

Robust Performance Improvement of Automotive Powertrain Control Systems Using Deep Reinforcement Learning Considering Parameter Variations

Hiroki Shibata¹⁾ Heisei Yonezawa¹⁾ Shota Sato²⁾ Hiroya Kikuchi²⁾ Takashi Hatano²⁾
Shuichi Kondo²⁾ Shigeki Hiramatsu²⁾ Itsuro Kajiwara¹⁾

*1) Graduate School of Engineering, Hokkaido University
Kita 13, Nishi 8, Kita-ku, Sapporo, Hokkaido 060-8628, Japan*

*2) Mazda Motor Corporation
3-1 Shinchu, Fuchu-cho, Aki-gun, Hiroshima 730-8670, Japan*

KEY WORDS: Electronics and control, Control simulation, Control system software [E1]

Automotive powertrain systems contain nonlinear elements such as gear backlash and friction, which can induce vibration during power transmission and degrade control performance. Conventional model-based (MB) controllers require accurate modeling of these nonlinear characteristics, which can complicate controller design. In contrast, reinforcement learning (RL) can acquire nonlinear control laws through interaction with the controlled system. However, RL controllers are generally trained in simulation environments, and their performance may degrade owing to the sim-to-real gap caused by discrepancies between simulation models and actual systems. Therefore, acquiring a control policy that is robust against parameter variations is important for practical implementation. In this study, domain randomization (DR) and its extended method, curriculum domain randomization (CDR), were introduced into a Twin Delayed Deep Deterministic Policy Gradient (TD3) controller for a simplified powertrain model that explicitly considers gear backlash. Their effectiveness in improving robustness against parameter variations was investigated.

The controlled system was represented as a mass–spring–damper model with gear backlash modeled as a dead zone. In DR, the vehicle body mass, actuator mass, backlash length, and reference displacement were randomized at the beginning of each training episode. In contrast, CDR gradually expanded the randomization range as training progressed. The conceptual difference between DR and CDR is illustrated in Fig. 1.

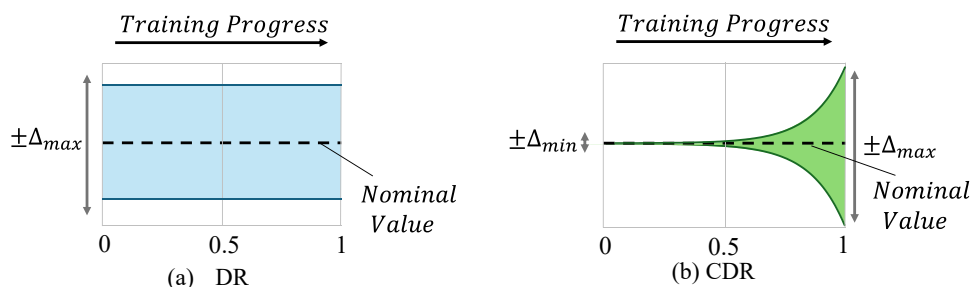


Fig. 1 Conceptual comparison of DR and CDR

To evaluate the robustness of the trained controllers, four control methods were compared: a conventional TD3 controller trained using only the nominal model, TD3-DR incorporating DR, TD3-CDR incorporating CDR, and an MB controller. The vehicle body mass, actuator mass, backlash length, and reference displacement were varied within prescribed ranges, and 200 evaluation models were generated using Latin hypercube sampling. For each model, the 2-norm of the response error between the reference value and vehicle body displacement was calculated. Robust performance was evaluated in terms of the mean, standard deviation, and worst-case value of the response error. The evaluation results are summarized in Table 1.

Table 1 Statistical comparison of control methods

Method	TD3	TD3-DR	TD3-CDR	MB
Mean	0.757	0.834	0.753	0.851
Std	0.180	0.143	0.164	0.159
Worst	1.250	1.250	1.268	1.272

TD3-DR and TD3-CDR reduced the standard deviation of the response error compared with conventional TD3. In particular, TD3-CDR achieved a mean response error comparable to that of conventional TD3 while suppressing performance variation under parameter uncertainty.