

DOI:10.3969/j.issn.1001-4551.2023.07.013

基于遗传算法优化的深度强化学习-PI 空气舵伺服 系统控制策略*

洪子祺¹, 许文波^{2*}, 吕 晨¹, 欧阳权¹, 王志胜¹

(1. 南京航空航天大学 自动化学院, 江苏 南京 210016; 2. 北京精密机电控制
设备研究所 航天伺服驱动与传动技术研究室, 北京 100076)

摘要: 针对传统比例积分控制难以选定控制性能更好参数的问题, 以空气舵伺服系统为研究对象, 提出了一种基于遗传算法优化的强化学习-PI 的控制方法。首先, 建立了空气舵伺服系统的数学模型; 然后, 采用遗传算法优化了 PI 控制器的初始参数; 采用深度确定性策略梯度算法对当前 PI 控制器进行了实时整定, 从而实现对空气舵伺服系统进行位置指令控制的功能; 最后, 在 Simulink 中通过仿真分析, 对所采用的方法应用于空气舵伺服系统的效果进行了验证。研究表明: 改进的算法在参数扰动时, 具备一定的在线稳定性; 在空载情况下, 所需要的调节时间要小于遗传算法-PI、DDPG-PI 与传统 PI 算法, 至少缩短了 20%; 同时, 在负载情况下, 相比其他 3 种方法, 改进算法的波动幅值与负载结束后回到稳态时间至少缩短了 15%, 证明了所使用方法在空气舵伺服系统里的有效性。

关键词: 伺服系统; 比例积分 (PI) 控制器; 遗传算法; 深度确定性策略梯度算法; 参数优化; Simulink

中图分类号: TH-39; TJ765

文献标识码: A

文章编号: 1001-4551(2023)07-1071-08

Deep reinforcement learning-PI control strategy of air servo system based on genetic algorithm optimization

HONG Zi-qi¹, XU Wen-bo², LV Chen¹, OUYANG Quan¹, WANG Zhi-sheng¹

(1. School of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;

2. Laboratory of Aerospace Servo Actuation and Transmission, Beijing Institute of Precision

Mechatronics and Controls, Beijing 100076, China)

Abstract: Aiming at the problem that traditional proportional integral (PI) control was difficult to select parameters with better control performance, taking the air rudder servo system as the research object, a control method of reinforcement learning-PI based on genetic algorithm optimization was proposed. Firstly, the mathematical model of the air rudder servo system was established. Then, the initial parameters of PI controller were optimized by genetic algorithm. The current PI controller was adjusted in real time using the deep deterministic policy gradient (DDPG) algorithm to realize the position command control of the air rudder servo system. Finally, the effect of the method used in the air rudder servo system was verified in Simulink through simulation analysis. The results show that the improved algorithm has certain online stability when the parameters are perturbed. In the case of no load, the required adjustment time is less than that of genetic algorithm - PI, DDPG-PI and traditional PI algorithm, and it is increased by at least 20%. At the same time, in the case of load, the fluctuation amplitude of the improved algorithm is at least 15% better than that of the other three methods compared with the time to return to steady state after the end of load, which proves the effectiveness of the method used in the air rudder servo system.

Key words: servo system; proportional integral (PI) controller; genetic algorithm; deep deterministic policy gradient (DDPG) algorithm; parameter optimization; Simulink

收稿日期: 2022-11-07

基金项目: 航天伺服驱动与传动技术实验室开放基金资助项目 (LASAT-20210502)

作者简介: 洪子祺 (1999-), 男, 浙江杭州人, 硕士研究生, 主要从事电机伺服系统控制方面的研究。E-mail: hongziqi@nuaa.edu.cn

通信联系人: 许文波, 男, 博士, 研究员, 高级工程师。E-mail: webel@163.com

式中: i_d 为电流直轴的分量, A ; i_q 为交轴分量, A ; L_d 为定子电感的直轴分量, H ; L_q 为交轴分量, H ; ψ_d 为定子磁链的直轴分量, Wb ; ψ_q 为交轴分量, Wb ; ψ_r 为转子磁链, Wb 。

在同步旋转坐标系下, $i_d = 0$ 时电机的电磁转矩如下式所示:

$$T_e = i_q K_t \quad (2)$$

式中: T_e 为电机的电磁转矩, $N \cdot m$; K_t 为转矩常数, $N \cdot m/A$ 。

在实际的空气舵伺服系统当中, 由于机械结构存在一些非线性因素与传感器一般不会安装在舵面位置的问题, 若直接使用舵面摆角作为控制量会很难控制实际空气舵系统, 导致控制效果不佳, 影响整体系统的控制效果。因而笔者在建模过程中, 没有直接使用舵面摆角作为位置环的输入, 而是使用了电机转角作为输入。

笔者对于该模型的总体数学建模的运动方程如下:

$$\begin{cases} \frac{K_1 K_2}{K_1 + K_2} (\theta_m P_h - \theta_1 R) R - T_l - B \omega_1 = J_1 \dot{\omega}_1 \\ T_e - \frac{K_1 K_2}{K_1 + K_2} (\theta_m P_h - \theta_1 R) P_h = J_m \dot{\omega}_m \\ \theta_1 = \int \omega'_1 dt \\ \omega'_1 = F(\omega_1, x_d) \end{cases} \quad (3)$$

式中: ω'_1 为经过死区后的负载角速度, rad/s ; $F(\omega_1, x_d)$ 为 Simulink 里的死区模型; x_d 为死区宽度。

2 基于遗传算法优化的强化学习-PI 控制算法

学术界采用过许多控制算法用于伺服系统, 希望提升控制性能, 如鲁棒控制^[10-14]、自抗扰控制^[15-18]、模糊控制^[19-22]以及神经网络^[23-26]控制等。

为了让空气舵伺服系统能够精确快速地跟踪参考输入信号, 笔者将人工智能算法与传统 PID 算法^[27]相结合, 提出了一种基于遗传算法优化 PI 控制器初值的深度确定性策略梯度-PI (DDPG-PI) 的控制算法。

该控制算法如下式所示:

$$u = [e \int edt] \times [K'_p + \Delta K_p, K'_i + \Delta K_i] \quad (4)$$

式中: e 为电机摆角参考值与实际值的误差, rad ; u 为 PI 控制器的输出; θ_{ref} 为参考输入, rad ; θ_m 为空气舵伺服系统中电机的实际摆角, rad ; θ_1 为舵面实际摆角, rad ; (K_p, K_i) 为控制器遗传算法优化后参数; (K'_p, K'_i) 为 PI 控制器原始参数; $(\Delta K_p, \Delta K_i)$ 为 DDPG 算法

在线优化参数变化值。

2.1 遗传算法优化 PI 控制器参数

笔者采用遗传算法, 对人工选取的 PI 控制器的初值进行优化选取, 从而得到控制性能更好的 PI 控制器, 并采用 DDPG-PI 控制策略进行优化, 以追求更好的控制效果。

遗传算法优化流程图如图 2 所示。

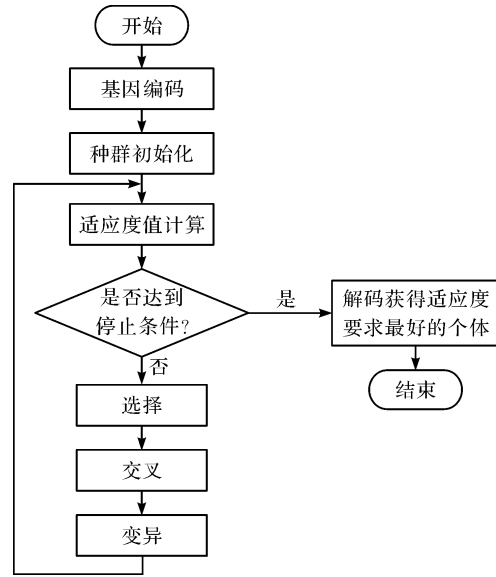


图2 遗传算法优化流程图

笔者采用的遗传算法优化 PI 控制器参数, 主要分为以下几个步骤:

1) 编码。将目标数据参数表示为基因型数据串结构, 用于遗传算法当中。笔者采用的是实值编码的方式, 原因在于实值编码适用精度较高, 且适用的范围较广;

2) 初始种群生成。在人工选取的 PI 控制器初值的基础上, 选取了合适的优化比例参数与积分参数的范围, 初始种群会在该范围内随机产生 n 个初始数据串, 表示对应 PI 控制器的参数值, 给予算法一个输入;

3) 适应度评估。在每次流程结束之后, 算法需要对种群的数据串进行适应度分析, 然后得到适应度最好的个体。

笔者所采用的适应度函数如下:

$$\text{fitness} = 0.99 \times \int |e(t)| dt + 0.005 \times t_s + 0.005 \times M_p \quad |e(t)| \leq 0.01 \quad (5)$$

式中: $e(t)$ 为 t 时刻参考值与实际值的误差, rad ; t_s 为在该个体数据作为 PI 控制器参数情况下输出的调节时间, s ; M_p 为超调量, $\%$ 。

可以看出: 当适应度值越小的时候, 对应种群个体的数据效果更好;

4) 选择。算法通过适应度值的选择,从当前种群中选取最优秀个体作为父代,以繁殖下一代个体;

5) 交叉。模拟自然界的繁殖特点,在一定交叉概率的情况下,算法交叉组合父代数据的特点,从而在保证种群的稳定性的情况下,可以得到性能更优秀的个体;

6) 变异。在较为低的概率条件下,用算法对个体数据串中某一个数据进行替换,避免在训练过程中局部最优情况的发生,使其最后可以得到更好的结果。

每次训练循环步骤 3) ~ 6), 当达到停止要求时,算法对得到的最优个体数据进行解码操作,得到优化后的 PI 控制器参数。

2.2 强化学习

强化学习^[28-31]是一种学习如何从状态映射到行为以使得获取的奖励最大的学习机制。它是一种需要不断在环境中进行实验,通过环境给予的反馈来不断优化状态-行为的对应关系。

强化学习的整个训练过程都是基于马尔科夫决策 (Markov decision process, MDP) 来进行,其核心在于下一步系统状态只与当前状态、当前状态所将要采取的行动有关。

笔者的研究对象—空气舵伺服系统运行状态满足马尔科夫性质,适用于与强化学习智能体不断交互的强化学习过程。

强化学习智能体与作为环境模型的空气舵伺服系统的强化学习流程,如图 3 所示。

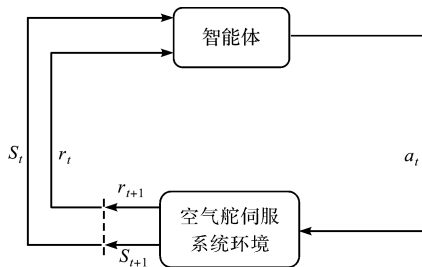


图 3 空气舵伺服系统的强化学习流程

在强化学习训练的每一回合中, t 时刻时,智能体从空气舵伺服系统中观测到系统的当前状态为 S , 并通过智能体中的策略函数得到智能体的动作行为 a 作用回环境当中,同时得到奖励信号 r_t , 最后环境给出新的状态 S_{t+1} 进入新的循环当中。

强化学习的最终目标是最大化每一个状态的期望回报,即未来的奖励总和,这样便能得到最后最好的优化结果。

2.3 DDPG-PI 控制策略

DDPG 算法是强化学习领域的一种经典算法。该算法结合神经网络学习的优点,适用于空气舵伺服系统这种连续动作环境的在线训练。

DDPG 算法满足执行者-评价者 (Actor-Critic) 结构。笔者所采用的 DDPG-PI 控制算法首先通过策略初始化状态变量 S_t 。其中, S_t 如下式所示:

$$S_t = [e \int edt \quad \theta_m]^T \quad (6)$$

在初始条件下,执行者通过策略函数 $\mu(S)$ 与探索噪声 N , 可以在智能体中得到对应的动作输出 a_t 。

在该方法中,空气舵伺服系统位置环的 PI 控制器对应增益代表着执行者动作 a_t 。

执行者动作的求取过程如下:

$$a_t = \mu(S | \lambda^Q) + N \quad (7)$$

式中: λ^Q 为评价者网络 $Q(S, a | \lambda^Q)$ 中的权重。

接着,空气舵伺服系统执行动作 a_t , 并返回出奖励值 r_t 与下一时刻的状态 S_{t+1} 到智能体中。

执行者网络将这一转换过程保存下来。待系统完成采样之后,目标网络的 Q 与 μ 值用于计算得到目标网络值 y_t , 如下式所示:

$$y_t = r_t(S_t, a_t) + \gamma Q(S_{t+1}, \mu(S_t | \lambda^\mu) | \lambda^2) \quad (8)$$

式中: λ^μ 为执行者网络 $\mu(S | \lambda^\mu)$ 的权重; γ 为折扣因子。

然后,执行者通过最小化损失函数 $L(\lambda)$, 以更新评价者网络参数。 $L(\lambda)$ 求取如下式所示:

$$L(\lambda^2) = E_{(S,a)} [(Q(S_t, a_t | \lambda^2) - y_t)^2] \quad (9)$$

执行者网络则通过梯度函数 $\nabla_{\lambda^\mu} J^{\lambda^\mu}$ 来更新网络参数。该函数如下式所示:

$$\nabla_{\lambda^\mu} J^{\lambda^\mu} \approx E [\nabla_a Q(S, a | \lambda^Q) |_{a=\mu(S)} \nabla_{\lambda^\mu} \mu(S | \lambda^\mu)] \quad (10)$$

在空气舵伺服系统中,出于快速性与稳态性的考虑,以及实际反馈值与误差对系统控制的影响,笔者采用了如下的奖励函数:

$$r_t = r_{1t} + r_{2t} \quad (11)$$

$$r_{1t} = -|e(t)| \quad (12)$$

$$r_{2t} = \begin{cases} -40u_1 - 40u_2, & t < 0.1 \text{ s} \\ -40u_1 - 10u_2, & 0.1 \text{ s} < t < 0.2 \text{ s} \\ -10u_1 - 10u_2, & t > 0.2 \text{ s} \end{cases}$$

$$u_1 = \begin{cases} 1 & \theta_m > 1.003 \frac{R}{P_h} \theta_{\text{ref}} \\ 0 & \text{else} \end{cases}$$

$$u_2 = \begin{cases} 1 & \theta_m < 0.997 \frac{R}{P_h} \theta_{\text{ref}} \\ 0 & \text{else} \end{cases} \quad (13)$$

3 仿真分析

3.1 仿真介绍

笔者采用 MATLAB/Simulink 来搭建空气舵机电

伺服系统的仿真模型进行实验,并在此基础上用遗传算法优化的 DDPG-PI 算法取代多环控制中的位置环的传统 PI 控制,然后进行仿真对比,以此来验证文中算法对于提升控制性能的有效性。

仿真模型图如图 4 所示。

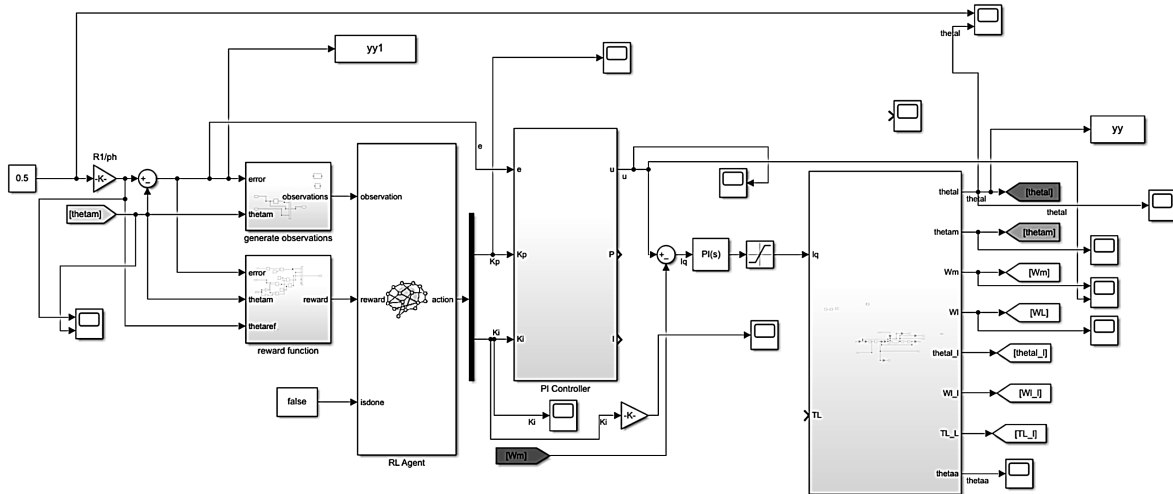


图 4 仿真模型图

笔者所选取的位置环 PI 的 P 与 I 初值分别为 10, 0.1。

其中,空气舵伺服系统的相关参数如表 1 所示。

表 1 空气舵伺服系统相关参数

序号	参数/单位	数值
1	电机转动惯量/(N·m ²)	2.4 × 10 ⁻⁴
2	负载转动惯量/(N·m ²)	0.103
3	机电作动器刚度/(N·m)	3.384 × 10 ¹⁵
4	机电作动器安装 支承刚度/(N·m)	1.28 × 10 ⁷
5	滚珠丝杠导程/(m/rad)	6.37 × 10 ⁻⁴
6	转矩系数/(Nm/A)	0.16
7	空气舵摇臂长度/m	0.073

在人工选取 PI 控制器参数之后,笔者采用遗传算法对 PI 控制器的参数进行优化,以获得控制效果更好的 PI 控制器参数。

遗传算法参数如表 2 所示。

表 2 遗传算法参数

参数	数值
种群大小	30
进化代数	100
交叉率	0.9
变异率	0.1

笔者在优化好的 PI 控制器基础上,采用 DDPG 算法对 PI 控制器进行动态参数优化。

DDPG 超参数设置如表 3 所示。

表 3 DDPG 超参数设置

超参数	数值
折扣因子	1
评价者学习率	0.001
执行者学习率	0.000 1
最小批量	128
噪声标准差	0.6
噪声标准差衰减率	1 × 10 ⁻⁵

3.2 仿真结果

笔者采用遗传算法,对人工选取的 PI 控制器参数进行迭代优化,观察每代种群的最佳适应度,最后得到 PI 控制器优化结果的最优解。

遗传算法适应度曲线如图 5 所示。

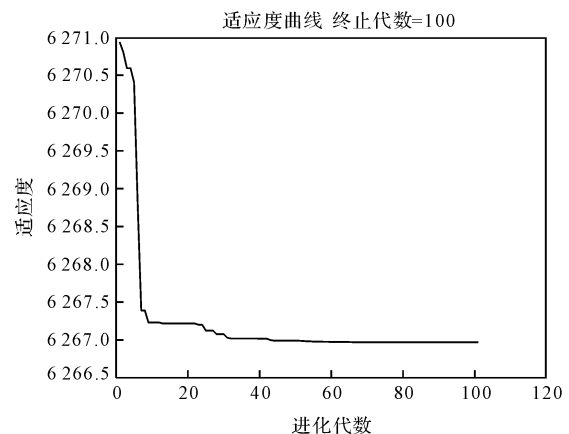


图 5 遗传算法适应度曲线

观察图 5 可以发现:遗传算法训练在结束时适应

度值已经趋于稳定;算法训练结束之后,PI 控制器初值 P 优化至约为 19.84, I 初值优化至约为 0.025。

在优化好的 PI 控制器的基础上,笔者采用 DDPG 智能体对其进行强化学习训练,可以得到评价者的价值评估曲线、每回合训练的实际奖励曲线以及平均训练奖励曲线。

DDPG 智能体回报曲线如图 6 所示。

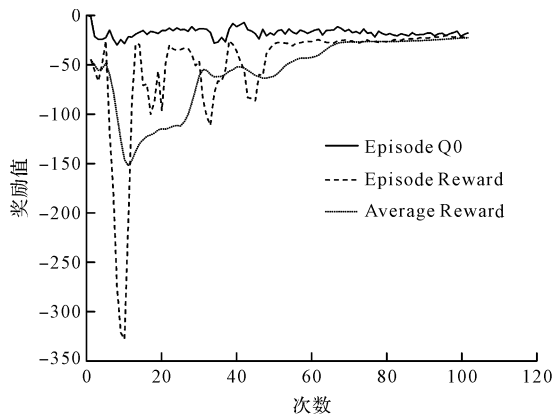


图 6 DDPG 智能体回报曲线

观察图 6 可以看出,DDPG 训练对于奖励值的估计已经比较接近实际奖励值,训练后期实际奖励已经趋于稳定。

笔者在 0.5 rad 摆角指令下,使系统在 0.1 s 时,出现对系统转动惯量与作动器相关刚度的参数变大 5 倍与变小 5 倍的摄动。

系统参数出现摄动时响应如图 7 所示。

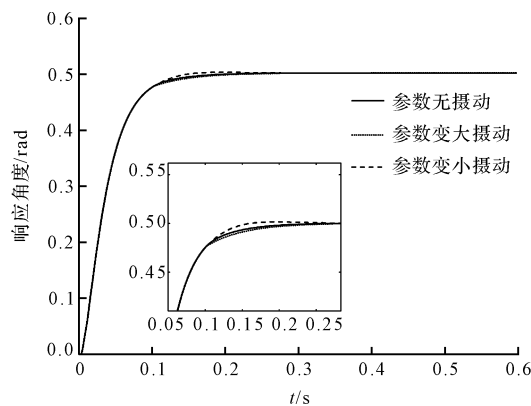


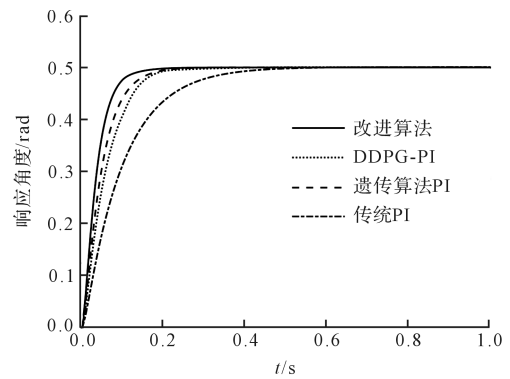
图 7 系统参数出现摄动时响应

观察图 7 可以看出,DDPG-PI 算法在系统参数摄动情况下,仍能保证系统控制的基本性能。

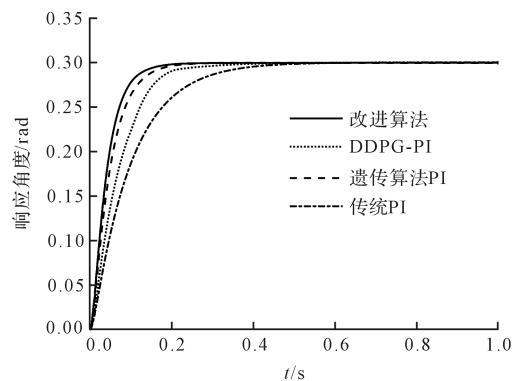
因此,在参数摄动的情况下,DDPG-PI 算法存在良好的在线稳定性。

在空载情况下,笔者对空气舵伺服系统分别给予 0.5 rad,0.3 rad 的摆角指令信号,并比较改进后算法、传统 PI、遗传算法-PI 及 DDPG-PI 算法的效果。

空载情况下响应如图 8 所示。



(a) 0.5 rad 响应图



(b) 0.3 rad 响应图

图 8 空载情况下响应

在空载情况下,笔者对 4 种算法的控制效果进行分析,得到其 4 种方法的调节时间对比,如表 4 所示。

表 4 4 种方法的调节时间对比

控制方法	0.5 rad 响应 调节时间/s	0.3 rad 响应 调节时间/s
改进后算法	0.104 1	0.113 4
DDPG-PI	0.151 2	0.175 3
遗传算法-PI	0.133 5	0.133 4
传统 PI	0.280 6	0.280 6

由表 4 可以看出:相比于遗传算法-PI、传统 PI 算法及 DDPG-PI 算法,笔者所使用的算法有着更短的调节时间;改进后算法在两种控制指令的情况下,调节时间均小于 0.12 s,相比于其他 3 种算法,至少提升了 20%,控制效果提升明显。

笔者在给予 0.5 rad 摆角指令的情况下,在 0.6 s 时,给予时长为 0.5 s 的 200 Nm 的负载干扰力矩。负载干扰响应如图 9 所示。

在负载实验情况下,笔者对 4 种方法进行对比,可以得到负载干扰下控制效果对比,如表 5 所示。

根据表 5 可知:在负载干扰情况下,改进后算法相比于其他 3 种算法,在负载干扰时,系统响应所受到的波动影响更小;在负载干扰结束后,系统响应能更快地回到稳态;

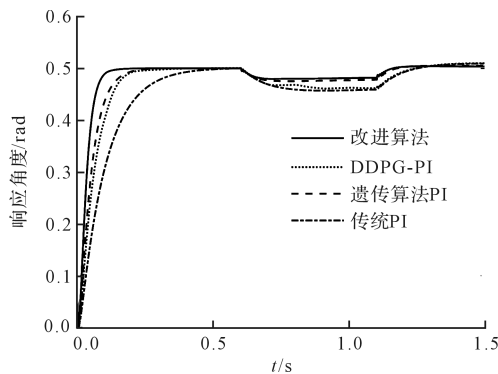


图9 负载干扰响应

表5 负载干扰下控制效果对比

控制方法	200 Nm 负载干扰 最大波动幅值/rad	200 Nm 负载干扰结束 后的稳态时间/s
改进后算法	0.020 2	0.058
DDPG-PI	0.037 3	0.142 9
遗传算法-PI	0.022 6	0.074 5
传统 PI	0.042 5	0.148 2

改进后算法在该干扰下系统响应波动率小于4%,同时在负载干扰结束后,系统响应回到稳态值的时间可以达到0.06 s以下。

相比于其他3种算法,改进后算法至少都优化了15%,说明改进后算法在负载干扰的情况下,控制效果也更加优秀。

从仿真分析结果可以说明,笔者所提出的算法对于空气舵伺服系统来说具有良好的控制性能,其有效性与优越性得到了验证。

4 结束语

笔者以空气舵伺服系统作为研究对象,在传统PI控制的基础上,采用遗传算法对PI控制器的初始参数进行优化,得到新的PI控制器,再采用DDPG算法对新的PI控制器参数进行实时调节。

仿真分析结论如下:

1)改进后算法在系统参数摄动的情况下,控制效果波动不明显,证明算法具备一定的在线稳定性;

2)相比于传统的PI控制、遗传算法优化的PI算法及DDPG-PI算法,基于遗传算法改进的DDPG-PI算法在空气舵伺服系统空载时,通过不同的指令信号控制,都能使其调节时间有明显降低,且达到0.12 s以下,至少提升了20%,控制性能优秀;

3)在受到200 Nm负载干扰时,相比于其他3种算法,改进后算法系统响应有着更小的波动幅值,波动率小于4%;同时,在负载干扰结束后能更快回到稳态值,可以达到0.06 s以下,都至少优化了15%。该结

果证明,笔者的控制算法有着更好的控制性能。

笔者在智能控制方法与传统控制方法相结合的基础上,加入了扩展性优秀的遗传算法以进一步提升控制器性能,这也为未来伺服系统结合一些新兴算法的研究提供了思路与想法。

同时,未来笔者会将改进算法用于实验测试,为实际应用做准备。

参考文献(References):

- [1] GHEISARNEJAD M, KHOOBAN M H. An intelligent noninteger PID controller-based deep reinforcement learning: implementation and experimental results [J]. IEEE Transactions on Industrial Electronics, 2021, 68 (4): 3609-3618.
- [2] QI X W, LUO Y D, WU G Y, et al. Deep reinforcement learning enabled self-learning control for energy efficient driving [J]. Transportation Research Part C: Emerging Technologies, 2019, 99(1): 67-81.
- [3] 张茂盛,段杰,肖息,等.基于深度强化学习-PI控制的机电作动器控制策略[J].应用科技,2022,49(4):5.
- [4] 张德胜,张奇,赵睿杰,等.基于RBF神经网络和NSGA-II算法的海水淡化高压泵多工况优化设计[J].排灌机械工程学报,2022,40(12):1189-1196.
- [5] 刘延飞,彭征,王艺辉,等.基于改进的遗传算法的有刷直流电机PID参数整定[J].计算机应用,2022,42(5):8.
- [6] 王旭,蒋奇.基于改进遗传算法的故障电液伺服系统控制方法研究[J].机床与液压,2022,50(4):167-172.
- [7] 许文波,焦玮玮,潘龙.基于极点配置的PMSM电流环PI控制器设计[J].航天控制,2021,39(1):74-80.
- [8] 刘春庆,姜玉峰,李永宏,等.空气动力控制伺服系统位置控制精度研究[J].机电工程技术,2020,49(5):103-105.
- [9] CHEN S Y, HUNG Y H, GONG S S. Speed control of vane-type air motor servo system using proportional-integral-derivative-based fuzzy neural network [J]. International Journal of Fuzzy Systems, 2016, 18(6): 1065-1079.
- [10] 梅鲁海,刘哲纬.基于开环补偿与鲁棒控制的电液位置伺服加载系统研究[J].机电工程,2022,39(1):59-64,86.
- [11] 邢科新,倪伟琦,何德峰.荷载不确定移动机器人视觉伺服系统鲁棒预测控制[J].控制理论与应用,2022,39(2):327-335.
- [12] 王焯波,刘家辉,卫健,等.机电伺服系统积分鲁棒自适应控制[J].科技创新与应用,2021,11(23):7-11.
- [13] WANG D, ZHAO D, GONG M, et al. Research on robust model predictive control for electro-hydraulic servo active suspension systems [J]. IEEE Access, 2018, 6(10): 3231-3240.
- [14] RAHMAN M, RIAZ M T, MAHMUD M, et al. The Prescribed fixed structure intelligent robust control of an electrohydraulic servo system [J]. Mathematical Problems in Engineering, 2022, 1(1): 1-8.

- [15] 吴凌华,神英洪,王 静,等. 电液位置伺服系统降阶自抗扰控制[J]. 机床与液压,2022,50(15):175-180.
- [16] 黄小津,胡超迪,魏 巍,等. 直线光栅伺服系统自抗扰控制仿真分析与实验研究[J]. 机械设计与制造,2022,376(6):109-113.
- [17] LIU R J, WU M, LIU G P, et al. Active disturbance rejection control based on an improved equivalent-input-disturbance approach[J]. IEEE./ASME. Transactions on Mechatronics,2013,18(4):1410-1413.
- [18] WANG X J, LIU M Z, CHEN S, et al. Predictive function and sliding model controller of continuous rotary electro-hydraulic servo motor applied to simulator[J]. Journal of Jilin University: Engineering and Technology Edition, 2019,1(1):1-7.
- [19] 孙晓林,沈宏亮,杨立奎,等. 基于模糊控制的三电机伺服控制系统特性研究[J]. 电子设计工程,2022,30(1):117-120,126.
- [20] 鹿山山,王志胜,欧阳权,等. 飞机机电作动器模糊 PI-ANN 复合控制仿真[J]. 兵器装备工程学报,2021,42(10):158-164.
- [21] MU S, GOTO S, SHIBATA S, et al. Intelligent position control for pneumatic servo system based on predictive fuzzy control[J]. Computers & Electrical Engineering, 2019,75(1):112-122.
- [22] WANG L, TIAN M, GAO Y. Fuzzy Self-adapting PID Control of PMSM Servo System[C]//Electric Machines & Drives Conference, 2007. IEMDC'07. Antalya: IEEE, 2007:860-863.
- [23] 李瑞琦,边火丁,杨树炳,等. 基于 RBF 神经网络的永磁同步电机控制系统[J]. 轻工机械,2022,40(4):52-56.
- [24] 丁 豪,于海生,孟祥祥,等. 永磁同步电机的神经网络滑模与预测控制[J]. 控制工程,2022,4(1):1-10.
- [25] YAO Z K, YAO J Y, YAO F Y, et al. Model reference adaptive tracking control for hydraulic servo systems with nonlinear neural-networks[J]. ISA transactions,2020,100(1):396-404.
- [26] LI G, YUN Y. Flexible manipulator position control based on rbf neural network[J]. Journal of Hunan University of Technology,2014,28(3):6.
- [27] 韩镇锚. 永磁同步电机直驱式位置伺服系统控制策略研究[D]. 南京:南京航空航天大学自动化学院,2017.
- [28] PASSALIS N, TEFAS A. Continuous drone control using deep reinforcement learning for frontal view person shooting[J]. Neural computing and applications, 2020,32(9):4227-4238.
- [29] 苏萌韬,曾 碧. 基于渐进式神经网络的多任务强化学习算法[J]. 机电工程技术,2022,51(11):21-25.
- [30] 张继明,张 淳,胡金双. 内高压胀形设备的伺服控制[J]. 机械设计,2021,38(S1):313-316.
- [31] 董星宇,唐开强,傅汇乔,等. 基于强化学习的六足机器人动态避障研究[J]. 传感器与微系统,2022,41(1):19-23.

本文引用格式:

洪子祺,许文波,吕 晨,等. 基于遗传算法优化的深度强化学习-PI 空气舵伺服系统控制策略[J]. 机电工程,2023,40(7):1071-1078.

HONG Zi-qi, XU Wen-bo, LV Chen, et al. Deep reinforcement learning-PI control strategy of air servo system based on genetic algorithm optimization[J]. Journal of Mechanical & Electrical Engineering, 2023,40(7):1071-1078.

《机电工程》杂志:<http://www.meem.com.cn>

(上接第 987 页)

- [16] SHAO Si-yu, MCALEER S, YAN Ru-qiang, et al. Highly accurate machine fault diagnosis using deep transfer learning[J]. IEEE. Transactions on Industrial Informatics, 2018,15(4):2446-2455.
- [17] LI Yan-ting, JIANG Wen-bo, ZHANG Guang-yao, et al. Wind turbine fault diagnosis based on transfer learning and convolutional autoencoder with small-scale data[J]. Renewable Energy,2021,171(3):103-115.
- [18] GUO Jian-wen, WU Jia-peng, ZHANG Shao-hui, et al. Generative transfer learning for intelligent fault diagnosis of the wind turbine gearbox[J]. Sensors,2020,20(5):1361.
- [19] WEN Long, LI Xin-yu, GAO Liang. A transfer convolutional neural network for fault diagnosis based on ResNet-50[J]. Neural Computing and Applications,2020,32(10):6111-6124.
- [20] YAN Ru-qiang, GAO R X, CHEN Xue-feng. Wavelets for fault diagnosis of rotary machines: a review with applications[J]. Signal Processing,2014,96(3):1-15.
- [21] RIOUL O, VETTERLI M. Wavelets and signal processing[J]. IEEE. Signal Processing Magazine,1991,8(4):14-38.
- [22] RAWAT W, WANG Zeng-hui. Deep convolutional neural networks for image classification: a comprehensive review[J]. Neural Computation,2017,29(9):2352-2449.
- [23] HE Kai-ming, ZHANG Xiang-yu, REN Shao-qing, et al. Deep Residual Learning for Image Recognition [C]// Proceedings of the IEEE. Conference on Computer Vision and Pattern Recognition,2016:770-778.
- [24] HUANG Gao, HUANG Guang-bin, SONG Shi-ji, et al. Trends in extreme learning machines: a review[J]. Neural Networks,2015,61(1):32-48.
- [25] Case Western Reserve University. (Aug. 2020). Case Western Reserve University (CWRU) Bearing Data Center. [Online]. Available: <https://csegroups.case.edu/bearingdatacenter/pages/download-data-file/>.