Available online at www.eccomasproceedia.org
Eccomas Proceedia SMART (2023) 519-525

ECCOMAS Proceedia

X ECCOMAS Thematic Conference on Smart Structures and Materials
SMART 2023
D.A. Saravanos, A. Benjeddou, N. Chrysochoidis and T. Theodosiou (Eds)

NOVEL ACTIVE STRUCTURAL VIBRATION CONTROL STRATEGY BASED ON DEEP REINFORCEMENT LEARNING

YI-ANG ZHANG*, SONGYE ZHU†

* Department of Civil and Environmental Engineering
The Hong Kong Polytechnic University
Kowloon, Hong Kong, China
e-mail: yi-ang.zhang@connect.polyu.hk

† Department of Civil and Environmental Engineering
The Hong Kong Polytechnic University
Kowloon, Hong Kong, China
e-mail: ceszhu@polyu.edu.hk

Abstract. Numerous control systems have been devised for addressing structural responses under dynamic loading, especially when inherent structural resistance is inadequate. The Neural Network (NN) controller, an active control method, offers simple solutions for complex structural vibration problems in uncertain environments. However, prior NN-based control strategies often failed to achieve optimal control performance. This research introduces a cutting-edge deep reinforcement learning driven active vibration control strategy, utilizing NN controllers' learning potential while delivering control performance comparable to model-based optimal controllers. The learning algorithm forms control strategy through environmental interactions, without requiring dynamic system model knowledge. This model-free approach provides optimal control performance across various excitations. The strategy was tested on a 6-story shear-building model using full-state feedback and expanded to partially observed feedback scenarios. The results reveal that the control performance closely matches a traditional linear quadratic regulator. Additionally, the learned controller outperforms conventional output feedback controllers in partially observed systems.

Key words: active control, deep reinforcement learning, neural network control, model-free algorithm, optimal active control

1 INTRODUCTION

Structural control is a crucial aspect of mitigating the vibration of buildings and civil structures induced by external forces, including wind and seismic loads. Despite the development of numerous structural control systems, relying solely on strength design often falls short in reducing the structural response under various dynamic loading [1]. Structural control systems are generally classified into three groups based on operational mechanisms: passive, active and semi-active. While passive and semi-active methods offer reliability and cost-effectiveness, their adaptability to different types of excitations is restricted [2]. Consequently, active control techniques, including active mass dampers and active brace systems, have emerged. To meet

Peer-review under responsibility of the organizing committee of SMART 2023.
doi: 10.7712/150123.9806.444204

various performance requirements, precise forces need to be determined based on sensor measurements.

Initial efforts to employ neural networks structural vibration control [3-5] is to train the controller network through minimizing various error functions in relation to the desired response. While these targets can lower structural responses, they do not provide optimal results. This led to a demand for innovative approaches that combine the learning capabilities of NN control with performance equal to or surpassing classical control methods.

Reinforcement learning (RL) is a potent technique that finds optimal policies directly from its experiences with the environment, and it has been extensively utilized by researchers to discover control policies in feedback control. The advent of deep learning has enhanced RL by incorporating deep neural networks, resulting in deep reinforcement learning. Ground breaking algorithms, such as Deep Q-Network (DQN) have demonstrated the potential to address complex control issues through deep RL strategies, laying the foundation for deep RL [6]. Recently developed algorithms, including deep deterministic policy gradient (DDPG) [7], trust region policy optimization (TRPO) [8], and proximal policy optimization (PPO) [9], have exhibited exceptional performance in robotics control tasks within the action-state domain. A modified DQN is used for structural control by optimizing a straightforward reward function [10]. Nonetheless, the control signal continues to be discrete, preventing accurate determination.

In this study, the innovative DRL method DDPG is proposed for active structural vibration control, with its performance tested on a multi-degree of freedom system. The training reward is selected as the quadratic cost function, and the performance is compared to the conventional linear control method LQR for full-state feedback control. Additionally, the analysis extends to partial observable state control, and DDPG's effectiveness is compared with output feedback control. The outcomes illustrate DDPG's capability to discover optimal strategies unattainable by standard NN methods, rendering it a promising instrument for active vibration control.

2 CONTROL PROBLEM FOMULATION

In this research, a 6-story shear building system with the actuator located on the first floor is selected as the target structure. The system's structural properties are outlined in Table 1, with several assumptions made. Floors are modeled as rigid, without rotation, the mass of each floor is considered as lumped mass, and a damping ratio of 5% is applied to all modes.

Table 1: Shear building properties

Story	Floor Mass ($\times 10^3$ kg)	Floor Stiffness ($\times 10^6$ N/m)
6	15	6.06
2–5	20	6.06
1	25	6.82

The corresponding equation of the in-plane vibration can be written as

$$\mathbf{m}\ddot{\mathbf{z}} + \mathbf{c}\dot{\mathbf{z}} + \mathbf{k}\mathbf{z} = -\mathbf{m}\mathbf{b}_w\mathbf{w} + \mathbf{b}\mathbf{u}, \tag{1}$$

In this context, **m**, **c**, and **k** represent the mass, damping, and stiffness matrices of the MDOF, respectively. The vector $\mathbf{z} = [z_1 \dots z_6]^T$ encompasses the floor displacement, with **w** correspond to the ground excitation and $\mathbf{u}$ indicating the force. Furthermore, $\mathbf{b} = [1\ 0\ 0\ 0\ 0\ 0]^T$ and $\mathbf{b}_w = [1\ 1\ 1\ 1\ 1\ 1]^T$. The state-space representation form is

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{B}_w\mathbf{w}, \tag{2}$$

$\mathbf{x}$ represents the full-state vector. The state matrix $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{m}^{-1}\mathbf{k} & -\mathbf{m}^{-1}\mathbf{c} \end{bmatrix}. \tag{3}$$

The excitation input matrix $\mathbf{B}_w$

$$\mathbf{B}_w = \begin{bmatrix} \mathbf{0} \\ \mathbf{b}_w \end{bmatrix}. \tag{4}$$

The control input matrix $\mathbf{B}$

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{m}^{-1}\mathbf{b} \end{bmatrix}. \tag{5}$$

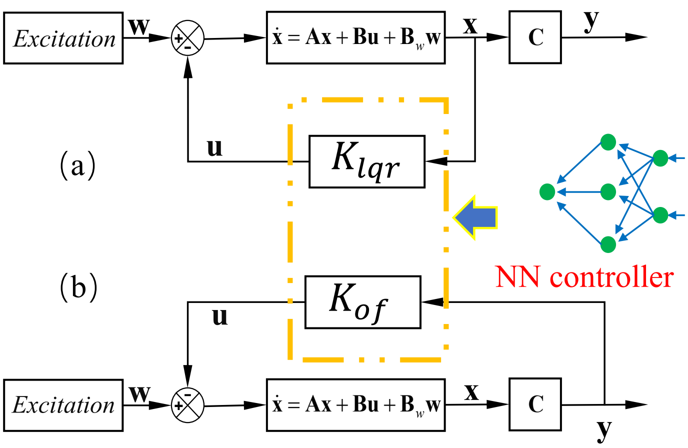


Figure 1: Classical control diagrams: (a) full state feedback controllers and (b) an output feedback controller

This study examines two scenarios. For the first one, the feedback is the complete state $\mathbf{x}$, while for the second, only a portion of the state is observable. For comparison with the proposed controller, two model-based controllers are chosen. The control action is expressed as $\mathbf{u} = -\mathbf{K}_{lqr}\mathbf{x}$, where $\mathbf{K}_{lqr}$ represents the optimal control gain that minimizes the quadratic cost function

$$J_x = \int_0^{\infty} (\mathbf{x}^T\mathbf{Q}_x\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u})dt, \tag{6}$$

Figure 1 showcases the traditional optimal controllers for both scenarios.

3 DRL DRIVEN VIBRATION CONTROL

In this research, the DDPG algorithm was adopted to update the agent's policy. DDPG is an online, model-free, and off-policy RL learning technique utilizing a deterministic policy, unlike traditional stochastic policies. The DDPG agent's actor and critic are characterized by deep neural networks. The actor generates the relevant action from the observation, while the critic estimates rewards by observation and action. To create the RL environment, the dynamic system model is discretized, and the setting is structured as a Markov Decision Process model with (***S***, ***A***, ***P***, ***G***). ***S*** represents the set of states, which are either the system state x in full-state control or the observation state y in partial observable state control. ***A*** represents the action set, while ***P*** and ***G*** designate the transition probability distribution and reward function, respectively. At every time step t, the agent experiences the state $s_t \in \boldsymbol{S}$ and takes the action $a_t \in \boldsymbol{A}$ according to policy μ. The agent then receives the reward $r_t \in \boldsymbol{G}(s_t, a_t)$ and ends up in a new state s_{t+1} with the probability ***P***.

The primary objective of this study is to effectively suppress vibrations under different excitations. Specifically, at each step, the expected reward represents the future performance. By adopting DRL, the need for an analytical or numerical solution to find optimal strategies is eliminated, and instead, the policy is learned through interaction, without requiring a complete dynamics model. In full-state feedback, the step reward is the negative quadratic cost, guiding the learning process to optimize the controller's performance.

$$r_t = -\mathbf{x}_t^T \mathbf{Q}_x \mathbf{x}_t - \mathbf{u}_t^T \mathbf{R} \mathbf{u}_t \tag{7}$$

As a result, maximizing the total reward throughout the training process is equivalent to minimizing the target function associated with the vibration control problem. The trained actor can function as a offline controller, producing control forces based on the input states, similar to a conventional controller, as illustrated in Figure 1.

4 RESULTS

4.1 Full-state feedback

To assess the proposed DDPG controller's effectiveness, the state-space model of the MDOF is discretized by 0.01 s sampling interval. Performance indices $\mathbf{Q}_x$ and $\mathbf{R}$ are established as $\mathbf{Q}_x = 10 \times \mathbf{I}_{12\times12}$ and $\mathbf{R} = 10^{-4}$. During the training process, a step excitation is given at the start, allowing the MDOF to undergo free vibration. For each training episode, 5-second free vibration are used to help the model learn and improve its performance.

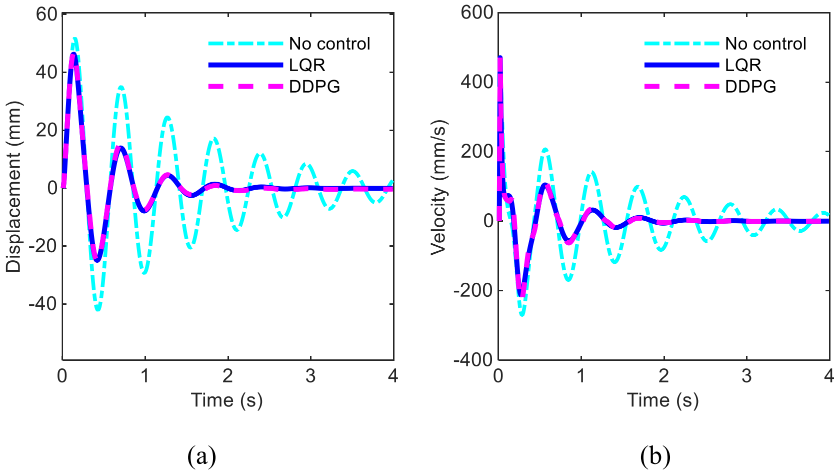
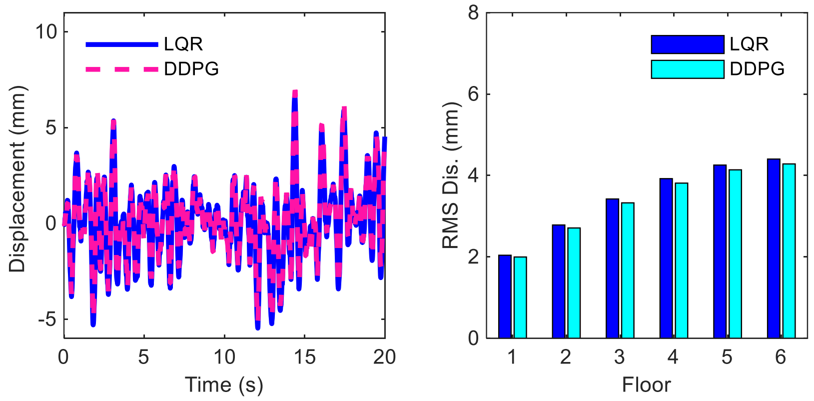


Figure 2: MDOF response of free vibration: (a) displacement of the top floor and (b) velocity of the first floor

The total cost for the DDPG agent and the LQR controller is 295 and 292, respectively. Even with the marginally higher cost for the DDPG agent, its performance remains competitive with the LQR controller. Figure 2a and 2b display the displacement and velocity time histories for two representative floors. The decay rates are impressively similar and exhibit the same phase for both controllers.

To thoroughly appraise the performance of the suggested DDPG controller, it undergoes testing under random excitations. Figure 3a presents the displacement response and 3b demonstrates the root-mean-square (RMS) displacement responses at various floors. The RMS displacement for the DDPG agent is less than that of the LQR, signifying marginally superior control performance.



(a) (b)

Figure 3: MDOF under white-noise excitations: (a) displacement of the top floor nd (b) RMS displacement of all floors

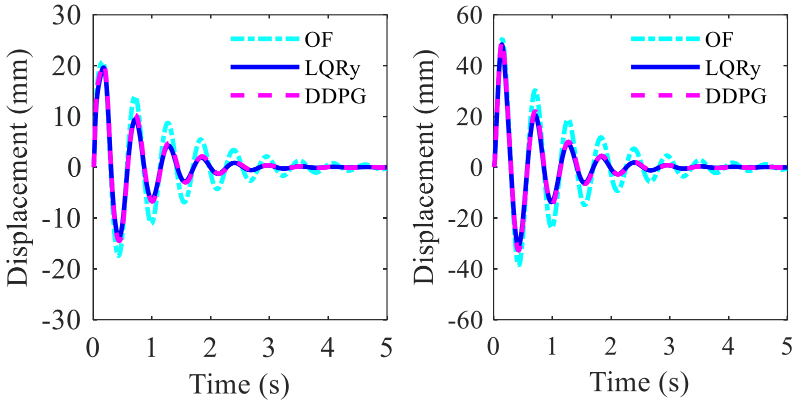
4.2 Partial-state feedback

In the partially observed state scenario, a limited number of states are observable, specifically the displacements and velocities of the first, third, and fifth floors. The agent structure for this situation remains identical to the full-state controller, with the sole distinction being the nodes of the input layers. Additionally, the reward calculation differs during training. The weighting matrix $\mathbf{Q}_y$ is chosen as $10 \times \mathbf{I}_{6\times6}$, while **R** remains unchanged.

Table 2: Cost of partially-observed MDOF

Controller	Feedback states	Cumulative Cost
LQRy	12	164
DDPG	6	168
Output feedback	6	203

Table 2 displays the free vibration results for the MDOF system employing output weighting LQR control (LQRy), DDPG agent, and the output feedback controller with an equal number of observable states. LQRy displays the most effective performance, as the control force relies on the fully observable state. In contrast, the trained DDPG agent still manages to attain a performance close to LQRy and even exceeds the output feedback controller. Figures 4a and 4b showcase the displacement responses of the first and the top floor, respectively.



(a) (b)

Figure 4: Free vibration displacement response of partially-observed MDOF: (a) the first floor and (b) the top floor

5 CONCLUSIONS

To sum up, this study presented a cutting-edge DRL-based vibration control approach using the DDPG algorithm for NN controller training. The effectiveness of this method was tested on a 6-story shear building model under different conditions. Findings revealed that the DRL-based controller could attain control performance akin to traditional model-based LQR controllers, necessitating knowledge of structural dynamic models. The trained agent maintained its control performance under random excitations. Furthermore, the proposed DRL controller surpassed the output feedback controller when feedback was not fully available. The DRL controller's control performance did not significantly decline compared to the full-state feedback controller LQRy, even though it used fewer states in the feedback. Consequently, the proposed DRL-based vibration control strategy shows great potential for active structural vibration control in various linear CT vibration control problems.

REFERENCES

[1] G. W. Housner *et al.*, "Structural control: Past, present, and future," *Journal of Engineering Mechanics,* vol. 123, no. 9, pp. 897-971, Sep 1997, doi: Doi 10.1061/(Asce)0733-9399(1997)123:9(897).

[2] F. Y. Cheng, H. Jiang, and K. Lou, *Smart structures: innovative systems for seismic response control*. CRC press, 2008.

[3] K. Bani-Hani and J. Ghaboussi, "Nonlinear structural control using neural networks," *J Eng Mech-Asce,* vol. 124, no. 3, pp. 319-327, Mar 1998, doi: Doi 10.1061/(Asce)0733-9399(1998)124:3(319).

[4] Y. A. He and J. J. Wu, "Control of structural seismic response by self-recurrent neural network (SRNN)," *Earthquake Engineering & Structural Dynamics,* vol. 27, no. 7, pp. 641-648, Jul 1998. [Online]. Available: <Go to ISI>://WOS:000074819300001.

[5] Y. Tang, "Active control of SDF systems using artificial neural networks," *Computers & Structures,* vol. 60, no. 5, pp. 695-703, Jul 10 1996, doi: Doi 10.1016/0045-7949(95)00438-6.

[6] V. Mnih *et al.*, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602,* 2013.

[7] T. P. Lillicrap *et al.*, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971,* 2015.

[8] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *International conference on machine learning*, 2015: PMLR, pp. 1889-1897.
[9] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347,* 2017.
[10] H. R. Rahmani, G. Chase, M. Wiering, and C. Konkea, "A framework for brain learning-based control of smart structures," *Advanced Engineering Informatics,* vol. 42, p. 100986, Oct 2019, doi: ARTN 100986 10.1016/j.aei.2019.100986.