International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: 260

VISION-AIDED DEEP REINFORCEMENT LEARNING FOR ENERGY MANAGEMENT OF HYBRID ELECTRIC VEHICLES

Yong Wang[1], Yuankai Wu[1*], Jiankun Peng[1*], Huachun Tan[2], Dechong Zeng[3], Hongwen He[1]
1 School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China,
2 School of Transportation, Southeast University, Nanjing 211102, China,
3 HanKaiSi Intelligent Technology Company, Guiyang 550081, China,

ABSTRACT

This paper introduces an energy management strategy that combines visual perception and deep reinforcement learning (DRL) algorithms to minimize fuel consumption. The proposed method is capable of autonomously learning the optimal control policy without any prediction efforts. We used a monocular camera in the windshield of a car to catch visual information as inputs. Next, we used state-of-the-art convolutional neural networks based object detection methods to detect and classify traffic light. The traffic light information is used as a state input for a model-free deep reinforcement learning based energy management system with continuous control action. Hence, the traffic light information is incorporated into the energy management system. The experimental results indicate that the fuel economy of the proposed vision-aided strategy achieves 94.5% of dynamic programming-based method's, and is 6.8% better than that of the original DRL algorithm without traffic light information under a real-world driving cycle.

Keywords: hybrid electric vehicle, energy management strategy, visual perception, deep reinforcement learning, traffic light

NONMENCLATURE

Abbreviations	
EMS	Energy management strategy
HEV	Hybrid electric vehicle
DP	Dynamic programming
SOC	State of Charge
DRL	Deep reinforcement learning
CNN	Convolutional neural network
DDPG	Deep deterministic policy gradients
MDP	Markov decision process

Symbols	
s_t, a_t	State, Action
T	Time
R	Reward

1. INTRODUCTION

Cleaner, more efficient cars can make a big difference to our society in terms of environmental benefits. Hybrid electric vehicles (HEVs) are fuel efficient, able to overcome range anxiety and friendly to the environment. Additionally, automobiles are beginning to be equipped with more and more on-board sensors as vehicles become more intelligent. These sensors provide Internet connectivity, vehicle condition monitoring, higher efficiency and assistance for drivers to enhance both road safety and travel comfort. With developments of vehicle electrification and intelligence, it is possible to improve fuel economy of HEVs by using intelligent sensing and control algorithms.

As one of key technologies in HEVs, energy management is critical to achieve higher fuel efficiency, which has been studied extensively. Many of the existing energy management strategies (EMSs) require predicted information prior to the trip, which is of utmost importance for many existing EMS in HEVs [1]. Another necessary requirement for an optimal EMS is the currently available trip information, which can be acquired through existing instrumentation installed on-board. For almost a decade, various types of on-board perception sensors, such as radar sensors, cameras and ultrasonic sensors, have been widely used on production vehicles. Among these sensors, camera is the most popular sensor due to its low-cost and capability to capture rich visual information.

Meanwhile, recent developments in deep learning approaches have greatly advanced the visual processing technologies, which includes lane detection, object detection, drivable road segments, traffic lights detection and so on.

In recent years, DRL-based architecture has been applied to learn the optimal control strategy based on the driving data [3]. He et al. [4] explored a deep Q learning based energy management strategy, and demonstrated that the fuel economy of the DRL-based strategy is close to the global optimum. However, these DRL-based approaches tend to assume that the trip information or the environment is merely related to the current speed and acceleration. The variations in trip characteristics and prevailing traffic conditions are always ignored. However, the trip information is not only dependent on current vehicle states, but also related to various factors such as driving styles, route choices and traffic environments. A recent study [5] suggested that the incorporation of traffic information is able to improve the performance of HEVs energy management systems.

In this paper, the traffic information collected from on-board visual perception sensors are incorporated into the energy management system. We developed a novel energy management system that combined optimizing perception and decision-making algorithms. The image information collected from the camera and vehicle states collected from on-board sensors are used as inputs for the EMS. The EMS contains a CNN based image processing module to collect visual information and a DRL based control model to split power. We evaluate the performance of the proposed system on a Toyota Prius II HEV. In order to make the proposed methodology more understandable, a simple schematic overview of our work is provided in Fig 1.

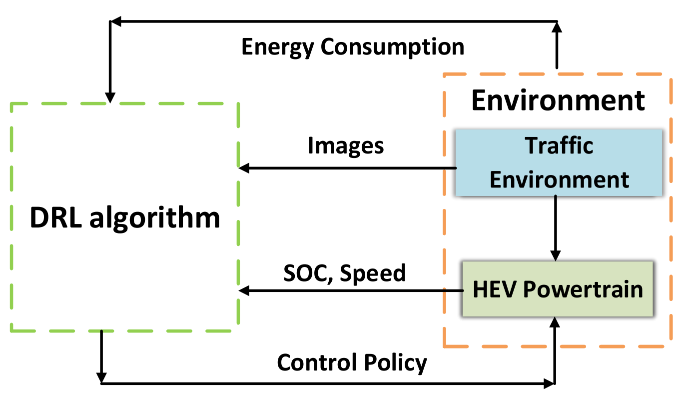

Fig 1 The interaction between the DRL algorithm and environment.

2. DEEP REINFORCEMENT LEARNING FOR ENERGY MANAGEMENT

2.1 Deep Reinforcement Learning

In this paper, the energy management of a HEV is formulated as an Markov decision process (MDP) problem. The solution of an MDP problem is a policy π which is a distribution of actions given states that maximizes the expected discounted sum of the rewards. The action value function, denoted as $Q_\pi(s,a)$, is defined as the expected sum of future rewards for selecting action a in state s and following policy. The action function can be expressed by the Bellman Equation as follows:

$$Q_\pi(s,a) = \mathbb{E}[R_{t+1} + \gamma Q_\pi(s_{t+1}, a_{t+1}) | s_t, a_t] \quad (1)$$

where s_{t+1}, a_{t+1} and R_{t+1} are the state, action and reward of next time step, and $\gamma \in [0,1]$ is the discount factor. At each time step t, the agent receives some representation of the environment's state s_t, and selects an action a_t according to the state signal and policy π. In this paper, we selected a simple continuous action domain model-free reinforcement learning algorithm: deep deterministic policy gradients (DDPG) [6], to show that it is capable of solving the MDP.

2.2 DDPG-based Strategy

The MDP problem for HEV energy management is resolved by DDPG. The exact definition of the state space S, action space A and reward function R are defined as following.

State space: the state space is the definition of the observations that the algorithm receives at each time step. In this paper, we use the image of the current observation, together with the observed vehicle speed, SOC and acceleration are as state variables.

Action space: we use output power of engine as action variable. To simplify the model of power-split HEV in DDPG and reduce the amount of action variables, we set that engine works in a specified area.

Reward function: in order to guarantee the regular iteration of the network, the deep reinforcement learning single-step reward function R needs to be defined, reward function is one of the key factors in determining the performance of the DDPG. Traditionally, reward function is defined as the energy cost at each step:

$$R_1(s,a) = \int_0^T cost\, dt \quad (2)$$

where cost represents the engine fuel consumption, $t \in [0,T]$ is the specific time horizon. We want SOC to keep in a range and to make the performance of SOC stable. The reward function of our work is defined as:

$$R_2(s,a) = \int_0^T [\text{cost} + \beta(\text{SOC}(t) - \text{SOC}_r)^2]dt \qquad (3)$$

where β is a positive weighting factor and SOC_r is a preassigned constant to maintain charge-sustaining constraints. Therefore, both low fuel consumption and SOC stability are expected to be achieved by defining the reward function.

$$R(s,a) = R_1(s,a) + R_2(s,a) \qquad (4)$$

3. VISION-AIDED ENERGY MANAGEMENT STRATEGY

3.1 Data Collection

For getting a more practical evaluating evaluation of the proposed DRL model, we conducted our experiments in a real-world environment from Chinese roads instead of a standard driving cycle. Data are collected from endemic dataset archived from roads in Guiyang downtown. The data are collected from Changling South Road to Chinese Academy of Sciences Guizhou Technology Innovation Park and the route is given in Fig 2. The collected data include speed and image collected from the camera. Fig 3 shows the on-board computer vision system. We used a monocular camera in the windshield of the car to collect visual inputs, and the vehicle speed is recorded via the CAN bus. The collected images are downsampled into 80 × 160 pixel and 1Hz.

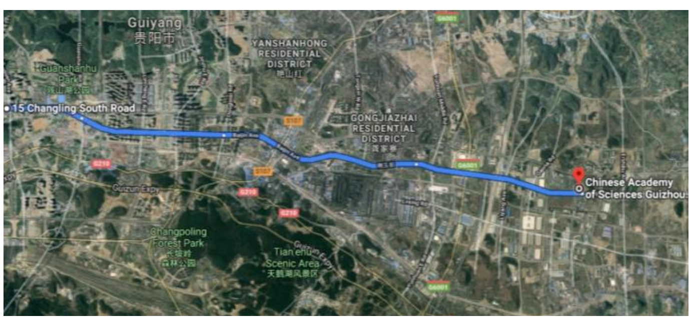

Fig 2 Example trip from Google Map.

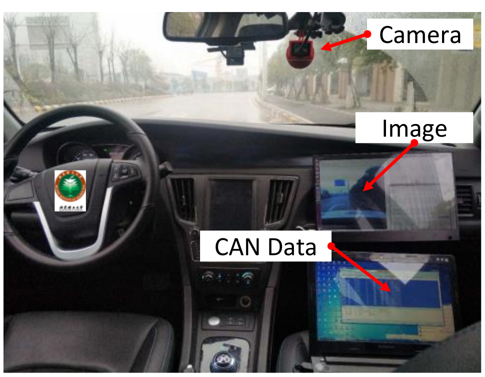

Fig 3 The structure of the real equipment.

3.2 Image Processing

The raw image could be fed directly into the DRL agent without any mediated perception. But it is very inefficient for training neural network models. Driving cycle is constrained by the traffic. The image covers rich traffic information, we focus on a special case: the detection and classification of traffic signals on urban driving cycle. Because traffic signals provide important information for driving, and a wide variety of CNN-based approaches have been developed to detect and recognize traffic lights. In this paper we apply the state-of-the-art, real-time object detection system You Only Look Once (YOLO) [6], for traffic light detection and classification. Detection results from the YOLO V3 detector are shown in Fig 4.

Fig 4 Results from the YOLO V3 model

4. EXPERIMENTS

4.1 Experiment Setup

We build Deep Reinforcement Learning model by TensorFlow, which is an open source deep learning platform. Meanwhile, A GeForce GTX 2080Ti GPUs are used to assist accelerating training phase, and the hyper parameters of the DDPG agent are listed in Table 1.

Table 1. DRL-based algorithm hyper parameters.

Hyper Parameters	Value
Learning rate for Actor	0.001
Learning rate for Critic	0.001
Gamma discount	0.99
Batch size	64
Replay memory size	10000
Initial exploration	1.5

4.2 Results

The collected velocity trajectory is showed in Fig 5. We use CNN based image processing technique to classify the state of the traffic lights. If the lights are green, it outputs signal 1; If the lights are red, it outputs signal 2. Fig 6 provides the state of the traffic lights.

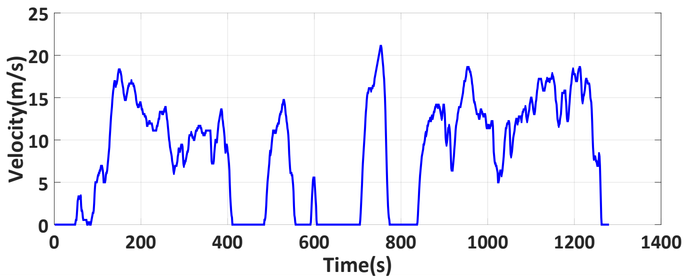

Fig 5 Synthesized velocity profile of the example trip.

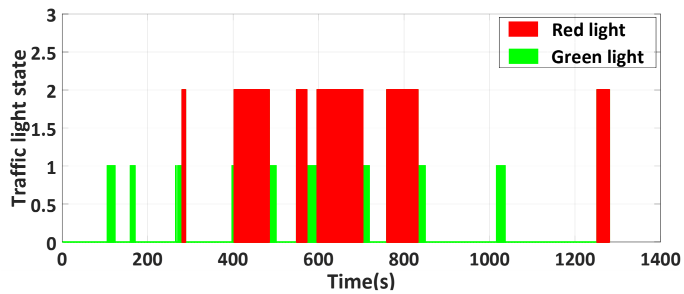

Fig 6 The states of the traffic lights in the example trip.

DP based HEV power Management is used as the benchmark to validate the effectiveness of the vision-aided DDPG method. In order to ensure the fairness of the comparison, a same hyper parameter setting is applied in all the three methods and the initial SOC is set to 0.62. The SOC curves corresponding to different control policies are shown in Fig 7.

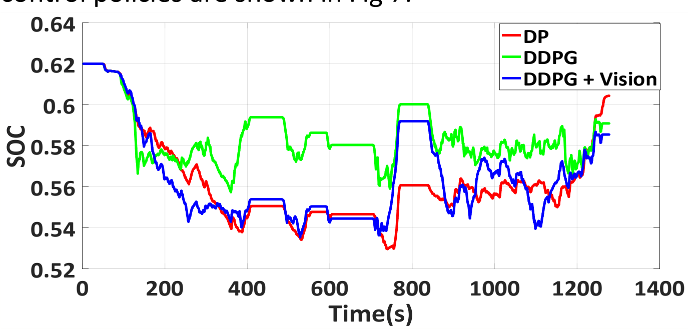

Fig 7. SOC curves for different control policies in the trip.

The fuel consumption of DP, DDPG and DDPG+Vision can be seen in Table 2. The Vision-aided DDPG strategy can achieve 94.5% fuel optimality of the DP benchmark in real driving cycle and is 6.8% better than original DDPG regarding the fuel optimality. The improvement is achieved by the incorporation of more environment information.

Table 2. Fuel economy comparison under the trip.

Control strategies	Fuel consumption(L)	Final SOC	Relative ratio to DP (%)
DP	0.328	0.60	100
DDPG	0.374	0.59	87.7
DDPG + Vision	0.347	0.59	94.5

5. CONCLUSION

In this paper, a system that combines visual perception and deep reinforcement learning algorithms is performed in order to minimize fuel consumption of a HEV. The results show that traffic lights information collected from on-board camera can improve the fuel economy of HEV. This is because the image provides important traffic information. In this paper, we only considered the traffic light information collected from image data. A future direction is to exploit more information from visual inputs to improve EMS. Additionally, the research on application of other sensors such as radar, lidar, ultrasonic and Global Positioning System (GPS) for EMS will be a valuable future work.

ACKNOWLEDGEMENT

This work was supported by National Key R&D Program of China (Grand No.2018YFB0105900), National Natural Science Foundation of China (Grant No.51705020 & No.61620106002) and the Project funded by China Postdoctoral Science Foundation (Grant No. 2016M600933). The authors would also like to thank Dr. Qin Li and Dr. Hailong Zhang for their corrections and helpful suggestions.

REFERENCE

[1] Martínez CM, Hu X, Cao D, et al. Energy management in plug-in hybrid electric vehicles: recent progress and a connected vehicles perspective. IEEE Transactions on Vehicular Technology 2017;66:4534–4549.

[2] X Qi, Y Luo, G Wu, K Boriboonsomsin, M Barth. Deep reinforcement learning enabled self-learning control for energy efficient driving. Transp. Res. Part C 2019; 99:67-81.

[3] Wu J, He H, Peng J, Li Y, Li Z. Continuous reinforcement learning of energy management with deep q network for a power split hybrid electric bus. Applied Energy 2018; 222:799–811.

[4] Wu Y, Tan H, Peng J, H Zhang, He H. Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus. Applied Energy 2019; 247:454–466.

[5] Silver D, Lever G, Heess N, Degris T, Wierstra D, Riedmiller M. Deterministic policy gradient algorithms. ICML. 2014.

[6] Redmon J, Farhadi A. YOLOv3: An incremental improvement, 2018, arXiv preprint arXiv: 1804.02767.