

Acceleration Control Design of HEVs with Comfortability Evaluation based on IRL

Shohei Narita* Jiangyan Zhang** Weidong Zhang***
Tielong Shen*

* *Department of Engineering and Applied Sciences, Sophia University,
Tokyo 102-8554 Japan (e-mail: s-narita-7wq@eagle.sophia.ac.jp).*

** *College of Mechanical and Electronic Engineering, Dalian Minzu
University, Dalian 116600, China (e-mail:
zhang-jiangyan@dlmu.edu.cn)*

*** *Department of Automation, Shanghai Jiao Tong University,
Shanghai 200240, China*

Abstract: For passenger cars, comfortability is an important issue to consider in powertrain control. However, it is not easy to take comfortability into account when designing a powertrain control strategy because of its subjective nature and difficulty in being quantified. This paper presents a solution by using the inverse reinforcement learning (IRL) method. An acceleration scenario of a hybrid electric powertrain is considered to show this design approach. With a sample acceleration profile scored by an expert evaluating module, a reward function is obtained by training an extreme learning machine. Using the analytical representation of comfortability, an estimation of distribution algorithm (EDA) is used to seek the optimal acceleration reference. Given the reference acceleration signal, the control law for the electric motors that provide power assistance during the acceleration phase is obtained by solving a optimal tracking control problem. A numerical example is shown to evaluate the design approach.

Copyright © 2021 The Authors. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: Hybrid electric vehicle, ride comfort, acceleration, inverse reinforcement learning, optimal control

1. INTRODUCTION

Comfortability, which is an important index for evaluating the performance of passenger cars, can be influenced strongly by dynamic control of the powertrain system. For hybrid electric vehicles (HEVs), controlling the powertrain to achieve desired comfortability is a challenging issue since different power sources such as the internal combustion engine and electric motors operate cooperatively to generate power. Especially during the acceleration and braking modes that are required in common maneuvers such as merging, passing, and deceleration, the electric motor and internal combustion engine generally require on-off control to switch between power production and energy regeneration. The energy management problem accounts for most of the literature in the field of HEV powertrain optimization, for instance, Lee (2020) and Hu (2019), which provide insight on effective energy management strategies for the cooperative operation of the internal combustion engine and electric motors.

Much of the existing literature focuses on finding the optimal control strategy for some given route. However, a strategy that is applicable in the real world must function without a priori knowledge of a vehicle's path. To this end, Jiang (2017) and Zhang (2021) propose a real-time adaptive optimal control scheme to minimize the energy

consumption of an automated HEV. While many papers discuss long-term optimization for HEV powertrains, only a few address the issue of short-term optimal control that is required in specific driving scenarios. For instance, a merging scenario is addressed by Xu (2021), in which the torque distribution problem is formulated as an optimal control problem with the cost function containing velocity and energy consumption terms, and consideration given to the dynamics of the vehicle. An acceleration scenario is addressed by Zhang (2020), in which the torques of the motor and combustion engine are controlled to minimize energy consumption while also tracking the vehicle's speed in real time. However, neither paper considers comfortability. Considering comfortability raises challenges, in that it is subject to an individual's senses, and it cannot easily be quantified. To this end, recent papers such as Shen (2020) and Zhu (2020) aim to increase comfortability by placing a limit on a vehicle's jerk signal. Although a correlation is seen between comfortability and jerk, giving an exact description of said relation poses additional challenges.

This paper addresses the powertrain control problem of HEVs while also considering comfortability. A major obstacle in optimizing comfortability via dynamic control lies in the need to explicitly define a cost function. In this paper, this challenge is overcome by employing an inverse reinforcement learning (IRL) technique. Suppose that a human's perception of comfortability is quantized by some score. Using this score as training data, the

* This work is supported in part by National Natural Science Foundation (NNSF) of China under Grant 61973053.

relation between a vehicle's response and the cost function is approximated by an extreme learning machine (ELM). Using the result of IRL together with an estimation of distribution algorithm (EDA), the optimal acceleration profile is obtained. In turn, the optimal power assistance law for the electric motor is designed by solving an optimal control problem, assuming that the engine is operated at full throttle. Finally, the proposed design scheme is validated by numerical simulation.

2. BACKGROUND AND CONTROL ISSUE

Consider the powertrain shown in Fig. 1, composed of two electric motors MG1 and MG2, a battery package, an internal combustion engine with a turbocharger, a clutch, an automatic transmission (AT) system with a 10-speed gearbox, differential gears, and a distortion component. Suppose the vehicle is initially cruising in EV mode at a

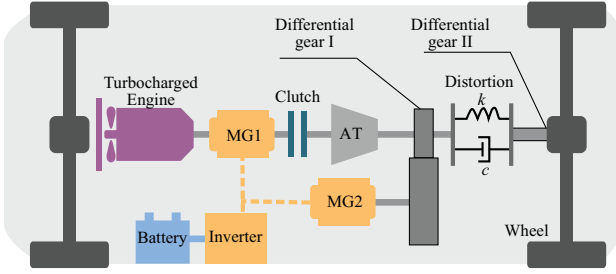


Fig. 1. Configuration of HEV with two motors

constant speed v_0 , with the driving torque provided solely by motor MG2. Thereafter, the vehicle enters acceleration mode targeting a final speed v_f . Due to the electric power constraint, the combustion engine is switched on with full throttle operation. While the vehicle is accelerating, the gear must be managed promptly. Achieving high comfortability by controlling the powertrain in this acceleration scenario raises the following issues:

- 1) Determining when to engage and release the clutch.
- 2) Managing the power generated by the electric motor to in turn control the total driving torque.
- 3) Evaluating comfortability using a cost function, according to the dynamical behavior of the powertrain during the acceleration mode.

Of these issues, the third can be solved provided the answers to the first two. That is, given the optimal controls resulting from 1) and 2), the speed and acceleration profiles may be obtained in accordance with the dynamics of the powertrain. However, solving issues 1) and 2) comes with multiple difficulties resulting from there being switching actions in the system, the power transformation having varying efficiency, and the power generation by the combustion engine with a turbocharger having strong nonlinearity. Above all, designing a control law via model-based optimization is difficult in the absence of an analytic formulation of comfortability. The only available knowledge regarding comfortability is the scores given by humans. Expert models are used in industry to evaluate comfortability.

Fig. 2 shows three acceleration responses under different power settings, where v_i and a_i ($i = 1, 2, 3$) represent the

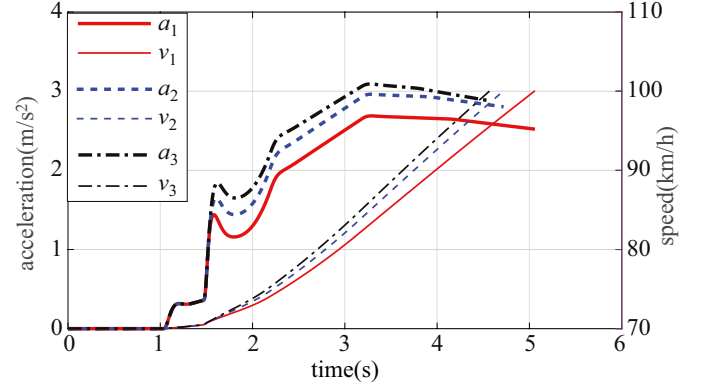


Fig. 2. Three examples of comfortability score

Table 1. Test conditions and comfortability scores from expert model

	Assistant motor torques	Score
Test 1 (a_1, v_1)	$\tau_{m1} = \tau_{m2} = 0$	27.9
Test 2 (a_2, v_2)	$\tau_{m1} = 0, \tau_{m2} = 20(\text{Nm})$	25.7
Test 3 (a_3, v_3)	$\tau_{m1} = 0, \tau_{m2} = \tau_{m2}^{max}$	26.1

speed and corresponding acceleration signals, respectively. The power settings of the motors as well as the corresponding comfortability scores are shown in Table 1.

We begin with an analytic representation of the comfortability score to find the optimal acceleration profile. First, the inverse reinforcement learning technique is used to construct this analytic function. Then, using this function as a measure for optimality, the optimal acceleration signal is found via EDA. Finally, the optimal control law is sought by solving the optimal control problem regarding acceleration tracking.

3. DESIGN SCHEME

3.1 Inverse Reinforcement Learning

Since the comfortability score is based on human feeling which, in turn, is a dynamical response of the acceleration signal, we use an extreme learning machine (ELM) to approximate this score. To this end, the Distance Minimized Inverse Reinforcement Learning (DMIRL) approach proposed by Burchfiel (2016) is used with a set of training data.

Let a_k^i ($k = 0, 1, \dots, M^i \leq M$) denote a discretized signal of acceleration, where $i = 1, 2, \dots, S$ is the index of the acceleration data. The corresponding score of a_k^i is represented by J^i . Suppose the unknown reward function R on time step k is given by

$$R(x_k) = w^\top \phi(x_k), \quad (1)$$

where $x_k = [a_k \ k]^\top$ is the state of the system, $\phi(x_k) = [\phi_1(x_k) \ \phi_2(x_k) \ \dots \ \phi_N(x_k)]^\top$ maps the state x_k onto a N -dimensional feature vector, and $w \in \mathbb{R}^N$ is the vector of weighting coefficients, i.e.,

$$w = [w_1 \ w_2 \ \dots \ w_N]^\top.$$

Following the standard notation of extremal learning machines (ELMs), the neuron function $\phi_j(x_k)$ is defined as follows:

$$\phi_j(x_k) = \tanh(p_j a_k + h_j k + q_j).$$

The values of the weighting factors p_j, h_j, q_j ($j = 1, 2, \dots, N$) are assigned randomly. The structure of the ELM is shown in Fig. 3.

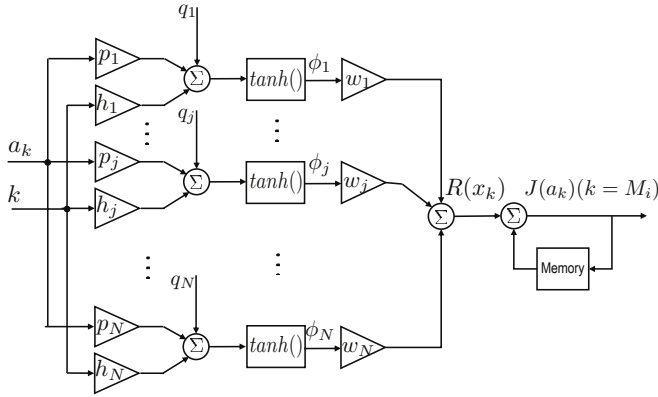


Fig. 3. ELM-based model structure

Given the output of the ELM, the cost function for representing the comfortability score is constructed by

$$J_E^i(a_k^i) = \sum_{k=1}^{M_i} R(x_k). \quad (2)$$

Note that J_E^i is linear in the weighting coefficient w , that is,

$$\begin{aligned} J_E^i(a_k) &= \sum_{k=1}^{M_i} R(x_k) \\ &= w^T \sum_{k=1}^{M_i} \phi(x_k) \\ &= w^T \begin{bmatrix} \sum_{k=1}^{M_i} \tanh(p_1 a_k^i + h_1 k^i + q_1) \\ \sum_{k=1}^{M_i} \tanh(p_2 a_k^i + h_2 k^i + q_2) \\ \vdots \\ \sum_{k=1}^{M_i} \tanh(p_N a_k^i + h_N k^i + q_N) \end{bmatrix}. \end{aligned} \quad (3)$$

Benefiting from this linearity, we are able to determine the optimal coefficient w by minimizing the the following cost function:

$$\min_w \sum_{i=1}^S \|J(a_k^i) - J_E^i(a_k^i)\|^2. \quad (4)$$

3.2 Seeking Optimal Acceleration profile

An advantage of the ELM representation of the comfortability score is that it enables us to seek the extremal acceleration signal by constructing an analytic algorithm, i.e., by finding an acceleration signal $a_k^*, k = 1, 2, \dots, M$ that maximizes the cost function $J_E(a_k)$. In this section, we discuss the EDA to reach this goal.

To simplify the problem of seeking the optimal acceleration profile, we represent the acceleration signal with four parameters, a_1, a_2, t_1 , and t_2 , which capture the fundamental

characteristics of the signal as shown in Fig. 4. Using this representation, the extremum seeking problem can be formulated as follows. Note that the optimization problem is discretized with time step Δt .

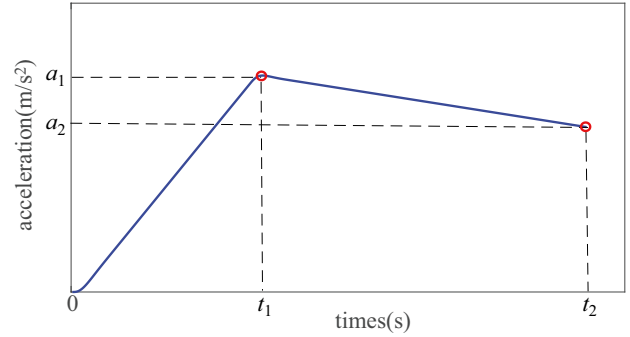


Fig. 4. Shape of the acceleration

$$\begin{aligned} &\max_u \sum_{k=1}^M R(a, t) \\ &\text{subject to} \\ &\begin{cases} u = [a_1 \ a_2 \ t_1 \ t_2] \\ a(t) = \frac{a_1}{t_1} t \quad [t \leq t_1] \\ a(t) = \frac{a_2 - a_1}{t_2 - t_1} (t - t_1) + a_1 \quad [t_1 < t \leq t_2] \\ a_{min} \leq a(t) \leq a_{max} \\ 0 \leq t_1 \leq t_2 \leq 6 \\ 0 \leq t \leq t_2 \\ M = \frac{t_2}{\Delta t} \end{cases} \end{aligned} \quad (5)$$

An estimation of distribution algorithm, which is a stochastic optimization method, is used to solve this optimization problem. It searches for the optimum by building and sampling explicit probabilistic models of promising candidate solutions. Simply put, EDA is an extension of the genetic algorithm (GA). The difference between EDA and other meta-heuristic optimization methods is that it builds a probabilistic model that captures the probability distribution from the population of good candidate solutions. The flowchart of EDA is shown in Fig. 5.

The algorithm begins with an initial population of randomly generated individuals given a range of parameters. Then, the performance of each individual in the initial population is evaluated. Based on these evaluations, the top individuals are selected from the population. A probabilistic model is estimated based on these selected individuals. The next generation of individuals is generated by sampling the estimated probability model. The better the evaluation value, the higher the chance an individual makes its way into the next generation. Finally, the each individual in the next generation is evaluated. This process is repeated until the optimal solution is reached.

3.3 Powertrain Control Law

The acceleration of a vehicle is generated by control of the powertrain dynamics. As is shown in Fig. 6, the powertrain system considered in this paper has a combustion engine

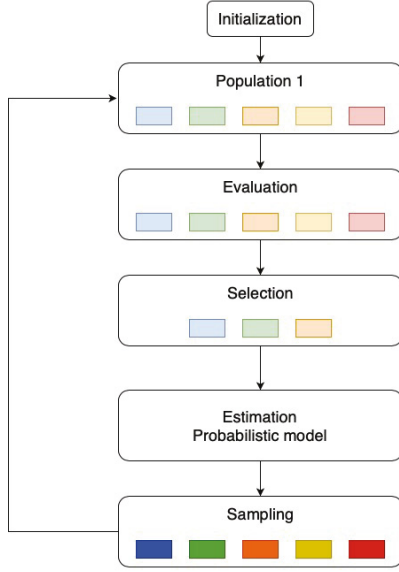


Fig. 5. Flowchart of EDA

and two electric motors which generate power. As mentioned in the Section 2, the combustion engine is assumed to be operating at full throttle. Therefore, the only way to control the total propulsion power of the powertrain is to control the motor torques. In what follows, a dynamic control law for tuning the motor torque is designed by solving an optimal acceleration tracking problem.

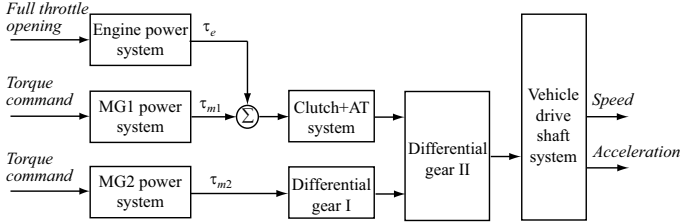


Fig. 6. Block diagram of the HEV driveline system

First, we summarize the model of the powertrain dynamics. Since the engine with a turbocharger is operating at full throttle, the torque generated by the engine is given by

$$\dot{\tau}_e(t) = f_{\tau_e}(\omega_e(t)), \quad (6)$$

where ω_e denotes the engine speed and $f_{\tau_e}(\cdot)$ is a nonlinear function that represents the torque raising rate with respect to engine speed. Note that the maximum torque during this engine starting period is constrained by a pre-calibrated bounding line $\tau_e^{max}(\omega_e)$. The dynamics of the motor are modeled by

$$\dot{\tau}_{m_i}(t) = \frac{1}{c_{m_i}}(u_{m_i}(t) - \tau_{m_i}(t)), \quad (i = 1, 2), \quad (7)$$

where u_{m_i} represents the torque commands and c_{m_i} represents time constants. The maximum torque limits τ_{m1}^{max} and τ_{m2}^{max} of MG1 and MG2, respectively, are also given by their respective rotational speeds ω_{m1} and ω_{m2} .

With consideration of the power transmission efficiency denoted by η_t , the total torque τ_t acting on the differential gear is given by

$$\tau_t(t) = \eta_t i_g(t)(\tau_e(t) + \tau_{m1}(t)), \quad (8)$$

where i_g denotes the gear ratio, including efficiency-loss. Taking efficiency-loss into account, the gear ratio during the transient period of gear changing is represented by

$$i_g(t) = f_g(t, i_g(t^*), i_g^f), \quad t \in [t^*, t^* + 0.7], \quad (9)$$

where t^* denotes the timing of gear shifting. Combining these two models, the transmission efficiency is given by

$$\eta_t(t) = f_\eta(t, \eta_t(t^*), \eta_t^f), \quad t \in [t^*, t^* + 0.7], \quad (10)$$

where f_η is a nonlinear continuous function and η_t^f is the efficiency at the desired gear position after shifting gears. Moreover, during the gear shifting process, the gear ratio variation is approximated using a time-varying parameter t , where f_g is a linear continuous function of time t and i_g^f denotes the ratio of the desired gear position.

Furthermore, when the clutch is engaged, considering the relation (8), the inertia dynamics of the powertrain system can be represented as

$$\begin{aligned} \dot{v} &= f_v(t, \tau_e, \tau_{m1}, \tau_{m2}, v) \\ &= \frac{1}{I_1 + I_2 \rho(t)} [C_2 \eta_t i_g (\tau_e + \tau_{m1}) + C_1 \tau_{m2} - F(v)]. \end{aligned} \quad (11)$$

In Equation (11), v represents vehicle speed, I_1 and I_2 represent the equivalent moment of inertia from the MG2 shaft to the vehicle driving shaft and the equivalent moment of inertia from the engine shaft to the wheel of the vehicle, respectively, $F(v)$ represents the road load that is a nonlinear function of vehicle speed, and $C_1 = \eta_f G_m G_f / R_w$ and $C_2 = \eta_f G_f / R_w$. η_f represents the final transmission efficiency, R_w represents the radius of the vehicle wheel, and G_m and G_f represent the reduction ratio of the differential gears I and II, respectively.

Moreover, the ω_{m2} is related to the vehicle speed by

$$\omega_{m2} = \frac{G_m G_f}{R_w} v. \quad (12)$$

Note that when the clutch is released, the vehicle operates in EV mode driven by MG2 while MG1 operates to start the engine. In summary, the engine speed has the following relations

$$\begin{cases} (I_e + I_{m1})\dot{\omega}_e = \tau_e + \tau_{m1}, & \text{clutch released} \\ \omega_e = \omega_{m1} = \frac{i_g(t)G_f}{R_w} v, & \text{clutch engaged} \end{cases} \quad (13)$$

where I_e and I_{m1} denote the inertias of the engine shaft and the MG1 shaft, respectively.

Thus, for a given optimal acceleration profile a^* that is synthesized with the extremum seeking result u^* given in Section 3.2, the optimal torque assistant control law can be found by solving the following optimal acceleration tracking problem subject to the system dynamics constraint.

$$\begin{aligned} \min_u J(u) &= \min_u \int_{t_0}^{t_f} (a^*(t) - a(t))^2 dt \\ \text{subject to} & \begin{cases} \dot{v}(t) = f_v(t, \tau_e, \tau_{m1}, \tau_{m2}, v) \\ \dot{\tau}_e(t) = f_{\tau_e}(\omega_e(t)) \\ \dot{\tau}_{m_i}(t) = \frac{1}{c_{m_i}}(u_{m_i}(t) - \tau_{m_i}(t)) \\ 0 \leq \tau_e \leq \tau_e^{max}(\omega_e) \\ \tau_{m_i}^{min}(\omega_{m_i}) \leq \tau_{m_i} \leq \tau_{m_i}^{max}(\omega_{m_i}) \\ v_{min} \leq v \leq v_{max} \end{cases} \end{aligned} \quad (14)$$

The control signal includes the torque command for both electric motors, that is, $u = [u_{m1} \ u_{m2}]$. t_0 is the starting time of the acceleration mode, and $[t_0, t_f]$ denotes the acceleration period.

4. SIMULATION VALIDATION

In summary, a control design approach was proposed to achieve better comfortability as evaluated by an ELM-based cost function obtained via inverse reinforcement learning. The first step is to formulate a method for comfortability evaluation using inverse reinforcement learning. Using the result of IRL, an ELM-based cost function is constructed, and EDA is used to determine an optimal acceleration profile. This profile is used as a reference signal for the powertrain control, which is realized by solving an optimal acceleration tracking problem subject to the powertrain dynamics and other physical constraints. This proposed design approach is validated using numerical simulation.

300 acceleration data are used as training data for inverse reinforcement learning. The training data is generated by randomly varying the characteristic parameters (see Fig. 7), where the corresponding scores are 79.9, 66.8 and 58.0, respectively. To evaluate the ELM-based cost function, the cost function was tested using a set of 90 data points. The root mean square error (RMSE) for this set was found to be 3.71.

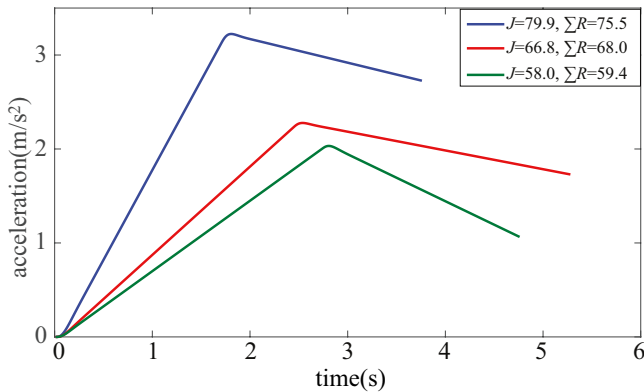


Fig. 7. Three samples of the training data

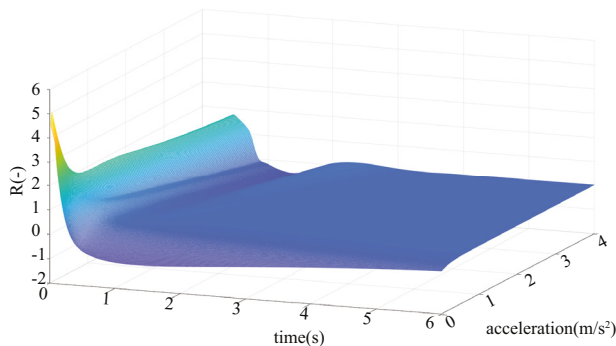


Fig. 8. The obtained reward function by IRL

The result of IRL is shown in Fig. 8, where the value of $R(a, k)$ is sketched. Observe that the performance during

Table 2. Physical parameter of powertrain

symbol	value	symbol	value
I_e	0.22(kg·m ²)	I_{m1}	0.06 (kg·m ²)
I_1	2.975×10^3	I_2	17.52
C_1	41.42	C_2	7.38
c_{m1}	0.05	c_{m2}	0.05
G_m	5.75	G_f	3.307
R_w	0.39(m)		

the initial raising period is mostly focused on evaluating the score. However, it should be noted that this reward function is nothing but an approximate representation with ELM of the evaluation score, hence, has no physical meaning. This implies that searching for the most comfortable acceleration profile $a^*(t)$ using this reward function or directly using this reward function as a cost function in optimizing the torque commands with the powertrain dynamics might not equate to the highest score. Moreover, there is no guarantee of the convexity of the extremum seeking or optimal control problem. Indeed, at the second step, the score of the obtained optimal acceleration profile is 71.4, with the parameters given by

$$a_1 = a_2 = 2.60, \quad t_1 = 2.5, \quad t_2 = 4.5.$$

An interpolation curve of the signal is shown in Fig. 4. Using this signal as the desired acceleration reference $a^*(t)$, the optimal control input u of the powertrain, i.e., the torque command for the electric motors, is obtained by solving the optimal problem formulated in Section 3.3 using a dynamic programming algorithm. The physical parameters of the powertrain are shown in Table 2, and the constraint conditions are defined as follows:

$$v_{min} = 70(\text{km/h}), \quad v_{max} = 100(\text{km/h})$$

Also, the clutch is engaged at $t_0 = 0.47[\text{sec}]$ with $v(t_0) = 70.76(\text{km/h})$, $\tau_e(t_0) = 302.1(\text{Nm})$, and $\tau_{m2}(t_0) = 44.03(\text{Nm})$. The optimal control problem is solved for the period with $t_f = 4.5(\text{sec})$. Fig. 9 shows the optimal torque commands. It is seen that motor MG1 initially provides motoring torque to start the engine and MG2 provides assistant torque since the engine is not yet switched on. However, once the engine torque acts on the shaft, MG2 turns to reject the total torque by regenerative breaking. The actual acceleration response is as shown in Fig. 10, where an overshoot appears when the engine and the motors are switched on or off. The response of the optimized control is reasonable and the score calculated with the sum of the reward function is 72.4. The score given by the expert model is 69.3. The gap between these scores is caused by learning error.

5. CONCLUSION

For HEVs, especially in acceleration mode, the torque of the motor and the timing of gear shifting should be managed with consideration given to comfortability. When the combustion engine is operated at full throttle, if the vehicle begins accelerating in EV mode, the desired acceleration torque is too large. In this paper, a control design approach that gives rise to a motor assistant control law achieving comfortable acceleration behavior is proposed. The key contribution of this design approach is the use of inverse reinforcement learning for evaluating comfortability. It is shown that using IRL enables us to design a control law by optimizing an analytically formulated cost function.

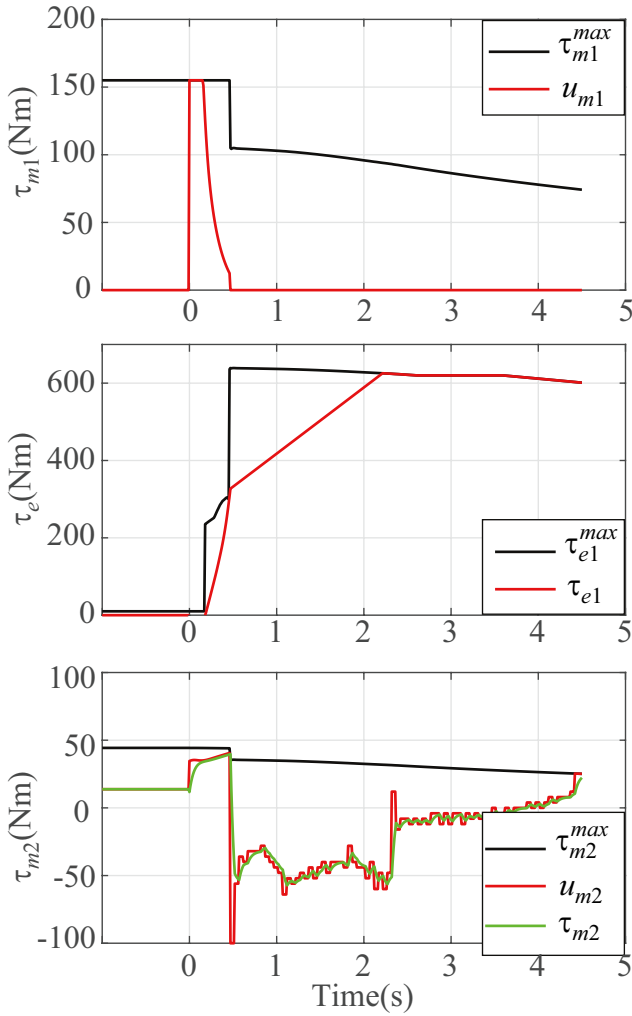


Fig. 9. Optimal torque commands

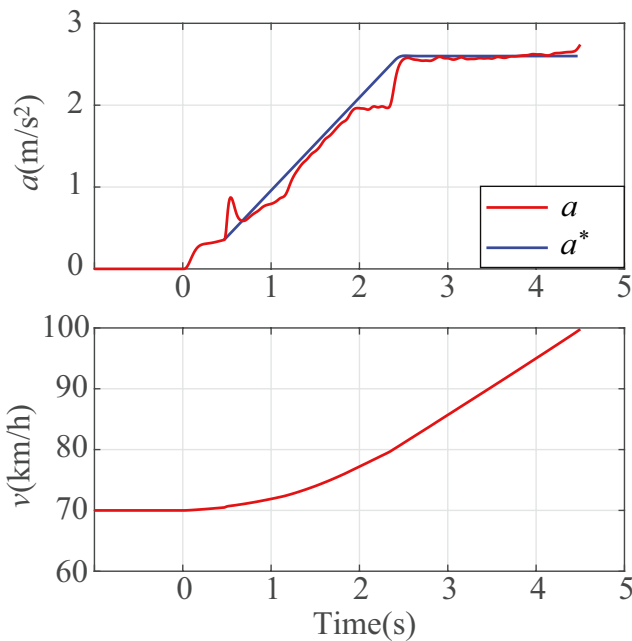


Fig. 10. Outputs of acceleration and vehicle speed

It should be noted that evaluation of comfortability is not an easy task. This is why the subject is challengeable when the comfortability is taken into power control account. Most existing approach such as shown in Shen (2020) the comfortability is described as a specification of the acceleration profile which enables us to directly apply the dynamical control method. This paper focuses on the human feeling which is eliminated as a non-analytic module. An inverse reinforcement learning approach is exploited to translate the comfortability feeling to an acceleration profile, then dynamical optimal control design is followed to obtain a comfortable acceleration profile.

REFERENCES

- H. Lee, C. Song, N. Kim, et al. Comparative analysis of energy management strategies for HEV: dynamic programming and reinforcement learning. *IEEE Access*, 8: 67112-67123, 2020.
- X. Hu, T. Liu, X. Qi, et al. Reinforcement learning for hybrid and plug-in hybrid electric vehicle energy management: recent advances and prospects. *IEEE Industrial Electronics Magazine*, 13(3): 16–25, 2019.
- Q. Jiang et al., Comparative study of real-time HEV energy management strategies. *IEEE Trans. on Vehicular Technology*, 66(12):10875–10888, 2017.
- F. Zhang, X. Hu, R. Langari, et al. Adaptive energy management in automated hybrid electric vehicles with flexible torque request. *Energy*, 214: 118873, 2021.
- F. Xu and T. Shen. Decentralized optimal merging control with optimization of energy consumption for connected hybrid electric vehicles. *IEEE Trans. on Intelligent Transportation Systems*, DOI:10.1109/TITS.2021.3054903, 2021.
- J. Zhang, T. Shen and J. Kako. Short-term optimal energy management of power-split hybrid electric vehicles under velocity tracking control. *IEEE Transactions on Vehicular Technology*, 69(1): 182–193, 2020
- X. Shen, X. Zhang, T. Ouyang, et al. Cooperative Comfortable-driving at signalized intersections for connected and automated vehicles. *IEEE Robotics and Automation Letters*, 5(4): 6247–6254, 2020.
- M. Zhu, Y. Wang, Z. Pu, et al. Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving. *Transportation Research Part C: Emerging Technologies*, 117:102662, 2020.
- B. Burchfiel, C. Tomasi, and R. Parr, et al. Distance Minimization for Reward Learning from Scored Trajectories. *AAAI Conf. on Artificial Intelligence*, 30(1), 2016.
- A. Y. Ng and S. Russell, et al. Algorithms for Inverse Reinforcement Learning. *the 17th International Conference on Machine Learning*, 663-670, 2000.
- P. Abbeel and A. Y. Ng, et al. Apprenticeship Learning via Inverse Reinforcement Learning. *the 21st International Conference on Machine Learning*, 1-8, 2004.
- Guang-Bin Huang and Qin-Yu Zhu and Chee-Kheong Siew Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. *Neurocomputing*, vol. 70, pp. 489-501 2006.
- P Larrañaga and R Etxeberria and J A Lozano and J M Peña et al. Optimization by learning and simulation of Bayesian and Gaussian networks. *Servicio de Publicaciones de la Facultad de Informática, UPV-EHU. EHU-KZAA-1K-4-99*, 1999.