Application of Artificial Intelligence (AI) Technologies to Driving Robots

Kento Yoshida,
Hironobu Fukai,
Makoto Niwakawa,
Masayasu Kanke

Keywords Driving robot control, AI technology, Deep learning, Deep reinforcement learning

Abstract

We are developing driving robots to be used as automatic operation equipment for automobile testing. At present, there is currently a pressing demand for better perfect control performance of driving robots. In order to realize improved target speeds and a human-simulated pedal operation, we are working on the introduction of a control system based on Artificial Intelligence (AI) technologies.

According to the prediction of feedforward manipulating variables with accelerator pedals through supervised learning, the ability to follow a target speed was found to be higher than that of conventional control methods when knowledge about vehicle dynamics including transient status was obtained through learning the driving history data. In addition, by substituting the control method with reinforcement learning, we could realize not only the ability to follow the target at a high speed, but also a predictive pedal selection.

1 Preface

We manufacture and sell a chassis dynamometer test system to reproduce road load driving at the test bench. The driving robot is equipment for automatic operation used to examine the operation of the accelerator pedal, brake pedal, clutch, and power transmission of a completed vehicle on the chassis dynamometer. Fig. 1 shows a driving robot. Compared with operation by a human test driver, operation performed by a driving robot is superior in terms of extended endurance driving and reproducibility of driving. There are a variety of tests for evaluating vehicle performance. In the case of exhaust gas emission and fuel consumption tests, vehicle performance is measured during follow-up driving ("driving in operating mode" hereafter) based on a vehicle speed pattern defined by the relevant standard ("mode" hereafter). Fig. 2 shows the WLTC (Worldwide-harmonized Light vehicles Test Cycle) mode. The mode used for testing is determined for specific countries and regions. The WLTC mode has been actively introduced as the international standard. The ability to follow vehicle speed patterns is demanded in the market as the basic functional requirements of the driving robot is

Fig. 1 Driving Robot

The driving robot is loaded on the driver's seat of a completed car so that the robot can manipulate the accelerator, brake pedals, and other devices.

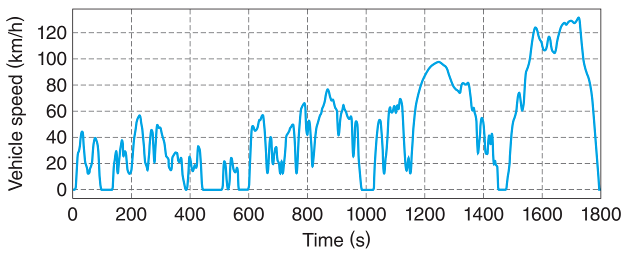

Fig. 2 WLTC Mode

The axis of abscissas is used to express time and coordinates indicating vehicle speed. During testing, driving conforming to such a vehicle speed pattern is required.

ever-increasing. This paper introduces some examples of Artificial Intelligence (AI) technologies applied to the controls for a driving robot. This is to achieve better following for driving at the high level attained only by a professional test driver or realize human-simulated driving performance.

2 AI Technologies

2.1 Background of AI Technologies

AI is a technology to substitute a computer's activities for a human's intellectual activities. The first concept of this technology was born in the 1950s. Fig. 3 shows the AI technological system. Since the early days of 1950s, approaches of reasoning, investigation, and knowledge representation was the mainstream. The current AI boom, however, is brought by a machine learning technology of a neural network, particularly by deep learning. Neural network was known to exist since it's inception, but is becoming recognized with the growth of computers based on the recognition of its high learning capability.

Regarding learning method and data handling, machine learning is roughly classified into supervised learning, unsupervised learning, and reinforcement learning. Supervised learning is a learning method to obtain a rule between input and desired output (teacher) data. Reinforcement learning is a learning method that acquires an action decision policy to realize a desirable state in a series of actions.

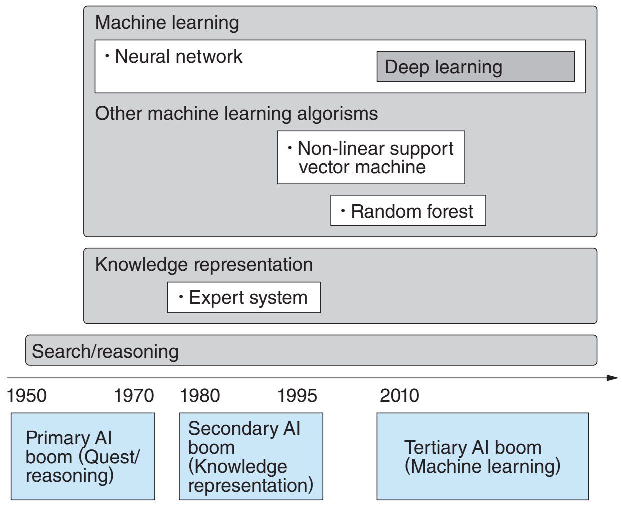

Fig. 3 AI Technological System

Trending technologies vary year after year. The basic technologies currently available, however, can be found in those of early years of AI development.

Deep learning has not only been used as supervised learning, but has also been proposed for use as a superb function approximator in combination with the framework of reinforcement learning. Recently, this deep reinforcement learning has been increasingly noticed in the AI field. In the following section, we introduce our efforts in applying supervised learning and reinforcement learning by deep learning to the controls of driving robots.

2.2 Supervised Learning

In supervised learning, the machine learning model is trained so that the predicted output for certain input data comes close to the teacher data. By using a trained model, it is possible to obtain a predicted output as a response to the new input data.

In order to carry out supervised learning, it is necessary to prepare a set of input and teacher data. For an application to the controls of driving robots, a practical vehicle running data is used as leaning data to obtain the vehicle dynamics between the target speed and accelerator pedal operation.

2.3 Reinforcement Learning

Fig. 4 shows the basic process of reinforcement learning. In the reinforcement learning system, there are two prominent factors, "Environment" and "Agent." The Agent determines an action based on the state of the environment and executes that action in the environment. The state of the environment changes according to the action of the agent, and the agent observes the new state and determines the next action. Simultaneously, when the status has been observed, the agent receives a "reward." This reward is expressed by a scholar value that evaluates the state after an action. In reinforcement learning, an action policy is trained so that more rewards can be obtained as a result of a series of actions chosen by the policy. When

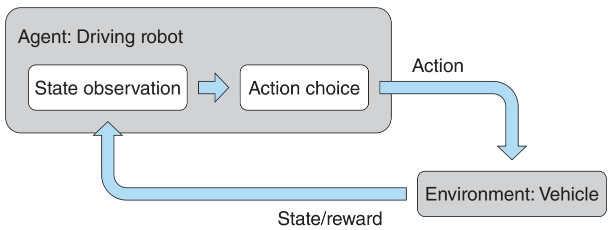

Fig. 4 Basic Process of Reinforcement Learning

A proper way to take reasonable actions is obtained by repeating experiences.

learning of the control done by the driving robot, the agent is a driving robot and the environment is a vehicle. In this case, a reward is designed such that a larger value can be obtained when the deviation between detected vehicle speed and the target speed becomes smaller. For the reinforcement learning, it is possible to provide an obvious training objective, like "human-simulated driving."

3 Application of AI Technologies to Control

3.1 Application of Supervised Learning

For one of the AI technology applications to the control of a driving robot, we tried to make a substitution of driving force map in the neural network.

Fig. 5 shows a schematic diagram of a conventional control system by driving robots. In the conventional control of a driving robot, a control method, which is a combination of the Feedforward (FF) system called the driving force map and the FB system according to the vehicle speed deviation, is used to determine the accelerator pedal manipulating value for driving in the operating mode. The driving force map specifies the relationship between the required driving force and the accelerator pedal value to realize the target speed of the mode. This map is established by recording the steady vehicle characteristics before the test. When the target speed and required driving force are given to this driving force map, the accelerator pedal manipulating variable is obtained. Since transient vehicle dynamics are not reflected on the establishment of the driving force map, an error may appear in the accelerator pedal manipulating variable that is required during actual driving. As a solution, the FB system generates an output compensation variable for accelerator pedal manipulation in order to compensate for the vehicle speed deviation.

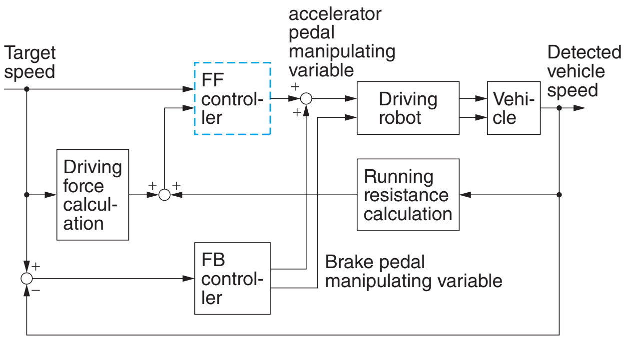

Fig. 5 Schematic Diagram of a Conventional Control System by Driving Robots

The accelerator pedal control consists of a combination of the FF system (driving force map) and the FB system.

During our development activities, we substituted a neural network for the driving force map that in a part of a whole control system. Practically, the FF system indicated by dotted lines in Fig. 5 was replaced by the neural network. If the vehicle dynamics are learned encompassing the transient conditions during driving, improved ability to follow for driving in the operating mode can be expected as a result of attaining a high-precision prediction for the accelerator pedal manipulating value.

To evaluate deep learning control, we carried out a simulation. For the neural network, actual driving data based on the driving force map was used as training data and various inputs such as the detected vehicle speed, required driving force, number of engine rotations, and target speed for the future constant time were entered to train the required output for the accelerator pedal manipulating value. For upper-grade manipulation such as accelerator pedal/brake pedal changeover and brake pedal manipulation, the same method as for conventional control was used.

Table 1 shows the driving result in WLTC mode by respective control methods. For the driving in the operating mode test, a permissible error range of ±1.0 s and ±2.0 km/h is stipulated for a com-

Table 1 Driving Result in WLTC Mode by Respective Control Methods

The result of driving for 1800 seconds in the WLTC mode is shown. Regarding the ability to follow the target speed, the result is verified acceptable when no foul is recorded, and the mean and maximum vehicle speed errors are minimal. Frequency of pedal operation is regarded as one of the indexes for human-simulated driving.

	Foul time (s)	Mean vehicle speed error (km/h)	Max. vehicle speed error (km/h)	Accelerator pedal operating frequency (times)	Brake pedal operating frequency (times)
Conventional control	0.00	0.44	4.27	38	41
Deep learning control FF + FB	0.00	0.28	2.48	42	43
Deep reinforcement learning control FF	0.00	0.37	2.56	51	52
Deep reinforcement learning control FF + FB	0.00	0.30	1.72	48	49

manded vehicle velocity at a specific time. In this case, the foul time is defined as a total deviation time from the permissible error range. Even in the case of a conventional control system, the foul time was 0 s, but both the mean and maximum vehicle speed errors were smaller than the values of the conventional control when a deep learning control approach was applied.

Fig. 6 shows a partial section driving result in the WLTC mode. Some improvements can be perceived in the ability to follow where vehicle speed deviation was substantial in the case of a conventional control. Generally, the dependence on the FB system for accelerator pedal operation seemed to be small. Consequently, the ability to follow the target speed is improved because the prediction accuracy of the FF system has been improved by virtue of the deep learning control.

According to Table 1, the number of pedal operation times is increased compared with the conventional control. An increase in the number of pedal operation times implies that a changeover between accelerator pedal and brake pedal tends to be frequent. Although the ability to follow and pedal changeover times can be a trade-off, the unnecessary increase in the number of pedal operation times results in departing from human-simulated driving. According to waveforms of accelerator pedal manipulating values in Fig. 6, variation in pedal operation is large in the case of deep learning control. A sudden change in pedal operation can also result in departing from human-simulated driving. In the case of supervised learning, the control performance obtained by learning tends to be dependent on the nature of the training data. Learning is, therefore, considered indispensable to be based on humanly-driving achievement data to realize human-simulated driving.

3.2 Application of Reinforcement Learning

In our activities for the application of AI technologies to the driving robot control, we actively utilized deep reinforcement learning. In this way, we tried to use learning for optimal selection and the manipulating method of the accelerator and brake pedals in the task of driving in the operating mode.

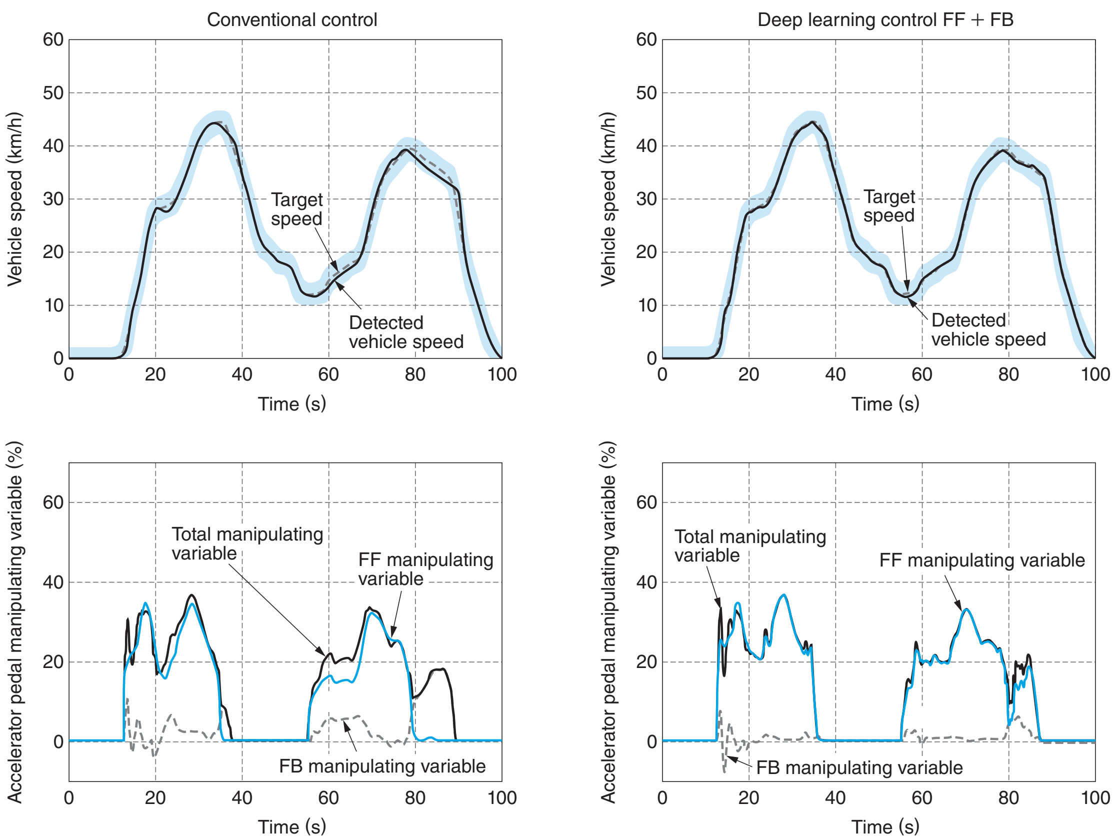

Fig. 6 Partial Section Driving Result in WLTC Mode

The upper row shows vehicle speed, lower row shows accelerator pedal manipulating value, left shows conventional control, and right shows deep learning control. The shaded vehicle speed waveform shows the permissible error range.

In the application of deep learning control as supervised learning, partial substitution of machine learning was adopted based on conventional control methods. In this case, however, the control method was drastically modified. For a conventional control method, manual work by engineers was needed, such as parameter adjustment for each vehicle and control designing. In the case of reinforcement learning system, however, the characteristics per vehicle were recognized in the common learning processes and even subsidiary control rules, never devised in conventional controls, became obtainable.

For the evaluation of deep reinforcement learning control, we conducted a simulation test. In the algorism of deep reinforcement learning, a processing from state observation to the action choice is shown in **Fig. 4**, is attempted by using the neural network. The observed states are, an accelerator pedal detection variable accumulated during the constant period in the past, a brake pedal detection variable, number of engine rotations, the detected vehicle speed, target speed, and target speed for a constant future time. These data are inputted in the neural network where the accelerator and brake pedal manipulating values are outputted. The "reward" that is the basis of status evaluation for deep reinforcement learning is defined be the standard itemized below.

(1) Larger when the ability to follow the target speed is higher.

(2) Larger when pedal manipulation is smoother.

(3) Larger when the pedal changeovers are less frequent.

As shown in **Table 1**, even deep reinforcement learning control could realize driving in the operating mode with a zero-second foul time irrespective of whether the FB system was present or not. Both the mean and maximum vehicle speed errors were smaller than those of the conventional control and a high ability to follow performance to the target speed was obtained. When combined with the FB system, a higher ability to follow performance than the result of driving based on supervised learning was attained. **Fig. 7** shows partial section driving result in the WLTC mode by the deep reinforcement learning

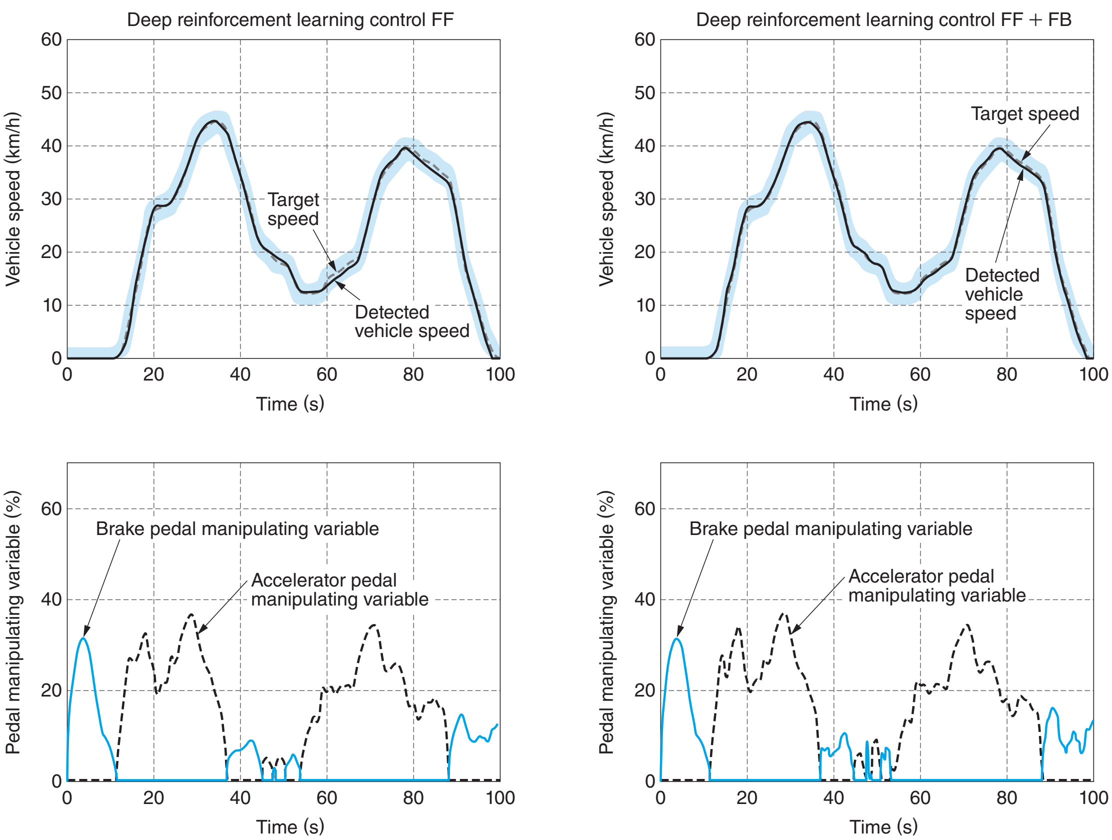

Fig. 7 **Partial Section Driving Result in WLTC Mode by Deep Reinforcement Learning Control**

The upper row shows vehicle speed, lower row shows pedal manipulating variable, left shows deep reinforcement learning control FF, and right shows deep reinforcement learning control FF + FB. The shaded vehicle speed waveform shows the permissible error range.

ing control. In the case of the conventional control as shown in **Fig. 6**, the FB system becomes too sensitive during acceleration around the time of 10 seconds, thus making the accelerator pedal manipulation too violent and causing overshooting as a result. In the case of deep reinforcement learning control, on the other hand, a target speed seems to be followed up as a result of smooth treading on the accelerator pedal. Judging from these factors, any pedal to be manipulated can be properly predicted and the feasibility of effective application to the control system is suggested, although there is still a margin of improvement for the prediction of pedal manipulating values.

Meanwhile, compared with the result of deep learning controls supervised learning, the frequency of pedal manipulation was found to be higher. As shown in **Fig. 7**, brake pedal manipulation for about 0.5 seconds can be seen around the time lapse of 48 seconds. In this way, short-time pedal changeovers can be seen several times. When the FB system is combined, smooth pedal manipulation tends to be degraded. In order to realize human-simulated driving, we will focus our future challenges on reviewing the reward design and learning method.

4 Postscript

This paper introduced two kinds of challenges regarding the application of AI technologies to driving robot controls. According to the prediction of the accelerator pedal manipulating values by deep supervised learning, vehicle dynamics including transient conditions can be obtained through actual driving data learning and the ability to follow the target speed can be obtained exceeding the achievement of the conventional control method. In the case of control by deep reinforcement learning, not only the ability to follow the target speed but also adequate pedal selection is obtained to follow up the target speed.

Together with further improvement of the ability to follow, our future challenges are the reflection of human-simulated driving such as the reduction of pedal changeover frequency. We will work on examining the usability issue including the learning method needed at the time of real testing.