

基于深度强化学习的永磁同步电机 控制算法研究

范慧妍, 王 爽*

(上海大学 机电工程与自动化学院, 上海 200444)

Research on Control Algorithm of Permanent Magnet Synchronous Motors Based on Deep Reinforcement Learning

FAN Huiyan, WANG Shuang*

(School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China)

Abstract: [Objective] As an emerging intelligent control technology, deep reinforcement learning (DRL) has demonstrated remarkable potential in the field of motor drive system control. In this regard, this paper researches and designs an advanced DRL-based drive control architecture for permanent magnet synchronous motor (PMSM), aiming to achieve high-precision, model-free robust control without relying on the accurate identification of motor physical parameters. [Methods] This paper combined the deep Q-network with finite control set torque control, directly outputting the switching states of the inverter through online learning, thus enabling the agent to determine the optimal switching states of the inverter directly via continuous online learning and interaction with the motor environment. Firstly, a comprehensive multi-level reward function was designed to reflect the complex characteristics of the PMSM, simultaneously accommodating multiple optimization objectives including high-fidelity torque tracking, stator current amplitude minimization, and overall energy efficiency maximization. Secondly, to bridge the gap between theoretical exploration and practical safety requirements, a novel safety protection and evaluation mechanism based on current constraints was established. This mechanism ensured that the inherent random exploration process of DRL did not lead to system overcurrent or hardware damage. Finally, the convergence and control performance of the algorithm were effectively improved by introducing the Q-learning structure and an automated hyperparameter optimization method. [Results] The simulation results showed that the average reward value stabilized at approximately 1 after 400 training episodes, which verified the excellent convergence of the proposed algorithm. The algorithm accurately tracked the

torque commands and maintained fast response speeds with minimal steady-state errors under various speed and load step conditions. With the valid weight coefficients, the system successfully balanced torque precision and operational efficiency. Furthermore, the safety protection mechanism effectively truncated the expected future returns of high-risk states via the done signal, ensuring that the stator current was strictly confined within the safety threshold, which validated the robustness of the model even in small-sample scenarios. [Conclusion] The proposed scheme achieves high-performance model-free torque control, and its integrated safety assessment mechanism provides a scientific foundation and novel insights for preventive operation and maintenance alongside the application of reinforcement learning in the power electronics field, as well as a new research direction for intelligent motor control.

Key words: deep reinforcement learning; permanent magnet synchronous motor; finite control set torque control; multi-level reward function

摘要: [目的]深度强化学习(DRL)作为一种新兴的智能控制技术,在电机驱动系统控制领域展现出显著潜力。对此,本文研究并设计了一种先进的基于DRL的永磁同步电机(PMSM)驱动控制架构,旨在不依赖电机物理参数精确辨识的情况下,实现高精度、无模型的鲁棒控制。[方法]本文将深度Q网络与有限控制集转矩控制结合,通过在线学习直接输出逆变器的开关状态,使智能体能够通过电机环境的持续在线学习与交互,直接确定逆变器的最优开关状态。首先,设计了一个综合性多层次奖励函数以反映PMSM的复杂特性,同时兼顾了高保真转矩跟踪、定子电流幅值最小化以及系统整体能量效率最大化等多个优化目标。其次,为了弥补理论探索与实

际安全需求之间的差距,建立了一种基于电流约束的新型安全保护与评估机制。该机制确保了 DRL 固有的随机探索过程不会导致系统过流或硬件损坏。最后,通过引入 Q 学习结构和自动化超参数优化方法,有效提高了算法的收敛性和控制性能。【结果】仿真结果表明,在训练 400 个回合后平均奖励值稳定于 1 附近,证明了算法优异的收敛性。所提算法能够精准跟踪转矩指令,在不同转速及负载阶跃工况下均保持了较快的响应速度与极小的稳态误差。通过合理的权重配置,系统有效实现了转矩精度与运行效率的平衡。此外,安全保护机制通过 done 信号实时截断高风险状态的预期收益,确保定子电流始终严格约束在安全阈值内,验证了模型在小样本场景下的稳健性。【结论】所提方案实现了无模型的高性能转矩控制,其引入的安全评估机制为强化学习在电力电子领域的应用提供了科学依据与预防性运维的新思路,为电机智能控制提供了新的研究方向。

关键词: 深度强化学习;永磁同步电机;有限控制集转矩控制;多层次奖励函数

0 引言

电驱动系统及其控制是学术和工业研究中的一个重要课题,电机控制技术在现代工业和日常生活中扮演着至关重要的角色。从电动汽车到工业自动化,从家用电器到航空航天,高效、精准的电机控制系统是确保这些应用正常运行的关键。随着技术的不断进步和应用需求的日益复杂化,传统的电机控制方法已经难以满足高性能、高效率以及高可靠性的要求^[1-3]。

先进的电机控制方法,如线性二次型调节器 (Linear Quadratic Regulator, LQR)^[4]、模型预测控制 (Model Predictive Control, MPC)^[5] 以及比例积分 (Proportional Integral, PI) 控制^[6],都需要精确的驱动模型。虽然 PI 控制由于积分反馈,其鲁棒性高于 MPC 和 LQR,但高性能 PI 控制仍然需要精确的模型^[7-9]。由于温度变化、磁饱和等多种原因,实际驱动系统与驱动模型之间可能出现严重偏差。在某些情况下,模型甚至可能是完全未知的,例如,当新电机连接到电力电子变换器进行自调试状态。相比之下,无模型的强化学习 (Reinforcement Learning, RL) 控制完全基于数据驱动,并直接从驱动器的响应中学习最优控制策略^[10-11]。铁饱和、铁损耗、集肤效应和逆变器非线性

行为等影响,可以通过 RL 控制学习并直接补偿。与 MPC 相比,RL 控制允许更多的实现灵活性,因为学习控制策略可以与控制推理异步执行^[12]。

深度强化学习 (Deep RL, DRL) 是将深度学习与 RL 相结合的一种机器学习方法^[13]。DRL 的核心思想是通过与环境的交互来学习最优策略,这种试错的学习方式使其特别适合于控制问题。文献[14]首次提出了基于 DRL 的永磁同步电机 (Permanent magnet synchronous motor, PMSM) 电流控制方案,验证了 DRL 设计架构的合理性及有效性,展现了数据驱动控制器设计的潜力。文献[15-16]发布了一个开源的电机 Python 工具箱,用于加速电机电流控制 RL 代理的开发和培训。该工具箱包含多种直流、三相交流电机模型,便于进行模拟,并可用于将经过训练的 RL 代理与其他先进控制方法进行比较。近年来,研究人员还探索了许多先进的 DRL 算法,如深度 Q 网络 (Deep Q-Network, DQN)、深度确定性策略梯度和软行动者—评论家^[17-19]等,并取得了良好的效果。文献[20-21]通过调整电压源逆变器有限数量的不同开关状态和 DQN 有限控制集框架,设计了永磁电机 DQN 转矩控制器。

DQN 是最成熟的 DRL 算法之一,用于具有实值特征向量的场景,该特征向量来源于系统地测量,并且具有离散动作,因此,将该算法应用于电机的有限集控制领域,可以省去复杂的调制算法的设置^[22-25]。但是,目前针对电机有限集控制算法的研究仍具有一定局限性,如何构建适合电机系统特性的 DRL 框架是关键问题之一。同时,电机控制中的多目标优化难以通过单一奖励函数有效表达,并且电机系统固有的安全约束条件也给探索策略设计带来了额外复杂性^[26-29]。此外,现有研究多将 RL 视为“黑盒”解决方案,缺乏对学习过程和控制策略演化的深入分析,限制了其在电机控制领域的透明度和可解释性。

针对上述问题,本文进行了深入研究。首先,通过分析 PMSM 的数学模型和逆变器特性,设计了适合电机控制的状态—动作表示方法,建立了 RL 与电机控制的有效映射。其次,提出了一种基于 DQN 的改进学习框架,解决了值函数过估计问题,同时采用经验回放和目标网络技术增强了训练稳定性。然后,设计了阶段递进式奖励函数,既

确保了系统安全运行,又能引导智能体逐步学习从基本控制到精细优化的控制策略。最后,基于 gym-electric-motor 仿真库^[13]构建了完整的验证环境,对所提方法在不同工作条件下的性能进行了系统评估,证明了该方法在转矩控制精度和电流利用效率方面的优势。

1 PMSM 系统建模与 RL 设计

1.1 PMSM 系统数学模型

PMSM 在同步旋转 d - q 坐标系下的数学模型为

$$\begin{bmatrix} u_d \\ u_q \end{bmatrix} = \begin{bmatrix} R_s + pL_d & -\omega_e L_q \\ \omega_e L_d & R_s + pL_q \end{bmatrix} \begin{bmatrix} i_d \\ i_q \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_e \psi_f \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} \psi_d \\ \psi_q \end{bmatrix} = \begin{bmatrix} L_d i_d \\ L_q i_q \end{bmatrix} + \begin{bmatrix} \psi_f \\ 0 \end{bmatrix} \quad (2)$$

$$T_e = 3p(\psi_d i_q - \psi_q i_d) \quad (3)$$

式中: u_d 、 u_q 、 i_d 、 i_q 和 ψ_d 、 ψ_q 分别为电压、电流和磁链的 d 、 q 轴分量; T_e 为电磁转矩; R_s 为定子电阻; L_d 、 L_q 分别为 d 、 q 轴电感; ω_e 为电角频率; ψ_f 为永磁体磁链; p 为电机极对数。

本文所研究的 PMSM 由三相两电平逆变器供电,如图 1 所示。该逆变器有三个桥臂,每个桥臂由上下两个开关管组成。 $S_i = 0$ ($i = A, B, C$) 表示对应桥臂的下开关管导通、上开关管关闭; $S_i = 1$ 表示上开关管导通、下开关管关闭,共获得 $2^3 = 8$ 个开关状态。根据直流母线电压值和逆变器开关管的开关状态,可得到电机定子绕组的相电压为

$$\begin{bmatrix} u_A \\ u_B \\ u_C \end{bmatrix} = \frac{U_{dc}}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} S_A \\ S_B \\ S_C \end{bmatrix} \quad (4)$$

式中: U_{dc} 为直流母线电压。

$$\mathbf{o}_k = \begin{bmatrix} \frac{i_{d,k}}{i_{lim}} & \frac{i_{q,k}}{i_{lim}} & \frac{\omega_k}{\omega_{lim}} & \kappa \cos(\varepsilon_{el,k}) & \kappa \sin(\varepsilon_{el,k}) & \frac{u_{d,k-1}}{U_{dc}} & \frac{u_{q,k-1}}{U_{dc}} & \frac{T_{e,k}}{T_{lim}} & \frac{T_{e,k}^*}{T_{lim}} & 2 \frac{i_{s,k}}{i_{lim}} - 1 & \frac{T_{e,k}^* - T_{e,k}}{T_{lim}} \end{bmatrix} \quad (6)$$

式中: \mathbf{o}_k 为 k 时刻的观测向量; i_{lim} 为电流极限值,用于对 k 时刻的实际电流 $i_{d,k}$ 和 $i_{q,k}$ 进行归一化; ω_{lim} 为电机转速最大值; κ 为系数,用于调整角度

根据 Clarke 变换理论可得两相静止 α - β 坐标系下的电压为

$$\begin{bmatrix} u_\alpha \\ u_\beta \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ 0 & \sqrt{3} & -\sqrt{3} \end{bmatrix} \begin{bmatrix} u_A \\ u_B \\ u_C \end{bmatrix} \quad (5)$$

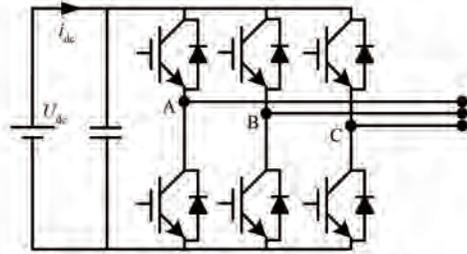


图 1 三相桥式电压源逆变器电路图

Fig. 1 Circuit of three-phase bridge voltage source inverter

1.2 状态空间表示

在 PMSM 控制系统中,选择合适的状态变量对于全面反映系统动态特性至关重要。本文选择 i_d 、 i_q 和转子角速度 ω 三个关键变量作为构建状态空间的基础。

i_d 主要影响电机的磁链,在表贴式 PMSM 中,通常将 i_d 控制为零以最小化铜损;在嵌入式 PMSM 或需要弱磁控制的情况下, i_d 的动态变化不可忽略。 i_q 直接决定电机产生的电磁转矩, i_q 的瞬时值和动态变化对转矩控制的精度和响应速度有着决定性的影响。 ω 反映了 PMSM 的机械动态特性, ω 不仅是速度控制的直接反馈,还影响反电动势的大小,进而影响电流动态特性。

为了能够全面描述电机的电磁和机械状态,转子位置 θ 、 u_d 、 u_q 以及负载转矩估计值 T_L 也应加入观测向量中。考虑到在后续奖励函数的设置中,定子电流以及转矩给定值为关键数据,也将其一同并入被观测的状态变量中。进行归一化处理后,最终所设计的状态空间为

$$\mathbf{o}_k = \begin{bmatrix} \frac{i_{d,k}}{i_{lim}} & \frac{i_{q,k}}{i_{lim}} & \frac{\omega_k}{\omega_{lim}} & \kappa \cos(\varepsilon_{el,k}) & \kappa \sin(\varepsilon_{el,k}) & \frac{u_{d,k-1}}{U_{dc}} & \frac{u_{q,k-1}}{U_{dc}} & \frac{T_{e,k}}{T_{lim}} & \frac{T_{e,k}^*}{T_{lim}} & 2 \frac{i_{s,k}}{i_{lim}} - 1 & \frac{T_{e,k}^* - T_{e,k}}{T_{lim}} \end{bmatrix} \quad (6)$$

信号的幅值; $\varepsilon_{el,k}$ 为 k 时刻的电角度; T_{lim} 为转矩极限值,用于归一化转矩指令; $T_{e,k}^*$ 为 k 时刻的给定电磁转矩,即控制系统的目标值。

归一化处理后的状态变量将作为 RL 算法的输入。通过电流传感器和位置编码器可以直接测量或估算这些变量的值,将这些变量作为输入,无需电机参数信息,实现了无模型控制。

1.3 动作空间定义

动作空间的定义直接影响控制策略的精度和灵活性。更大的动作空间(更多的离散动作)可以提供更精细的控制,但会增加学习的难度和时间;较小的动作空间有利于快速学习和收敛,但可能限制控制精度。因此需要权衡控制精度与学习效率。

在传统直接转矩控制中,逆变器的开关状态通常被离散化为有限数量的电压矢量。这种离散化方法在 RL 中同样适用。本文使用的方案源于传统二电平逆变器直接转矩控制,共有 8 种开关动作组合 a ,包括 6 个有源矢量($V1 \sim V6$),对应逆变器的 6 个非零开关状态;2 个零矢量($V0, V7$),对应两种使所有相电压为零的开关状态,其映射关系如表 1 所示。每个动作的运行时间与采样时间一致,确保精细化。

表 1 开关动作、开关状态和电压之间的对应关系

Tab.1 The correspondence between switch action, switch status and voltage

a	S_A	S_B	S_C	u_α	u_β
0	-	-	-	0	0
1	+	-	-	$2U_{dc}/3$	0
2	+	+	-	$U_{dc}/3$	$U_{dc}/3$
3	-	+	-	$-U_{dc}/3$	$U_{dc}/3$
4	-	+	+	$-2U_{dc}/3$	0
5	-	-	+	$-U_{dc}/3$	$-U_{dc}/3$
6	+	-	+	$-U_{dc}/3$	$-U_{dc}/3$
7	+	+	+	0	0

1.4 奖励函数设计

在电机控制中应用 RL 时,探索性动作可能导致系统过载或损坏,因此需要安全层来确保所有操作都在安全限制内进行。本文设计了安全保护评估机制,当电流大小超过安全范围时,学习中断,如式(7)所示:

$$\begin{cases} \text{if } i_s > i_{lim} \\ r_{total} = -1 \\ term = 1 \end{cases} \quad (7)$$

式中: i_s 为定子电流; r_{total} 为综合奖励函数; $term$ 为

中止信号, $term = 0$ 时学习过程正常进行, $term = 1$ 时学习过程异常并中止。

对于嵌入式 PMSM,即 $L_q > L_d$,采用弱磁控制,需要满足 $i_d < 0$:

$$\begin{cases} \text{if } i_s < i_{lim} \text{ and } i_d > 0 \\ r_{total} = -1 + w_1 \frac{i_{lim} - i_d}{i_{lim}} \\ term = 0 \end{cases} \quad (8)$$

式中: w_1 为弱磁约束权重系数。

转矩跟踪性能通常是 PMSM 转矩控制系统的主要目标之一,转矩脉动过大可能导致电机运行过程中产生振动、噪声问题,并影响电机寿命。设计奖励函数时,需要鼓励系统快速、准确地跟踪给定转矩参考值 T_e^* ,所以当转矩脉动超出正常脉动范围时,设置关于转矩跟踪的奖励函数 r_{torque} ,如式(9)所示:

$$r_{torque} = -w_2 \left| \frac{T_e^* - T_e}{T_{lim}} \right| \quad (9)$$

式中: w_2 为转矩跟踪权重系数。

将能效因素纳入奖励函数可以帮助系统学习到更加节能的控制策略。最直接的方法是尽可能减小运行时的总电流,设置关于电流最小化的奖励函数 $r_{current}$,如式(10)所示:

$$r_{current} = -w_3 \frac{i_s}{i_{lim}} = -w_3 \frac{\sqrt{i_d^2 + i_q^2}}{i_{lim}} \quad (10)$$

式中: w_3 为电流最小化权重系数。

将上述所有因素组合成一个综合奖励函数,如式(11)所示:

$$\begin{cases} \text{if } i_s < i_{lim} \\ r_{total} = 1 + r_{torque} + r_{current} \\ term = 0 \end{cases} \quad (11)$$

式中:1 为基础奖励。

各部分的权重需要根据控制目标进行调整。此时,奖励函数的变化范围为 $(1-2w_2-w_3, 1)$,同时 $1-2w_2-w_3 \geq -1+w_1$ 。

本文基于优先级分层原则确定权重系数为: $w_1 = 0.5, w_2 = 0.7, w_3 = 0.2$ 。转矩跟踪作为核心任务,其权重设为最高,以保证动态响应速度;弱磁约束确保嵌入式 PMSM 在高速区的运行稳定性,权重居中;电流最小化作为效率优化项,权重设为最低。

2 DQN 算法原理

2.1 DQN 算法

Q 学习是一种无模型的 DRL 算法,旨在学习最优动作价值函数 $Q^*(s, a)$ 。Q 学习的更新规则为

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \quad (12)$$

式中: $Q(s, a)$ 为当前状态 s 和动作 a 的估计 Q 值; α 为学习率, $0 < \alpha \leq 1$; r 为即时奖励; γ 为折扣因子, $0 \leq \gamma < 1$; $\max_{a'} Q(s', a')$ 为下一时刻状态 s' 和动作 a' 的最大 Q 值。

在特定条件下, Q 学习可以收敛到 $Q^*(s, a)$ 。但在复杂问题中,状态空间可能非常大或连续,无法用表格表示 Q 值。此时,需要使用函数近似方法,将强化 Q 学习与深度学习结合,利用 DQN 进行非线性函数近似,如式(13)所示:

$$Q(s, a; w) \approx Q^*(s, a) \quad (13)$$

式中: w 为神经网络的权重系数。

在 DQN 中, Q 值的更新使用了最大化操作,其目标是最大化回报 y ,即随时间累积的折扣奖励,如式(14)所示:

$$y = Q(s, a; w) = r + \gamma \max_{a'} Q(s', a'; w') \quad (14)$$

式中: w' 为目标神经网络的权重参数。

式(14)递归地定义了动作价值函数 $Q(s, a; w)$,通过即时奖励 r 和折扣后的未来动作价值来计算。因此,动作价值代表了基于当前状态-动作对可预期的回报。对于给定的动作价值函数,最优动作 a^* 可通过式(15)确定:

$$a^* = \arg \max_a Q(s, a; w) \quad (15)$$

整个网络采用软更新的策略,每步使用小比例更新:

$$w' \leftarrow \tau w + (1 - \tau) w' \quad (16)$$

式中: τ 为一个很小的正数,如 0.001。

2.2 损失函数

DQN 的损失函数 L 可以表示为

$$L = E \left\{ \left[\underbrace{r + \gamma \max_{a'} Q(s', a'; w')}_{\text{目标值}} - \underbrace{Q(s, a; w)}_{\text{当前预测值}} \right]^2 \right\} \quad (17)$$

式中: $E(\cdot)$ 为期望。

训练过程中,对损失函数关于 w 求导,并使

用梯度下降法更新 w 。

由于 Q 值的频繁更新,神经网络可能难以收敛。同时,连续采样的样本之间存在强相关性,影响学习效果。为解决此问题,DQN 采用了一种经验回放机制。

经验回放在 DQN 及其变体的一个重要组成部分。智能体与环境交互时,将每一步的经验 (s, a, r, s') 存储在回放缓冲区的数据结构中。在训练时,从回放缓冲区中随机采样一批经验,而不是使用最近的经验。这有助于打破经验之间的时间相关性,使学习过程更加稳定,其存储格式为 $(s, a, r, s', done)$ 。

安全保护评估机制通过控制 $done$ 信号,在底层优化了目标 Q 值的计算逻辑。当 $done = 1$ 时,目标 Q 值的计算将截断未来收益的回报。这种截断效应在损失函数中形成了显著的梯度惩罚,使得神经网络能够建立起“动作-过流-训练终止”的强关联映射。相较于仅通过负奖励提示风险,基于 $done$ 的强制终止机制能显著增强 Q 值空间在安全边界处的区分度,确保控制策略在收敛过程中对边界工况具有极高的敏感性。

在选择动作时,采用 ϵ -greedy 探索。以 ϵ 的概率随机选择动作,以 $1 - \epsilon$ 的概率选择 Q 值最大的动作,如式(18)所示:

$$P(a | s) = \begin{cases} \epsilon / |A| + (1 - \epsilon), & a = \arg \max_{a'} Q(s, a') \\ \epsilon / |A|, & \text{其他} \end{cases} \quad (18)$$

式中: $P(a | s)$ 为在状态 s 下选择动作 a 的概率; $|A|$ 为动作空间的大小; ϵ 为探索率。

ϵ 通常随时间衰减,开始时较高以鼓励探索,后期降低以更多地利用学到的知识,如式(19)所示:

$$\epsilon = \max \left[\epsilon_{\text{end}}, \epsilon_{\text{start}} - \frac{t}{T} (\epsilon_{\text{start}} - \epsilon_{\text{end}}) \right] \quad (19)$$

式中: ϵ_{start} 、 ϵ_{end} 分别为初始探索率、最终探索率; T 为总迭代步数。

DQN 训练过程如图 2 所示。

3 仿真分析

基于 gym-electric-motor 库对 PMSM 驱动系统进行建模与仿真,具体电机模型参数如表 2 所示。DQN 的构建与训练通过 TensorFlow 2.0 框架结合

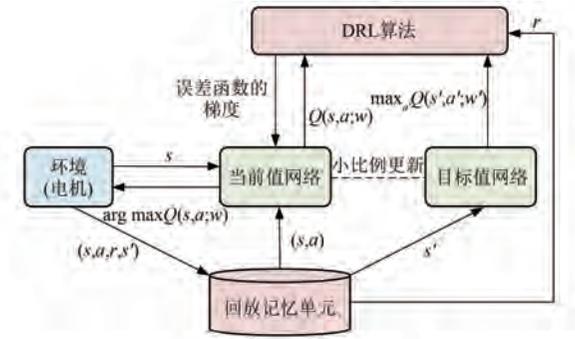


图 2 DQN 训练过程

Fig. 2 Training process of DQN

Keras-RL2 库实现,网络权重优化采用 Adam 优化器,采用 Leaky ReLU 作为激活函数。

整个 DRL 训练过程采用回合式更新策略,将电机系统离散为有限个控制时刻,具体超参数如表 3 所示。

表 2 电机模型参数

Tab. 2 Motor model parameters

参数名称	参数值
定子电阻/ Ω	0.23
d 轴电感/mH	0.197
q 轴电感/mH	0.257
极对数	5
额定功率/W	200
额定转速/ $(r \cdot \text{min}^{-1})$	3 000
额定转矩/ $(N \cdot \text{m})$	0.64
最大电流/A	17
直流母线电压/V	24
最大参考转矩/ $(N \cdot \text{m})$	1.23
采样频率/kHz	20

表 3 关键超参数

Tab. 3 The key hyperparameters

超参数名称	参数值
折扣因子 γ	0.95
初始学习率 α_{start}	2.3×10^{-4}
最终学习率 α_{end}	5.7×10^{-5}
初始探索率 ϵ_{start}	0.3
最终探索率 ϵ_{end}	0.05
网络层数 layer	8
每层神经元数 neurons	512
批大小 batch size	64

图 3~图 5 展示了一个典型训练过程中不同阶段的代表性回合表现,清晰地反映了智能体学习控制策略的演化过程。

早期训练回合表现如图 3 所示。在训练初期,网络权重随机初始化,导致智能体对电机执行基本上是随机动作。智能体无法有效跟踪参考转矩,控制表现不稳定, d 轴和 q 轴电流波动较大,存在明显的超调现象。且 q 轴电流从 0 上升到 15 A 以上,明显超出了安全运行范围,并很快结束训练回合。

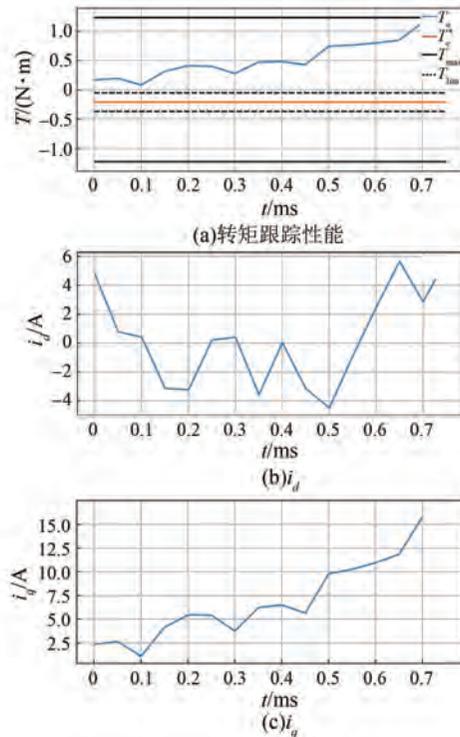


图 3 早期训练回合表现 (转速: -170 rad/s)

Fig. 3 Performance in early training episodes (rotational speed: -170 rad/s)

中期训练回合表现如图 4 所示。随着训练的深入,智能体逐渐学习到如何更好地跟踪参考转矩,转矩跟踪性能有了明显改善,尽管仍有波动,但已经能够基本跟随参考转矩的阶跃变化。 d 轴和 q 轴电流的控制也更加合理,整体保持在可控范围内。此时智能体已经初步掌握了转矩控制策略,但控制精度仍有提升空间。

后期训练回合表现如图 5 所示。在训练后期,智能体已经能够精确跟踪参考转矩,并在参考转矩发生阶跃变化时迅速响应,展现出良好的动态性能。虽然 d 轴和 q 轴电流波动仍然存在,但

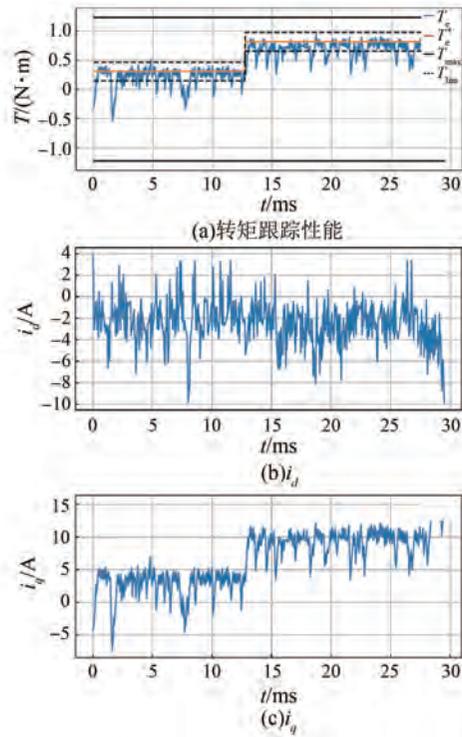


图 4 中期训练回合表现 (转速: 130 rad/s)
Fig. 4 Performance in mid-term training episodes (rotational speed: 130 rad/s)

整体受控且能够适应转矩变化需求。此阶段智能体已经学会了如何在避免系统极限违反的同时, 实现高精度、快响应的转矩控制。

值得注意的是, 由于 DRL 中的探索机制, 即使在训练后期, 系统仍会偶尔执行随机动作以探索更优策略。此外, 为了进一步激励控制回路的学习, 系统在任意时间步都有一定概率重新采样参考转矩, 如图 5 所示的转矩阶跃变化。并且最大参考转矩设定超过了 PMSM 的额定转矩, 这种设置迫使控制器在靠近过流区的边缘运行, 同时又要避免进入这些区域以防止获得低奖励, 从而使系统稳定性成为训练的重要组成部分。

DQN 训练过程中奖励值的变化趋势如图 6 所示。随着训练的进行, 奖励值逐渐提升并趋于稳定, 这表明控制策略正在不断改善, 平均奖励从初始的 -3 提升到了 1, 并在后续训练中进一步优化和稳定。

为验证控制器性能, 设置电机转速变化如图 7 所示。该测试工况设计了多个不同速度水平的阶梯变化, 以全面评估控制器在各种运行条件下的表现。

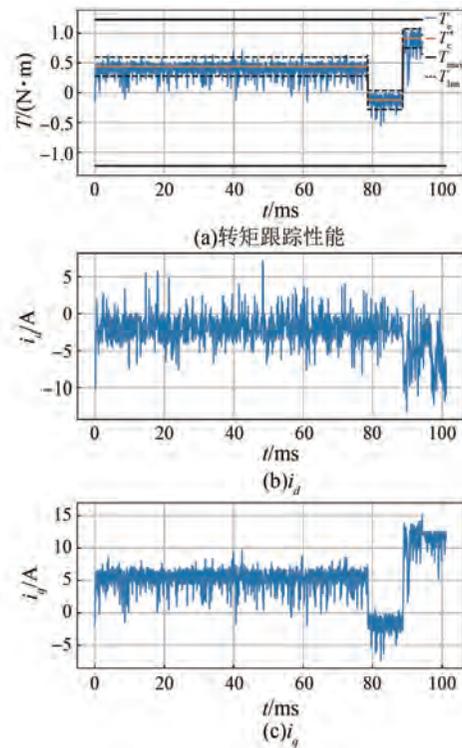


图 5 后期训练回合表现 (转速: 85 rad/s)
Fig. 5 Performance in late training episodes (rotational speed: 85 rad/s)

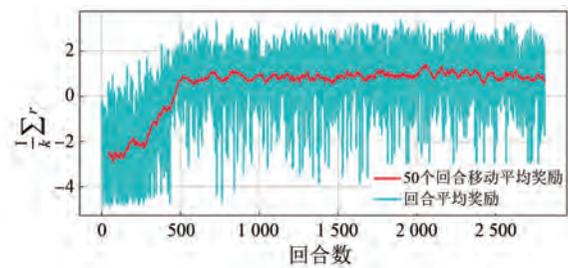


图 6 奖励值变化趋势
Fig. 6 Trend of reward value

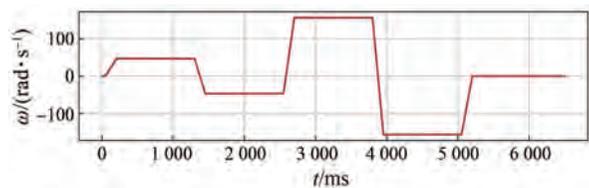


图 7 电机转速变化

Fig. 7 Variation of motor rotational speed

在图 7 所示的测试工况下, 得到电机转矩跟踪性能如图 8 所示, 此时探索率 ϵ 为 0。由图 8 可知, 从 $-0.6 \text{ N}\cdot\text{m}$ 到 $0.6 \text{ N}\cdot\text{m}$, 控制器能够准确地跟踪阶跃变化的转矩指令, 即使在转速发生剧烈变化的区域, 控制器仍能保持良好的转矩跟踪性能,

表明算法具有较强的鲁棒性。橙色标记的偏差较小区域显示了控制精度,大部分工作点的转矩误差控制在允许范围内,表明算法具有良好的稳态特性。

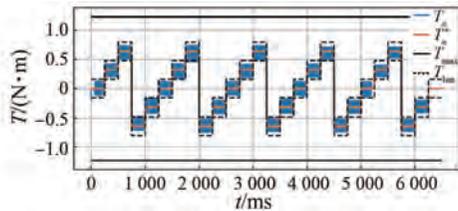


图8 转矩跟踪性能

Fig. 8 Torque tracking performance

在图7所示的测试工况下,得到 d 、 q 轴电流轨迹如图9所示。从图9可知,控制器成功地将电流控制在安全范围内,大部分工作点都保持在额定电流圆内,有效避免了过流风险。不同转矩水平下,电流矢量都集中在各自对应的最优区域周围,证明了控制器能够自动学习并接近最优的控制性能,在保证转矩精度的同时实现了电流最小化。

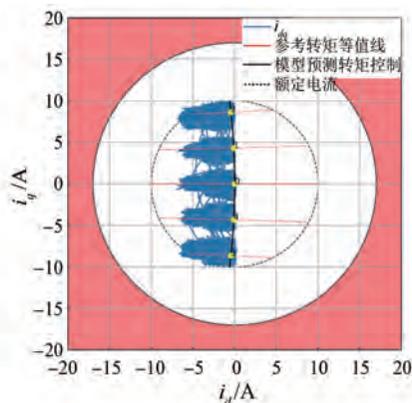


图9 电流轨迹

Fig. 9 Current trajectory

4 结语

本文研究了一种基于DRL的PMSM智能控制方法,通过结合DQN与有限控制集转矩控制,实现了高性能的无模型电机控制。首先,基于DRL的控制策略可以在无需精确电机参数模型的前提下,通过在线学习实现对PMSM的高精度转矩控制。其次,所设计的多层次奖励函数有效平衡了转矩跟踪精度、电流最小化等多个控制目标。通过分阶段奖励机制,系统首先保证运行安

全性,然后逐步优化控制性能,最终实现了接近理论最优点的控制特性。最后,通过仿真验证了所提方法能够在不同转速和负载条件下,保持良好的转矩跟踪精度和动态响应特性。

利益冲突声明

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

作者贡献

范慧妍进行了方案设计、内容总结与论文撰写,王爽参与了方案设计、稿件的审核与修改。所有作者均阅读并同意了最终稿件的提交。

Fan Huiyan carried out the scheme design, content summary and paper writing. Wang Shuang participated in the scheme design, as well as the review and revision of the manuscript. All authors have read and approved the final version of the manuscript for submission.

参考文献

- [1] 毛帅,刁晓飞,王晓,等. 具有非惯量负载转矩补偿功能的伺服永磁同步电机控制方法[J]. 电机与控制应用, 2023, 50(9): 20-27.
Mao S, Diao X F, Wang X, et al. Control method of servo permanent magnet synchronous motor with non-inertia load torque compensation function [J]. Electric Machines & Control Application, 2023, 50(9): 20-27.
- [2] 李耀华,刘东梅,赵承辉,等. 基于CNN的MPTC与DTC自适应切换的表贴式永磁同步电机控制策略[J]. 电机与控制应用, 2022, 49(5): 8-13.
Li Y H, Liu D M, Zhao C H, et al. Adaptive switching control strategy between MPTC based on CNN and DTC for surface-mounted PMSM [J]. Electric Machines & Control Application, 2022, 49(5): 8-13.
- [3] 曹美禾,宋琳,孙军,等. 深度学习及其在电机控制中的研究现状与展望[J]. 电机与控制应用, 2023, 50(5): 1-9.
Cao M H, Song L, Sun J, et al. Research and prospect on deep learning and its application in motor control [J]. Electric Machines & Control

- Application, 2023, 50(5): 1-9.
- [4] Xia C L, Liu N, Zhou Z Q, et al. Steady-state performance improvement for LQR-based PMSM drives [J]. IEEE Transactions on Power Electronics, 2018, 33(12): 10622-10632.
- [5] Kazmierkowski M P. Model predictive control of high power converters and industrial drives [J]. IEEE Industrial Electronics Magazine, 2018, 12(3): 55-56.
- [6] Lee S B. Closed-loop estimation of permanent magnet synchronous motor parameters by PI controller gain tuning [J]. IEEE Transactions on Energy Conversion, 2006, 21(4): 863-870.
- [7] 吴崇凯, 王寅, 潘佳炜, 等. 交流伺服系统速度环变参数 PI 及负载转矩补偿控制研究[J]. 电机与控制应用, 2025, 52(11): 1193-1204.
Wu C K, Wang Y, Pan J W, et al. Research on variable parameter PI and load torque compensation control of AC servo system speed loop [J]. Electric Machines & Control Application, 2025, 52(11): 1193-1204.
- [8] 王伟炳, 宋伟豪. 基于 FOCA 的永磁游标电机 PI 参数自整定[J]. 电机与控制应用, 2024, 51(7): 94-102.
Wang W B, Song W H. PI parameter self-tuning of permanent magnet vernier motor based on FOCA [J]. Electric Machines & Control Application, 2024, 51(7): 94-102.
- [9] 张曦, 王涛, 陈致君, 等. 永磁伺服电机复合模型预测控制[J]. 电机与控制应用, 2023, 50(4): 8-15.
Zhang X, Wang T, Chen Z J, et al. Composite model predictive control for permanent magnet servo motor [J]. Electric Machines & Control Application, 2023, 50(4): 8-15.
- [10] 朱俊威, 顾曹源, 王鼎, 等. 多伺服电机智能化协调容错轨迹跟踪控制系统设计[J]. 控制理论与应用, 2021, 38(7): 1023-1032.
Zhu J W, Gu C Y, Wang D, et al. Design of an intelligent cooperative fault-tolerant trajectory tracking control system for multi-servo motors [J]. Control Theory & Applications, 2021, 38(7): 1023-1032.
- [11] 陈刚. 基于强化学习的工业电机多模式智能控制策略研究[J]. 自动化应用, 2025, 66(2): 109-111+115.
Chen G. Research on multi-mode intelligent control strategy for industrial motors based on reinforcement learning [J]. Automation Application, 2025, 66(2): 109-111+115.
- [12] Sutton R S, Barto A. Reinforcement learning: An introduction [M]. 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
- [13] Book G, Traue A, Balakrishna P, et al. Transferring online reinforcement learning for electric motor control from simulation to real-world experiments [J]. IEEE Open Journal of Power Electronics, 2021, 2: 187-201.
- [14] Schenke M, Kirchgässner W, Wallscheid O. Controller design for electrical drives by deep reinforcement learning: A proof of concept [J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4650-4658.
- [15] Traue A, Book G, Kirchgässner W, et al. Toward a reinforcement learning environment toolbox for intelligent electric motor control [J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 33(3): 919-928.
- [16] Balakrishna P, Book G, Kirchgässner W, et al. Gym-electric-motor (GEM): A python toolbox for the simulation of electric drive systems [J]. Journal of Open Source Software, 2021, 6(58): 1-3.
- [17] 卓博群, 杨文骏, 李建康. 基于深度强化学习的飞行控制方法研究现状及发展[J]. 航空科学技术, 2025, 36(8): 1-10.
Zhuo B Q, Yang W J, Li J K. Research status and development of flight control methods based on deep reinforcement learning [J]. Aerospace Science and Technology, 2025, 36(8): 1-10.
- [18] 于荣, 郑福. 基于深度 Q 网络改进交通信号灯控制设计与实现[J]. 渤海大学学报, 2025, 46(2): 134-141.
Yu R, Zheng F. Research and implementation of traffic signal control based on deep Q-network algorithm optimization [J]. Journal of Bohai University, 2025, 46(2): 134-141.
- [19] Ke Z M, Li Z B, Cao Z H, et al. Enhancing transferability of deep reinforcement learning-based variable speed limit control using transfer learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(7): 4684-4695.
- [20] Li L J, Xiang Q Y, Xu X W, et al. An optimized control strategy based on multidimensional feature

- operation pattern [J]. IEEE Transactions on Control Systems Technology, 2024, 32(4): 1226-1239.
- [21] Yang D J, Miao C Y, Liu Y, et al. Improved DQN-based intelligent trajectory control for coal gangue sorting robotic manipulators [J]. IEEE Sensors Journal, 2025, 25(21): 40632-40650.
- [22] Zeng B S, Zhong Y, Niu X H. Efficient exploration through bootstrapped and Bayesian deep Q-networks for joint power control and beamforming in mmwave networks [J]. IEEE Communications Letters, 2023, 27(2): 566-570.
- [23] 钟林涛, 宋冬梅, 张衡镜, 等. 基于DQN和DDPG算法的多智能体泵系统节能控制优化研究[J]. 机械, 2025, 52(8): 14-22.
- Zhong L T, Song D M, Zhang H J, et al. Research on optimization of energy-saving control for pump system based on multiple-agent of DQN and DDPG algorithms [J]. Machinery, 2025, 52(8): 14-22.
- [24] Schenke M, Wallscheid O. A deep Q-learning direct torque controller for permanent magnet synchronous motors [J]. IEEE Open Journal of the Industrial Electronics Society, 2021, 2: 388-400.
- [25] Guan Q, Yao X L, Lin Z F, et al. A robust control scheme for PMSM based on integral reinforcement learning [J]. IEEE Transactions on Transportation Electrification, 2025, 11(1): 4214-4223.
- [26] Li M J, Zhou P, Wang L Y, et al. Data-driven multiobjective predictive optimal control of refining process with non-Gaussian stochastic distribution dynamics [J]. IEEE Transactions on Industrial Informatics, 2021, 17(11): 7269-7278.
- [27] Zhao W D, Wang X J, Gerada C, et al. Multi-physics and multi-objective optimization of a high speed PMSM for high performance applications [J]. IEEE Transactions on Magnetics, 2018, 54(11): 8106405.
- [28] Sun M X, Xu Y L, Han K. Structure and optimization design of cup winding permanent magnet synchronous machine in flywheel energy storage system [J]. IEEE Transactions on Magnetics, 2023, 59(5): 8100805.
- [29] Liu A, Huang X Y, Chen Z, et al. Nonlinear analytical model-embedded optimal design for ER-PMSM [J]. IEEE Transactions on Magnetics, 2024, 60(3): 8100305.
-
- 收稿日期:2025-09-11
收到修改稿日期:2026-01-04
作者简介:
范慧妍(2000—),女,硕士研究生,研究方向为永磁同步电机驱动与控制,fanhuiyan@shu.edu.cn;
*通信作者:王爽(1977—),男,博士,副教授,研究方向为智能控制理论及其在新能源汽车电力电子和伺服控制系统中的应用,wang-shuang@shu.edu.cn。