体根据最优价值动作(贪婪策略)调整控制器参数
- (11) 结束
- (12) 根据智能体选择出的动作计算调整 PI 控制器参数
- (13) 在新的控制器参数下运行一个采样循环,并接收燃气轮机反馈
- (14) 依据反馈信息与奖励函数计算并得到本步即刻奖励
- (15) 根据 QL 算法更新状态动作策略库参数
- (16) 更新状态向量
- (17) 如果当前状态为终止状态则跳出当前循环
- (18) 结束本回合
- (19) 结束本次训练

伪代码可以清晰地展现整个算法的运行训练过程,主要实现了马尔科夫决策过程中的决策更新过程以及每次进行训练前的初始化过程,体现了贪婪策略在策略迭代过程中起到避免智能体陷入局部最优的作用。初始化的策略库是一张全为零值的表格,经过策略迭代后可得到对应目标转速与控制器参数下的动作价值。回合的含义为从燃气轮机转速状态波动或偏离到转速回稳的过程,回合结束的标志由算法开始时定义的仿真最大时长与燃气轮机转速稳定性共同决定。

2 强化学习控制算法实现

在设计中算法部分为强化学习算法,环境为采用 PI 控制器的燃气轮机转速控制回路,通过观测状态模块、奖励值计算模块和终止信号采集模块分别计算出输入智能体的状态、即时奖励与回合终止信号。

训练开始前,需要对强化学习算法中的必要参数进行赋值,在每个回合开始前需要对环境参数配

置进行初始化,被控三轴燃气轮机的转速安全范围在 3 400 r/min 以下,以保证燃气轮机的运行。表 2 为训练参数设置。

表 2 训练参数设置

Tab.2 Main parameter setting table for training

参数	数值
转速最小值 $n_{\min}$	2 500
转速最大值 $n_{\max}$	3 500
比例系数调整步长 P	0.1
积分系数调整步长 I	0.05
学习率 α	0.5
贪婪因子最小值 ε_0	0.1
贪婪因子衰减系数 ε_1	0.5
最小贪婪系数 $\varepsilon_{\min}$	0.1
衰减因子 γ	0.9
每回合最大步数	40
训练最大回合数	1 000
回合最大奖励平均值	500

其中,比例系数与积分系数调整步长可以根据控制细腻程度的需要在合理值范围内调整取值,分别取值为 0.1 和 0.05。衰减因子、贪婪因子及其衰减系数、最小贪婪因子均可根据实际训练的速度和效果进行调整。

训练开始前需要创建空白的 Q-table(状态动作价值 $Q(S^i,A^i)$ 全为零)作为初始化的状态动作策略库,表格大小为 $m \times n$,共 $m \times n$ 个状态动作价值,从 $Q(S^1,A^1)$ 到 $Q(S^j,A^i)$ 一直到 $Q(S^m,A^n)$ 。随着训练进行,此策略库会逐渐迭代至收敛。

将训练过程整体的结束标志设置为回合最大。奖励平均值大于 500 时,取近 30 回合内的奖励值平均值作为判断,如图 4 所示。由图 4 可知,训练开始时累计奖励值很小,这是因为在智能体开始训练时动作随机性强,处于对环境的探索阶段,选择到低奖励值的概率较大。随着训练进行,累计平均奖励值在逐渐增加,说明策略在逐渐趋近于收敛。在训练进行到累计平均奖励值达到 500 时停止,此时训练还未进行到 1 000 步,但已经达到结束条件,即训练终止,得到成熟的状态动作价值表作为动作价值策

略库,可以在燃气轮机的运行过程中出现转速失稳时对 PI 参数进行快速调整。

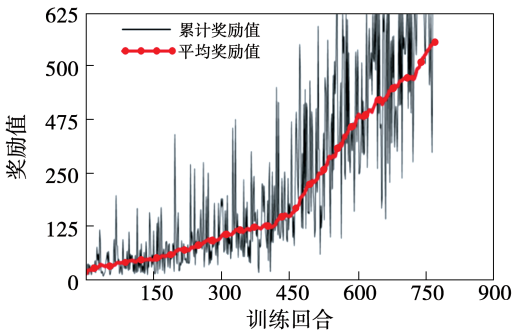


图 4 Q-Learning 算法可视化训练过程

Fig.4 Visualization training process of Q-learning algorithm

3 代码生成与半物理仿真验证

3.1 燃气轮机在环仿真

建立燃气轮机控制系统及主要控制策略模型,使用代码自动生成机器配套软件得到燃气轮机控制策略嵌入式 C 代码,从而建立了实验所需的硬件在环仿真系统(HIL)。之后将转速控制算法进行半物理闭环测试,得到算法整定过程并对其进行分析。

本实验基于微控制器 NXP MPC5674F 实现,此步骤需要在 Simulink 环境下进行详细的设置,在线标测参数过程需要用到 MeCa 软件,它是与 ECU-Coder 匹配的一款实时仿真标定测量软件,可以实现对微控制器底层程序的参数更改。将 Simulink 仿真步长设置为 0.005 s,总时长为 200 s。之后即可将 Simulink 程序通过代码自动生成编译到控制器中^[6]。

硬件在环仿真(Hardware-in-the-loop, HIL)是一种以全面、低成本和可重复的方式对嵌入式系统进行系统级测试的技术。这种技术最常用于嵌入式系统的开发,这些系统多用于不能在其实机操作环境中进行测试^[7]。

在 HIL 仿真过程中,控制系统的对象需要用具有发动机数学模型的实时机(Real-time machine, RTM)来表示。该 RTM 可用于自动控制系统的调控,以估计发动机的工作能力和检测发动机故障^[8-10]。

建立如图 5 所示的硬件在环仿真构架。在此构架中 TestBase 上位机使用 Simulink 对燃气轮机实时仿真模型基于 TLC 文件进行编译,生成 .dll 文件,用于上位机向 TestBase 下载实时仿真模型。TestBase 中的程控电源管理模块对微控制器提供 24V 供电,控制系统上位机利用 MeCa 软件将基于 ECUCoder 编译生成的燃气轮机控制策略嵌入式 C 代码中的 .mot 文件,与 .a21 文件烧写至微控制器,并且在上位机建立 MeCa 标定测量工程文件。微控制器此时可以实现对燃料喷射执行器的 PWM 驱动信号输出,TestBase 中的 TB6101 板卡采集执行器输出的电流信号经过信号调理后将所转化的电压信号通过 TB256 板卡发送至 TB4700 板卡,微控制器模拟燃气轮机燃油调节器的位置(开度)信号。至此整个硬件实现闭环回路。

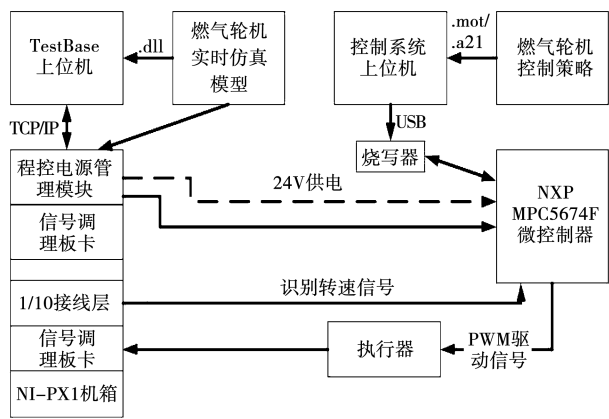


图 5 HIL 系统硬件回路图

Fig. 5 HIL system hardware loop diagram

将燃气轮机实时仿真模型下载至 TestBase 下位机中,需要先对模型进行编译,生成可以下载的 .dll 文件。然后使用 NI VeriStand 软件将其导入下位机,在 NI VeriStand 中建立控制平台,即可在半物理仿真中实现对转速的监测,以及其他必要功能如设定目标转速、停机等。至此已经完成实验所需软硬件环境。

3.2 燃气轮机控制策略闭环测试

在燃气轮机运行的目标转速区间 2 800 ~ 3 200 r/min 内,对其目标转速进行加减速,在多个转速下手动调整 K_p , K_i 参数,得到多种转速失稳情况,完成

对燃气轮机实际运行条件的较大范围模拟。自整定功能使能区间在图中可以清楚看到,且在控制策略介入下最终回稳。对测试结果进行分类讨论,区间 1,3 体现了主要由于 K_i 参数过大导致的转速出现震荡情况,区间 4,6 体现了由于 K_p 参数过大导致的转速失稳现象,区间 2,5 体现了在目标转速变化的情况下由于 K_p , K_i 参数不合适导致出现静差现象,如图 6 所示。

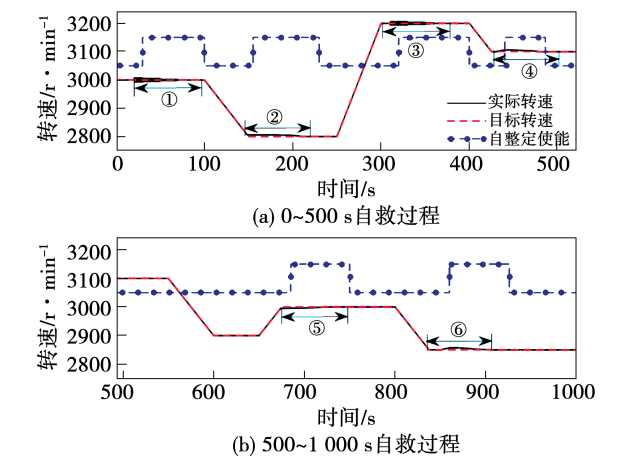


图 6 稳态工况下转速失稳自救过程

Fig. 6 Self-rescue process of speed instability under steady condition

3.2.1 K_p 参数过小及 K_i 参数过大自救过程

对图 6 的区间 1 进行局部放大处理,并以仿真时间对应的方式附上 K_p , K_i 参数调整过程图,如图 7 所示。在目标转速 3 000 r/min 的工况下,分别手动调整 K_p 和 K_i 参数至 0.5 和 0.9。此时燃气轮机的转速波动程震荡情况,无渐进回稳现象出现,判断其原因是 K_i 项参数过大,系统对燃气轮机运行转速进行采集,两个步长后得到误差大于判断自整定开启误差,此时算法调参触发,开始对 K_p , K_i 参数自整定。在整定过程中, K_p 参数在前 3 次调整过程中的动作为按步长 0.1 上升,停止 4 个步长后继续上升,这是因为在本次调整过程中 K_p 对转速影响较小,智能体在选择动作输出时有概率选择不调整 K_p 参数; K_i 参数在调整过程中一直保持以步长 0.05 下降的趋势,这是因为本次调整过程中引起转速失稳的主要原因

是 K_i 参数过大,所以智能体在输出动作时的方向性十分明确。最终调整结果为 $K_p = 1.1, K_i = 0.35$ 。

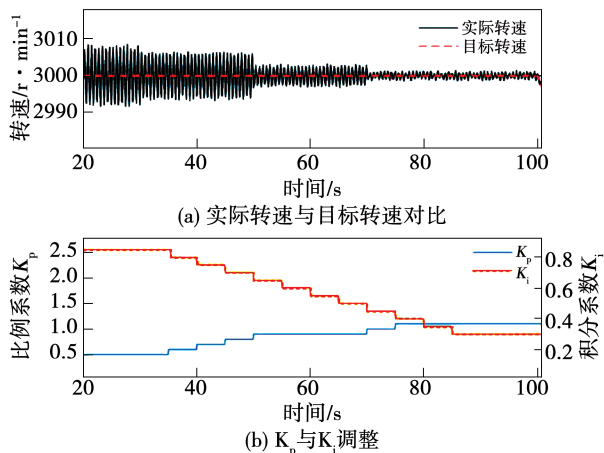


图 7 区间 1 内 K_p, K_i 参数调整过程

Fig. 7 K_p and K_i parameter adjustment process in interval 1

对图 6 的区间 3 进行局部放大处理,并以仿真时间对应的方式附上 K_p, K_i 参数调整过程图,如图 8 所示。在目标转速 3 200 r/min 的工况下,分别手动调整 K_p 和 K_i 参数至 0.4 和 2.7。此时燃气轮机的转速波动作震荡情况,无渐进回稳的现象出现,判断其原因是 K_i 项参数过大,系统对燃气轮机运行转速进行采集,两个步长后得到误差大于判断自整定开启误差,此时算法调参触发,开始对 K_p, K_i 参数自整定。在整定过程中 K_p 参数动作选择有上升和保持不变两种,这是因为在本次调整过程中 K_p 对转速影响较小,智能体在选择动作输出时有概率选择不调整 K_p 参数, K_i 参数在调整过程中一直保持步长 0.05 下降的趋势,这是因为本次调整过程中引起转速失稳的主要原因是 K_i 参数过大,所以智能体在输出动作时的方向性十分明确。最终调整结果为 $K_p = 1.1, K_i = 0.2$ 。

3.2.2 K_p 和 K_i 参数过小导致出现静差的自救过程

对图 6 的区间 2 进行局部放大处理,并在下方以仿真时间对应的方式附上 K_p, K_i 参数调整过程图,如图 9 所示。分别手动调整 K_p 和 K_i 参数至 0.6 和 0.05,此时燃气轮机的转速出现静差为 5 r/min 的情况,判断其原因是 K_i 项参数过小,系统对燃气

轮机运行转速进行采集,两个步长后得到误差大于判断自整定开启误差,此时算法调参触发,开始对 K_p, K_i 参数自整定。在整定过程中,对 K_p 和 K_i 参数的调整分别按步长 0.1 和 0.05 同时出现上升的动作。4 次调整周期后,最终调整结果为 $K_p = 0.9, K_i = 0.25$ 。

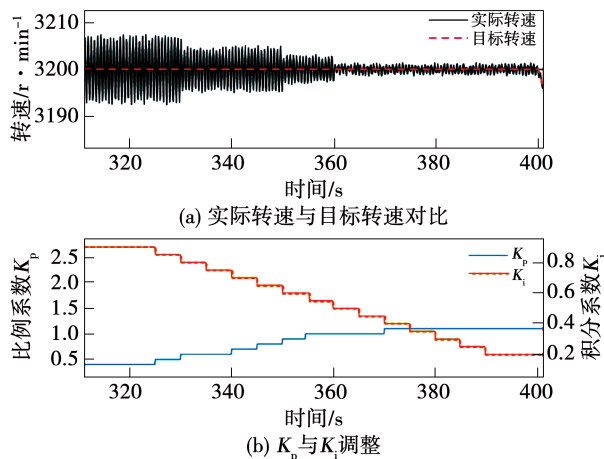


图 8 区间 3 内 K_p, K_i 参数调整过程

Fig. 8 K_p and K_i parameter adjustment process in interval 3

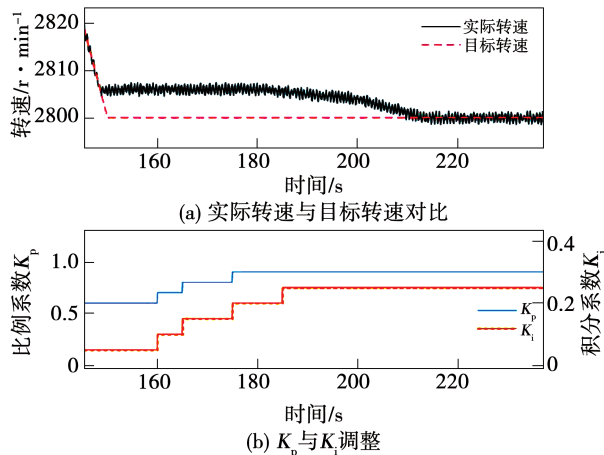


图 9 区间 2 内 K_p, K_i 参数调整过程

Fig. 9 K_p and K_i parameter adjustment process in interval 2

对图 6 的区间 5 进行局部放大处理,并在下方以仿真时间对应的方式附上 K_p, K_i 参数调整过程图,如图 10 所示。分别手动调整 K_p 和 K_i 参数至 0.4 和 0.05,此时燃气轮机的转速出现静差为 4 r/min

的情况。判断其原因是 K_i 项参数过小,系统对燃气轮机运行转速进行采集,两个步长后得到误差大于判断自整定开启误差,此时算法调参触发,开始对 K_p 、 K_i 参数自整定。在整定过程中,对 K_p 和 K_i 参数的调整分别按步长 0.1 和 0.05 同时出现上升的动作。4 次调整周期后,最终调整结果为 $K_p = 0.8$, $K_i = 0.15$ 。

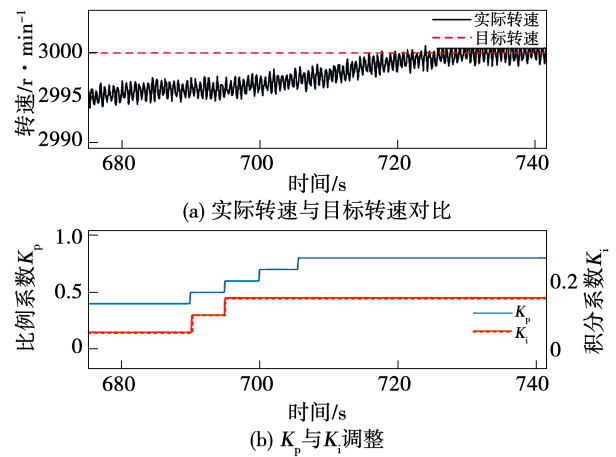


图 10 区间 5 内 K_p 、 K_i 参数调整过程

Fig. 10 K_p and K_i parameter adjustment process in interval 5

3.2.3 K_p 参数过大和 K_i 参数过小自救过程分析

对图 6 的区间 4 进行局部放大处理,并在下方以仿真时间对应的方式附上 K_p 、 K_i 参数调整过程图,如图 11 所示。在目标转速 3 100 r/min 的工况下,分别手动调整 K_p 和 K_i 参数至 2.7 和 0.2,此时燃气轮机的转速波动情况较大,转速曲线无渐进回稳的现象。判断其原因, K_i 项参数较小,系统对燃气轮机运行转速进行采集,两个步长后得到误差大于判断自整定开启误差,此时算法调参触发,开始对 K_p 、 K_i 参数自整定。在整定过程中, K_p 项参数持续降低, K_i 参数在 3 次连续升高后保持不变。最终调整结果为 $K_p = 1.1$, $K_i = 0.35$ 。

对图 6 的区间 6 进行局部放大处理,并在下方以仿真时间对应的方式附上 K_p 、 K_i 参数调整过程图,如图 12 所示。在目标转速 2 850 r/min 的工况下,分别手动调整 K_p 和 K_i 参数至 2.8 和 0.1,此时燃气轮机的转速波动情况较大,转速曲线无渐进回稳

的现象,判断其原因是 K_i 项参数较小,系统对燃气轮机运行转速进行采集,两个步长后得到误差大于判断自整定开启误差,此时算法调参触发,开始对 K_p 、 K_i 参数自整定。在整定过程中, K_p 项参数持续降低, K_i 参数在 3 次连续升高后保持不变。最终调整结果为 $K_p = 1.2$, $K_i = 0.25$ 。

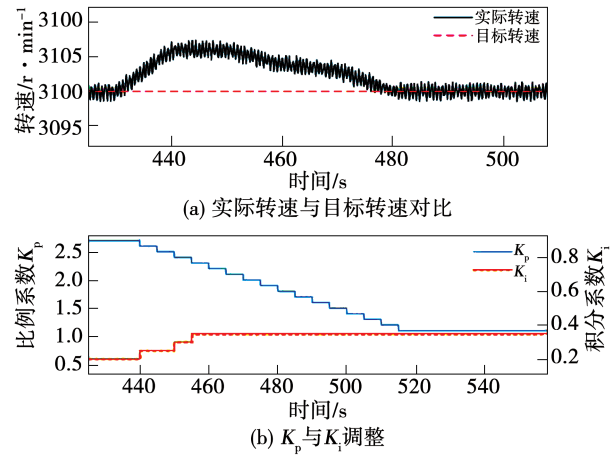


图 11 区间 4 内 K_p 、 K_i 参数调整过程

Fig. 11 K_p and K_i parameter adjustment process in interval 4

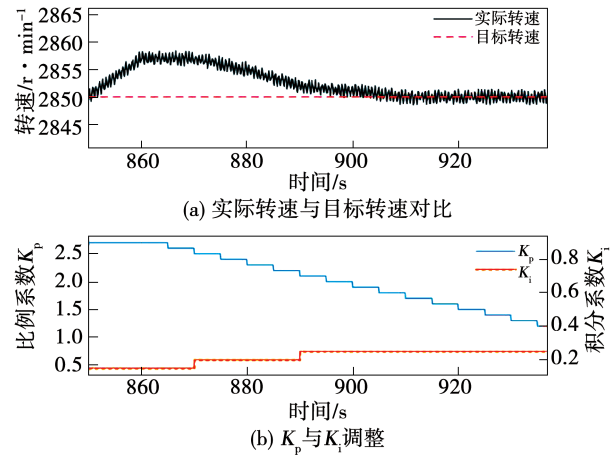


图 12 区间 6 内 K_p 、 K_i 参数调整过程

Fig. 12 K_p and K_i parameter adjustment process in interval 6

通过以上 3 种情况下的燃气轮机转速失稳现象的自救过程分析可知,基于强化学习的燃气轮机转速控制策略有很好的实际效果,在燃气轮机转速发生失稳现象时可以及时判断失稳情况,并且以直接

的决策输出相应对 K_p, K_i 参数值的调整动作,不再出现在输出动作时的参数变化升降反复,实现自救,算法稳定可靠。

4 结 论

针对燃气轮机非线性、时变性等复杂特点,转速的稳定控制是本文主要研究的重点。以新兴的人工智能算法为启发,本文成功设计了一种以 Q-Leaening 算法为基础的燃气轮机转速 PI 自整定控制器,通过自动代码生成技术以及快速控制原型完成了从 Simulink 环境到嵌入式代码的自动生成,最后通过搭建硬件在环系统实现对整个控制策略的实际效果验证。得到以下结论:

(1) 相比于常规的人工调参方法,本文设计的算法具有更强的便捷性以及更快的转速失稳自救反应速度,鲁棒性更强,其调参过程模拟人工,依据训练后的成熟状态动作策略库输出调参动作,对目标燃气轮机的参数整定具有更好的适应效果。

(2) 基于 HIL 系统的仿真试验,说明本文设计的 Q-Learning 算法调参的控制器在实际物理环境下同样具有很好的控制效果,证明了研究的价值以及实用性。

参考文献:

[1] WILLIAM O. Gas turbine control[J]. US,1948.

[2] JEFFS E. Gas turbine development[J]. Generating Power at High Efficiency,2008,37:57-86.

[3] XUE Yin-chun, SUN Jian-guo. A survey of gas turbine control technique[J]. Journal of Aerospace Power, 2005, 20 (6): 168-173.

[4] SANUSI I, MILLS A, TRODDEN P, et al. Reinforcement learning for condition-based control of gas turbine engines[C]//2019 18th European Control Conference (ECC),2019.

[5] LU Zhao-hui. Simulation and control of gas turbine[J]. Jet Engines (III),2014.

[6] KUPER R I. Automatic code generation [P]. US: US6357039 B1,2002.

[7] JUNG D, TSIOTRAS P. Modelling and hardware-in-the-loop simulation for a small unmanned aerial vehicle [C]//AIAA Infotech,2007.

[8] ZHANG Y, YIN C L, ZHANG J W. Matlab based human and hardware-in-loop simulation for the study on vehicle stability control [J]. Journal of Shanghai Jiao Tong University (Science Edition), 2006,11(4):498-505.

[9] YU Yue, ZHANG Tian-hong, SHENG Han-lin, et al. Experimental study on hardware-in-loop of rapid prototyping controller for micro gas turbine generator set[J]. Aeroengine,2019,45(1):70-75.

[10] ZHANG H, MING S, WENG S. The hardware-in-the-loop simulation study on the control strategy of gas turbine[C]//ASME International Mechanical Engineering Congress and Exposition. American Society of Mechanical Engineers,2002.

(姜雪梅 编辑)