

考虑驾驶风格的混合动力汽车 强化学习能量管理策略

施德华^{1,2}, 袁超¹, 汪少华^{1,2}, 周卫琪¹, 陈龙¹

(1. 江苏大学汽车工程研究院, 212013, 江苏镇江;

2. 江苏省新能源汽车电驱动系统与智能控制工程研究中心, 212013, 江苏镇江)

摘要: 为了提升混合动力汽车能量管理策略对不同风格驾驶员的适应性,基于深度强化学习和等效燃油消耗最小策略(equivalent consumption minimization strategy, ECMS),提出一种考虑驾驶风格的混合动力汽车能量管理策略。通过实车试验采集驾驶员行驶数据,基于采集数据进行驾驶员驾驶风格的聚类分析,建立驾驶风格识别模型;构建基于强化学习和 ECMS 的能量管理策略,将驾驶风格系数作为强化学习状态变量,利用多种驾驶风格的组合工况训练深度确定性策略梯度智能体,获取不同工况和驾驶风格下 ECMS 等效因子,采用 ECMS 求解最优发动机、电机转矩分配以及变速箱挡位;搭建硬件在环测试平台,并基于实际采集的不同驾驶员驾驶数据构建测试工况,验证所提出控制策略的有效性。研究结果表明,相较于基于规则策略、基于等效因子比例修正的自适应 ECMS 以及 DRL-SAC 策略,提出的考虑驾驶风格的强化学习能量管理策略使整车能量消耗分别降低 16.35%、11.11%和 7.56%,所提控制策略的有效性得到了验证。

关键词: 混合动力汽车;能量管理策略;驾驶风格;强化学习;等效燃油消耗最小策略

中图分类号: U463.2 **文献标志码:** A

DOI: 10.7652/xjtuxb202410005 **文章编号:** 0253-987X(2024)10-0051-12

Reinforcement Learning-Based Energy Management Strategy Considering Driving Style for Hybrid Electric Vehicle

SHI Dehua^{1,2}, YUAN Chao¹, WANG Shaohua^{1,2}, ZHOU Weiqi¹, CHEN Long¹

(1. Automotive Engineering Research Institute, Jiangsu University, Zhenjiang, Jiangsu 212013, China;

2. Jiangsu Province Engineering Research Center of Electric Drive System and Intelligent Control for Alternative Vehicles, Zhenjiang, Jiangsu 212013, China)

Abstract: A considering driving style energy management strategy is proposed to enhance the adaptability of energy management strategies for different driving styles in hybrid electric vehicles. The strategy combines deep reinforcement learning with the equivalent consumption minimization strategy (ECMS). Real vehicle experiments are conducted to collect driving data, which is then subjected to clustering analysis to identify distinct driving styles. A driving style recognition model is developed based on this data. The energy management strategy is built using reinforcement learning and ECMS, with driving style coefficients serving as the reinforcement learning state variables. A deep deterministic policy gradient agent is trained using various

combinations of driving styles and operating conditions to determine ECMS equivalent factors for different driving styles and conditions. The ECMS is employed to optimize the engine, motor torque allocation, and gearbox gear selection. To validate the effectiveness of the proposed control strategy, a hardware in the loop testing platform is constructed, and test scenarios are generated using real driving data from different drivers. The research findings demonstrate that the reinforcement learning-based energy management strategy considering driving style reduces overall vehicle energy consumption by 16.35%, 11.11%, and 7.56% compared with rule-based strategy, equivalent factor ratio correction-based adaptive ECMS, and DRL-SAC strategy, respectively. The effectiveness of the proposed control strategy is successfully validated.

Keywords: hybrid electric vehicle; energy management strategy; driving style; reinforcement learning; equivalent consumption minimization strategy

随着国家“双碳”目标的实施,节能减排迫在眉睫,汽车产业电动化转型是实现“双碳”目标的重要途径^[1]。纯电动汽车的电池技术尚未解决里程焦虑的问题,混合动力汽车多动力源协同工作兼顾了动力性和燃油经济性,正在成为各大车企的研究热点,其中能量管理策略是保证混合动力汽车高效运行的关键^[2]。

混合动力汽车能量管理策略主要可分为基于规则和基于优化算法的策略两类^[3]。基于规则的策略根据专家经验预先定义逻辑门限值,算法简单,实时性强^[4]。在基于优化算法的能量管理策略中,基于动态规划的全局优化策略能够实现全局最优,但计算量大,并且需预知工况信息,难以满足车辆的在线应用需求^[5]。等效燃油消耗最小策略(equivalent consumption minimization strategy, ECMS)通过等效因子将电池电量消耗等价为燃油消耗,通过实时寻优确定使总等效消耗最小的控制策略^[6]。不同风格驾驶行为造成的能耗差异较大,为了提高 ECMS 对不同驾驶风格的适应性,国内外学者提出了考虑驾驶风格的等效因子优化机制。Chen 等基于实车传感器采集的汽车纵向和横向行驶数据,结合调查问卷,采用监督式机器学习的方法研究了不同特征参数对驾驶行为建模准确性的影响^[7]。Zhang 等基于一种新型电液混合动力汽车,将驾驶员与驾驶风格解耦,建立了一种用于行驶工况预处理的速度子集提取算法,根据采集的实际车辆速度,提出驾驶风格的特征参数和评估策略,基于驾驶风格识别设计了模糊逻辑能量管理策略^[8]。Tian 等基于即时驾驶条件设计了一种用于驾驶风格识别的无监督和监督相结合的算法,基于多维高斯分布对影响驾驶员驾驶风格的因素进行分析,将驾驶风格识别引入模型预测控制中,以制定面向不同驾驶员的预测能量管理策

略^[9]。Guo 等将行驶工况与驾驶风格解耦,在对驾驶工况进行分类识别的基础上,分析不同驾驶工况下油门踏板开度及其变化率,建立模糊逻辑识别器来识别驾驶风格,采用混合粒子群优化-遗传算法优化了考虑不同驾驶风格的 ECMS 等效因子,并在新欧洲驾驶循环测试工况下通过实车测试验证了策略的有效性^[10]。

随着大数据、人工智能技术的快速发展,基于数据驱动的学习方法受到越来越多的重视,近年来被广泛应用于混合动力汽车(hybrid electric vehicle, HEV)能量管理策略的研究。深度强化学习算法具有强大的自学习能力,通过训练智能体与环境的交互,可以获取丰富的经验,并逐步优化决策策略^[11]。Lee 等基于 DQN 算法调节 ECMS 等效因子,有效提升了 ECMS 策略的泛化能力^[12]。Xu 等提出了一种基于启发式 Q-learning 算法的能量管理策略,通过引入专家经验,显著提升了算法训练的收敛速度,并取得优于基于规则策略的控制效果^[13]。Hu 等研究了网联环境下基于深度强化学习的能量管理策略,引入交通信息和驾驶员经验,显著提升了燃油经济性^[14]。Wang 等对比研究了 13 种主流的基于强化学习算法的能量管理策略,讨论了不同深度强化学习算法的奖励性能、计算成本和学习收敛性,结果表明,连续动作空间学习过程中的智能体比离散动作空间中的智能体更稳定^[15]。

总体而言,国内外学者为了提高 ECMS 对驾驶员驾驶行为的适应性,通常建立不同驾驶风格与等效因子间的联系,这种对应关系将等效因子描述为与驾驶风格相关的若干离散值,难以实现等效因子的连续调节^[16]。在基于强化学习的能量管理策略方面,大多研究在进行强化学习智能体训练时仅考虑需求转矩转速以及电池荷电状态(state of

charge,SOC)的影响,较少考虑驾驶员驾驶风格对能量管理策略的作用,导致强化学习能量管理策略难以实现不同驾驶员驾驶行为下的最优。如何在深度强化学习智能体的设计与训练过程中考虑驾驶员行为的影响,对于提高混合动力汽车在不同驾驶员作用下的整车燃油经济性具有重要意义。

针对上述问题,本文结合深度确定性策略梯度(deep deterministic policy gradient,DDPG)强化学习算法与 ECMS,提出考虑驾驶风格的 DDPG-ECMS 策略。基于实车试验数据获取不同驾驶员驾驶行为数据,建立驾驶风格识别模型,将驾驶风格系数作为状态变量引入 DDPG 智能体,以获取不同工况和驾驶风格下的 ECMS 等效因子,并由 ECMS 实现动力总成各动力源输出转矩的优化决策,提高能量管理策略在不同驾驶员作用下的适应性,为混合动力汽车能量管理策略优化提供新的方案。

1 混合动力系统建模

1.1 混合动力构型

本文研究的混合动力系统构型如图 1 所示,主要由发动机、电动机、离合器和电控机械自动变速器(automated mechanical transmission,AMT)组成,采用 P2.5 构型。该变速器包括 3 个同步器和用于各种传动比的多个齿轮。系统驱动过程运行模式包括发动机驱动、纯电驱动、混合动力驱动。该 AMT 在换挡过程中利用 3 个同步器在机械动力路径和电力路径之间无缝转换,在方便换挡的同时确保至少一个动力路径的连续运行。它能够在挡位和扭矩变化过程中实现扭矩耦合,在机械挡位和电动挡位之间交替换挡,其中包括 4 个纯电动挡位、6 个发动机挡位和 9 个混合动力挡位。该方法有效解决了换挡时动力中断的问题。系统关键部件参数如表 1 所示。

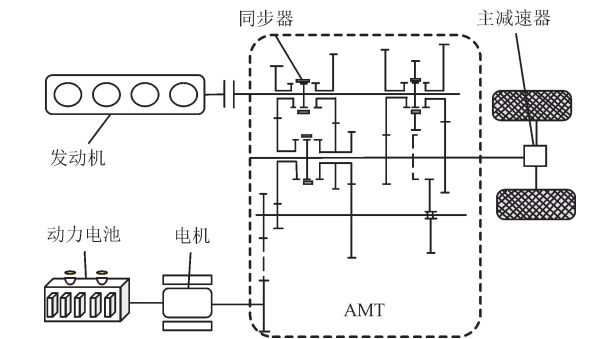


图 1 混合动力系统构型

Fig. 1 Configuration of the hybrid powertrain

表 1 混合动力汽车参数
Table 1 Specifications of the HEV

部件	参数	数值
整车	整备质量 M/kg	1 881
	车轮半径 R_w/m	0.336
	迎风面积 A_f/m^2	2.7
	空气阻力系数 C_d	0.36
发动机	最大转速 $\omega_{e_max}/(\text{r} \cdot \text{min}^{-1})$	6 000
	最大转矩 $T_{e_max}/(\text{N} \cdot \text{m})$	250
电机	最大转速 $\omega_{G_max}/(\text{r} \cdot \text{min}^{-1})$	12 000
	最大转矩 $T_{m_max}/(\text{N} \cdot \text{m})$	180
动力	额定电压 V_{rat}/V	326
电池	额定容量 $Q_{\text{brat}}/(\text{A} \cdot \text{h})$	10.2

1.2 混合动力系统建模

本文研究的配置为双轴并联混合动力汽车,其输入轴扭矩损失与输出轴累积的等效转动惯量和传动效率有关。变速箱传动关系如下

$$\begin{cases} T_{\text{loss}} = T_e i_e (1 - K_e) + T_m i_m (1 - K_m) + T_{\text{inertia}} \\ \omega_e = \omega_{gb} i_e \\ \omega_m = \omega_{gb} i_m \end{cases} \quad (1)$$

T_e 、 T_m 、 ω_e 、 ω_m 须满足如下关系

$$\begin{cases} T_{\text{req}} = T_e(k) i_e + T_m(k) i_m \\ T_{e_min} \leq T_e(k) \leq T_{e_max} \\ \omega_{e_min} \leq \omega_e(k) \leq \omega_{e_max} \\ T_{m_min} \leq T_m(k) \leq T_{m_max} \\ \omega_{m_min} \leq \omega_m(k) \leq \omega_{m_max} \end{cases} \quad (2)$$

式中: T_{req} 为需求转矩; T_{loss} 为传动扭矩损失; T_e 为发动机输出扭矩; K_e 为发动机挡位的传动效率; T_m 为电机输出扭矩; K_m 为电机挡位的传动效率; T_{inertia} 为传动系统中齿轮装置产生的惯性扭矩; ω_e 为发动机转速; ω_m 为电机转速; ω_{gb} 为变速器输出轴转速; i_e 为发动机速比; i_m 为电机速比。

车辆纵向动力学描述为

$$\begin{cases} F_{\text{req}} = F_a + F_g + F_r + F_w \\ F_a = \delta Ma \\ F_w = 0.5 C_d A \rho v^2 \\ F_r = Mg f_r \cos \theta \\ F_g = Mg \sin \theta \end{cases} \quad (3)$$

式中: F_{req} 为车辆所需驱动力; F_a 为加速阻力; F_w 为空气阻力; F_r 为滚动阻力; F_g 为道路坡度阻力; δ 为

旋转质量换算系数,本文取 1.05; a 为车辆纵向加速度; ρ 为空气密度; A 为车辆迎风面积; v 为车速; g 为重力加速度,本文取 9.81 m/s^2 ; θ 为道路坡度; f_r 为滚动阻力系数。

在 HEV 能量管理策略研究中,发动机和电机模型常表示为基于效率图的准静态模型,发动机燃油效率图、电机效率图及外特性数据可由台架试验结果插值拟合而得。

电池模型可以分为内阻模型、RC 模型、Thevenin 模型等。本文将电池模型简化,忽略电池温度对内阻的影响,建立等效内阻模型,且电池 SOC 估算采用安时积分法

$$\xi_{\text{SOC}} = \xi_{\text{SOC,ini}} - \frac{Q_{\text{bat}} - \int_0^T I/3600 dt}{Q_{\text{bat}}} \quad (4)$$

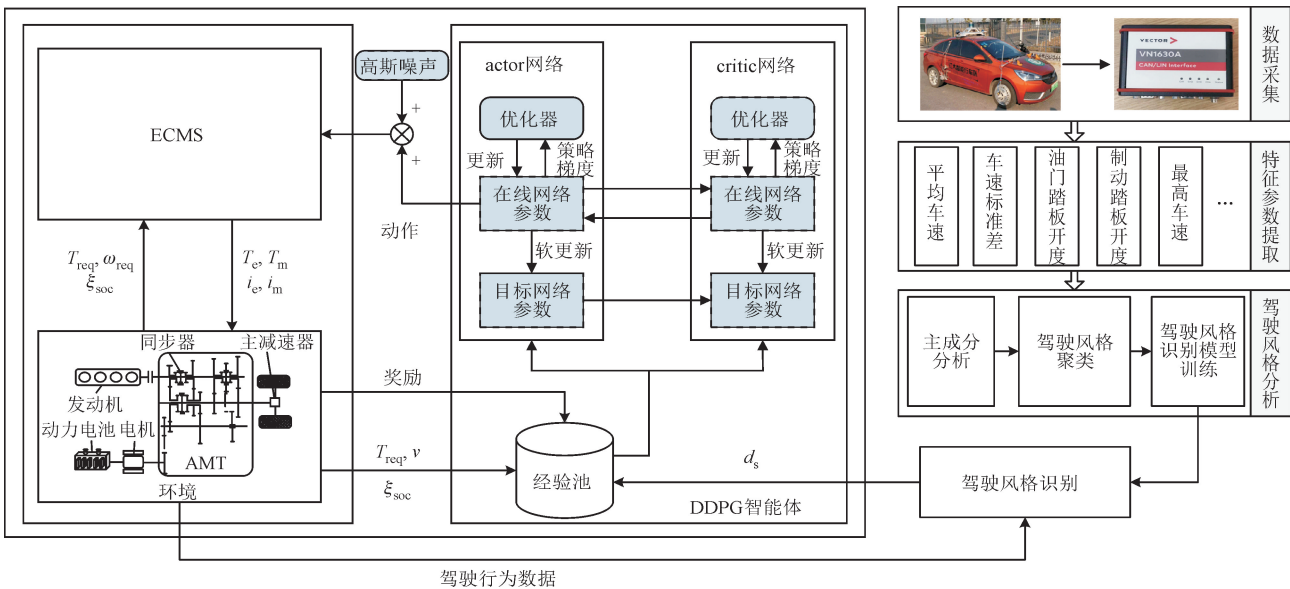
$$I = \frac{V_{\text{oc}} - \sqrt{V_{\text{oc}}^2 - 4R_{\text{int}}P_{\text{bat}}}}{2R_{\text{int}}} \quad (5)$$

式中: I 为电流; V_{oc} 为电池开路电压; R_{int} 为电池内阻;

P_{bat} 为电池功率; ξ_{SOC} 为电池荷电状态; $\xi_{\text{SOC,ini}}$ 为电池电量初始值; Q_{bat} 为电池容量。

2 DDPG-ECMS 自适应策略

本文提出的考虑驾驶风格的 DDPG-ECMS 能量管理策略架构如图 2 所示。为了建立驾驶员驾驶风格识别模型,基于实车试验采集不同驾驶员驾驶行为数据,提取不同特征参数,在此基础上进行驾驶风格的分类与驾驶风格识别模型的离线训练。采用 DDPG 智能体实现基于强化学习的能量管理策略优化,将驾驶风格系数作为智能体状态变量,结合工况需求转矩、转速以及电池 SOC 信息,离线训练得到 DDPG 智能体,训练好的 DDPG 智能体根据驾驶风格识别模块得到的驾驶风格系数以及电池 SOC 等状态信息在线获取 ECMS 最优等效因子,ECMS 模块基于不同驾驶风格和工况下的等效因子实现发动机和电机转矩、挡位的实时在线求解。



T_{req} —需求转矩; ω_{req} —需求转速; d_s —驾驶风格。

图 2 能量管理策略架构

Fig. 2 Schematic diagram of the energy management strategy

2.1 驾驶风格聚类与识别

驾驶风格按照驾驶员的驾驶激进程度一般分为保守型、正常型和激进型共 3 类^[17],即 $d_s = 1, 2, 3$ 。相较于保守型驾驶员,激进型驾驶员平均车速更大,更倾向于猛踩油门和刹车,频繁地急加减速,导致车辆的能耗更高^[18]。本文基于实车试验采集了 40 名驾驶员在不同道路工况下的行驶数据,驾驶员的选取考虑了性别、驾龄以及职业等因素。行驶工况包含

城市道路、高架道路和郊区道路,累计行驶里程达 1 400 km。所采集的数据来源于整车控制器局域网(controller area network, CAN)信号,采样频率为 100 Hz,采集信号包括车速、加速度、油门踏板开度、制动踏板开度。本文将油门踏板与制动踏板合二为一,统一为踏板开度,选取平均车速、车速标准差、最高车速、加速度绝对值均值、加速度绝对值标准差、加速度绝对值最大值、踏板开度变化率绝对值均值、

踏板开度变化率绝对值标准差、踏板开度变化率绝对值最大值为特征参数^[19]。

为了获取更多的样本数量,达到更好的聚类效果,需要对数据进行短行程的划分切片处理,以每次车辆启动到车辆静止为一个短行程,行程片段划分示意如图 3 所示。

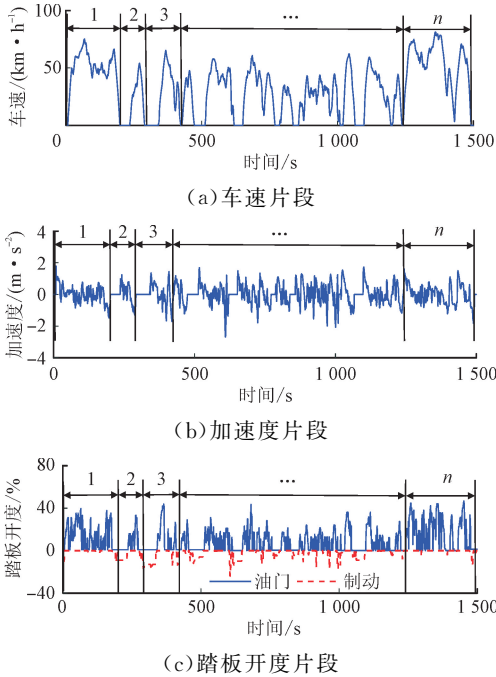


图 3 行程片段划分示意图

Fig. 3 Schematic diagram of trip segment division

驾驶风格聚类识别的具体步骤如下。

(1)选取表征驾驶风格的特征参数,基于主成分分析法对特征参数进行降维处理。根据特征值大于 1 以及累计贡献率大于 85%的准则,将所选 9 个特

征参数降维到 3 个成分,主成分贡献率如图 4 所示。

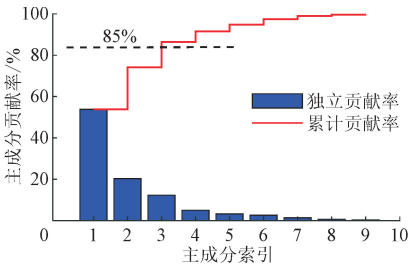


图 4 主成分贡献

Fig. 4 Principal component contribution

(2)计算主成分得分。利用 K-means 聚类方法^[20],对划分的 896 个样本进行聚类分析,最终聚类结果如图 5 所示。可以看出,第 1、2、3 类样本分别占比 39.0%、31.8%、29.2%。各类样本下的特征参数平均值如表 2 所示。其中:第 1 类样本数据的均值和反映数据波动的标准差都较小,定义为保守型;第 3 类样本数据的均值和标准差都较大,定义为激进型;第 2 类样本数据的均值和标准差介于两者之间,定义为正常型。

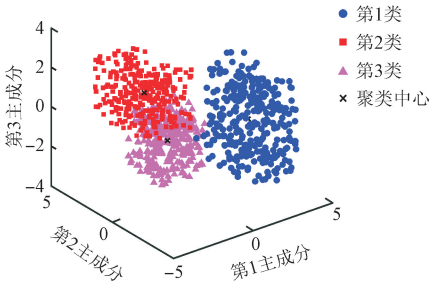


图 5 K-means 聚类结果

Fig. 5 K-means clustering results

表 2 各类驾驶风格样本的特征参数平均值

Table 2 Average value of characteristic parameters of samples of various driving styles

特征参数	平均值		
	第 1 类	第 2 类	第 3 类
平均车速/(km · h ⁻¹)	31.53	37.81	44.26
车速标准差/(km · h ⁻¹)	15.32	16.27	18.95
车速最大值/(km · h ⁻¹)	59.49	68.70	75.65
加速度绝对值均值/(m · s ⁻²)	0.371	0.569	0.712
加速度绝对值标准差/(m · s ⁻²)	0.334	0.482	0.547
加速度绝对值最大值/(m · s ⁻²)	2.471	2.890	3.356
踏板开度变化率绝对值均值/(% · s ⁻¹)	6.84	9.51	13.23
踏板开度变化率绝对值标准差/(% · s ⁻¹)	6.75	10.87	13.48
踏板开度变化率绝对值最大值/(% · s ⁻¹)	31.69	35.93	41.27

(3) 确定驾驶风格识别周期, 本文驾驶风格识别周期设定为 60 s, 每 60 s 计算特征参数值^[21]。

(4) 采用支持向量机进行驾驶风格的识别。首先载入训练集、测试集和相应类别标签, 然后对数据进行归一化处理, 选择径向基函数为核函数, 分别利用 svmtrain 函数和 svmpredict 函数训练模型。支持向量机识别结果如图 6 所示。可以看出, 测试集识别准确率达到 91.04%。

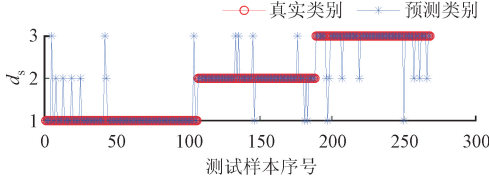


图 6 驾驶风格识别结果

Fig. 6 Recognition results of the driving style

2.2 等效燃油消耗最小策略

等效燃油消耗最小策略使用等效因子将电池的能耗等效为燃油消耗, 通过实时求解最小化发动机油耗和等效燃油消耗之和, 实现最优控制目标^[22]。总燃油消耗为

$$m_{\text{tot}} = m_f + m_{\text{ele}} \quad (6)$$

$$m_{\text{ele}} = S_{\text{eq}} \lambda \frac{P_L}{\eta_{\text{bat}} Q_{\text{lhv}}} + S_{\text{eq}} (1 - \lambda) \frac{\eta_{\text{bat}} P_L}{Q_{\text{lhv}}} \quad (7)$$

$$\lambda = \frac{1 + \text{sgn}(P_L)}{2} \quad (8)$$

式中: S_{eq} 表示等效因子; λ 用来判断充放电状态; Q_{lhv} 为发动机燃油热值; m_{tot} 为总的燃油消耗, 包括发动机油耗 m_f 和等效油耗 m_{ele} ; P_L 为负载功率; η_{bat} 为电池工作效率, 表达式为

$$\eta_{\text{bat}} = \begin{cases} \frac{V_{\text{oc}} (V_{\text{oc}} - \sqrt{V_{\text{oc}}^2 - 4P_L R_{\text{chg}}})}{2P_L R_{\text{chg}}}, & P_L < 0 \\ \frac{2P_L R_{\text{dis}}}{V_{\text{oc}} (V_{\text{oc}} - \sqrt{V_{\text{oc}}^2 - 4P_L R_{\text{dis}}})}, & P_L \geq 0 \end{cases} \quad (9)$$

其中, R_{chg} 、 R_{dis} 分别为电池充、放电状态下的内阻。

受到电机、电池以及发动机物理特性的约束, 系统优化求解时须满足以下条件

$$\begin{cases} \xi_{\text{SOC}_{\min}} \leq \xi_{\text{SOC}} \leq \xi_{\text{SOC}_{\max}} \\ \omega_{\text{e}_{\min}} \leq \omega_e \leq \omega_{\text{e}_{\max}} \\ T_{\text{e}_{\min}} \leq T_e \leq T_{\text{e}_{\max}} \\ \omega_{\text{m}_{\min}} \leq \omega_m \leq \omega_{\text{m}_{\max}} \\ T_{\text{m}_{\min}} \leq T_m \leq T_{\text{m}_{\max}} \end{cases} \quad (10)$$

选择 T_e 、 i_e 和 i_m 为控制变量, 发动机转矩按照

最大值进行等距离散化。为防止发动机和电机频繁换挡以及发动机的频繁启停, 需对上述 ECMS 策略加入惩罚

$$m_{\text{ie}} = \begin{cases} \alpha_{\text{ie}}, & i_e(t) \neq i_e(t-1) \\ 0, & i_e(t) = i_e(t-1) \end{cases} \quad (11)$$

$$m_{\text{im}} = \begin{cases} \alpha_{\text{im}}, & i_m(t) \neq i_m(t-1) \\ 0, & i_m(t) = i_m(t-1) \end{cases} \quad (12)$$

$$m_s = \alpha_s |T_e(t) - T_e(t-1)| \quad (13)$$

式中: m_{ie} 和 m_{im} 为挡位变化的惩罚函数; α_{ie} 和 α_{im} 为相应的惩罚系数; m_s 为发动机频繁启停的惩罚项; α_s 为发动机转矩波动的惩罚系数。

因此, ECMS 目标函数为

$$J^*(t) = m_{\text{tot}} + m_{\text{ie}} + m_{\text{im}} + m_s \quad (14)$$

2.3 考虑驾驶风格的 DDPG-ECMS 策略

ECMS 等效因子关系到能量管理策略的有效性, 而基于 SOC 实际值与参考值的偏差作为反馈补偿进行等效因子调节存在一定的不足之处, 对不同驾驶员驾驶风格的适应性较差^[23]。等效因子受到诸多因素的影响, 难以建立驾驶风格与等效因子之间精确的数学表达式。深度强化学习算法在处理复杂、高维度的状态空间和动作空间时表现出色, 智能体通过与环境的实时互动, 根据奖励信号调整动作输出, 学习最优决策, 这种基于数据驱动的学习方式适用于建立驾驶风格及复杂行驶工况与等效因子的非线性映射关系。

DDPG 是一种用于解决连续动作空间上的强化学习问题的算法。该算法包含两个神经网络: critic 网络用于估计动作价值函数, actor 网络用于生成动作。actor 网络根据当前状态输出动作, critic 网络使用该动作和当前状态作为输入来估计当前状态的价值。DDPG 算法使用 actor 和 critic 网络共同协作, 通过最小化动作价值函数的误差来同时更新 actor 和 critic 网络的参数。DDPG 算法训练过程包括采样阶段和学习阶段。在采样阶段, 智能体与环境交互, 存储新获得的经验数据。在学习阶段, DDPG 算法从经验池中采样数据进行学习, 优化 actor 和 critic 网络的参数。

本文设定 DDPG 智能体的状态空间为

$$S_0 = \{T_{\text{req}}, v, \xi_{\text{SOC}}, d_s\} \quad (15)$$

DDPG 智能体动作输出为

$$A_0 = s_{\text{eq}} \quad (16)$$

奖励函数设定如下

$$R_0 = - \int_{T_{t-1}}^{T_t} [\alpha \xi_{\text{SOC}_d}^2 + \beta m_f(t) + \gamma (s_{\text{eq}}(t) - \bar{s}_{\text{eq}})^2] dt \quad (17)$$

式中: ξ_{SOC_d} 为实际 SOC 与参考值的偏差; $\bar{s}_{\text{eq}}$ 是等效因子范围内的中值; α 、 β 和 γ 分别为各部分的权重系数。DDPG 算法的优化目标是使 SOC 维持在参考范围内,并使得燃油消耗最小。因此,奖励函数的设定同时考虑了 SOC 偏差、油耗和等效因子的波动范围。

3 仿真结果分析

3.1 对比策略

将本文提出的策略与基于规则的策略(rule)、等效因子比例修正的 AECMS(P-AECMS)策略以及未考虑驾驶风格的强化学习策略进行对比分析。基于规则的策略主要包括模式切换、挡位决策以及转矩分配共 3 个部分。首先,根据当前车速、需求功率以及 SOC,决策出运行模式;然后,根据车速、踏板开度查表决策挡位;最后,根据当前的模式和挡位,进行发动机和电机的转矩分配。

基于比例修正的 P-AECMS 策略的等效因子自适应调节方式^[24]如下

$$s_{\text{eq}}(k+1) = k_p(\xi_{\text{SOC}_{\text{ref}}} - \xi_{\text{SOC}(t)}) + s_{\text{eq}}(k) = k_p\xi_{\text{SOC}(t)} + s_{\text{eq}}(k)$$

(18)

式中: $s_{\text{eq}}(k+1)$ 和 $s_{\text{eq}}(k)$ 分别表示下一时刻和当前时刻的等效因子; $\xi_{\text{SOC}_{\text{ref}}}$ 为电池 SOC 参考值,本文设定为 0.6; $\xi_{\text{SOC}(t)}$ 为当前 SOC 实际值; k_p 为 SOC 参考值与实际值偏差的比例增益。

软演员-评论家算法(SAC)为深度强化学习(DRL)算法,同样用于连续动作空间,被证明可以实现较好的燃油经济性以及电池充放电状态的平衡^[25]。因此,本文进一步将所提出的策略与基于 SAC 算法的策略(DRL-SAC)进行对比分析。与 DDPG-ECMS 策略不同的是,DRL-SAC 算法直接以发动机转矩为动作输出,且状态输入不包含驾驶风格,奖励函数的设定不再考虑等效因子的波动。SAC 智能体的状态空间为

$$S_1 = \{T_{\text{req}}, v, \xi_{\text{SOC}}\}$$

(19)

SAC 智能体动作输出为

$$A_1 = T_{\text{eng}}$$

(20)

奖励函数设定如下

$$R_1 = - \int_{T_{t-1}}^{T_t} [\alpha \xi_{\text{SOC}_d}^2 + \beta m_f(t)] dt$$

(21)

3.2 DDPG 训练结果

以涵盖多种驾驶风格的组合工况作为 DDPG 智能体的训练数据,训练工况以及对应的驾驶风格如图 7 和图 8 所示。训练工况主要为基于实车试验

采集的不同驾驶员的驾驶数据样本组合。

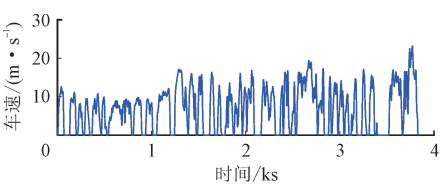


图 7 训练组合工况

Fig. 7 Combination cycle for the training

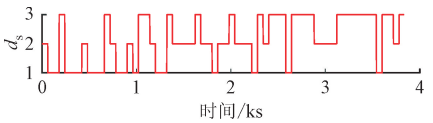


图 8 训练工况下的驾驶风格

Fig. 8 Driving style under training condition

DDPG 智能体超参数设定如表 3 所示。DDPG 算法的 actor 网络和 critic 网络均包含输入层、3 个隐藏层和 1 个输出层,其中每一隐藏层的神经元数量为 100,神经网络的各层之间都采用全连接结构。

表 3 DDPG 算法超参数设定

Table 3 Specifications of the DDPG algorithm

参数	数值	参数	数值或类型
经验池大小	10^6	探索率最小值	0.01
折扣率	0.9	探索率衰减	0.000 2
每次更新提取样本量	256	激活函数	ReLU
目标网络更新率	0.001	学习率	0.000 5

DDPG 智能体最终训练的奖励函数曲线如图 9 所示。图中,平均回合奖励为过去 5 个回合的奖励均值,期望奖励表示 critic 网络在每轮训练开始时对给定环境的初始观测长期折扣奖励的估计值。可以看出,DDPG 算法在 460 回合左右开始收敛,最终奖励收敛到 2 400 附近,训练效果较好。

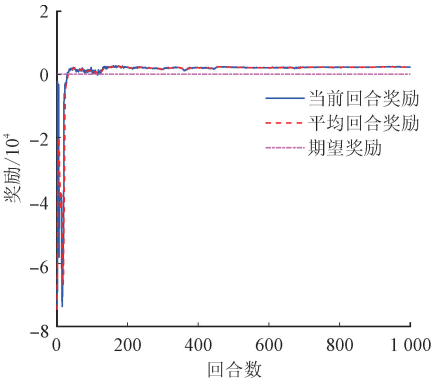


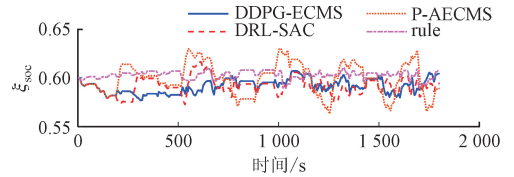
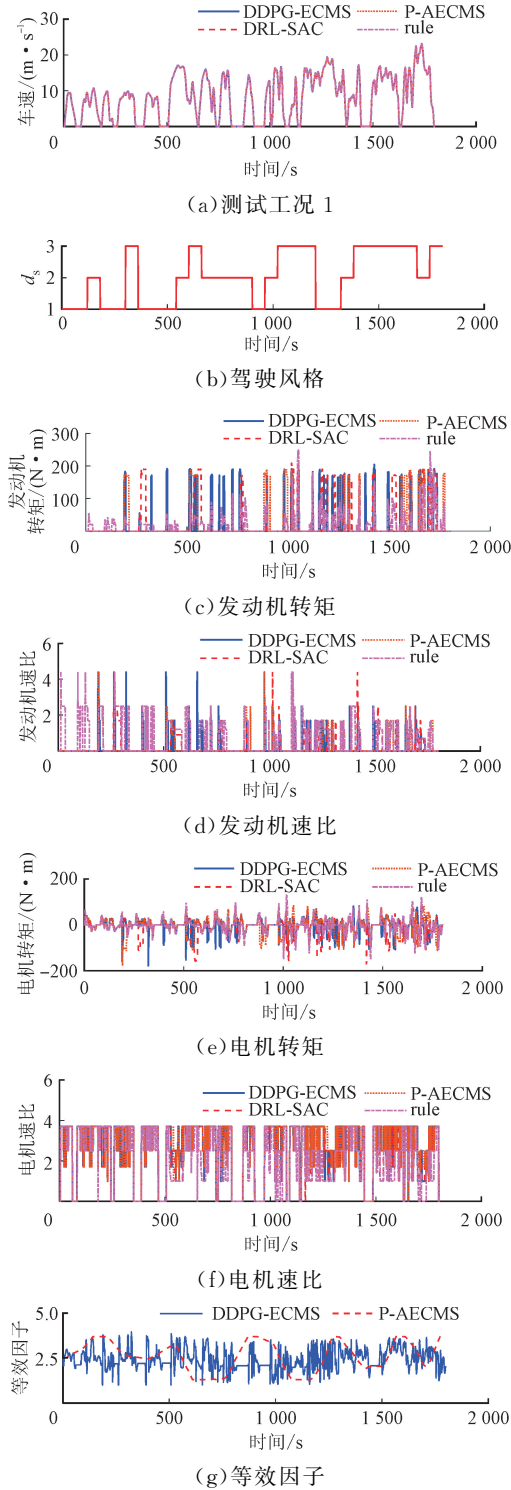
图 9 奖励曲线

Fig. 9 Reward curve

3.3 仿真结果对比分析

为了验证 DDPG-ECMS 算法的泛化能力以及优化效果,基于图 7 所示组合工况的训练结果,以包含 3 种驾驶风格的 1 800 s 的组合工况作为 DDPG-ECMS 算法的测试工况进行仿真验证。

测试工况下,4 种策略的仿真结果对比如图 10 和 11 所示。



(h) 电池 SOC

图 10 不同策略下的仿真结果对比

Fig. 10 Comparison of the simulation results with different strategies

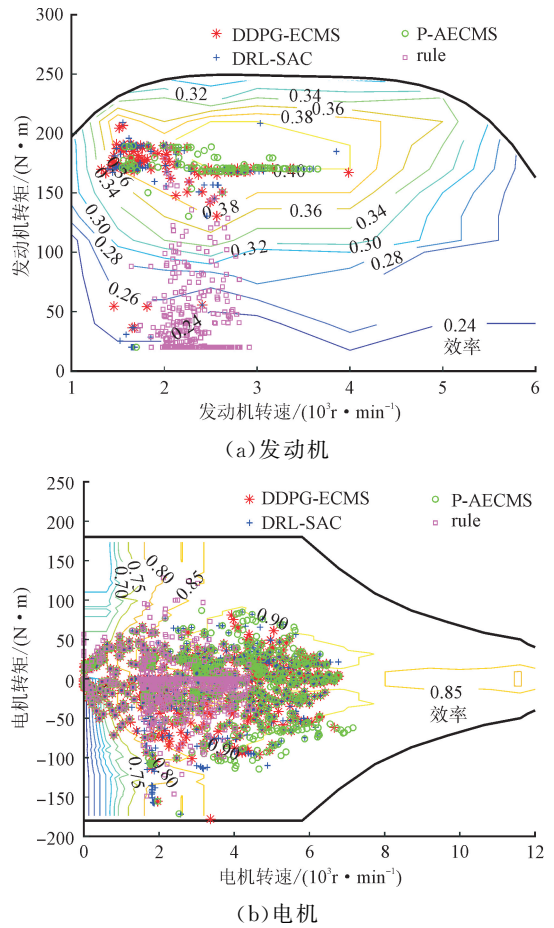


图 11 发动机和电机工作点

Fig. 11 Operation points of the engine and motor

由图 10(a)可以看出,不同策略均实现了整车对目标工况的准确跟踪,相对应的驾驶风格如图 10(b)所示。由图 10(c)和 10(e)可以看出,基于规则的 HEV 能量管理策略,更多地以发动机来驱动车辆,电机转矩作为补充,发动机转矩输出更加频繁。相较于 P-AECMS,考虑驾驶风格的 DDPG-ECMS 使发动机更多的参与工作,但工作点主要集中在高效区。由图 10(d)和 10(f)可以看出,本文所提出的策略可以有效抑制不同驾驶风格下的频繁换挡。由图 10(h)可以看出,虽然基于规则的策略将 SOC 维

持得较好,具有最小的波动,但发动机大多工作在非高效区,没有充分利用动力电池对发动机工作点的调节功能,燃油经济性最差。考虑驾驶风格的DDPG-ECMS策略可以根据车辆状态信息和驾驶风格实时调节等效因子,电机输出更多转矩驱动车辆行驶。虽然SOC波动相较于规则策略的较大,但最终SOC控制较好,实现了电池充放电平衡,并且有效降低了发动机燃油消耗。由图10(g)可以看出,由于P-AECMS策略仅根据电池SOC真实值与参考值的插值比例修正等效因子,等效因子调节对电池SOC具有滞后性,因此SOC波动较大,等效因子受到SOC较慢动力学特征的影响,其变化程度也较缓慢。由图10(g)可以看出,相较于P-AECMS策略,在本文提出的策略中,训练好的DDPG智能体根据多个状态量信息实时获取等效因子,使等效因子剧烈变化,以适应各时刻不同工况和驾驶风格的控制需求。DRL-SAC算法结果与P-AECMS接近,由于训练状态并未考虑驾驶风格,因此对不同驾驶风格的测试工况的泛化性较差。

由于不同策略导致SOC最终值不同,将最终电池充电或放电电量等效为相应的发动机油耗,结合实际发动机油耗进行等效燃油经济性的评估^[26]。不同策略下仿真结果如表4所示。可以看出,所提出的DDPG-ECMS策略使整车能量消耗相较于规则策略降低了17.05%,相较于P-AECMS降低了12.17%,相较于DRL-SAC降低了8.07%。并且,本文所提出的控制策略也有效地控制了电池最终SOC,验证了控制策略的有效性。

表 4 仿真结果对比

Table 4 Comparison of simulation results				
SOC 初值	策略	SOC 终值	100 km 发动 机油耗/L	100 km 等效 油耗/L
0.6	rule	0.607	5.29	5.22
	P-AECMS	0.590	4.84	4.93
	DRL-SAC	0.595	4.66	4.71
	DDPG-ECMS	0.604	4.36	4.33

为了进一步验证所提出策略的泛化性,本文增加了两组测试工况,测试工况同样来自实车采集数据的样本组合。测试工况如图12所示。

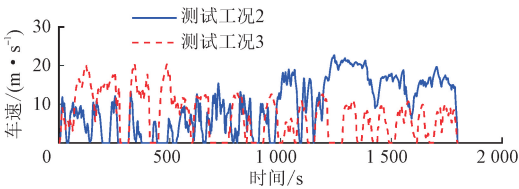


图 12 补充测试工况

Fig. 12 Additional test cycles

补充测试工况下的结果对比如表5所示。可以看出:在测试工况2下,DDPG-ECMS策略相对于规则策略、P-AECMS策略以及DRL-SAC策略,能量消耗分别降低了20.01%、7.39%和4.45%;在测试工况3下,DDPG-ECMS策略相对于规则策略、P-AECMS策略以及DRL-SAC策略,能量消耗分别降低了19.16%、6.54%和-0.23%。虽然在测试工况3下,DRL-SAC等效油耗略低,但SOC终值下降到了0.59,不能很好地维持充放电平衡。DRL-SAC直接以发动机转矩作为输出,动作空间更大,其泛化性比调节等效因子的DDPG-ECMS略差。

表 5 补充测试工况结果对比

Table 5 Comparative results under additional test cycles

SOC 初值	测试工况	策略	SOC 终值	100 km 发动 机油耗/L	100 km 等效 油耗/L	能量消耗 降低量/%
0.6	2	rule	0.605	5.69	5.64	20.01
		P-AECMS	0.602	4.89	4.87	7.39
		DRL-SAC	0.605	4.77	4.72	4.45
		DDPG-ECMS	0.596	4.48	4.51	
	3	rule	0.606	5.54	5.48	19.16
		P-AECMS	0.604	4.78	4.74	6.54
		DRL-SAC	0.590	4.33	4.42	-0.23
		DDPG-ECMS	0.594	4.37	4.43	

4 硬件在环试验

为了进一步验证所提出的 DDPG-ECMS 策略的实时性与有效性,本文基于 MATLAB 的 Simulink Real-Time 实时仿真系统搭建了硬件在环(hardware in the loop, HiL)试验平台,硬件在环试验原理及试验平台如图 13 所示。该平台主要由 PC 主机、实时仿真目标机以及嵌入式工控机共 3 部分组成。其中:PC 主机用于模型的编译与下载,以及输入输出信号的数据观测;实时仿真目标机用于被控对象模型的实时仿真;嵌入

式工控机用于运行能量管理控制策略。通过编写 DBC 文件对控制器与被控对象的输入输出 CAN 信号进行定义,将被控对象模型编译并下载到实时仿真目标机中运行。DDPG-ECMS 策略编译下载到控制器中运行,控制器模型中 DDPG 智能体在 Simulink 中不能直接编译,需根据 generatePolicyFunction(agent)函数进行策略部署,转化为 MATLAB Function 的形式。控制器与实时仿真目标机通过 CAN 通信相连,上位机与实时仿真目标机通过 TCP/IP 通信相连,实时记录控制器的输入输出数据。

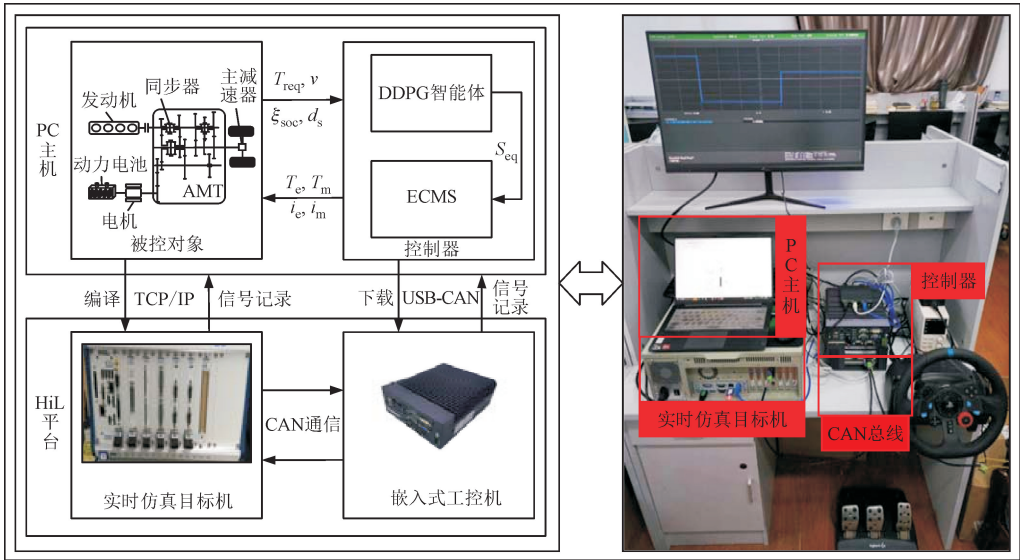
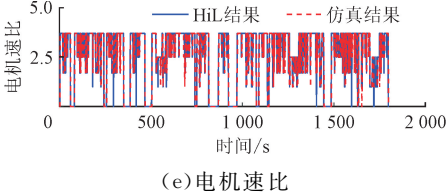
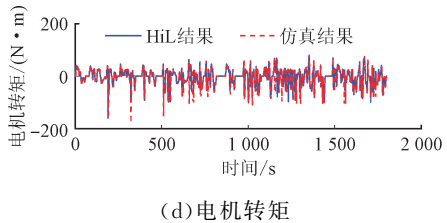
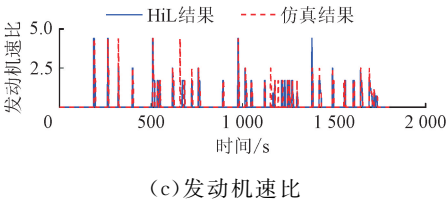
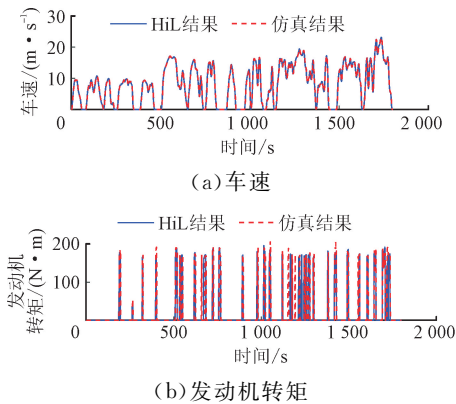


图 13 硬件在环试验平台
Fig. 13 Hardware in the loop platform

硬件在环试验结果如图 14 所示。可以看出,硬件在环试验结果与仿真结果基本一致,车速跟踪效果较好,实时仿真系统下,控制器与实时仿真机通信存在延迟,故输出转矩和挡位有一定的波动,但总体在可接收范围以内。试验表明,在 0.01 s 的仿真步长下,DDPG-ECMS 算法具有良好的实时性和有效性。



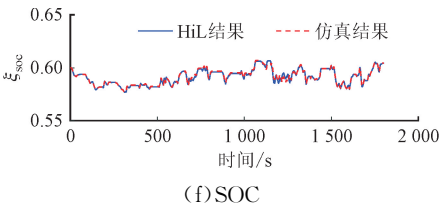


图 14 硬件在环试验结果

Fig. 14 Hardware in the loop test results

不同控制策略下,硬件在环测试的最终 SOC、发动机油耗及整车等效油耗如表 6 所示。可以看出,相较于规则策略、P-AECMS 以及 DRL-SAC 策略,所提出的 DDPG-ECMS 策略使整车能量消耗分别下降 16.35%、11.11% 和 7.56%,并且也能实现较好的电池 SOC 控制。

表 6 不同策略下的硬件在环对比结果

Table 6 HiL comparison results under different strategies

SOC 初值	策略	SOC 终值	100 km 发动机油耗/L	100 km 等效油耗/L
0.6	rule	0.603	5.29	5.26
	P-AECMS	0.587	4.83	4.95
	DRL-SAC	0.592	4.68	4.76
	DDPG-ECMS	0.606	4.46	4.40

5 结 论

(1)基于实车试验采集了 40 名驾驶员在不同行驶路况下的行驶数据,选取了平均车速、车速标准差等 9 个表征驾驶风格的特征参数,基于主成分分析法、K-means 聚类以及支持向量机识别完成了驾驶风格的聚类与识别,驾驶员驾驶风格识别精度达到 91%。

(2)针对一款具有多个发动机和电机挡位的并联式混合动力汽车,建立了混合动力系统模型。基于强化学习 DDPG 算法和 ECMS 策略构建了考虑驾驶风格的 HEV 能量管理策略。其中,DDPG 智能体根据驾驶风格和车辆状态信息求解 ECMS 等效因子,ECMS 用来获取发动机、电机转矩分配以及变速箱档位。基于涵盖多种驾驶风格的组合工况,完成了对 DDPG-ECMS 算法智能体的训练。

(3)基于实际采集的不同驾驶员驾驶数据构建测试工况,通过仿真分析验证了本文提出的控制策略性能。相较于基于规则的能量管理策略、等效因子比例修正的 P-AECMS 策略以及 DRL-SAC 策略,提出的考虑驾驶风格的 DDPG-ECMS 策略在测试工况 1 下整车能量消耗分别降低了 17.05%、12.17% 和 8.07%,在测试工况 2 下分别降低了 20.01%、7.39%

和 4.45%,在测试工况 3 下分别降低了 19.16%、6.54% 和 -0.23%。通过搭建的硬件在环试验平台,进一步验证了所提策略在测试工况 1 下的实时性和有效性,DDPG-ECMS 策略整车能量消耗相较于其他 3 种策略分别降低了 16.35%、11.11% 和 7.56%。

参考文献:

[1] 王跃飞,王志,孙睿,等. 基于驾驶意图多步预测的智能网联 HEV 等效排放最小控制策略 [J]. 机械工程学报, 2023, 59(18): 271-282.
WANG Yuefei, WANG Zhi, SUN Rui, et al. Equivalent emission minimization strategy of intelligent connected HEV based on multi-step prediction of driving intention [J]. Journal of Mechanical Engineering, 2023, 59(18): 271-282.

[2] 黄康,王强,邱明明,等. 考虑模式切换频率的多模式混合动力汽车参数优化 [J]. 西安交通大学学报, 2019, 53(7): 99-107.
HUANG Kang, WANG Qiang, QIU Mingming, et al. Parameter optimization of multi-mode hybrid vehicle considering mode-switching frequency [J]. Journal of Xi'an Jiaotong University, 2019, 53(7): 99-107.

[3] YANG Chao, DU Xuelong, WANG Weida, et al. Variable optimization domain-based cooperative energy management strategy for connected plug-in hybrid electric vehicles [J]. Energy, 2024, 290: 130206.

[4] 黄硕,李亮,杨超,等. 基于规则修正的同轴并联混合动力客车瞬时优化能量分配策略 [J]. 机械工程学报, 2014, 50(20): 113-121.
HUANG Shuo, LI Liang, YANG Chao, et al. Rule correction-based instantaneous optimal energy management strategy for single-shaft parallel hybrid electric bus [J]. Journal of Mechanical Engineering, 2014, 50(20): 113-121.

[5] DU Guodong, ZOU Yuan, ZHANG Xudong, et al. Deep reinforcement learning based energy management for a hybrid electric vehicle [J]. Energy, 2020, 201: 117591.

[6] 王文彬,田韶鹏,郑青星,等. 基于 FA 的等效燃油消耗最小控制策略优化 [J]. 江苏大学学报(自然科学版), 2022, 43(2): 147-153.
WANG Wenbin, TIAN Shaopeng, ZHENG Qingxing, et al. Optimization of equivalent fuel consumption minimization strategy based on firefly algorithm [J]. Journal of Jiangsu University (Natural Science Edition), 2022, 43(2): 147-153.

[7] CHEN Yao, WANG Ke, LU J J. Feature selection for driving style and skill clustering using naturalistic driv-

- ing data and driving behavior questionnaire [J]. *Accident Analysis & Prevention*, 2023, 185: 107022.
- [8] ZHANG Zhen, ZHANG Tiezhu, HONG Jichao, et al. Energy management strategy of a novel electric-hydraulic hybrid vehicle based on driving style recognition [J]. *Sustainable Energy & Fuels*, 2023, 7(2): 420-430.
- [9] TIAN Xiang, CAI Yingfeng, SUN Xiaodong, et al. Incorporating driving style recognition into MPC for energy management of plug-in hybrid electric buses [J]. *IEEE Transactions on Transportation Electrification*, 2023, 9(1): 169-181.
- [10] GUO Qiuyi, ZHAO Zhiguo, SHEN Peihong, et al. Adaptive optimal control based on driving style recognition for plug-in hybrid electric vehicle [J]. *Energy*, 2019, 186: 115824.
- [11] ZHU Zhaoxuan, LIU Yuxing, CANOVA M. Energy management of hybrid electric vehicles via deep Q-networks [C]//2020 American Control Conference (ACC). Piscataway, NJ, USA: IEEE, 2020: 3077-3082.
- [12] LEE W, JEOUNG H, PARK D, et al. A real-time intelligent energy management strategy for hybrid electric vehicles using reinforcement learning [J]. *IEEE Access*, 2021, 9: 72759-72768.
- [13] XU Bin, HOU Jun, SHI Junzhe, et al. Learning time reduction using warm-start methods for a reinforcement learning-based supervisory control in hybrid electric vehicle applications [J]. *IEEE Transactions on Transportation Electrification*, 2021, 7(2): 626-635.
- [14] HU Dong, ZHANG Yuanyuan. Deep reinforcement learning based on driver experience embedding for energy management strategies in hybrid electric vehicles [J]. *Energy Technology*, 2022, 10(6): 2200123.
- [15] WANG Hanchen, YE Yiming, ZHANG Jiangfeng, et al. A comparative study of 13 deep reinforcement learning based energy management methods for a hybrid electric vehicle [J]. *Energy*, 2023, 266: 126497.
- [16] HE Kun, QIN Dongchen, CHEN Jiangyi, et al. Adaptive equivalent consumption minimization strategy for fuel cell buses based on driving style recognition [J]. *Sustainability*, 2023, 15(10): 7781.
- [17] YANG Sen, WANG Wenshuo, ZHANG Fengqi, et al. Driving-style-oriented adaptive equivalent consumption minimization strategies for HEVs [J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(10): 9249-9261.
- [18] GONG Changchao, HU Minghui, LI Shuxian, et al. Equivalent consumption minimization strategy of hybrid electric vehicle considering the impact of driving style [J]. *Proceedings of the Institution of Mechanical Engineers: Part D Journal of Automobile Engineering*, 2019, 233(10): 2610-2623.
- [19] 王旭, 马菲, 廖小棱, 等. 基于多分类监督学习的驾驶风格特征指标筛选 [J]. *交通信息与安全*, 2022, 40(1): 162-168.
- WANG Xu, MA Fei, LIAO Xiaoleng, et al. Feature selection for recognition of driving styles based on multi-classification and supervised learning [J]. *Journal of Transport Information and Safety*, 2022, 40(1): 162-168.
- [20] 李经纬, 赵治国, 沈沛鸿, 等. 驾驶风格 K-means 聚类与识别方法研究 [J]. *汽车技术*, 2018(12): 8-12.
- LI Jingwei, ZHAO Zhiguo, SHEN Peihong, et al. Research on methods of K-means clustering and recognition for driving style [J]. *Automobile Technology*, 2018(12): 8-12.
- [21] LIN Xinyou, LI Kuiliang, WANG Liming. A driving-style-oriented adaptive control strategy based PSO-fuzzy expert algorithm for a plug-in hybrid electric vehicle [J]. *Expert Systems with Applications*, 2022, 201: 117236.
- [22] 施德华, 容香伟, 汪少华, 等. 基于功率比的混合动力汽车模糊自适应等效燃油消耗最小策略研究 [J]. *西安交通大学学报*, 2022, 56(1): 12-21.
- SHI Dehua, RONG Xiangwei, WANG Shaohua, et al. Fuzzy adaptive equivalent consumption minimization strategy for hybrid electric vehicle based on power ratio [J]. *Journal of Xi'an Jiaotong University*, 2022, 56(1): 12-21.
- [23] TIAN Xiang, CAI Yingfeng, SUN Xiaodong, et al. An adaptive ECMS with driving style recognition for energy optimization of parallel hybrid electric buses [J]. *Energy*, 2019, 189: 116151.
- [24] ONORI S, SERRAO L, RIZZONI G. Adaptive equivalent consumption minimization strategy for hybrid electric vehicles [C]//ASME 2010 Dynamic Systems and Control Conference. New York, USA: ASME, 2010: 499-505.
- [25] WANG Zexing, HE Hongwen, PENG Jiankun, et al. A comparative study of deep reinforcement learning based energy management strategy for hybrid electric vehicle [J]. *Energy Conversion and Management*, 2023, 293: 117442.
- [26] KONG Yan, XU Nan, LIU Qiao, et al. A data-driven energy management method for parallel PHEVs based on action dependent heuristic dynamic programming (ADHDP) model [J]. *Energy*, 2023, 265: 126306.

(编辑 陶晴)