

基于强化学习的 P2 结构混动车辆能量优化控制

胡作磊^{1†}, 童紫威^{1,2}, 刘平³, 施伟⁴

(1. 湖南大学 信息科学与工程学院, 湖南 长沙 410082; 2. 湖南锦和园林建设工程有限公司, 湖南 长沙 410148;

3. 中国(湖南)自由贸易试验区长沙片区临空管理委员会, 湖南 长沙 410137; 4. 长沙医学院 信息工程学院, 湖南 长沙 410219)

摘 要:针对 P2 结构混动车辆, 提出了一种基于强化学习的自适应等效燃油消耗最小策略(RL-ECMS), 通过两个智能体实现等效因子的自适应更新和车辆扭矩的动态分配, 以适应不断变化的驾驶需求。通过 MATLAB/Simulink 仿真平台对比了 RL-ECMS 与传统 ECMS 和基于规则的控制策略。结果表明, RL-ECMS 在 FTP75 和 FTP75-Highway 两种典型驾驶工况下均能实现更低的燃油消耗, 且不影响车辆性能。同时测试了未经训练的 ECE 典型工况, 结果表明本文所提算法同样具有良好的泛化性与鲁棒性。

关键词:混合动力车辆; 能量优化; 强化学习; 自适应控制

中图分类号:TP277

文献标识码:A

Reinforcement Learning-Based Energy Optimization Control for Hybrid Vehicles with P2 Structure

HU Zuolei^{1†}, TONG Ziwei^{1,2}, LIU Ping³, SHI Wei⁴

(1. College of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan 410082, China;

2. Hunan Jinhe Landscape Construction Engineering Co., Ltd. Limited, Changsha, Hunan 410148; China;

3. China (Hunan) Pilot Free Trade Zone, Changsha Area, Changsha, Hunan 410137, China;

4. School of Information Engineering, Changsha Medical University, Chang, Hunan 410219, China)

Abstract: A reinforcement learning-based adaptive equivalent fuel consumption minimization strategy (RL-ECMS) is proposed for P2-structured hybrid vehicles, which realizes adaptive updating of the equivalence factor and dynamic allocation of vehicle torque through two intelligences to adapt to the changing driving demands. The RL-ECMS is compared with the conventional ECMS and rule-based control strategies through the MATLAB/Simulink simulation platform. The results show that the RL-ECMS can achieve lower fuel consumption under both FTP75 and FTP75-Highway typical driving conditions without affecting vehicle performance. The untrained ECE typical conditions are also tested, and the result shows that the algorithm proposed in this paper also has good generalization and robustness.

Key words: hybrid vehicles; energy optimization; reinforcement learning; adaptive control

新能源汽车是纯电车辆(Pure Electric Vehicle, PEV)、混合动力汽车、燃料电池汽车(Fuel Cell Electric Vehicle, FCEV)的统称^[1]。纯电汽车和氢燃料电池汽车都可以实现车辆的零排放, 电

机、电控、电池为纯电汽车电驱系统中的关键核心技术。但由于电池受正极材料、环境温度等影响, 存在续航短、不安全等因素, 及汽车充电桩部署尚未普及、充电时间长等问题, 造成了一部分车主不

选择纯电动汽车。氢燃料电池是一种无污染、高效率的理想的车载动力电池,但由于技术门槛高,且存在诸多专利技术壁垒,同时制氢成本高昂、难存储等问题使得氢燃料电池汽车难以普及。能量管理算法是混合动力车辆的关键核心技术之一,直接影响着整车的燃油经济性,也能决定车辆排放是否能够达到国家标准^[1]。因此研究混合动力车辆能量优化控制方法有着很强的经济价值很大的与工程意义。

1 相关工作

目前应用于 HEV 的能量管理算法主要分为基于规则的能量管理策略、基于优化的能量管理策略、基于学习的能量管理策略三类。

基于规则的能量管理方法主要有确定规则的方法和模糊规则的方法两种。基于规则的能量管理策略是根据车辆发动机的车辆行驶状态参数,如发动机的工作效率、电池充电状态(SOC)、车辆速度等参数,该方法设计简单、易于实现,计算资源要求低,鲁棒性较好,主要依赖于有设计经验的工程师^[2,3]。但算法自适应性弱,在较复杂的车辆行驶工况下很难使得 HEV 的电池与发动机有最优的动力扭矩分配。为了提高基于确定规则方法的自适应性,基于模糊规则的方法被提出,从而达到更好的鲁棒性与自适应性。文献[4]采用模糊预测的控制算法,根据车辆当前行驶状态预测下一时刻的车辆状态参数,有效地提高了整车燃油经济性。

基于优化的能量管理策略相较于基于规则的能量管理测量,能够更好地解决 HEV 在行驶过程中车辆状态参数的时变非线性问题^[5]。基于优化的能量管理策略的前提条件是需要确定算法的成本函数与约束条件,在成本函数与约束条件下计算系统目标函数的极值^[6]。基于全局优化的能量管理测量主要依赖 HEV 行驶过程中的先验知识。局部优化的能量管理策略依据 HEV 车辆在行驶过程中的动态参数进行实时优化。文献[7]建立了一个以发动机、发电机的转矩和转速的动态规划模型,该模型可以在保证电池组的荷电状态在安全阈值内,达到整车行驶过程中的燃油消耗最小。模型预测的方法取决于模型的有效性,如果模型过于复杂,不但加大了计算的复杂度,还降低了预测的实时性,迟滞会影响能耗最优的求解。与全局优化对比,局部优化虽然计算量较小,但这是以不考虑车辆未来信息为代价的,因此基于局部优化的算法还

有很多的改进空间。

基于学习的能量管理策略可以避免复杂的规则制定和系统建模,该方法逐渐成为该领域的研究热点。文献[8]利用车辆的定位服务及地图导航信息来获取实时交通与路况信息,通过强化学习,对这些状态量进行优化计算,以降低整车的燃油消耗量。文献[9]针对目前强化学习大多为仿真实验的情况,提出一种通过对 Q-Table 空间的状态变量进行压缩的方法,搭建了硬件在环仿真平台,取得了较好的实验结果。但强化学习由于状态参数多,难以在实车上进行应用,许多科研学者提出了一类鲁棒性更高、实时性更强的深度强化学习方法。文献[10]采用深度强化学习,通过神经网络对参数进行训练,得到了最优的能量利用率,同时有效避免了车辆在能量管理边界条件下的车辆模式频繁切换。

2 HEV 模型

精准的混合动力车辆模型是评估本文所提算法有效性与先进性的前提,本节采用数据建模与机理建模相结合的方式对混合动力车辆各关键部件进行建模,并测试模型的可靠性。

2.1 车辆动力学模型

车辆工作理想条件下,忽略车辆横向作用力的影响,在只考虑车辆前进方向的力的情况下,HEV 的需求功率可以表示为:

$$P = \frac{1}{\eta_t} \left(\frac{m \cdot g \cdot f \cdot v \cdot \cos \alpha}{3600} + \frac{C_D A_a v^3}{76140} + \frac{m \cdot g \cdot v \cdot \sin \alpha}{3600} + \frac{m \cdot \delta \cdot v}{3600} \left(\frac{dv}{dt} \right) \right) \quad (1)$$

式(1)中, P 表示插电混合动力车辆的整车需求功率, η_t 表示车辆传动系统的能量传递效率; m 表示混动车辆的车身质量; g 表示地球重力加速度; f 表示路面滚动摩擦系数; v 表示车辆行驶速度; α 表示道路的坡度值; C_D 表示插电混合动力车的风阻系数; A_a 表示整车的迎风面积; δ 表示车辆旋转质量转换系数; dv/dt 表示车辆行驶过程中的加速度。

2.2 发动机模型

发动机作为 HEV 动力源之一,是整车的核心部件。基于数学方法建模的优点是能够实时反应发动机的动态特性,但是建模过程复杂,并且计算过程长且烦琐。本节采用基于实验数据的建模方法,能有效提高模型运行效率。基于实验数据建模

的发动机油耗可以用式(2)进行描述:

$$b = \sum_{k=1}^M \alpha_{pk, qk} n_e^{pk} T_e^{qk}, p, q \in N^M, \alpha \in R^M \quad (2)$$

其中, b 为整车的发动机油耗; $\alpha_{pk, qk}$ 为该表达式的拟合系数; n_e^{pk} 为车辆发动机的转速; T_e^{qk} 为整车发动机扭矩; M 表示该模型的最高次数。

2.3 电机模型

并联式 P2 架构 HEV 中的电机既可以作为电动机驱动车辆,也可以作为发电机对电池进行充电或回收制动能量。本文选择基于实验数据的方法对电机进行建模,电机的工作效率可用式(3)表示:

$$\eta_m = f(T_m, n_m) \quad (3)$$

建模中忽略电机的温度和电磁的作用对电机效率产生的影响,电机的输出功率可用式(4)表示:

$$P_m = \begin{cases} \frac{T_m \cdot n_m}{9550 \eta_m}, T_m > 0 \\ -\frac{T_m \cdot n_m \cdot \eta_m}{9550}, T_m < 0 \end{cases} \quad (4)$$

$T_m > 0$ 表示电机参与整车驱动下的输出功率; $T_m < 0$ 表示电机处于发电工况。

2.4 电池模型

常见的电池模型有电化学模型、等效电路模型及基于数据驱动模型,本文选择二阶戴维宁等效电路模型来描述电池内部参数的变化,该模型计算简单、模型准确度高,二阶 RC 等效电路模型如图 1 所示。图 1 中, U_α 表示电池组的开路电压, U_t 表示电池的端电压, I 表示电池的工作电流, R_0 表示电池内阻,两个 RC 回路中的 R_1 、 R_2 、 C_1 、 C_2 分别表示电池的极化电阻和极化电容。

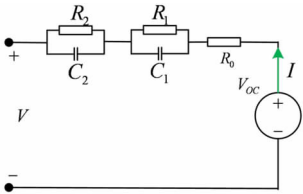


图 1 二阶 RC 等效电路模型

根据基尔霍夫定律及欧姆定律,二阶戴维宁等效电路动力学方程可用式(5)表达。

$$\begin{cases} \frac{dI_{R_1}}{dt} = \frac{1}{R_1 C_1} I - \frac{1}{R_1 C_1} I_{R_1} \\ \frac{dI_{R_2}}{dt} = \frac{1}{R_2 C_2} I - \frac{1}{R_2 C_2} I_{R_1} \\ \frac{dSOC(U_\alpha)}{dt} = \frac{\eta I}{Q} \\ SOC(U_\alpha) = U_t + R_0 I + R_1 I_{R_1} + R_2 I_{R_2} \end{cases} \quad (5)$$

式中, I 表示工作电流,充电为正值,放电为负值; U_1 、 U_2 、 I_{R_1} 、 I_{R_2} 分别表示两个 RC 回路的电压与 R_1 、 R_2 的电流; η 表示库伦效率; Q 表示额定容量。

3 能量优化算法设计

3.1 等效燃油消耗最小策略

等效消耗最小化策略(ECMS)是一种启发式算法。算法在每个瞬间仅使用基于动力系统实际能量流的参数来解决瞬时等效燃油最小化问题。混合动力汽车能源管理领域的庞特里亚金最小原理条件在本质上等同于 ECMS 方法。

ECMS 的关键思想是,在充电和放电中,等效的燃料消耗可以与电能的使用相关联;未来(或过去)的等效燃料消耗量 $\dot{m}(t)_{ress}$ (g/s),可加到当前的实际燃料消耗量 $\dot{m}(t)_f$ (g/s),从而得到瞬时等效燃料消耗量 $\dot{m}(t)_{f,eqv}$ 。表达式如式(6)。

$$\dot{m}(t)_{f,eqv} = \dot{m}(t)_f + \dot{m}(t)_{ress} \quad (6)$$

发动机瞬时燃料消耗量为:

$$\dot{m}(t)_f = \frac{P_{eng}(t)}{\eta_{eng}(t) Q_{lhv}} \quad (7)$$

其中, Q_{lhv} (MJ/kg) 为燃油低热值(单位质量的能量含量); $\eta_{eng}(t)$ 为发动机效率; $P_{eng}(t)$ 为发动机以一定效率运行时所产生的功率。电机运行时的产生虚拟燃油消耗量为:

$$\dot{m}(t)_{ress} = sfc_{eq}(t) \cdot P_{batt}(t) \quad (8)$$

虚拟油耗可以通过使用虚拟比油耗 $sfc_{eq}(t)$ (g/kWh) 来评估。虚拟比油耗与等效系数 $s(t)$ 成正比,该系数在电池充电或放电时不同。

$$\dot{m}(t)_{ress} = \frac{s(t)}{Q_{lhv}} \cdot P_{batt}(t) \quad (9)$$

等效因子 $s(t)$ 是一个值向量, $s(t) = [s_{chg}(t), s_{dis}(t)]$ 。作用是为电力的使用分配成本,将电力转化为等效的燃料消耗。实际上,等效系数 $s(t)$ 代表了燃料转化为电能的效率链。等效系数在 ECMS 的最初推导中是一个常数,可以解释为电能循环路径的平均总效率。

实现 ECMS 时,常用惩罚函数使电池荷电量在允许的范围,即 $SOC_{min} \leq SOC \leq SOC_{max}$ 。因此,通过使用适当构造的乘法惩罚函数 $p(SOC)$ 对公式(15)进行改进:

$$\dot{m}_{f,eqv}(t) = \dot{m}_f(t) + \frac{s(t)}{Q_{lhv}} \cdot P_{batt}(t) \cdot p(SOC) \quad (10)$$

其中 $p(SOC)$ 的计算方法为:

$$p(\text{SOC}) = 1 - \left(\frac{\text{SOC}(t) - \text{SOC}_{\text{target}}}{(\text{SOC}_{\text{max}} - \text{SOC}_{\text{min}})/2} \right)^a \quad (11)$$

式(11)中, a 一般取值为 1、3、5 等奇数。如果 SOC 等于目标值 $\text{SOC}_{\text{target}}$, 则惩罚函数为 1, 但其值会变化(更高或更低)以适当地补偿与 $\text{SOC}_{\text{target}}$ 的偏差。当 $\text{SOC} > \text{SOC}_{\text{target}}$ 时, $p(\text{SOC}) < 1$, 电池能量的成本较低, 从而使 SOC 高于参考值时更容易放电。当 $\text{SOC} < \text{SOC}_{\text{target}}$ 时, $p(\text{SOC}) > 1$, 电池的能量成本增加, 降低电池的输出功率。

ECMS 算法思想源于工程直觉, 利用庞特里亚金最小值原理可以得到等效油耗的解析推导。哈密顿量函数可表示为式(12)。

$$H = P_{\text{fuel}}(t) + \lambda(t) \cdot P_{\text{ech}}(t) \quad (12)$$

若在 ECMS 中以功率的形式重写式(15), 将所有项乘 Q_{hiv} , 则瞬时成本变为:

$$P_{\text{eqv}}(t) = P_{\text{fuel}}(t) + s(t) \cdot P_{\text{batt}}(t) \quad (13)$$

对比式(18)和式(19), 表明优化问题的哈密顿量可以被视为等效燃料消耗或者等效燃料功率, 其中 $\lambda(t)$ 是将电池功率转换为燃料功率的加权因子。

3.2 基于强化学习的 ECMS

由于 ECMS 对等效因子(equivalence factor, EF) $S(t)$ 的敏感性和固定的 EF 不能适应复杂的真实驾驶条件, 本节开发了一种基于 Q-learning 的强化学习(RL)方法与常规 ECMS 相结合的能源管理算法。RL-ECMS 方法基于 RL 智能体与驾驶环境之间的交互更新等效因子, 从而获得更好的

SOC 轨迹和更强的鲁棒性以及更好的燃油经济性。并将所提出方法应用于并联混合动力汽车, 通过两种典型驾驶工况对算法进行测试, 并与基于规则的能量管理方法和标准 ECMS 进行比较。强化学习解决控制问题的结构框架如图 2 所示。

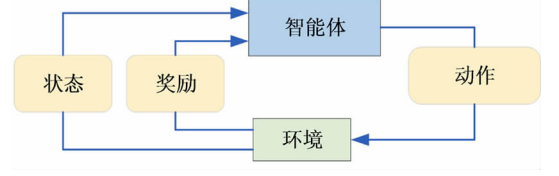


图 2 强化学习算法结构

在 Q 学习中, 使用动作价值函数(action value) π , 它可以定义为即时奖励和下一时刻价值函数的总和:

$$Q_{\pi}(s, a) = J_{\pi}(s) = g(s, a) + \gamma J_{\pi}(s') \quad (14)$$

式中, $g(s, a)$ 为在 k 时刻执行动作 u_k 后的瞬时代价; $J_{\pi}(s')$ 为 $k+1$ 时刻的动作价值函数; γ 表示折扣因子; s 为状态变量; a 为采取的动作; s' 为下一时刻的状态

最优代价函数可以被定义为:

$$J^{\times}(s) = \min_a (Q^{\times}(s, a)) \quad (15)$$

最优控制策略计算方式为:

$$\pi^{\times}(s) = \arg \min_a (Q^{\times}(s, a)) \quad (16)$$

Q 函数值通过式(17)更新为:

$$Q(s, a) \leftarrow Q(s, a) + \alpha(g(s, a) + \gamma \max_a Q(s', a) - Q(s, a)) \quad (17)$$

本文设计同时训练两个智能体来实现混合动力车辆能量优化分配。算法结构如图 3 所示。

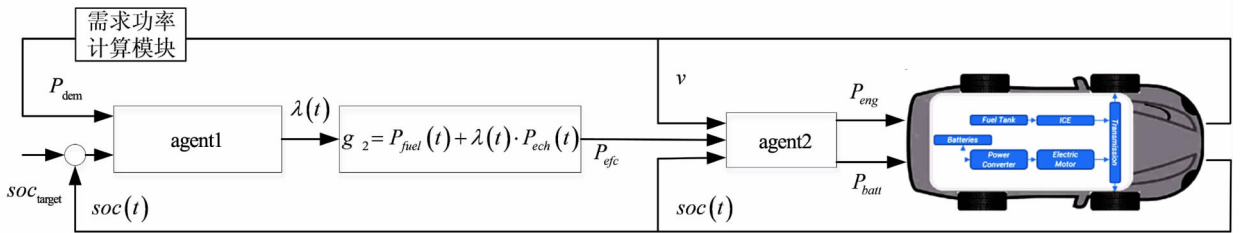


图 3 RL-ECMS 控制策略结构图

智能体 1 瞬时代价函数定义为式(18):

$$g_1(t) = \dot{m}_{\text{fuel}}(t) + F(\text{soc}(t)) \quad (18)$$

其中, $\dot{m}_{\text{fuel}}$ 为燃油消耗质量变化量。

$$F(\text{soc}(t)) =$$

$$\begin{cases} C_1 \cdot (\text{soc}(t) - \text{soc}_{\text{target}})^2, & \text{if } \text{soc} > \text{soc}_{\text{min}} \\ C_2, & \text{if } \text{soc} \leq \text{soc}_{\text{min}} \end{cases} \quad (19)$$

其中, C_1, C_2 为正数; soc_{min} 为 soc 最小值。智能体输出为等效因子的值, 则其动作价值函数可定义为式(20):

$$J_1(s(t), a) = \sum_{t=i}^{\infty} \gamma^{t-i} g_1[s(t), a(t)] \quad (20)$$

智能体 2 将等效燃油消耗量作为学习过程中

的瞬时代价函数,其瞬时代价函数定义为:

$$g_2 = P_{fuel}(t) + \lambda(t) \cdot P_{ech}(t) \tag{21}$$

则智能体动作价值函数可定义为:

$$J_2(s(t),a) = \sum_{t=i}^{\infty} \gamma^{-i} g_2[s(t),a(t)] \tag{22}$$

最优动作价值函数为:

$$J_2 \times (s(t)) = \min \{ g_2[s(t),a(t)] + \gamma J_2 \times (s(t+1)) \} \tag{23}$$

最优执行动作为:

$$a \times (t) = \arg \min_{a(t)} (g_2(s(t),a(t)) + \gamma J_2 \times (s(t+1))) \tag{24}$$

其中,状态 $s = [SOC, P_{dem}, v]$ 。

4 仿真实验

4.1 实验环境

本节基于第二章所述的景区混合动力车辆仿真模型验证所提出的控制策略的有效性。车辆仿真模型在基于 MATLAB/Simulink 搭建中进行,RL-ECMS 算法与对比算法均是基于 S-function 进行开发。仿真环境是在 Windows 操作系统上执行。电脑配置为:Core(TM) i7-8700k CPU,主频 3.70 GHz,内存 16 GB。车辆数据如表 1 所示。

表 1 仿真车辆基本参数

部件类型	参数名称	参数值
整车参数	整车整备质量/kg	11500
	迎风面积/m ²	6.5
	轮胎滚动半径/m	0.512
	空气阻力系数	0.6
	滚动阻力系数	0.012
发动机	峰值功率/kW	147
	峰值扭矩/Nm	800
电机	峰值功/kW	130
	峰值扭矩/Nm	850
电池	电池电/V	384
	电池容量/Ah	50
传动系	变速箱传动比	4.18;2.53;1.54;0.78
	主减速比	5.57

强化学习算法参数值见表 2。

在表 2 中,C1、C2 的值分别确定为 0.01 与 10,用来平衡式(24)中智能体 1 瞬时代价函数的燃料消耗与电量消耗。在物理意义上,C1 的值可以根据电池荷电状态的变化量来确定,C2 相对于 C1

给出了一个比较大的值,可以保障电池荷电状态保持在最小值以上。另外,α 是一个必须设置的超参数,用来确定智能体的学习速度。γ 表示未来成本的折扣率,假设其值接近 1。

表 2 强化学习算法参数表

参数名称	参数值
C1	0.01
C2	10
α	0.95
γ	0.99996
SOC target	0.600
SOC min	0.180

4.2 实验结果

本节采用对比算法为基于 PI 控制器的等效燃油消耗最小控制策略与基于规则(Rule)的控制策略,PI 控制器调节等效因子的方法为式(25):

$$\lambda(t) = \lambda(0) + K_p(soc_{target} - soc(t)) + K_i \int (soc_{target} - soc(t)) dt \tag{25}$$

其中,λ(0) 为初始等效因子;K_p、K_i 分别为控制器的比例系数与积分常数。模拟实验选取了 FTP75 与 FTP75-Highway 两种典型基础工况。PI 控制器参数分别设置为 25 和 0.06,并与自适应的 ECMS 策略进行比较。FTP75 工况实验结果如图 4~图 7 所示。

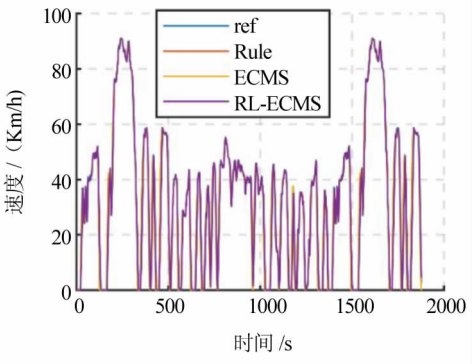


图 4 FTP75 工况车辆行驶速度

由图 4 可看出,3 种算法均能满足车辆正常行驶的性能需求,即均能很好地跟随给定工况速度曲线,这是控制策略设计的前提。

从图 5 可以看出,基于强化学习的 ECMS 控制策略平均值比另外两种算法高,说明车辆行驶过程中回收了较多的动能从而对电池进行充电,有效避免了能源浪费。

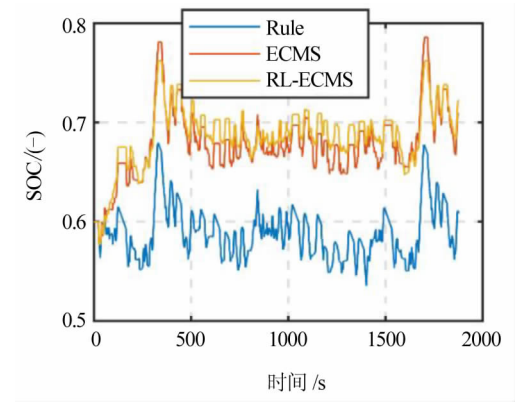


图 5 FTP75 工况车辆电池 soc

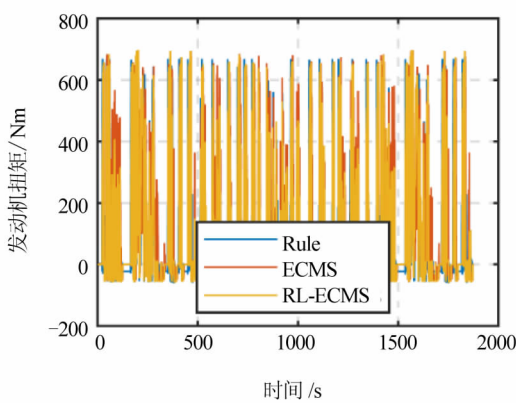


图 6 FTP75 工况发动机扭矩

由图 6 可以看出,RL-ECMS 策略的发动机扭矩在车辆运行过程中出现比另外两种工况更小的情况,这是由于整车的实际需求扭矩一部分由电机承担,即消耗电池的电量,在不损失车辆动力性能的同时实现节油效果。

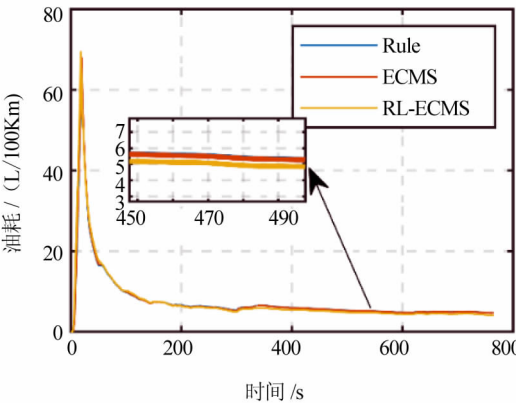


图 7 FTP75 工况油耗对比

由图 7 能明显看出,RL-ECMS 策略油耗更低。算法节能优势更加明显。图 8 展示了两种不同策略的等效因子,ECMS 采用恒定的等效因子

3.43,而 RL-ECMS 由于采用基于强化学习的自适应等效因子,通过学习,能够实现根据行驶环境的变化,因此能获得比恒定等效因子更好的节能效果,且不损失车辆的动力性能。

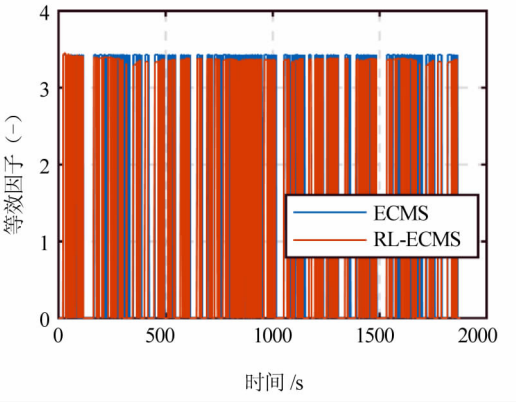


图 8 FTP75 工况等效因子

FTP75-高速工况如图 9~图 12 所示。

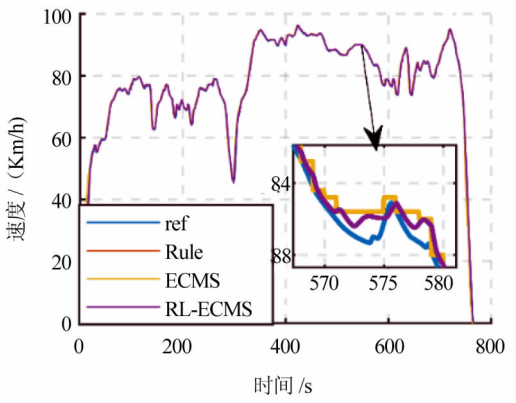


图 9 FTP75-高速工况车辆行驶速度

由图 9 同样可看出,3 种算法均能满足车辆正常行驶的性能需求,即均能很好地跟随给定工况速度曲线。

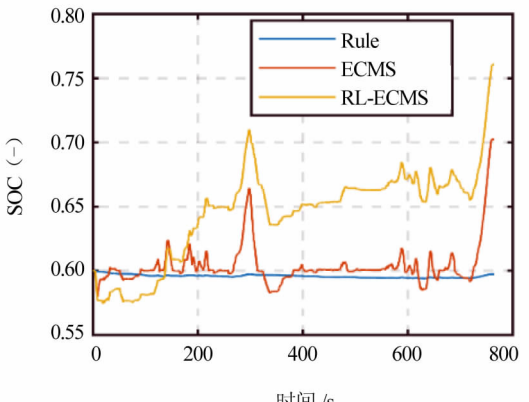


图 10 FTP75-高速工况车辆电池 soc

高速工况下,RL-ECMS 策略依然能够保持更高的 SOC 值,说明该策略能更有效地进行动能回收。

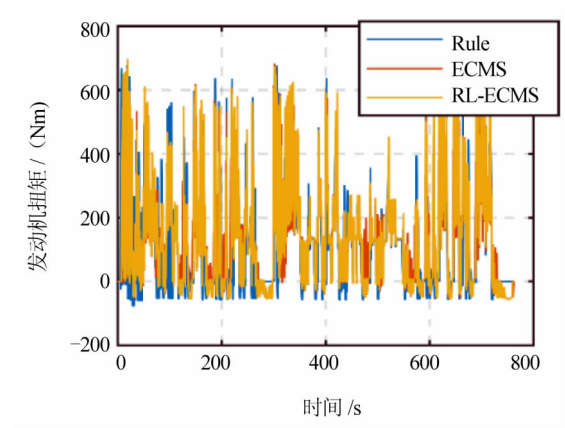


图 11 FTP75-高速工况发动机扭矩

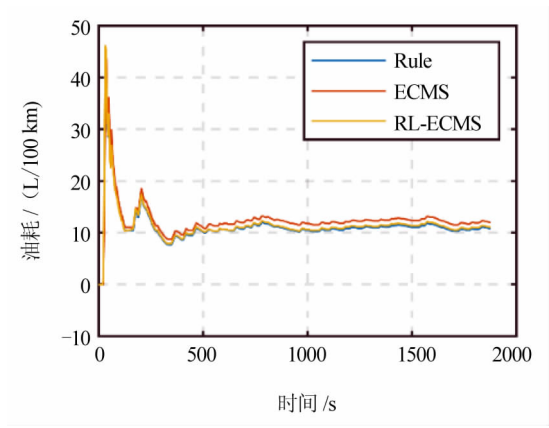


图 12 FTP75-高速工况燃油消耗

由图 12 明显看出,FTP75-高速工况下 RL-ECMS 也能实现更低的百公里燃油消耗。

为验证本文所提算法的鲁棒性,将测试工况设

置为一种新的 ECE 工况,并将电池起始 SOC 值设置为 0.19,实验结果对比见表 3。

表 3 三种工况对应的算法结果对比

仿真工况	控制策略	平均油耗/L	平均节油率/%	SOC初值	SOC终值
FTP75-high	Rule	12.219	—	0.600	0.609
	ECMS	11.478	6.1	0.600	0.709
	RL-ECMS	11.247	8.0	0.600	0.723
FTP75	Rule	8.531	—	0.600	0.597
	ECMS	7.713	9.6	0.600	0.702
	RL-ECMS	7.607	10.8	0.600	0.761
ECE	Rule	14.364	—	0.189	0.310
	ECMS	12.984	9.6	0.189	0.330
	RL-ECMS	12.125	15.6	0.189	0.544

表 3 为不同算法在 3 种不同工况下的性能表现对比。通过表 3 容易看出,RL-ECMS 具有更好的燃油经济性,且算法对新工况的适应性更好。通过数据分析计算,本文所提算法相较于基于规则的能量优化算法,三种工况下的节油率分别为8.0%,10.8%与 15.6%,均高于 ECMS 的 6.1%、9.6%与 9.6%,且算法在 ECE 工况下也能保持较好的节油率,说明算法具有较强的鲁棒性与泛化性。

5 结 论

针对 P2 结构混合动力车辆的能量管理问题,提出了一种基于强化学习的自适应等效燃油消耗最小策略(RL-ECMS)。通过两个智能体实现等效因子

的自适应更新和车辆扭矩的动态分配,有效适应了不断变化的驾驶需求。与常规 ECMS 和基于规则的控制策略的对比仿真实验结果表明,RL-ECMS 在 FTP75 和 FTP75-Highway 两种典型驾驶工况下,均展现出更低的燃油消耗,相较于传统控制策略,具有显著的节能效果。在未经训练的 ECE 典型工况下进行的测试表明,RL-ECMS 策略不仅在已知工况下表现良好,同时在未知工况下也能保持优秀的性能,证明了算法的鲁棒性和泛化性。本研究不仅为混合动力车辆的能量管理提供了有效的策略,也可为强化学习在汽车能量管理领域的应用提供有价值的参考。

(下转第 30 页)