

基于深度强化学习的无刷双馈电机空间 矢量调制直接转矩控制

施 隆^{1,2*}, 鉴 航¹, 彭云建¹, 仲兆峰², 郭伟文²

[1. 华南理工大学 自动化科学与工程学院, 广东 广州 511442;

2. 日立楼宇技术(广州)有限公司, 广东 广州 510660]

Space Vector Modulation Direct Torque Control Based on Deep Reinforcement Learning for Brushless Doubly-Fed Machine

Shi Long^{1,2*}, Jian Hang¹, Peng Yunjian¹, Zhong Zhaofeng², Guo Weiwen²

[1. School of Automation Science and Engineering, South China University of Technology,
Guangzhou 511442, China;

2. Hitachi Building Technology (Guangzhou) Co., Ltd., Guangzhou 510660, China]

Abstract: [Objective] In the space vector modulation direct torque control (SVM-DTC) of brushless doubly-fed machine (BDFM), inner-loop torque and flux proportional-integral (PI) controllers are commonly used. However, fixed PI parameters are difficult to adapt to the system's nonlinearity and frequently varying operating conditions, making it challenging to maintain optimal performance across all operating points. To address this problem, a deep reinforcement learning (DRL)-based SVM-DTC strategy is proposed to eliminate the dependence on manual parameter tuning and improve the system's dynamic performance and robustness. [Methods] The twin delayed deep deterministic policy gradient (TD3) algorithm was adopted to construct and train an Actor-Critic neural network, which replaced the inner-loop torque and flux PI controllers. The state space consisted of estimated torque, estimated flux, torque error, flux error, integral of torque error, integral of flux error, speed reference, and speed feedback. The action space was defined as the d - q axis components of the control winding reference voltage vector. Meanwhile, a reward function was designed to integrate torque ripple suppression, and flux deviation penalty, guiding the TD3-agent to learn optimal voltage decisions in a continuous action space. [Results] Simulations were carried out under typical operating conditions such as speed step change and sudden load variation to compare the performance of the proposed DRL-

based SVM-DTC strategy and the traditional SVM-DTC strategy. The simulation results showed that the proposed method achieved basically consistent speed control performance with the traditional SVM-DTC strategy.

[Conclusion] The well-trained Actor network can directly generate the control winding reference voltages according to the system's real-time states, obviating the parameter tuning process of the inner-loop torque and flux PI controllers. The TD3-based SVM-DTC strategy can achieve high-quality control of torque and flux for BDFM, exhibiting excellent dynamic response and robustness.

Key words: brushless doubly-fed machine; space vector modulation direct torque control; deep reinforcement learning; twin delayed deep deterministic policy gradient

摘 要: 【目的】在无刷双馈电机(BDFM)的空间矢量调制直接转矩控制(SVM-DTC)系统中,内环通常采用转矩与磁链比例积分(PI)控制器。然而固定的PI参数难以适应系统非线性特性以及频繁变化的运行工况,难以在全工作点维持最优控制性能。为此,本文提出一种基于深度强化学习(DRL)的SVM-DTC策略,摆脱对人工参数整定的依赖,提升系统动态性能与鲁棒性。【方法】采用双延迟深度确定性策略梯度(TD3)算法,构建并训练Actor-Critic神经网络代替内环转矩和磁链PI调节器。状态空间由转矩、磁链、转矩偏差、磁链偏差、转矩偏差积分、磁链偏差积分、转速参考值以及转速反馈值构成;动作空间为控制绕组参考电压矢量的 d 、 q 轴分量。设计一种综合转矩脉动抑制和磁链偏差惩罚的奖励函数,引导TD3智能体在连续动作空间中学习最优电压决策。【结果】在

基金项目: 国家自然科学基金项目(62173151)

National Natural Science Foundation of China (62173151)

转速阶跃、负载突变等典型运行工况下开展仿真,对本文所提基于 DRL 的 SVM-DTC 策略与传统 SVM-DTC 策略进行性能分析。仿真结果表明,所提策略与传统 SVM-DTC 策略的转速控制效果基本一致。【结论】经充分训练的 Actor 网络能够依据系统实时状态直接生成控制绕组参考电压,省去内环转矩与磁链 PI 调节器的参数整定环节。基于 TD3 的 SVM-DTC 策略能够实现对 BDFM 转矩与磁链的高品质控制,具有优异的动态响应能力与鲁棒性。

关键词: 无刷双馈电机;空间矢量调制直接转矩控制;深度强化学习;双延迟深度确定性策略梯度

0 引言

无刷双馈电机 (Brushless Doubly-Fed Machine, BDFM) 因其无刷化结构、所需变流器容量小以及可靠性高等优点,成为风机、泵类应用领域潜在的优秀机型^[1-2]。BDFM 转子结构通常为磁阻和笼型形式,分别称为无刷双馈磁阻电机 (Brushless Doubly-Fed Reluctance Machine, BDFRM) 和无刷双馈感应电机 (Brushless Doubly-Fed Induction Machine, BDFIM)。相较于 BDFIM, BDFRM 具有效率高、建模简单以及转矩和无功率可解耦控制的优点^[3]。

BDFRM 定子上有两套极数不同的绕组,分别称为功率绕组 (Power Winding, PW) 和控制绕组 (Control Winding, CW)。PW 直接连接电网;CW 连接变流器,该变流器仅提供或回馈滑差功率。BDFRM 转子极数设计为两套定子绕组极数总和的一半,以实现两套不同极数的定子绕组间的电磁耦合。这种特殊的结构和运行机理,使得与其控制的相关研究相较于常见的感应电机和永磁同步电机 (Permanent Magnet Synchronous Machine, PMSM) 更加困难,且发展相对滞后。经过多年探索,磁场定向控制^[4-6]、直接转矩控制 (Direct Torque Control, DTC)^[7-8]、转矩与无功功率控制^[9-10]、直接功率控制^[11]以及空间矢量调制直接转矩控制 (Space Vector Modulation DTC, SVM-DTC)^[12]等控制方法已被应用于 BDFRM 控制系统。

以反向传播神经网络 (Back Propagation Neural Network, BPNN) 为代表的浅层或深层前馈网络^[13-15]在电机控制中展现出强大的非线性映射能力,常被用于参数辨识、无传感器位置估计以及控制器参数的自整定等。然而,从学习范式的视角来

看,这类应用本质上属于监督学习 (Supervised Learning, SL),其需要大量包含“标准答案”的标注数据,例如用实测的转子位置作为标签,或通过某种优化准则生成最优比例积分 (Proportional-Integral, PI) 参数作为训练目标来离线训练一个静态映射网络。这种“数据驱动拟合”的模式,虽能逼近复杂的非线性函数,但带来了其自身难以克服的局限:网络只能复现训练数据中蕴含的控制律,无法自主探索出超越标签策略的更优解;一旦训练完成,网络参数便固定,泛化能力有限,难以在线持续优化;每次映射都是瞬时的,不具备处理序列决策和跨时间尺度性能折中的能力。

上述局限以及计算机技术的发展,推动了强化学习 (Reinforcement Learning, RL) 在电机控制中的应用。与 BPNN 等 SL 范式不同,RL 不再试图学习一个输入到输出的标准映射,而是将一个完整的智能体置于运行环境中,通过试错交互直接学习一种状态到动作的策略,使长期累积奖励最大化。这一范式转变,给电机控制应用带来了几个关键突破:摆脱对标签数据的依赖,具备在线自学习与自适应潜力以及可实现全局与多目标优化。因此,近些年有学者将 RL 引入 PMSM 控制系统,旨在从交互决策层面突破传统智能方法的性能边界^[16]。例如,文献[17]采用 RL 算法求解 PMSM 改进模型预测控制系统的权重系数;文献[18-20]利用深度强化学习 (Deep RL, DRL) 来优化 PMSM 自抗扰控制 (Active Disturbance Rejection Control, ADRC) 系统的非线性 ADRC 参数。文献[21-22]针对由两电平逆变器供电的 PMSM,设计了一种无模型 DRL 转矩控制方法,利用逆变器开关状态有限这一特点,采用深度 Q 学习算法构建有限控制集,让智能体直接学习从电机状态到最优开关状态的映射。文献[23-24]提出了基于深度确定性策略梯度 (Deep Deterministic Policy Gradient, DDPG) 的 PMSM 电流控制策略,并建立了一套完整的快速控制原型工具链。文献[25]采用 RL 算法补偿 PMSM PI 控制器的输出。

上述研究初步展现了 RL 在电机控制中的应用价值,但将其用于 BDFRM 的 SVM-DTC,特别是实现连续电压空间矢量的智能决策与多控制目标的协调,仍是一个有待深入探索的课题。为此,本文将 DRL 引入 BDFRM 的 SVM-DTC 系统,提出一

基于双延迟深度确定性策略梯度(Twin Delayed Deep Deterministic Policy Gradient, TD3)算法的控制策略。TD3 算法适合处理连续动作空间且能有效缓解价值函数过估计问题,本文利用其构建 Actor-Critic 神经网络,以转矩、磁链及其偏差和偏差积分等系统状态为输入,由 Actor 网络直接输出 CW 的 d 、 q 轴参考电压,从而完全替代内环转矩和磁链 PI 调节器。同时,设计综合转矩脉动抑制和磁链偏差惩罚的奖励函数,引导智能体在转速阶跃、负载突变等工况下快速、平稳地跟踪给定转速。所提策略旨在摆脱繁琐的人工参数整定,提升系统的动态响应性能和鲁棒性。

1 BDFRM 数学模型

由于 BDFRM 中通以不同频率电流的 PW 和 CW 均处于定子上,因此将 PW 和 CW 各变量表示在 r1r2 双旋转坐标系下。r1、r2 分别为 PW、CW 同步旋转坐标系,将与 PW 电流矢量具有相同角频率的空间矢量转换到 r1 坐标系下,将与 CW 电流矢量具有相同角频率的空间矢量转换到 r2 坐标系下,如式(1)所示:

$$\begin{cases} \mathbf{u}_p^{r1} = R_p \mathbf{i}_p^{r1} + p\boldsymbol{\psi}_p^{r1} + j\omega \boldsymbol{\psi}_p^{r1} \\ \mathbf{u}_c^{r2} = R_c \mathbf{i}_c^{r2} + p\boldsymbol{\psi}_c^{r2} + j(\omega_r - \omega) \boldsymbol{\psi}_c^{r2} \\ \boldsymbol{\psi}_p^{r1} = L_p \mathbf{i}_p^{r1} + L_{pc} \mathbf{i}_{pc}^{r1} \\ \boldsymbol{\psi}_c^{r2} = L_c \mathbf{i}_c^{r2} + L_{pc} \mathbf{i}_{cp}^{r2} \end{cases} \quad (1)$$

式中: R 、 L 分别为电阻、电感; $\mathbf{u}$ 、 $\mathbf{i}$ 和 $\boldsymbol{\psi}$ 分别为电压、电流和磁链矢量; ω_r 为转子机械角速度; ω 为 r1 坐标系旋转角速度; $\omega_r - \omega$ 为 r2 坐标系旋转角速度; p 为微分算子;下标 p、c 分别为 PW 和 CW; L_{pc} 为 PW 和 CW 之间磁场调制互感。

式(1)中的电流矢量具有如式(2)所示的关系:

$$\begin{cases} \mathbf{i}_{pc}^{r1} = \mathbf{i}_{pc}^s e^{-j\theta} \\ \mathbf{i}_{cp}^{r2} = \mathbf{i}_{cp}^s e^{-j(\theta_r - \theta)} \\ \mathbf{i}_{pc}^s = \mathbf{i}_c^{s*} e^{j\theta_r} = \mathbf{i}_c^s e^{j(\omega_1 t + \alpha)} \\ \mathbf{i}_{cp}^s = \mathbf{i}_p^{s*} e^{j\theta_r} = \mathbf{i}_p^s e^{j\omega_2 t} \end{cases} \quad (2)$$

式中: $\mathbf{i}_{pc}^s$ 、 $\mathbf{i}_{cp}^s$ 分别为经转子磁场调制作用产生的与 $\mathbf{i}_p^s$ 、 $\mathbf{i}_c^s$ 具有相同角频率的电流矢量, $\mathbf{i}_p^s$ 、 $\mathbf{i}_c^s$ 分别为定子静止坐标系下的 PW、CW 电流矢量; ω_1 、 ω_2 分别为 $\mathbf{i}_p^s$ 、 $\mathbf{i}_c^s$ 的角速度; θ_r 为转子位置角; θ 为 r1 坐标系的位置角; $\theta_r - \theta$ 为 r2 坐标系的位置角; α 为

$\mathbf{i}_{pc}^s$ 与 $\mathbf{i}_p^s$ 之间的空间夹角,也是 $\mathbf{i}_{cp}^s$ 与 $\mathbf{i}_c^s$ 之间的空间夹角;上标 * 为共轭。

需要注意的是, $\mathbf{i}_{pc}^{r1}$ 与 $\mathbf{i}_p^{r1}$ 具有相同旋转角速度 $\omega_1 - \omega$, $\mathbf{i}_{cp}^{r2}$ 和 $\mathbf{i}_c^{r2*}$ 的实、虚部相同; $\mathbf{i}_{cp}^{r2}$ 与 $\mathbf{i}_c^{r2}$ 具有相同旋转角速度 $\omega - \omega_1$, $\mathbf{i}_{cp}^{r2}$ 和 $\mathbf{i}_p^{r1*}$ 的实、虚部相同。

式(1)中的磁链矢量具有如式(3)所示的关系:

$$\begin{cases} \boldsymbol{\psi}_{pc}^{r1} = \boldsymbol{\psi}_{pc}^s e^{-j\theta} \\ \boldsymbol{\psi}_{cp}^{r2} = \boldsymbol{\psi}_{cp}^s e^{-j(\theta_r - \theta)} \\ \boldsymbol{\psi}_{pc}^s = \boldsymbol{\psi}_c^{s*} e^{j\theta_r} \\ \boldsymbol{\psi}_{cp}^s = \boldsymbol{\psi}_p^{s*} e^{j\theta_r} \end{cases} \quad (3)$$

式中: $\boldsymbol{\psi}_{pc}^s$ 、 $\boldsymbol{\psi}_{cp}^s$ 分别为经转子磁场调制作用产生的与 $\boldsymbol{\psi}_p^s$ 、 $\boldsymbol{\psi}_c^s$ 具有相同角频率的磁链矢量, $\boldsymbol{\psi}_p^s$ 、 $\boldsymbol{\psi}_c^s$ 分别为定子静止坐标系下的 PW、CW 磁链矢量。

$\boldsymbol{\psi}_{pc}^{r1}$ 与 $\boldsymbol{\psi}_p^{r1}$ 具有相同旋转角速度 $\omega_1 - \omega$, $\boldsymbol{\psi}_{cp}^{r2}$ 和 $\boldsymbol{\psi}_c^{r2*}$ 的实、虚部相同; $\boldsymbol{\psi}_{cp}^{r2}$ 与 $\boldsymbol{\psi}_c^{r2}$ 具有相同旋转角速度 $\omega - \omega_1$, $\boldsymbol{\psi}_{cp}^{r2}$ 和 $\boldsymbol{\psi}_p^{r1*}$ 的实、虚部相同。

BDFRM 的电磁转矩 T_e 可以表示为

$$T_e = \frac{3p_r L_{pc}}{2L_p} \mathbf{i}_{pc}^{r1} \times \boldsymbol{\psi}_p^{r1} \quad (4)$$

式中: p_r 为转子极数。

r1r2 双旋转坐标系下 BDFRM 的空间矢量图如图 1 所示,假设 BDFRM 处于电动机模式超同步运行状态。图中, δ 为 $\boldsymbol{\psi}_c^{r2}$ 与 $\boldsymbol{\psi}_{cp}^{r2}$ 之间的空间夹角,也是 $\boldsymbol{\psi}_p^{r1}$ 与 $\boldsymbol{\psi}_{pc}^{r1}$ 之间的空间夹角,定义为 BDFRM 的负载角。需要明确的是:当 BDFRM 以发电机模式运行时, $\boldsymbol{\psi}_{cp}^{r2}$ 超前 $\boldsymbol{\psi}_c^{r2}$, $\boldsymbol{\psi}_{pc}^{r1}$ 超前 $\boldsymbol{\psi}_p^{r1}$, $\delta < 0$;当 BDFRM 以电动机模式运行时, $\boldsymbol{\psi}_{cp}^{r2}$ 滞后 $\boldsymbol{\psi}_c^{r2}$, $\boldsymbol{\psi}_{pc}^{r1}$ 滞后 $\boldsymbol{\psi}_p^{r1}$, $\delta > 0$ 。

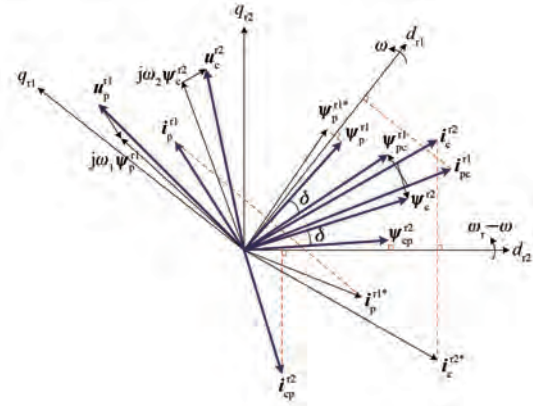


图 1 BDFRM 的空间矢量图

Fig. 1 Vector diagram of BDFRM

2 BDFRM SVM-DTC

BDFRM SVM-DTC 使用 CW 磁链和转矩 PI 调节器产生 CW 的 d 、 q 轴参考电压 $u_{cd_ref}^{r2}$ 、 $u_{cq_ref}^{r2}$ 。SVM-DTC 策略的实施基于 r1r2 双同步旋转坐标系: d_{r1} 轴沿 ψ_p^{r1} 定向, d_{r2} 轴沿 ψ_{cp}^{r2} 定向。磁场定向后, BDFRM 的电压和磁链方程简化为

$$\begin{cases} u_{pd}^{r1} = R_p i_{pd}^{r1} + p\psi_{pd}^{r1} \\ u_{pq}^{r1} = R_p i_{pq}^{r1} + \omega_1 \psi_{pd}^{r1} \\ u_{cd}^{r2} = R_c i_{cd}^{r2} + p\psi_{cd}^{r2} - \omega_2 \psi_{cq}^{r2} \\ u_{cq}^{r2} = R_c i_{cq}^{r2} + p\psi_{cq}^{r2} + \omega_2 \psi_{cd}^{r2} \end{cases} \quad (5)$$

$$\begin{cases} \psi_{pd}^{r1} = L_p i_{pd}^{r1} + L_{pc} i_{pd}^{r1} \\ 0 = L_p i_{pq}^{r1} + L_{pc} i_{pcq}^{r1} \\ \psi_{cd}^{r2} = L_c i_{cd}^{r2} + L_{pc} i_{cpd}^{r2} \\ \psi_{cq}^{r2} = L_c i_{cq}^{r2} + L_{pc} i_{cpq}^{r2} \end{cases} \quad (6)$$

式中: 下标 d 、 q 分别为矢量的 d 、 q 轴分量。

磁场定向后, 电磁转矩和 CW 磁链可表示为

$$T_e = \frac{3p_r L_{pc}}{2L_p} \psi_{pd}^{r1} i_{cq}^{r2} \quad (7)$$

$$\begin{cases} \psi_{cd}^{r2} = \frac{L_{pc}}{L_p} \psi_{pd}^{r1} + \sigma L_c i_{cd}^{r2} \\ \psi_{cq}^{r2} = \sigma L_c i_{cq}^{r2} \end{cases} \quad (8)$$

式中: σ 为漏磁系数, $\sigma = 1 - L_{pc}^2 / (L_p L_c)$ 。

由于 PW 直接连接电网, ψ_{pd}^{r1} 近似为常数, 由式(7)、(8)可以看出, T_e 和 ψ_{cd}^{r2} 分别仅受 i_{cq}^{r2} 和 i_{cd}^{r2} 控制。 i_{cq}^{r2} 为 CW 电流的转矩分量, i_{cd}^{r2} 为 CW 电流的励磁分量。

由式(5)~(8)可得 u_{cd}^{r2} 、 u_{cq}^{r2} 与 T_e 和 ψ_{cd}^{r2} 之间的关系, 如式(9)所示:

$$\begin{cases} u_{cd}^{r2} = \frac{R_c}{\sigma L_c} \psi_{cd}^{r2} + p\psi_{cd}^{r2} + u_{cd_com}^{r2} \\ u_{cq}^{r2} = \frac{2R_c L_p}{3p_r L_{pc} \psi_{pd}^{r1}} T_e + \frac{2\sigma L_c L_p}{3p_r L_{pc} \psi_{pd}^{r1}} pT_e + u_{cq_com}^{r2} \end{cases} \quad (9)$$

式中: $u_{cd_com}^{r2}$ 、 $u_{cq_com}^{r2}$ 分别为 CW 的 d 、 q 轴电压补偿量, 其表达式为

$$\begin{cases} u_{cd_com}^{r2} = -\frac{R_c L_{pc}}{\sigma L_c L_p} \psi_{pd}^{r1} - \omega_2 \frac{2\sigma L_c L_p}{3p_r L_{pc} \psi_{pd}^{r1}} T_e \\ u_{cq_com}^{r2} = \omega_2 \psi_{cd}^{r2} \end{cases} \quad (10)$$

根据式(9)、(10)可以方便地设计 CW 磁链和转矩 PI 调节器。BDFRM SVM-DTC 框图如图 2

所示。图中, ω_{r_ref} 、 T_{e_ref} 分别为转速、电磁转矩参考值; $u_{cd_pi}^{r2}$ 、 $u_{cq_pi}^{r2}$ 分别为 PI 调节器产生的 d 、 q 轴调节电压; $u_{c\alpha_ref}^s$ 、 $u_{c\beta_ref}^s$ 分别为 CW 的 α 、 β 轴参考电压; $\psi_{cd_ref}^{r2}$ 为 CW 的 d 轴磁链参考值; θ_p 为 ψ_p^s 的空间位置角。

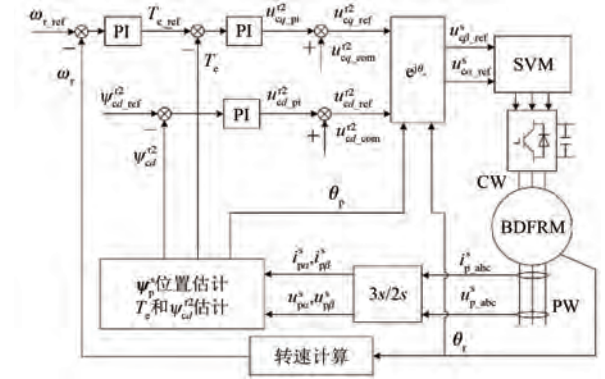


图 2 SVM-DTC 框图

Fig. 2 Block diagram of SVM-DTC

θ_p 可由式(11)、(12)估计得到:

$$\theta_p = \tan^{-1}(\psi_{p\beta}^s / \psi_{p\alpha}^s) \quad (11)$$

$$\begin{cases} \psi_{p\alpha}^s = \int_0^t (u_{p\alpha}^s - R_p i_{p\alpha}^s) dt \\ \psi_{p\beta}^s = \int_0^t (u_{p\beta}^s - R_p i_{p\beta}^s) dt \end{cases} \quad (12)$$

ψ_{cd}^{r2} 可由式(13)估计得到:

$$\psi_{cd}^{r2} = \frac{L_c}{L_{pc}} \psi_{pd}^{r1} - \frac{\sigma}{1 - \sigma} L_{pc} i_{pd}^{r1} \quad (13)$$

$\psi_{cd_ref}^{r2}$ 根据式(14)设定:

$$\psi_{cd_ref}^{r2} = k_1 \psi_{pd}^{r1} + k_2 \frac{Q_{p_ref}}{\psi_{pd}^{r1}} \quad (14)$$

式中: $k_1 = \frac{L_{pc}}{L_p} + \sigma$; $k_2 = -\frac{2\sigma L_c L_p}{3\omega_1 L_{pc}}$; Q_{p_ref} 为 PW 无功功率参考值。

若要实现单位功率因数运行, 则 $Q_{p_ref} = 0$; 若要实现 CW 的最大转矩电流比 (Maximum Torque Per Ampere, MTPA) 运行, 则式(8)中 $i_{cd}^{r2} = 0$, 此时 CW 电流中仅含有功分量, BDFRM 的励磁电流完全由 PW 提供。

3 基于 DRL 的 BDFRM SVM-DTC

3.1 TD3 算法

图 2 所示的 SVM-DTC 策略在通常情况下能取得较好的控制效果, 但其内环性能高度依赖 PI 调节器的参数整定与前馈补偿项的计算精度。固

定的 PI 参数难以在全工作点均保持最优,而补偿量 $\psi_{cd_com}^{r2}$ 、 $\psi_{cq_com}^{r2}$ 的计算又依赖于难以精确获取的电机参数,导致系统鲁棒性受限。上述问题的根源在于基于模型的设计方法难以完全适配 BDFRM 的非线性与多变性。近年来,RL 被引入电机控制领域,RL 的核心优势在于能够通过与环境的直接交互,以数据驱动的方式自主学习最优决策,从而摆脱对精确模型和固定控制律的依赖。RL 智能体与环境交互框图如图 3 所示。对于本文来说,图 3 中训练 RL 智能体所需的环境包括图 2 中除转矩和磁链 PI 调节器的所有环节。针对 SVM-DTC 需要连续电压控制的动作空间特性,本文引入 TD3 算法,该算法是面向连续动作空间的 DDPG 算法的代表性改进。

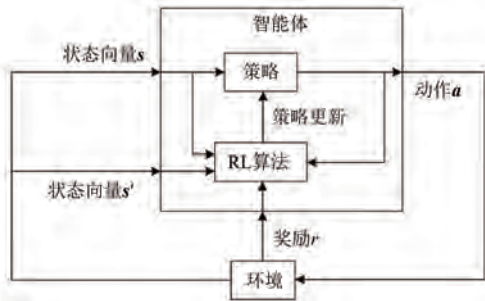


图 3 RL 智能体与环境交互框图

Fig. 3 Block diagram of RL-agent interacting with environment

TD3 算法从 DDPG 算法的过估计与训练不稳定问题出发,引入截断双 Q 学习、延迟策略更新以及目标策略平滑正则化三项核心机制。截断双 Q 学习通过维护两个独立的 Critic 网络并取其最小值构建时序差分 (Temporal Difference, TD) 目标,主动抑制 Q 值的系统性高估;延迟策略更新以较低频率更新 Actor 网络和目标 Actor 网络,使 Critic 的训练误差在干预策略前充分降低;目标策略平滑正则化向目标动作注入限幅噪声,强制 Q 函数在动作邻域内保持平滑,减少对 Critic 尖峰误差的过度依赖。三项机制互为补充,显著增强了训练的稳定性与样本效率,使 TD3 算法的性能在诸多连续控制基准上优于 DDPG 算法。

TD3 智能体由以下神经网络构成: Actor 网络 $\pi(\cdot|\phi)$ 、目标 Actor 网络 $\pi'(\cdot|\phi')$ 、Critic1 网络 $Q_{\theta_1}(\cdot|\theta_1)$ 、Critic2 网络 $Q_{\theta_2}(\cdot|\theta_2)$ 以及目标 Critic1 网络 $Q_{\theta'_1}(\cdot|\theta'_1)$ 、目标 Critic2 网络

$Q_{\theta'_2}(\cdot|\theta'_2)$ 。Actor 网络是策略的直接载体,通过输入状态 s ,输出确定性动作 a 。在与环境交互时,通过添加探索噪声来选择最终执行的动作;在更新阶段,Actor 网络参数 ϕ 通过最大化 Critic 的评估值来改进策略。目标 Actor 网络用于计算下一状态 s' 所对应的目标动作,是构造 TD 目标的重要组成部分。目标 Actor 网络本身不直接与环境交互,也不直接从损失函数更新,其参数 ϕ' 通过软更新逐步向 ϕ 靠拢。Critic1 网络和 Critic2 网络是两个结构相同但参数独立的 Q 函数,用于估计在状态 s 下采取动作 a 的期望回报,二者的参数 θ_1 、 θ_2 均通过最小化与 TD 目标之间的均方误差来更新。目标 Critic1 网络和目标 Critic2 网络用于参与构造 TD 目标的未来回报项,同样采用软更新方式。

3.2 基于 TD3 智能体的 SVM-DTC

基于 TD3 智能体的 SVM-DTC 框图如图 4 所示。 $\psi_{cd_ref}^{r2}$ 按 MTPA 设定, T_{e_ref} 由速度 PI 调节器产生。状态向量 s 为 8 维,包括 T_e 、 ψ_{cd}^{r2} 、电磁转矩偏差 ΔT_e 、磁链偏差 $\Delta\psi_{cd}^{r2}$ 、电磁转矩偏差积分 $\int\Delta T_e dt$ 、磁链偏差积分 $\int\Delta\psi_{cd}^{r2} dt$ 、 ω_{r_ref} 以及 ω_r 。动作向量 a 为 2 维,即 Actor 网络产生 $u_{cd_ref}^{r2}$ 和 $u_{cq_ref}^{r2}$,通过 Park 逆变换得到 $u_{\alpha_ref}^s$ 和 $u_{\beta_ref}^s$,进而由 SVM 获得三相逆变器中功率开关器件的开关信号。

Actor 网络和 Critic 网络结构需根据状态空间维度和动作空间维度进行设计。Actor 网络结构

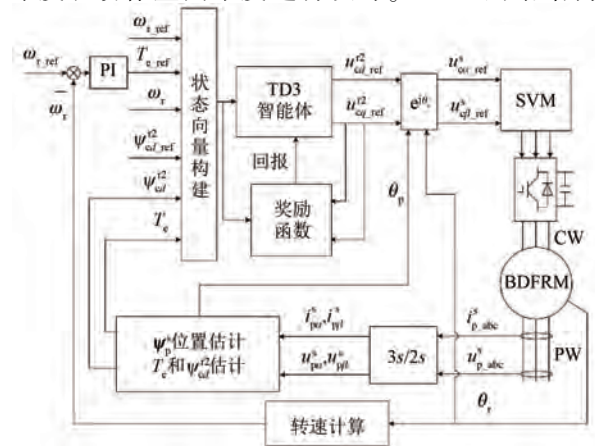


图 4 基于 TD3 智能体的 SVM-DTC 框图

Fig. 4 Block diagram of SVM-DTC based on TD3-agent

如图5所示。全连接层1包含64个节点,激活函数为 ReLU 函数;全连接层2包含32个节点,激活函数为 ReLU 函数;输出层包含2个节点,激活函数为 tanh 函数,输出的动作即为 $u_{cd_ref}^{r2}$ 和 $u_{cq_ref}^{r2}$ 。

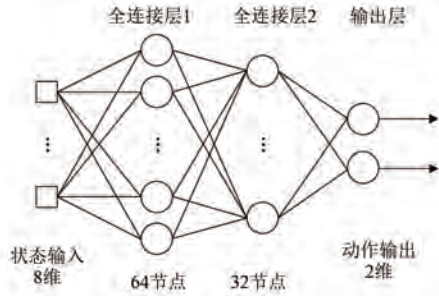


图5 Actor 网络结构

Fig. 5 Actor network architecture

Critic 网络结构如图6所示。状态输入所属的全连接层1包含64个节点;动作输入所属的全连接层1包含64个节点;加法层包含64个节点,激活函数为 ReLU 函数;全连接层2包含32个节点,激活函数为 ReLU 函数;输出层包含1个节点,输出的动作即为 Q 值。

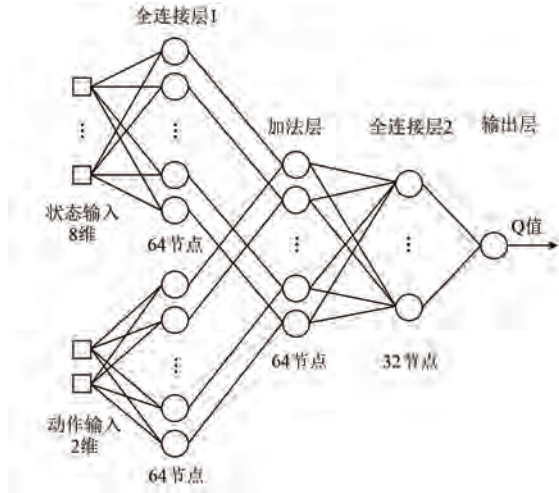


图6 Critic 网络结构

Fig. 6 Critic network architecture

TD3 算法中的奖励函数直接定义了价值网络的学习目标;Critic 网络通过最小化由奖励构造的 TD 误差来逼近真实 Q 值,随后 Actor 网络基于 Critic 网络的梯度方向更新策略以最大化长期累积奖励。因此,奖励的密度、尺度与形态不仅决定了策略优化的可行性与收敛速度,还通过影响价值函数的估计误差传递间接作用于 TD3 算法的截断双 Q 学习和目标平滑机制。合理的奖励设

计是保证训练稳定性和最终策略质量的关键前提。

为了引导 TD3 智能体根据转矩和磁链偏差跟踪转矩和磁链,将转矩和磁链跟踪奖励 r_T 和 r_ψ 设计为

$$\begin{cases} r_T = -a_1 |\Delta T_e| - a_2 |\Delta T_e|^2 \\ r_\psi = -a_3 |\Delta \psi_{cd}^{r2}| - a_4 |\Delta \psi_{cd}^{r2}|^2 \end{cases} \quad (15)$$

式中: a_1, a_2, a_3 和 a_4 为设计参数。

为避免转矩和磁链在控制过程中出现突变,将转矩和磁链跟踪的平滑性奖励 $r_{T,\psi}^{\text{smooth}}$ 设计为

$$r_{T,\psi}^{\text{smooth}} = a_5 e^{[-a_6(|\Delta T_e| + |\Delta \psi_{cd}^{r2}|)]} \quad (16)$$

式中: a_5, a_6 为设计参数。

为减轻逆变器饱和压力,在满足转矩和磁链跟踪的前提下使用更小的控制电压,将电压指令奖励 r_u 设计为

$$r_u = -a_7 [(u_{cd_ref}^{r2})^2 + (u_{cq_ref}^{r2})^2] \quad (17)$$

式中: a_7 为设计参数。

为了抑制电压指令抖动,减少电流或转矩高频纹波,将电压指令平滑性奖励 r_u^{smooth} 设计为

$$r_u^{\text{smooth}} = -a_8 [(\Delta u_{cd_ref}^{r2})^2 + (\Delta u_{cq_ref}^{r2})^2] \quad (18)$$

式中: a_8 为设计参数; $\Delta u_{cd_ref}^{r2}$ 、 $\Delta u_{cq_ref}^{r2}$ 分别为 $u_{cd_ref}^{r2}$ 、 $u_{cq_ref}^{r2}$ 在一个周期的增量。

综上所述,总奖励 r 为

$$r = r_T + r_\psi + r_u + r_{T,\psi}^{\text{smooth}} + r_u^{\text{smooth}} \quad (19)$$

TD3 智能体的训练过程如下:采集每次 TD3 智能体与环境交互时的 PW 三相电流和电压,估计出 T_e 和 ψ_{cd}^{r2} ,构建状态向量 s ;由 Actor 网络获得动作向量 a ,执行动作 a 后获得下一状态 s' ;计算奖励 r ;将经验 (s, a, s', r) 存入回放缓冲区;当缓冲区中的样本达到一定数量后,从回放缓冲区中采样一个小批量数据,根据此数据计算 TD 目标,更新 Critic1 和 Critic2 网络。

利用目标 Actor 网络计算状态 s' 下的动作 $a' = \pi'(s' | \phi')$,并在动作 a' 上加入正态分布噪声 ε ,得到 $\tilde{a}' = a' + \varepsilon$;然后根据目标 Critic1 网络和目标 Critic2 网络中的较小者计算奖励目标值 y ,即:

$$y = r + \gamma \min_{i=1,2} Q_{\theta'_i}(s', \tilde{a}' | \theta'_i) \quad (20)$$

利用梯度下降算法获得使 Critic1 网络和 Critic2 网络的评估值与目标值 y 之间误差最小化的参数 θ_1 与 θ_2 ,并更新 Critic1 网络和 Critic2 网络,即:

$$\theta_i = \operatorname{argmin}_{\theta_i} \frac{1}{N} \sum [y - Q_{\theta_i}(s, a | \theta_i)]^2 \quad (21)$$

式中: $i=1, 2; N$ 为小批量数据大小。

采用延迟频率更新 Actor 网络和目标 Actor 网络。首先利用 Actor 网络计算出状态 s' 的动作 a' , 即 $a' = \pi(s' | \phi)$; 然后利用 Critic1 网络计算动作价值 $Q = Q_{\theta_1}(s', a' | \theta_1)$; 最后利用梯度上升算法获得使 Q 最大的参数 ϕ , 并完成 Actor 网络的更新, 梯度策略为

$$\nabla_{\phi} J(\phi) = \frac{1}{N} \sum \nabla_a Q_{\theta_1}(s, a) \big|_{a=\pi_{\phi}(s)} \nabla_{\phi} \pi_{\phi}(s) \quad (22)$$

采用软更新方式对目标 Critic1 网络、目标 Critic2 网络以及目标 Actor 网络进行更新, 即:

$$\begin{cases} \theta'_i = \tau \theta_i + (1 - \tau) \theta'_i \\ \phi' = \tau \phi + (1 - \tau) \phi' \end{cases} \quad (23)$$

式中: τ 为软更新系数。

4 仿真分析

在相同仿真工况下, 对 BDFRM 传统 SVM-DTC 策略与本文所提基于 DRL 的 SVM-DTC 策略的控制效果进行对比分析。

4.1 SVM-DTC 策略仿真结果

传统 SVM-DTC 策略使用 CW 磁链和转矩 PI 调节器产生 $u_{cd_ref}^{r2}$ 和 $u_{eq_ref}^{r2}$ 。PW 直接与电网相连, 因此其三相电压固定为 380 V, 50 Hz; $\psi_{cd_ref}^{r2}$ 按 MTPA 设定为 0.783 Wb; 速度外环以 1 kHz 频率运行, 转矩与磁链控制内环以 10 kHz 频率运行。BDFRM 主要参数如表 1 所示。根据表 1 中 PW 和 CW 极数可以计算得到 BDFRM 的同步速为 500 r/min。考虑到 BDFRM 的典型运行范围通常为同步速的 $\pm 30\%$, 因此仿真时在以下范围内设定目标转速与负载转矩: 转速范围为 350~650 r/min, 转矩范围为 50~150 N·m。转速跟踪结果如图 7 所示, 相应的转矩响应波形如图 8 所示, CW 的 d 轴磁链响应波形如图 9 所示。

如图 7 和图 8 所示, 仿真工况涵盖转速大范围变化与负载突变。初始转速为 500 r/min, 0.5 s 时匀减速至 350 r/min 并维持; 1 s 时开始匀加速, 至 2 s 时升至 650 r/min 并维持。负载方面, 0~2.5 s 为空载, 2.5 s 时突加 150 N·m 负载转矩。

表 1 BDFRM 主要参数
Tab. 1 Main parameters of BDFRM

参数名称	参数值
额定功率/kW	25
PW、CW 极数	8、4
L_p /mH	47.66
L_c /mH	40.75
L_{pc} /mH	37.68
R_p/Ω	0.387 1
R_c/Ω	0.377 3

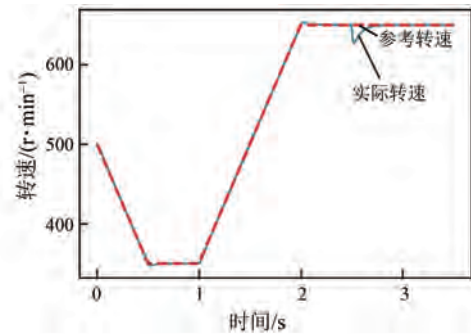


图 7 SVM-DTC 策略下的转速波形

Fig. 7 Speed waveforms under SVM-DTC strategy

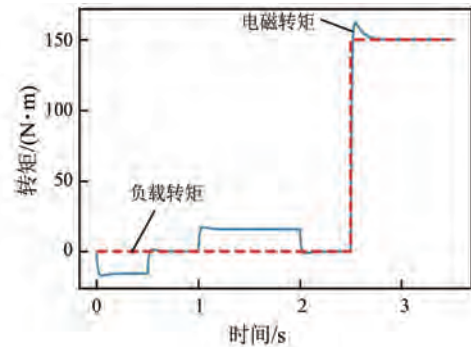


图 8 SVM-DTC 策略下的转矩波形

Fig. 8 Torque waveforms under SVM-DTC strategy

由 CW 磁链和转矩 PI 调节器产生的 $u_{cd_ref}^{r2}$ 和 $u_{eq_ref}^{r2}$ 波形如图 10 所示。

4.2 基于 DRL 的 SVM-DTC 策略仿真结果

为训练 TD3 智能体, 基于 Python 搭建交互式仿真环境。该环境包括 BDFRM 的动态数学模型、坐标变换以及速度外环 PI 控制器等必要环节, 以模拟 BDFRM 在实际工况下的连续响应。

环境根据 TD3 智能体输出的 CW 参考电压实时解算 BDFRM 状态, 并依据所设计的奖励函数计算即时奖励与下一时刻的状态向量, 形成完

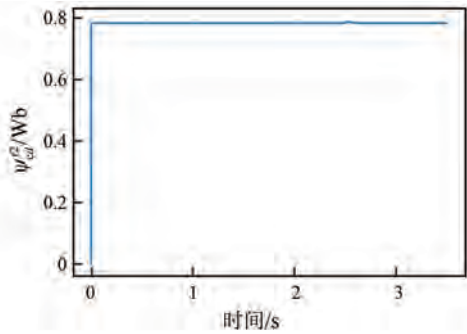


图9 SVM-DTC策略下CW的d轴磁链波形

Fig. 9 d-axis flux waveform of CW under SVM-DTC strategy

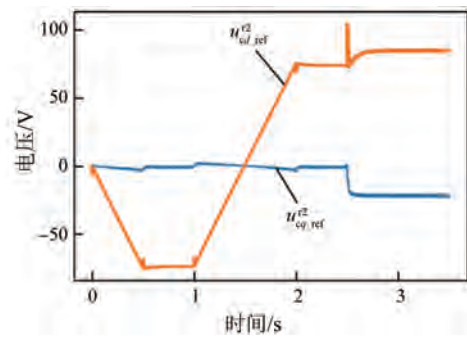


图10 SVM-DTC策略下CW的dq轴参考电压波形

Fig. 10 dq-axis voltage reference waveforms of CW under SVM-DTC strategy

整的马尔可夫决策过程交互闭环。TD3 智能体采用 PyTorch 深度学习框架实现,分别构建用于输出连续电压指令的 Actor 网络和用于评估动作价值的 Critic 网络,并配置相应的目标网络、优化器与经验回放池,以支撑 TD3 算法的稳定训练。为增强 TD3 智能体对不同工况的泛化能力,训练时在以下范围内设定目标转速与负载转矩:转速范围为 350~650 r/min,转矩范围为 50~150 N·m。

训练过程共包含 100 个回合,每个回合从工况范围内选取一组目标转速与负载转矩。控制系统采用双时间尺度架构,速度外环以 1 kHz 频率运行,转矩与磁链控制内环以 10 kHz 频率运行,后者也是 TD3 智能体与环境的交互频率。在交互与更新机制上,TD3 智能体每与环境交互 4 次,Critic 网络更新 1 次;延迟策略更新间隔设为 3,即 Critic 网络更新 3 次后,Actor 网络与目标网络更新 1 次。TD3 算法的其余超参数配置如表 2 所示。

在训练过程中,需根据奖励曲线的收敛趋势,

对 Actor 学习率、Critic 学习率及探索噪声等关键超参数进行多轮迭代调整,直至累计奖励稳定收敛。图 11 为累计奖励及其滑动平均值曲线。

表 2 TD3 算法超参数配置

Tab. 2 Hyperparameter configuration for the TD3 algorithm

参数名称	参数值
Actor 学习率	5e-5
Critic 学习率	3e-5
缓冲区大小	100 000
小批量数据大小	512
折扣因子	0.99
软更新系数	0.001
策略噪声	0.1
探索噪声	0.04
探索噪声衰减率	0.995
最小探索噪声	0.01

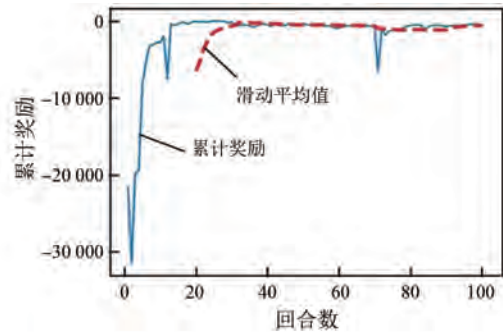


图 11 训练累计奖励

Fig. 11 Cumulative reward during training

由图 11 可知,整体学习曲线呈稳定上升态势并趋于收敛;训练期间偶有小幅下跌后迅速恢复,属于 DRL 训练的正常现象。这是因为 TD3 算法作为一种面向连续控制的不确定性策略算法,其训练过程本身并不完全平滑。探索噪声的阶段切换、Critic 网络在特定状态区域的估值偏差、环境固有的随机性或稀疏奖励特性以及经验回放池中不同采样数据的混合效应均可能引起奖励的短期波动,但这些波动并不影响策略最终收敛至最优。

为验证所提方法的控制效果,将训练收敛后的 Actor 网络参数固定,仅以系统状态向量 s 作为输入,直接输出 u_{d-ref}^{r2} 和 u_{q-ref}^{r2} ,构成端到端的电压决策,不再依赖转矩和磁链 PI 调节器。转速跟踪结果如图 12 所示。当转速指令由斜坡转为恒值

时,实际转速仅产生 2 r/min 的超调;在 2.5 s 突加 150 N·m 负载的瞬间,转速跌落 20 r/min,随后迅速恢复稳定。

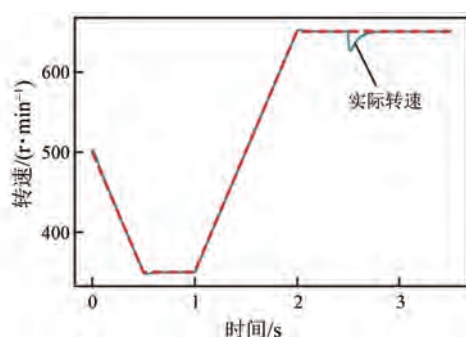


图 12 基于 DRL 的 SVM-DTC 策略下的转速波形

Fig. 12 Speed waveforms under DRL-based SVM-DTC strategy

对比图 12 和图 7 可知,所提基于 DRL 的 SVM-DTC 的转速控制效果与基于 CW 磁链和转矩 PI 调节器的传统 SVM-DTC 的转速控制效果基本一致,证明了所提策略在宽转速范围和负载扰动下均具备良好的转速跟踪与抗扰动能力。

基于 DRL 的 SVM-DTC 策略下的转矩响应波形如图 13 所示,其变化趋势与转速工况完全对应。0~0.5 s,电磁转矩小于负载转矩,电机处于减速状态;0.5~1 s,电磁转矩与负载转矩达到平衡,电机维持 350 r/min 空载匀速运行;1~2 s,电磁转矩大于负载转矩,电机加速至 650 r/min;2~2.5 s,转矩再次平衡,电机保持高速空载匀速;2.5 s 突加 150 N·m 负载后,电磁转矩迅速响应,经过短暂的动态调节后与负载转矩达到新的平衡,电机随即进入带载匀速稳态。整个过程中,实际电磁转矩展现出良好的动态响应与稳态精度。

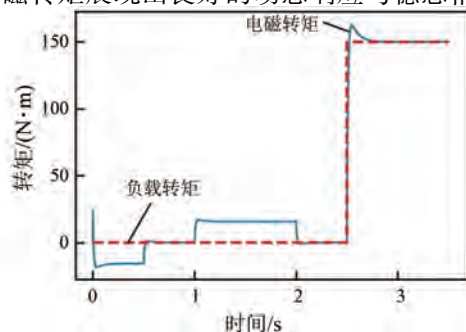


图 13 基于 DRL 的 SVM-DTC 策略下的转矩波形

Fig. 13 Torque waveforms under DRL-based SVM-DTC strategy

对比图 13 和图 8 可知,图 13 中的电磁转矩在仿真初始时刻有瞬间超调,这可能是因为基于 DRL 的内环对于工作点变化更加敏感。

基于 DRL 的 SVM-DTC 策略下 CW 的 d 轴磁链响应波形如图 14 所示。对比图 14 和图 9 可知,图 14 中的 d 轴磁链在速度变化时小幅度偏离参考值,而图 9 中的 d 轴磁链始终稳定在参考值。这是由于 Actor 网络的状态向量包含了转速参考值 ω_{r_ref} 和转速反馈值 ω_r 。

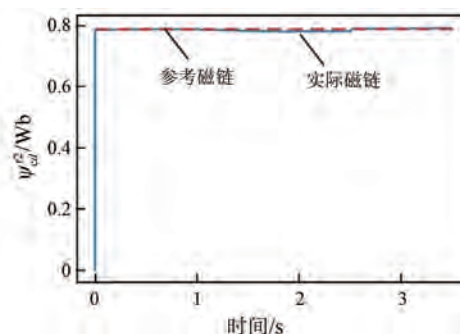


图 14 基于 DRL 的 SVM-DTC 策略下 CW 的 d 轴磁链波形

Fig. 14 d -axis flux waveforms of CW under DRL-based SVM-DTC strategy

基于 DRL 的 SVM-DTC 策略下 CW 的 A 相电流波形如图 15 所示。由于 $\psi_{cd_ref}^2$ 按 MTPA 设定, CW 不承担励磁功能,空载时仅提供微小的有功电流以克服机械损耗,因此电流幅值很小。2.5 s 突加 150 N·m 负载后, CW 电流迅速响应,稳态幅值为 22.2 A,与负载转矩的增量相匹配。

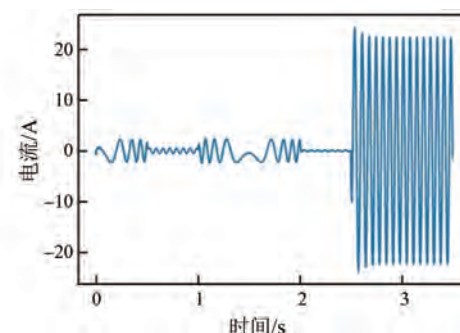


图 15 基于 DRL 的 SVM-DTC 策略下 CW 的 A 相电流波形

Fig. 15 A-phase current waveform of CW under DRL-based SVM-DTC strategy

为进一步分析加减速过程中 CW 电流的频率与相序特性,图 16 给出了 1~2 s 区间的 CW 三相电流波形。1~1.5 s, BDFRM 处于亚同步 (<500 r/min), CW 三相电流相序为 C-B-A; 1.5~

2 s, BDFRM 进入超同步 (>500 r/min), CW 三相电流相序反转为 A-B-C。两种运行状态下, CW 电流频率均为 15 Hz, 与当前转速相对于同步速的转差率相对应。CW 三相电流相序随运行状态的自然切换, 正是 BDFRM 可在亚同步与超同步下实现能量双向流动的典型特征, 也验证了所提控制策略能够正确驱动 BDFRM 在宽转速范围内稳定工作。

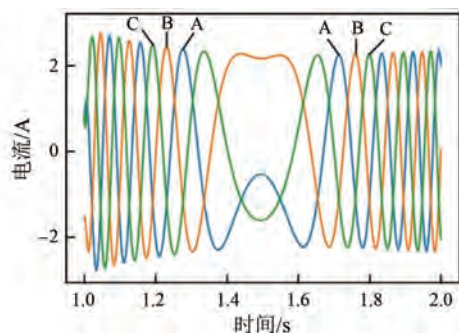


图 16 空载运行时 CW 三相电流波形

Fig. 16 Three phase current waveforms of CW when no-load operation

PW 三相电流波形如图 17 所示。由于 $\psi_{cd_ref}^{r2}$ 按 MTPA 设定, CW 不承担励磁功能, BDFRM 所需的全部励磁电流均由 PW 承担。因此即使 BDFRM 处于空载状态 (0~2.5 s), PW 仍需提供幅值为 21.5 A 的无功励磁电流。2.5 s 突加 150 N·m 负载后, PW 电流幅值增至 24.2 A。

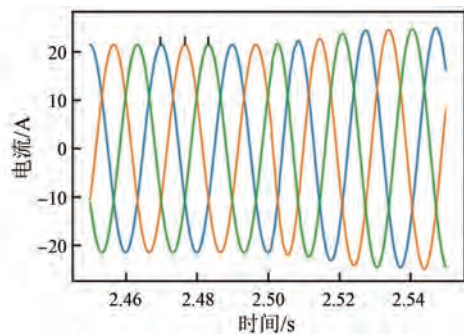


图 17 空载运行时 PW 三相电流波形

Fig. 17 Three phase current waveforms of PW when no-load operation

由状态向量 s 经训练后的 Actor 网络推理所得的 $u_{cd_ref}^{r2}$ 和 $u_{cq_ref}^{r2}$ 波形如图 18 所示。当 BDFRM 空载运行时, $u_{cd_ref}^{r2}$ 趋近于 0, $u_{cq_ref}^{r2}$ 波形与转速变化趋势相似。由式 (5) 可知, 在恒磁链和忽略 CW 电阻压降的条件下, u_{cq}^{r2} 近似与 ω_2 成正比, 而 $\omega_1 +$

$\omega_2 = \omega_r$, 因此加减速阶段 ω_2 变化直接导致电压幅值相应改变。电压指令从系统层面验证了 Actor 网络已成功学习到 BDFRM 的物理运行规律。对比图 18 和图 10 可知, 图 18 中的 $u_{cq_ref}^{r2}$ 在仿真初始时刻有瞬间超调, $u_{cd_ref}^{r2}$ 在突加转矩时有瞬间超调, 这可能是因为基于 DRL 的内环对于工作点变化更加敏感。

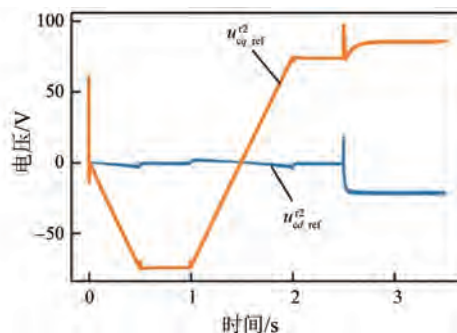


图 18 基于 DRL 的 SVM-DTC 策略下的 CW dq 轴参考电压波形

Fig. 18 dq-axis reference voltage waveforms of CW under DRL-based SVM-DTC strategy

5 结语

本文提出一种基于 DRL 的 SVM-DTC 策略, 为 BDFRM 控制提供了一种数据驱动的智能决策新范式。该策略将 BDFRM 的转矩和磁链调节问题建模为马尔可夫决策过程, 采用适合连续控制的 TD3 算法训练智能体, 以 Actor 网络直接映射出 CW 最优参考电压, 实现了端到端的电压决策。所设计的奖励函数融合了转矩脉动、磁链偏差及电压约束等多项, 使智能体在无需人工调参的前提下自动学习动态响应与稳态精度之间的最优折中策略。

仿真结果初步验证了所提策略的有效性, 但该策略仍存在训练超参数敏感、仿真到实物迁移差距、嵌入式计算资源受限以及黑箱输出缺乏稳定性保证等局限性。后续工作应围绕域随机化训练以增强策略鲁棒性、模型轻量化与安全监控机制设计以适应嵌入式部署、结合控制理论提供可解释的稳定性评估, 以及搭建试验台架开展实物闭环验证等方向深入推进, 以促进该数据驱动方法在电机控制领域的实际应用。

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

施隆进行了方案设计、仿真研究、内容总结与论文撰写, 鉴航参与了仿真研究、论文的修改与审核, 彭云建、仲兆峰和郭伟文参与了论文的审核。所有作者均阅读并同意了最终稿件的提交。

Shi Long was responsible for scheme design, simulation study, content summary and paper writing. Jian Hang participated in simulation study, paper revision and review. Peng Yunjian, Zhong Zhaofeng and Guo Weiwen conducted the paper review. All authors have read the last version of paper and consented for submission.

参考文献

- [1] Cheng M, Han P, Buja G, et al. Emerging multiport electrical machines and systems: Past developments, current challenges, and future prospects [J]. *IEEE Transactions on Industrial Electronics*, 2018, 65 (7): 5422-5435.
- [2] 于克训, 陈曦, 谢贤飞, 等. 无刷双馈电机研究综述与展望[J]. *电工技术学报*, 2023, 39(2): 397-422.
Yu K X, Chen X, Xie X F, et al. Overview and prospect of the brushless doubly-fed machine research [J]. *Transactions of China Electrotechnical Society*, 2023, 39(2): 397-422.
- [3] Ademi S, Jovanovic M G, Hasan M. Control of brushless doubly-fed reluctance generators for wind energy conversion systems [J]. *IEEE Transactions on Energy Conversion*, 2015, 30(2): 596-604.
- [4] Xu L Y, Li Z, Kim E H. Field-orientation control of a doubly excited brushless reluctance machine [J]. *IEEE Transactions on Industry Applications*, 1998, 34(1): 148-155.
- [5] Ademi S, Jovanovic M G. Vector control methods for brushless doubly fed reluctance machines [J]. *IEEE Transactions on Industrial Electronics*, 2015, 62 (1): 596-604.
- [6] Attya A B, Ademi S, Jovanovic M G, et al. Frequency support using doubly fed induction and reluctance wind turbine generators [J]. *International Journal of Electrical Power & Energy Systems*, 2018, 101: 403-414.
- [7] Jovanovic M G, Yu J, Levi E. Direct torque control of brushless doubly fed reluctance machines [J]. *Electric Power Components and Systems*, 2004, 32 (10): 941-958.
- [8] Jovanovic M G, Yu J, Levi E. Encoderless direct torque controller for limited speed range applications of brushless doubly fed reluctance motors [J]. *IEEE Transactions on Industry Applications*, 2006, 42 (3): 712-722.
- [9] Chaal H, Jovanovic M G. Practical implementation of sensorless torque and reactive power control of doubly fed machines [J]. *IEEE Transactions on Industrial Electronics*, 2012, 59(6): 2645-2653.
- [10] Chaal H, Jovanovic M G. Toward a generic torque and reactive power controller for doubly fed machines [J]. *IEEE Transactions on Power Electronics*, 2012, 27(1): 113-121.
- [11] Chaal H, Jovanovic M G. Power control of brushless doubly-fed reluctance drive and generator systems [J]. *Renewable Energy*, 2012, 37(1): 419-425.
- [12] Shi L, Jin S. Direct torque control and space vector modulation-based direct torque control of brushless doubly-fed reluctance machines [J]. *IET Electric Power Applications*, 2023, 17(8): 1069-1080.
- [13] 耿建平, 闫俞佰, 熊光阳, 等. 基于 BP 神经网络的永磁同步电机转矩观测器设计[J]. *电机与控制应用*, 2020, 47(1): 78-83.
Geng J P, Yan Y B, Xiong G Y, et al. Design of torque observer based on BP neural network for permanent magnet synchronous motor [J]. *Electric Machines & Control Application*, 2020, 47(1): 78-83.
- [14] 赵剑飞, 丁朋飞, 翟雪松. 基于模糊神经网络的纯电动车永磁同步电机矢量控制[J]. *电机与控制应用*, 2018, 45(12): 108-112.
- [15] 霍召晗, 许鸣珠. 基于小波神经网络 PID 的永磁同步电机转速控制[J]. *电机与控制应用*, 2019, 46(11): 1-6.
- Huo Z H, Xu M Z. Speed control of permanent

- magnet synchronous motor based on wavelet neural network PID [J]. *Electric Machines & Control Application*, 2019, 46(11): 1-6.
- [16] 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.
- [17] Gao J, Zhang J, Fan M Y, et al. Model predictive control of permanent magnet synchronous motor based on state transition constraint method [J]. *Mathematical Problems in Engineering*, 2021, 11, 3171417.
- [18] Wang Y C, Fang S H, Hu J X, et al. Multiscenarios parameter optimization method for active disturbance rejection control of PMSM based on deep reinforcement learning [J]. *IEEE Transactions on Industrial Electronics*, 2023, 70 (11): 10957-10968.
- [19] Wang Y C, Fang S H, Hu J X. Active disturbance rejection control based on deep reinforcement learning of PMSM for more electric aircraft [J]. *IEEE Transactions on Power Electronics*, 2023, 38 (1): 406-416.
- [20] Wang Y C, Fang S H, Hu J X, et al. A novel active disturbance rejection control of PMSM based on deep reinforcement learning for more electric aircraft [J]. *IEEE Transactions on Energy Conversion*, 2023, 38 (2): 1461-1470.
- [21] 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.
- [22] 范慧妍, 王爽. 基于深度强化学习的永磁同步电机控制算法研究[J]. *电机与控制应用*, 2026, 53 (3): 269-278.
- Fan H Y, Wang S. Research on control algorithm of permanent magnet synchronous motors based on deep reinforcement learning [J]. *Electric Machines & Control Application*, 2026, 53(3): 269-278.
- [23] Schenke M, Kirchgassner 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.
- [24] 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.
- [25] Peng Y Y, Luo X, Wu Z X, et al. Reinforcement learning-based feedforward compensation control in PMSM [C]//2023 IEEE 6th Student Conference on Electric Machines and Systems, HuZhou, China, 2023: 1-4.

收稿日期:2026-05-25

收到修改稿日期:2026-07-01

作者简介:

施 隆(1992—),男,博士,博士后研究员,研究方向为机器学习技术在电机控制领域的应用,s_longsy@163.com;

* 通信作者:施 隆(1992—),男,博士,博士后研究员,研究方向为机器学习技术在电机控制领域的应用,s_longsy@163.com。