

引用格式: 李雅楠, 杨森, 李晓峰. 基于强化学习的低速转动 PMSM 高性能伺服控制 [J]. 电子科技大学学报, 2025, 54(2): 266-273.
LI Y N, YANG S, LI X F. High-performance servo control of PMSM under low-speed rotation conditions based on reinforcement learning[J]. Journal of University of Electronic Science and Technology of China, 2025, 54(2): 266-273.

基于强化学习的低速转动 PMSM 高性能伺服控制



李雅楠, 杨森, 李晓峰*

(电子科技大学 航空航天学院, 成都 611731)

摘要: 实现对永磁同步电机 (PMSM) 在低速转动工况下的高精度伺服控制是建立高性能星载激光通信链路的前提, 其技术关键是对永磁同步电机转动状态的非线性特性进行精确描述。该文设计了一种具有能够处理连续动作空间特点的 DDPG 非线性控制器, 采用梯度下降法分别训练评价神经网络和动作神经网络, 实现了对非线性映射的精确拟合。Simulink 仿真结果表明: 和传统的比例-积分线性控制器相比较, DDPG 控制器在跟踪参考低速信号时响应时间和稳定时间更短、跟踪误差更小; 在施加扭矩时 q 轴电流响应更快, d 轴电流波动更小, 低速工况条件下的 PMSM 伺服控制性能得到了有效提高。

关键词: 永磁同步电机; 低速转动; 深度确定性策略梯度; 非线性; 伺服控制性能

中图分类号: TM3 文献标志码: A DOI: 10.12178/1001-0548.2024028

High-performance servo control of PMSM under low-speed rotation conditions based on reinforcement learning

LI Yanan, YANG Sen, and LI Xiaofeng*

(School of Astronautics and Aeronautics, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: Achieving high-precision servo control performance of permanent magnet synchronous motors (PMSM) under low-speed rotation conditions is a prerequisite for establishing high-performance satellite laser communication links. The key technology lies in the accurate description of the nonlinear characteristics of the PMSM rotational state. This paper designs a deep deterministic policy gradient (DDPG) nonlinear controller capable of handling continuous action spaces, utilizing gradient descent methods to train both the critic and actor neural networks separately, thereby achieving precise fitting of nonlinear mappings. Through Simulink simulations, the results demonstrate that compared to traditional proportional-integral linear controllers, the DDPG control exhibits shorter response and settling times, smaller tracking errors when following reference low-speed signals; it also shows faster q -axis current response and smaller d -axis current fluctuations when torque is applied, effectively improving the PMSM servo control performance under low-speed operating conditions.

Key words: permanent magnet synchronous motor; low-speed rotation; deep deterministic policy gradient; nonlinearity; servo control performance

与直流电机和感应电机相比, 永磁同步电机 (permanent magnet synchronous motor, PMSM) 具有免维护、高效率、大惯量比等优点^[1], 已在星地激光通信系统中得到应用^[2]。在中高轨卫星对地光通信链路信标光捕获、对准及跟踪阶段以及建链后的通信阶段, 实现对 PMSM 在低速转动工况下的高精度伺服控制是通信端机之间视轴中心对准的必

要条件^[3], 是建立和维持高性能星载激光通信链路的前提。由于机械轴系预紧产生的摩擦力矩具有强烈的非线性特性^[4], 低速工况条件下 PMSM 伺服控制容易导致性能变差, 表现为跟踪参考速度信号时响应时间和稳定时间较长且跟踪误差较大; 施加扭矩时 q 轴电流响应较慢, d 轴电流波动较大等现象。

收稿日期: 2024-02-02

作者简介: 李雅楠, 博士生, 主要从事通信与信息系统方面的研究。

*通信作者 E-mail: lxf3203433@uestc.edu.cn

PMSM 的电气元件响应速度比机械元件快, 是典型的具有复杂电气模型的双时间尺度系统^[5], 对于这样复杂的呈现非线性特征的电气系统, 在低速运行以及参数不确定的实际应用场景中, 采用传统的比例积分 (proportion integral, PI) 控制方法, 能达到的外环转速跟踪精度有限^[6]。为解决低速运行条件下的高精度伺服控制问题, 文献 [7] 设计了扩展卡尔曼滤波以提高速度估计的精度。文献 [8] 设计了基于改进滑模控制器的矢量控制系统, 可以使 PMSM 拥有更好的动态性能和抗扰动能力。文献 [9] 通过在二次电流基准上叠加适当的补偿信号来改进传统的 PMSM 控制器, 以减少低速时的速度波动。文献 [10] 提出了一种期望补偿自适应鲁棒电流控制器和一种基于扰动观测器的速度控制器, 能够实现定子电流和转子转速的鲁棒调节, 同时可估计模型参数。文献 [11] 提出了一种利用模糊逻辑方法逼近系统中的非线性函数和干扰的新方法, 保证系统在有限时间内的跟踪误差小于指定精度, 并通过 Lyapunov 方法分析了闭环系统的稳定性。然而, 上述控制方法通常都忽略了最优性, 不能保证闭环系统的响应性能 (如暂态期间更少的超调和更小的波动等)。

强化学习 (reinforcement learning, RL) 不依赖于被控对象的数学模型和内部信息, 它可以在无系统信息的情况下制定和改进控制规律^[12-13]。通过使用 Q 函数来表示控制策略执行的预期累积奖励, 根据最优目标进行实时决策^[14-15]。

基于 RL 方法具有适用于存在大量不确定性、脉动干扰和未知系统模型的 PMSM 控制的特点, 所以基于 RL 的 PMSM 控制方法在近年得到了广泛研究。文献 [16] 基于 RL 实现了离线异步电机参数辨识。文献 [17] 和文献 [18] 提出了针对不确定动力学的最优转矩控制设计, 实现快速的电流响应和较小的超调。文献 [19] 提出了一种基于 Actor-Critic 标识网络的最优控制方法, 以获得无参数不确定性和复合干扰的鲁棒跟踪性能。但是, 目前鲜见低速运行工况下基于 RL 的 PMSM 控制技术的相关研究报道。

围绕提高低速运行工况条件下的 PMSM 伺服控制性能的设计目标, 本文设计了一种低速运行条件下基于 RL 的 PMSM 电流环控制器。采用 LuGre 模型描述低速运行条件下摩擦力矩的强非线性和脉动性; 将面向磁场定向控制 (field oriented control, FOC) 电流环中 dq 轴的 PI 控制器替换为 RL 控制器, 构建 RL 训练环境, 引入深度确定性策略梯度 (deep deterministic policy gradient, DDPG) 决策机

制, 对 Actor 和 Critic 网络中的参数进行训练。在 MATLAB/Simulink 环境中采用数值仿真方法验证了低速运行工况下基于 DDPG 控制器提升 PMSM 伺服控制性能 (快速响应、稳定性及电流跟踪误差等) 的有效性。

1 PMSM 低速运行数学模型

PMSM 的基本部件是永磁极转子和三相绕组固定定子。当三相交流电源作用于定子绕组时, 会产生旋转磁场。转子上的磁极与旋转磁场相互作用, 产生电磁转矩, 激励转子同步旋转。PMSM 的原理如图 1 所示。

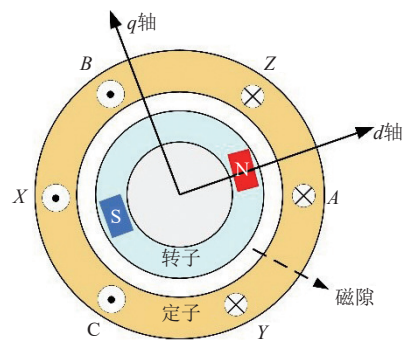


图 1 PMSM 旋转示意图

假设忽略定、转子铁心磁阻, 不计涡流和磁滞损耗; 磁体产生的励磁磁场和三相绕组产生的电枢反应磁场在气隙中均为正弦分布, 那么在转子同步旋转 dq 坐标系中, PMSM 的电气数学模型为:

$$\begin{cases} u_d = R_s i_d + L_d \frac{di_d}{dt} - \omega L_q i_q \\ u_q = R_s i_q + L_q \frac{di_q}{dt} + \omega(L_d i_d + \psi) \\ T_e = \frac{3}{2} P_0 [\psi i_q + (L_d - L_q) i_d i_q] \end{cases} \quad (1)$$

式中, u_d 、 u_q 分别为定子电压在 d 、 q 轴的分量; i_d 、 i_q 分别为定子电流在 d 、 q 轴的分量; L_d 、 L_q 分别为 d 、 q 轴电感的分量; R_s 为定子电阻; P_0 为电机极对数; ω 为电角速度; T_e 为电磁转矩; ψ 为永磁体产生的磁链。当 d 轴电流 i_d 为 0 时, d 轴不再对转矩有贡献, 仅通过改变 q 轴电流 i_q 就可改变电机的转矩。

对于 PMSM 伺服系统, 非线性摩擦、未建模不确定性、测量噪声和外界干扰等未知非线性因素会严重影响系统的动态性能。其中, 传动齿轮产生的摩擦力矩是降低系统性能的主要因素^[20]。

LuGre 摩擦模型能较为完善地描述摩擦的动态和静态特性, 其表现为在 2 个物体的微观接触面通

过弹性刚毛（具有刚度的硬毛）接触，当它们发生相互切向力产生相对位移时，接触面的刚毛会发生弹性形变，并在这个过程中产生摩擦力。LuGre 模型涉及参数较少，具有无源性，容易与实验数据匹配，目前已广泛应用于多种控制系统以达到使闭环控制渐近稳定的目的^[21-22]。本文基于 LuGre 模型完成对低速运行工况下的电机非线性力矩建模，PMSM 的动力学模型为：

$$\begin{cases} T_e = J\dot{\omega}_r + T_f + T_L \\ T_f = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 \omega_r \\ g(\omega_r) = T_c + (T_s - T_c) \exp(-\omega_r^2 / \omega_s^2) \\ \dot{z} = \omega_r - \frac{\sigma_0}{g(\omega_r)} z |\omega_r| \end{cases} \quad (2)$$

式中， J 为转动惯量； ω_r 为接触面之间的相对机械角速度； T_f 为摩擦力矩； T_L 为负载转矩； z 为刚毛的平均形变程度； ω_s 为 Stribeck 速度； T_c 为库伦摩

擦力矩； T_s 为最大静摩擦转矩； σ_0 为刚度系数， σ_1 为阻尼系数； σ_2 为黏滞系数。其中， T_c 、 T_s 、 ω_s 、 σ_2 属于静态参数， σ_0 和 σ_1 属于动态参数。稳态时， $\dot{z} = 0$ ，LuGre 模型可以写为静态模型：

$$T_f = [T_c + (T_s - T_c) \exp(-\omega_r^2 / \omega_s^2)] \text{sgn}(\omega_r) + \sigma_2 \omega_r \quad (3)$$

式中，右端第 2 项 $(T_s - T_c) \exp(-\omega_r^2 / \omega_s^2)$ （即 Stribeck 效应）在低速时不可忽略，低速工况下的摩擦力矩具有强烈的非线性特性。

整个 FOC 控制系统由 PMSM 模型、逆变器模型和控制器模型 3 部分组成，如图 2 所示，其中 PMSM 模型如本节所描述；逆变器模型利用 PWM 信号来调节功率半导体开关产生所需的转矩和磁通；控制器包括外环速度和内环电流 2 个控制环，速度环仍使用 PI 控制器，电流环使用 RL 代理模块，取代传统的 dq 轴 2 个 PI 控制器。

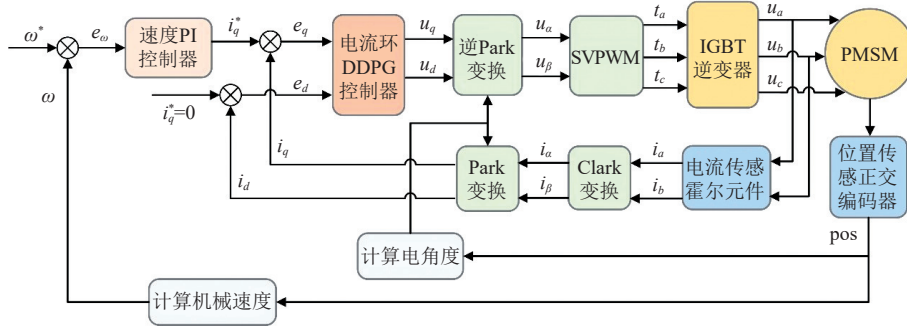


图 2 FOC 框架方框图

2 DDPG 控制策略

RL 是机器学习的一种范式和方法，用于描述和解决智能体在与环境交互过程中使用学习策略实现收益最大化或特定目标的问题，智能体是一个自迭代的网络集成模块，学习如何从状态映射到行为，以获得最大化的奖励。2013 年 DeepMind 提出深度 Q 网络（deep Q-network, DQN）技术，为深度网络提供目标值^[23]。DQN 用到了 2 个关键技术，一个是用来打破样本间关联性的样本池，另一个是使训练稳定性和收敛性更好地固定目标网络，但 DQN 只能输出离散动作空间。因此，DeepMind 随后提出了具有处理连续动作空间的 DDPG^[24]。它包含一个策略网络用来生成动作，一个价值网络用来评判动作的好坏，在 DQN 设计方案基础上，DDPG 使用样本回放池和固定目标网络，每个网络又细分为目标网络和现实网络，但目标网络的更新有所不同，其网络结构如图 3 所示。

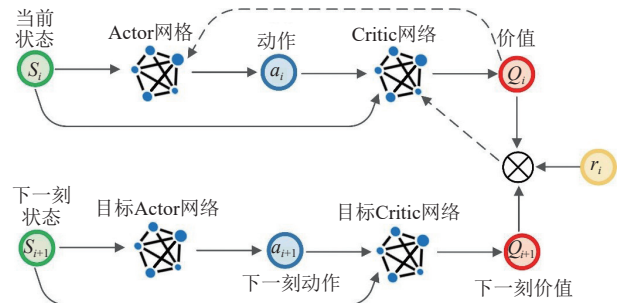


图 3 DDPG 结构图

DDPG 由 4 个网络组成。当前 Actor 网络（网络参数为 θ ）根据当前状态 S_i 选择当前动作 a_i ，用于与环境交互获得奖励 r_i 及下一时刻状态 S_{i+1} ，在经验回放池的缓冲区中存储当前时刻的元组数据 (S_i, a_i, r_i, S_{i+1}) 。目标 Actor 网络（网络参数为 θ' ）根据经验回放池中的下一时刻状态 S_{i+1} 选择最优的下一个动作 a_{i+1} ：

$$\begin{cases} a_i = U(S_i | \theta) + \varepsilon_i \\ a_{i+1} = U(S_{i+1} | \theta') + \varepsilon_{i+1} \end{cases} \quad (4)$$

式中, $\varepsilon_i, \varepsilon_{i+1}$ 表示随机噪声, 以使 DDPG 策略能够更好地探索。当前 Critic 网络 (网络参数为 ω) 根据当前状态 S_i 和当前动作 a_i 计算 Q 值。目标 Critic 网络 (网络参数为 ω') 根据下一时刻状态 S_{i+1} 和下一个动作 a_{i+1} 计算 Q 值:

$$y_i = r_i + \gamma Q(S_{i+1}, a_{i+1} | \omega') \quad (5)$$

式中, γ 为折扣因子, $0 < \gamma \leq 1$ 。神经网络使用梯度反向传播来更新 Actor 和 Critic 网络的参数。采用离线策略训练方法, 根据蒙特卡罗方法推导出的策略梯度更新算法, 当从回放池缓冲区中随机抽取 mini-batch 数据运用到策略梯度公式中时, 可以认为是期望值的无偏估计:

$$\nabla J(\theta) = \frac{1}{m} \sum_{i=1}^m \left[\nabla_a Q(S_i, a_i | \omega) \Big|_{a_i=U(S_i|\theta)} \cdot \nabla_{\theta} U(S_i | \theta) \right] \quad (6)$$

网络中的损失函数的计算采用类似监督学习的方法, 通常采用均方误差 (mean square error, MSE) 计算:

$$J(\omega) = \frac{1}{m} \sum_{i=1}^m [y_i - Q(S_i, a_i | \omega)]^2 \quad (7)$$

通过梯度下降更新当前 Actor 网络参数 θ 和当前 Critic 网络 ω 。然后每隔一定迭代次数对目标网络参数进行软更新:

$$\begin{cases} \tau\omega + (1-\tau)\omega' \rightarrow \omega' \\ \tau\theta + (1-\tau)\theta' \rightarrow \theta' \end{cases} \quad (8)$$

式中, $0 < \tau < 1$ 为更新系数, 这样算法能够保持一定的稳定性, 目标网络参数变化不大, 在训练中相对稳定, 易于收敛。

DDPG 根据观测向量输出动作, 观测向量由系统状态的测量值和相关计算值组成。针对电流环控制, 最直接的观测量为电流测量值, 根据此系统状态, 本文首先计算 d 轴和 q 轴上的参考电流跟踪误差:

$$\begin{cases} e_d = i_d^* - i_d \\ e_q = i_q^* - i_q \end{cases} \quad (9)$$

式中, i_d^*, i_q^* 分别为 d 轴和 q 轴的参考电流, i_d, i_q 分别为实际 d 轴和 q 轴测量电流。在此基础上, 进一步计算电流跟踪误差在 d 轴和 q 轴上的积分和导数以扩充观测向量, 具体利用前向欧拉方法在离散时域中实现数值计算。完整的系统观测向量由跟踪误差、误差的积分和误差的导数组成, 即:

$$S = \left[e_d, \int e_d, e'_d, e_q, \int e_q, e'_q \right] \quad (10)$$

本文设计的 DDPG 旨在跟踪定子电流。目标是找到一个最优的控制策略, 输出电压向量 $\mathbf{a} = [u_d, u_q]$, 使累计奖励最大化。参考模型预测控制中目标函数的设计思路^[25], 控制任务是希望能够在不消耗过多能量的情况下把状态误差控制到最小, 于是奖励函数定义如下:

$$r = -\alpha(u_d^2 + u_q^2) + \left[\beta \frac{|e_d|^{1/2} + |e_q|^{1/2}}{|e_d|^{1/2}|e_q|^{1/2}} - \delta(e_d^2 + e_q^2) \right] \quad (11)$$

式中, 右端第 1 项为电压的平方之和, 表示控制能量, $\alpha > 0$ 为能量系数, 由于负号的存在, 控制能量越小, 奖励值越大。右端第 2 项表示控制误差, 第 1 部分为误差平方的倒数之和, 第 2 部分为误差平方和的相反数, 这样组合的好处是, 训练初期误差的绝对值变化大, 第 2 部分在奖励值的变化上起主导作用; 随着误差趋近 0, 第 1 部分的影响越来越显著。 $\beta, \delta > 0$ 表示误差系数, 控制误差越小, 奖励值越大。

3 仿真实验结果与分析

3.1 仿真环境

本文在 Matlab/Simulink 中搭建 FOC 框架。模型的相关仿真参数如表 1 所示。需要说明的是, 速度正符号表示顺时针转, 负符号表示逆时针转。

表 1 仿真模型参数

参数项	数值
额定电压/V	24
额定转速/r·min ⁻¹	700
额定转矩/N·m	1.9
额定电流/A	4
额定功率/W	140
极对数 P_0	6
定子阻值 R_s/Ω	0.293
d 轴电感 $L_d/\mu\text{H}$	877
Q 轴电感 $L_q/\mu\text{H}$	777
反电势系数 $K_e/\text{V} \cdot \text{KRPM}^{-1}$	57.835
永磁体磁链 ψ/Wb	0.053
逆变器直流电压 V_{dc}/V	24
电流采样 ADC 位数/Bits	12
正交编码器位数/Bits	20
转动惯量 $J/\text{Kg} \cdot \text{m}^2$	0.04
库伦摩擦力矩 $T_c/\text{N} \cdot \text{m}$	1.02
最大静摩擦力矩 $T_s/\text{N} \cdot \text{m}$	1.48
Stribeck 速度 $\omega_s/\text{rad} \cdot \text{s}^{-1}$	0.1
刚度系数 σ_0	4.9
阻尼系数 σ_1	0.19
黏滞系数 σ_2	0.021

本文设计的 DDPG 控制器如图 4 所示。DDPG 训练参数如表 2 所示, Actor 网络和 Critic 网络的结构参数如表 3 所示。具体通过“分级学习”原则, 针对超参数和网络架构等参数, 先固定某类参数, 在此条件下改变其他参数并不断迭代, 最终找到综合优化的训练参数^[26]。Actor 网络最后一层激活函数选择 Tanh 函数, 是为了将输出动作(电压)规范到范围[-1,1]的标么值, 这样能够很大程

度上弱化不同电机在变量数值上的差别, 从而实现一个控制器在不同电机对象上的通用控制效果。此外, 在奖励函数中, 设置能量系数 α 为 0.05, 误差奖励系数 β 和 δ 分别为 0.001 和 0.1。在采集的 d 轴和 q 轴电流后添加无限脉冲响应滤波器(滤波系数为 0.2)以降低噪声带来的不利影响, 然后对模型进行离线训练。

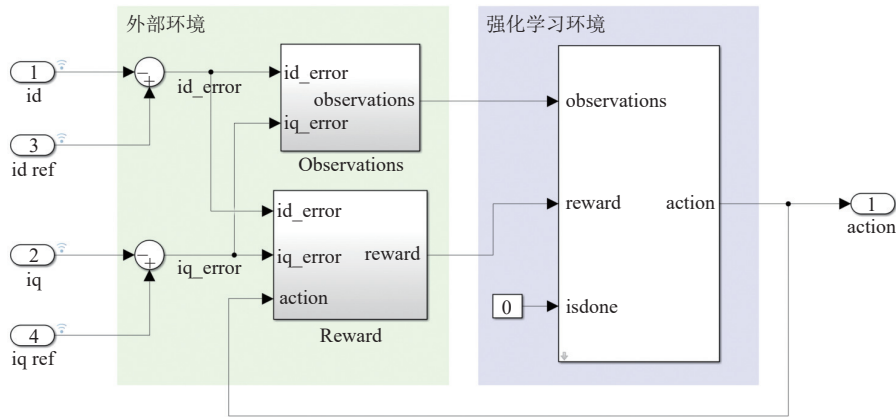


图 4 DDPG 控制器结构

表 2 DDPG 训练参数

参数项	数值
每回合最大步数	1 000
最大回合数	2 000
采样时间/ms	0.1
经验回放池长度	10^6
最小批量数据长度	128
折扣因子	0.995
探索方差	0.1
探索方差衰减率	10^{-5}

表 3 网络结构参数

网络	参数项	数值/类型
Actor	隐藏层数	2
	隐藏层神经元	128, 64
	激活函数	Relu, Relu, Tanh
	学习率	0.001
	梯度阈值	1
Critic	隐藏层数	3
	隐藏层神经元	128, 64, 32
	激活函数	Relu, Relu, Relu
	学习率	10^{-4}
	梯度阈值	1

利用 MATLAB 中的并行计算工具箱快速计算 DDPG 控制模型, 使程序和模型能够以交互和批处理的方式运行, 可大大缩短训练时间。

3.2 评价指标

为了评价训练网络在低速工况下的控制性能, 在仿真试验中将上述 DDPG 控制与 PI 控制^[6]进行了比较。文献 [6] 根据控制器工程设计方法, 将电流环整定为 I 型系统, 从而得到 PI 控制器的比例和积分系数。将电机仿真参数带入, 得到 d 轴电流 PI 控制器的比例系数为 0.365, 积分系数为 122.1; q 轴的比例和积分系数分别为 0.324 和 122.1。跟踪精度是根据均方根相对误差 (root mean square relative errors, RRE) 来评估的, 定义如下:

$$\text{RRE} = \sqrt{\frac{\sum_{i=1}^N (x_i^* - x_i)^2}{\sum_{i=1}^N x_i^2}} \quad (12)$$

式中, N 为采样数; x^* 和 x 分别为参考值和采集值。由于 d 轴参考电流恒为 0, 式 (12) 不再适用, 因此考虑另一种评价指标, 均方根误差 (root mean square errors, RMSE), 定义如下:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i^* - x_i)^2} \quad (13)$$

3.3 阶跃信号跟踪

首先让电机起动至给定转速顺时针 1 r/min, 0.15 s 后将转速变为逆时针 1 r/min, 0.4 s 时将转速

改变为逆时针 3 r/min, 0.6 s 时设定目标速度静止, 0.8 s 时给定转速顺时针 2 r/min, 1 s 时结束。两种控制器的速度环跟踪结果如图 5 所示。

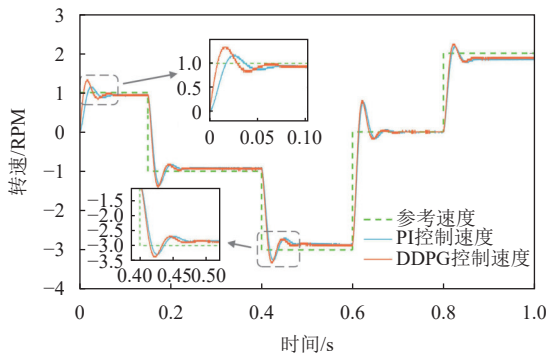
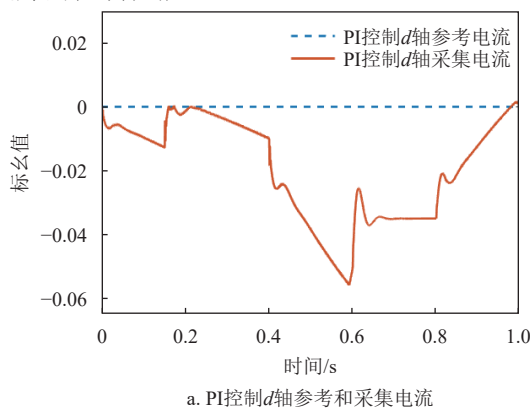


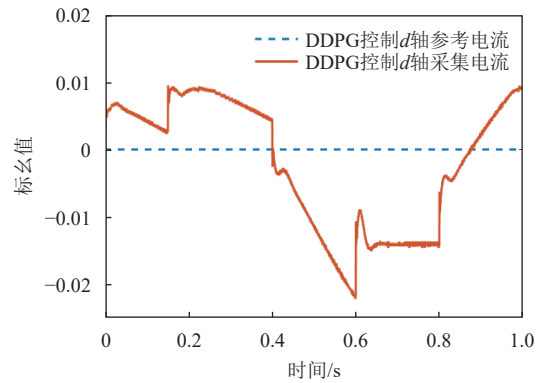
图 5 参考速度阶跃信号的速度环跟踪情况

可以观察到, 对于阶跃信号 DDPG 相较于 PI 控制器有更短的上升时间和稳定时间。具体地, DDPG 控制器的上升时间为 17 ms, 比 PI 控制减少 71%; DDPG 控制器的稳定时间为 51 ms, 比 PI 控制减少 23%。此外, DDPG 控制的速度环跟踪误差 RRE 为 0.22, 比 PI 控制器的跟踪精度提高 8.9%。这说明所提出的 DDPG 控制器在跟随参考阶跃速度时 (相较于 PI 控制) 具有更快的响应和更高的精度。需要说明的是, 非线性系统可能存在饱和现象 (或死区), 即输入信号在某些范围内不会引起输出的变化^[27]。这种情况下, 速度环 PI 控制无法适应非线性区域的变化, 从而导致两种电流控制器在跟踪一些阶跃信号时仍存在稳态误差。

图 6 和图 7 分别显示了两种控制器下对应的 d 轴和 q 轴电流跟踪情况。仿真结果表明, DDPG 控制器的 q 轴电流的跟踪性能明显优于 PI 控制器, DDPG 控制的 RRE 为 0.074, 仅为 PI 控制 RRE 的 21.6%。 d 轴电流的跟踪精度也有显著提升, DDPG 控制的 RMSE 为 0.098, 仅为 PI 控制 RMSE 的 39.3%。这表明 DDPG 控制器在施加扭矩时 q 轴电流响应速度更快, d 轴电流波动更小, 具有更好的控制性能。

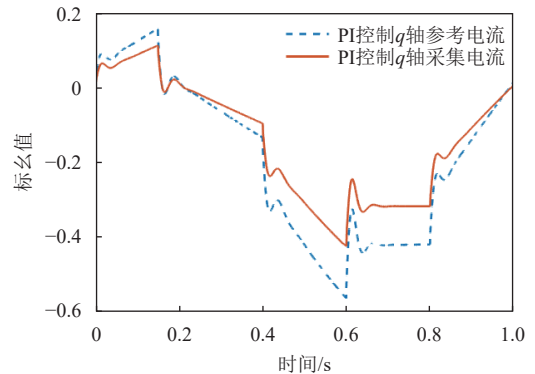


a. PI控制d轴参考和采集电流

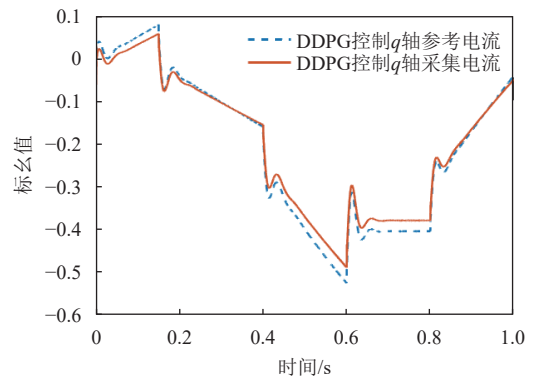


b. DDPG控制d轴参考和采集电流

图 6 参考速度阶跃信号的 d 轴电流跟踪



a. PI控制q轴参考和采集电流



b. DDPG控制q轴参考和采集电流

图 7 参考速度阶跃信号的 q 轴电流跟踪

3.4 正弦信号追踪

为了进一步验证所提出的控制器的有效性, 给定正弦参考转速信号的幅值在 $-3 \sim 5$ r/min 之间变化, 交变频率为 2 Hz。PI 和 DDPG 控制策略下的速度响应结果如图 8 所示。DDPG 控制的速度环跟踪误差 RRE 为 0.08, 比 PI 控制器的跟踪精度提高 28.8%, 这表明所提出的 DDPG 控制器在跟随参考正弦速度时具有更高的精度, 尤其是在启动和峰值阶段, DDPG 控制器具有更优良的伺服性能。

图 9 和图 10 分别显示了两种控制器下对应的 d 轴和 q 轴电流跟踪情况。DDPG 控制的 RRE 为 0.075, 仅为 PI 控制 RRE 的 21.3%。 d 轴电流的跟踪

精度也有显著提升, DDPG 控制的 RMSE 为 0.023, 为 PI 控制 RMSE 的 63%。这表明在正弦信号下, DDPG 也有更好的控制性能, d 轴电流波动更小, q 轴电流响应速度更快。

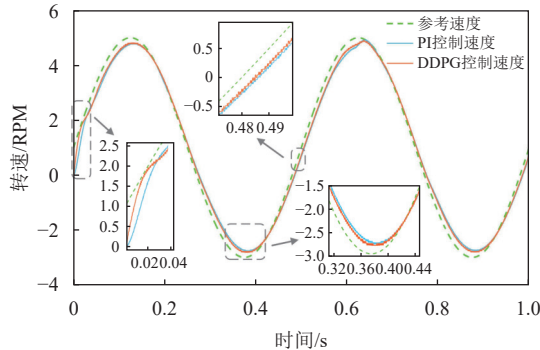
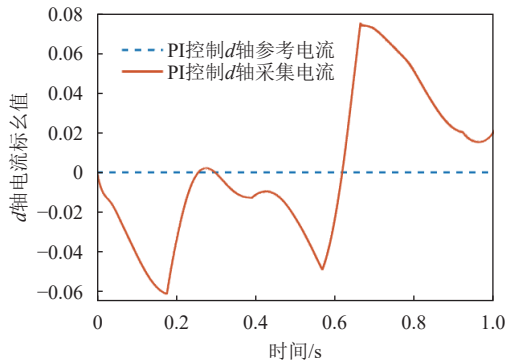
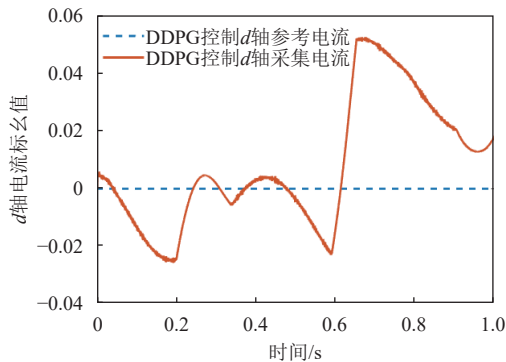


图 8 参考速度正弦信号的速度环跟踪情况

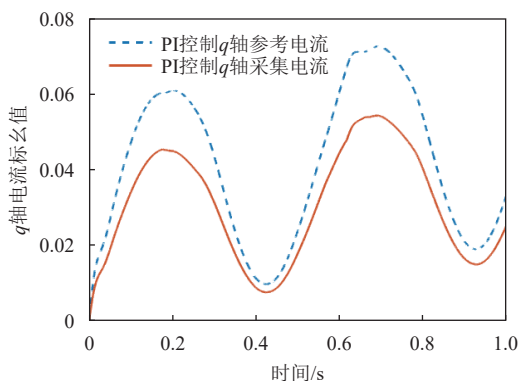


a. PI控制 d 轴参考和采集电流

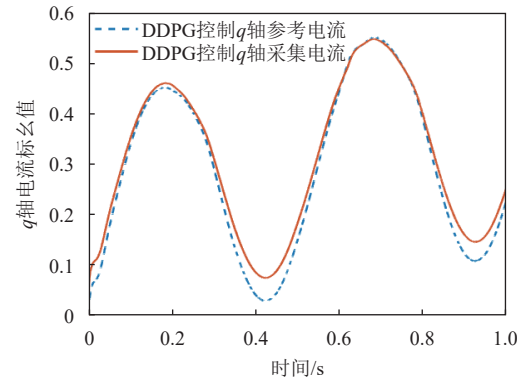


b. DDPG控制 d 轴参考和采集电流

图 9 参考速度阶跃信号的 d 轴电流跟踪



a. PI控制 q 轴参考和采集电流



b. DDPG控制 q 轴参考和采集电流

图 10 参考速度阶跃信号的 q 轴电流跟踪

4 结束语

本文针对 PMSM 低速运行工况设计了一种基于 DDPG 的控制器, 替代 FOC 电流环中传统的 PI 控制器。在 Simulink 平台上搭建 FOC 仿真模型, 比较了 PI 控制和 DDPG 控制在 PMSM 低速工况下的性能: 1) DDPG 控制在跟踪参考速度信号时, 上升时间和稳定时间更短, 跟踪误差更小; 2) DDPG 控制在施加扭矩时, q 轴电流响应更快, d 轴电流波动更小。表明 DDPG 具有在低速工况场景下处理不确定非线性力矩时的良好伺服性能。

参考文献

- [1] SAKUNTHALA S, KIRANMAYI R, MANDADI P N. A study on industrial motor drives: Comparison and applications of PMSM and BLDC motor drives[C]// Proceedings of the International Conference on Energy, Communication, Data Analytics and Soft Computing. New York: IEEE, 2017: 537-540.
- [2] CAO Y, WANG H T, JING F, et al. Research on high precision permanent magnet servo control system for laser communication[C]// Proceedings of the International Conference on Electronics Technology. New York: IEEE, 2018: 387-391.
- [3] 刘华青, 郝珂庆, 刘拥军, 等. 基于非线性扩张观测器的永磁同步电机无位置传感器低速控制策略研究[J]. 微电机, 2020, 53(10): 76-81.
LIU H Q, XI K Q, LIU Y J, et al. Research on sensorless low speed control strategy of PMSM based on nonlinear extended observer[J]. *Micromotors*, 2020, 53(10): 76-81.
- [4] LINARES-FLORES J, GARCÍA-RODRÍGUEZ C, SIRARAMÍREZ H, et al. Robust backstepping tracking controller for low-speed PMSM positioning system: Design, analysis, and implementation[J]. *IEEE Transactions on Industrial Informatics*, 2015, 11(5): 1130-1141.
- [5] CHE Z, YU H, MOBAYEN S, et al. Dual-time-scale sliding mode control for surface-mounted permanent

- magnet synchronous motors[J]. *Symmetry*, 2022, 14(9): 1835.
- [6] 仇国庆, 罗宣林, 王平, 等. PMSM 伺服系统的 PID 控制器设计及仿真[J]. *重庆大学学报*, 2008, 31(3): 259-262.
QIU G Q, LUO X L, WANG P, et al. Design and simulation of PID controller based on PMSM servo system[J]. *Journal of Chongqing University*, 2008, 31(3): 259-262.
- [7] XU D, ZHANG S G, LIU J M. Very-low speed control of PMSM based on EKF estimation with closed loop optimized parameters[J]. *ISA Transactions*, 2013, 52(6): 835-843.
- [8] 鹿文蓬, 廖文蕾, 潘霄, 等. 基于改进滑膜控制器的开绕组永磁同步电机调速系统[J]. *电机技术*, 2022(3): 18-22.
LU W P, LIAO W L, PAN X, et al. Speed-regulation of the opened winding PM synchronous motor based on improved sliding model controller[J]. *Electrical Machinery Technology*, 2022(3): 18-22.
- [9] HOUARI A, BOUABDALLAH A, DJERIOUI A, et al. An effective compensation technique for speed smoothness at low-speed operation of PMSM drives[J]. *IEEE Transactions on Industry Applications*, 2018, 54(1): 647-655.
- [10] XIA J H, LI Z, YU D H, et al. Robust speed and current control with parametric adaptation for surface-mounted PMSM considering system perturbations[J]. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 2021, 9(3): 2807-2817.
- [11] 吕霞, 陈健, 吕成兴, 等. 基于模糊逻辑的永磁同步电机自适应控制[J]. *控制工程*, 2022, 29(5): 837-843.
LYU X, CHEN J, LYU C X, et al. Adaptive control of permanent magnet synchronous motor based on fuzzy logic[J]. *Control Engineering of China*, 2022, 29(5): 837-843.
- [12] WIERING M A, VAN O M. Reinforcement learning[J]. *Adaptation, learning, and optimization*, 2012, 12(3): 729.
- [13] 刘全, 翟建伟, 章宗长, 等. 深度强化学习综述[J]. *计算机学报*, 2018, 41(1): 1-27.
LIU Q, ZHAI J W, ZHANG Z Z, et al. A survey on deep reinforcement learning[J]. *Chinese Journal of Computers*, 2018, 41(1): 1-27.
- [14] 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: 67-81.
- [15] HE W, GAO H J, ZHOU C, et al. Reinforcement learning control of a flexible two-link manipulator: An experimental investigation[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, 51(12): 7326-7336.
- [16] 漆星, 张倩. Actor-Critic 框架下的数据驱动异步电机离线参数辨识方法[J]. *电工技术学报*, 2019, 34(9): 1875-1885.
- QI X, ZHANG Q. Data-driven induction motor parameters offline identification method based on actor-critic framework[J]. *Transactions of China Electrotechnical Society*, 2019, 34(9): 1875-1885.
- [17] LI S H, WON H, FU X G, et al. Neural-network vector controller for permanent-magnet synchronous motor drives: Simulated and hardware-validated results[J]. *IEEE Transactions on Cybernetics*, 2020, 50(7): 3218-3230.
- [18] KHIABANI A G, HEYDARI A. Optimal torque control of permanent magnet synchronous motors using adaptive dynamic programming[J]. *IET Power Electronics*, 2020, 13(12): 2442-2449.
- [19] EL-SOUSY F F M, AMIN M M, AL-DURRA A. Adaptive optimal tracking control via actor-critic-identifier based adaptive dynamic programming for permanent-magnet synchronous motor drive system[J]. *IEEE Transactions on Industry Applications*, 2021, 57(6): 6577-6591.
- [20] CHAOUI H, SICARD P. Adaptive fuzzy logic control of permanent magnet synchronous machines with nonlinear friction[J]. *IEEE Transactions on Industrial Electronics*, 2012, 59(2): 1123-1133.
- [21] CANUDAS D W C, OLSSON H, ASTROM K J, et al. A new model for control of systems with friction[J]. *IEEE Transactions on Automatic Control*, 1995, 40(3): 419-425.
- [22] MASHAYEKHI A, BEHBAHANI S, NAHVI A, et al. Analytical describing function of LuGre friction model[J]. *International Journal of Intelligent Robotics and Applications*, 2022, 6(3): 437-448.
- [23] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing atari with deep reinforcement learning[EB/OL]. [2024-01-28]. <https://doi.org/10.48550/arXiv.1312.5602>.
- [24] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning[EB/OL]. [2024-01-28]. <https://doi.org/10.48550/arXiv.1509.02971>.
- [25] KARAMANAKOS P, LIEGMANN E, GEYER T, et al. Model predictive control of power electronic systems: Methods, results, and challenges[J]. *IEEE Open Journal of Industry Applications*, 2020, 1: 95-114.
- [26] SIRASKAR R. Reinforcement learning for control of valves[J]. *Machine Learning with Applications*, 2021, 4: 100030.
- [27] ZHAO S Y, PAN Y N, DU P H, et al. Adaptive control for non-affine nonlinear systems with input saturation and output dead zone[J]. *Applied Mathematics and Computation*, 2020, 386: 125506.