

# 스마트 TMD 제어를 위한 강화학습 알고리즘 성능 검토

## Performance Evaluation of Reinforcement Learning Algorithm for Control of Smart TMD

강 주 원\*  
Kang, Joo-Won

김 현 수\*\*  
Kim, Hyun-Su

### Abstract

A smart tuned mass damper (TMD) is widely studied for seismic response reduction of various structures. Control algorithm is the most important factor for control performance of a smart TMD. This study used a Deep Deterministic Policy Gradient (DDPG) among reinforcement learning techniques to develop a control algorithm for a smart TMD. A magnetorheological (MR) damper was used to make the smart TMD. A single mass model with the smart TMD was employed to make a reinforcement learning environment. Time history analysis simulations of the example structure subject to artificial seismic load were performed in the reinforcement learning process. Critic of policy network and actor of value network for DDPG agent were constructed. The action of DDPG agent was selected as the command voltage sent to the MR damper. Reward for the DDPG action was calculated by using displacement and velocity responses of the main mass. Groundhook control algorithm was used as a comparative control algorithm. After 10,000 episode training of the DDPG agent model with proper hyper-parameters, the semi-active control algorithm for control of seismic responses of the example structure with the smart TMD was developed. The simulation results presented that the developed DDPG model can provide effective control algorithms for smart TMD for reduction of seismic responses.

**Keywords :** Deep deterministic policy gradient, Reinforcement learning, Smart TMD, Structural control system, MR damper

### 1. 서론

지진과 같은 자연재해로부터 건축구조물의 피해를 줄이기 위한 다양한 연구가 수행되고 있다<sup>1),2)</sup>. 구조시스템만으로 지진하중에 저항하여 안전성을 확보하는 구조물을 설계할 경우 구조 부재의 크기 및 강도가 과도하게 커지는 결과를 초래할 가능성이 있다. 따라서 일상하중이 아닌 지진하중과 같은 재난 수준의 하중에 대해서는 추가적인 구조제어장치를 활용하는 사례가 증가하고 있다. 구조제어장치는 수동제어장치가 널리 사용되고 있지만 외부하중의 변화에 따라서 제어력을 가하여 구조응답을 저감시키는 능동제어장치에 대한 연구 및 적용도 다수 이루어지고 있다<sup>3),4)</sup>. 근래에는 능동제어장치보다 제어 성능은 우수하지 않지만 추가되는 제어력에 의해

구조물에 피해가 발생할 수 있는 위험을 제거하고, 동적 하중의 특성에 따라 제어장치의 특성을 실시간으로 변화시켜서 동적응답을 줄일 수 있는 준능동 혹은 스마트 제어장치에 대한 연구가 크게 증가하고 있다<sup>5),6)</sup>.

스마트 제어장치는 적용되는 형태에 따라서 스마트 TMD(Tuned Mass Damper)<sup>7)</sup>, 스마트 면진시스템<sup>8)</sup>, 스마트 아웃리거 댐퍼<sup>9)</sup> 등의 다양한 방법으로 적용될 수 있다. 그 중에서 스마트 TMD는 상대적으로 적용이 용이하기 때문에 더욱 활발히 연구가 진행되고 있다. 이러한 스마트 제어장치의 지진응답 제어 성능은 제어알고리즘에 의하여 크게 좌우된다<sup>9)</sup>. 스마트 제어장치는 강성이나 감쇠와 같은 장치의 특성을 변화시키는 준능동적인 특징이 있기 때문에 Skyhook이나 Groundhook와 같은 전통적인 준능동 제어알고리즘이 많이 사용된다<sup>10)</sup>. 또는 능동 제어이론에 기반하여 준능동 제어알고리즘으로 변환한 Clipped-optimal<sup>11)</sup> 등과 같은 제어알고리즘도 많이 사용되고 있다. 스마트 제어장치의 강한 비선형적 특성을 고려하여 최적의 성능을 나타낼 수 있는 알고리즘을 개발하기 위하여 소프트 컴퓨팅기법을 적용한

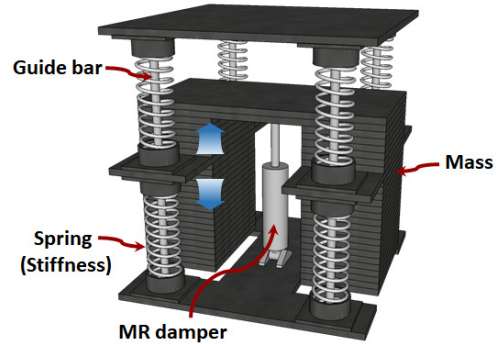
\* 정희원, 영남대학교 건축학부 교수, 공학박사  
School of Architecture, Yeungnam University  
\*\* 주·교신저자, 정희원, 선문대학교 건축학부 교수, 공학박사  
Division of Architecture, Sunmoon University  
Tel: 041-530-2315 Fax: 041-530-2839  
E-mail: hskim72@sunmoon.ac.kr

연구도 다수 수행되었다. 특히 퍼지제어알고리즘이나 유전자 알고리즘, 인공지능망 등을 함께 사용하여 매우 우수한 제어 성능을 안정적으로 나타내는 연구가 다수 수행되었다<sup>12),13)</sup>.

최근 기계학습을 기반으로 하는 인공지능을 다양한 제어 문제에 적용하여 우수한 결과를 나타내는 성과가 보고되고 있다. 특히 문제에 대한 특별한 정보를 주지 않고 에이전트와 환경의 상호작용에 의해서 최적의 해결방안을 찾아나가는 강화학습은 비디오 게임 플레이, 로봇 제어, 자동차 자율 주행 등 다양한 분야에서 성공적인 적용사례를 보여주고 있다<sup>14),15)</sup>. 특히 2015년 덤마 인드가 하나의 알고리즘<sup>16)</sup>으로 여러 개의 아타리 2600 게임을 사람보다 더 잘하는 수준으로 플레이하는 성과를 발표하면서 큰 주목을 받았다. 여기서 개발한 강화학습 알고리즘은 Deep Q-Network(DQN)이며, 이 알고리즘은 제어공학 내 다양한 분야에서의 적용 가능성을 보여주었다. DQN은 가치기반 강화학습의 대표적인 알고리즘으로 가치함수를 학습하여 최적의 가치함수를 구하고 이를 토대로 최적의 정책에 따라 제어명령을 수행한다. DQN은 아타리 2600 게임과 같이 왼쪽, 오른쪽, 정지 등의 이산적인(Discrete) 명령의 환경에서는 성공적인 제어 성능을 보여주지만 자동차의 운전대 조종, 스마트 제어장치의 제어 등과 같이 연속적인 제어명령이 필요한 환경에서는 적용하기 어렵다는 단점이 있다. 따라서 연속적인 제어명령을 적절한 간격(Step)을 두어 이산화시켜서 활용하는 방법도 있지만 이 경우 대상 시스템에 대하여 최적의 성능을 발휘하기 어렵게 된다.

본 연구에서는 기존의 DQN을 이용한 이산명령 스마트 TMD제어의 성능을 개선하기 위하여 연속제어가 가능한 강화학습 알고리즘을 이용한 스마트 TMD의 제어 성능 검토를 수행하였다. DQN과 달리 연속적인 제어명령을 출력할 수 있는 강화학습 기법으로는 다양한 정책 기반 알고리즘이 존재한다. 본 연구에서는 그 중 안정적인 학습능력으로 최근 다양한 분야에서 널리 사용되고 있는 Deep Deterministic Policy Gradient(DDPG)를 사용하여 스마트 TMD의 제어알고리즘을 개발하였다.

스마트 TMD 시제품 및 구조물 실험체를 대상으로 하여 수치해석 모델을 생성하였으며, 이를 예제구조물로 활용하여 DDPG에 의해서 생성된 알고리즘의 제어 성능을 평가하였다. 이를 위해 최적의 수동 TMD 및 널리 사



〈Fig. 1〉 Configuration of Smart TMD

용되는 준능동 제어알고리즘인 Groundhook 알고리즘과 성능을 비교하였다.

## 2. 스마트 TMD 해석 모델 및 하중

본 연구에서 수치해석에 사용할 스마트 TMD 실험체의 형태 및 구성을 〈Fig. 1〉에 나타내었다. 그림에서 보는 바와 같이 스마트 TMD의 질량은 수직방향으로 움직이며 수직방향 진동을 저감시키는데 사용한다. 그림에 나타낸 스마트 TMD는 강성 대공간구조물 지붕의 지진 응답을 저감시키기 위해서 설계되었으며, 이는 강성 대공간구조물의 지진응답이 지진하중의 종류에 관계없이 주로 수직방향으로 발생하는 특성을 고려한 것이다.

스마트 TMD 실험체는 주구조물의 아래쪽에 설치되어 진동제어 성능을 평가할 예정이다. 본 연구에서는 이러한 형태의 실험구조물을 예제로 선택하여 해석모델을 구성한다. 이때 주구조물의 질량은 750kg, 강성과 감쇠는 각각 266,670N/m 및 1.0%로 하였다. 주구조물의 고유진동수는 3Hz로 계산된다. 스마트 TMD의 질량은 100kg으로 하였다. 스마트 TMD의 제어 성능을 최적 수동 TMD와 비교하였으며, 이를 설계하기 위하여 식 (1)과 (2)를 사용하여 수동 TMD의 최적 감쇠비 및 진동수비<sup>17)</sup>를 계산하였다. 이때 수동 TMD의 질량을 스마트 TMD와 동일하게 100kg으로 하였다. 이 식에서  $\mu$ 는 주구조물에 대한 TMD의 질량비를 의미한다.

$$\sqrt{\frac{\mu(1-\mu/4)}{4(1+\mu)(1-\mu/2)}} \quad (1)$$

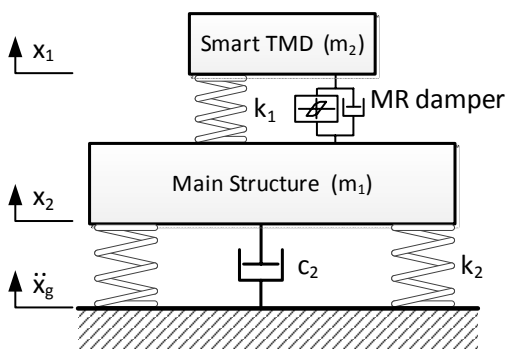
$$\frac{\sqrt{1-\mu/2}}{1+\mu} \quad (2)$$

스마트 TMD의 스프링 강성은 식 (2)로 계산되는 주 구조물과의 최적 진동수비를 통해서 얻었다. 스마트 TMD는 일반적인 수동 TMD에서 사용되는 점성감쇠기 대신에 스마트 재료로 제작된 제어장치를 사용한다. 본 연구에서는 <Fig. 1>에 나타난 바와 같이 MR 감쇠기를 사용하였다. MR 감쇠기의 최대 용량을 결정하기 위하여 파라미터 연구를 수행하였으며 400N으로 결정하였다. MR 감쇠기의 제어를 위하여 전달되는 명령전압은 0~1V 사이에서 결정된다. MR 감쇠기의 비선형 동적거동을 표현하기 위해서 현재 가장 널리 사용되고 있는 수치해석모델인 Bouc-Wen 모델<sup>18)</sup>을 사용하였다.

본 연구에서 사용한 구조물 및 스마트 TMD의 해석 모델을 <Fig. 2>에 나타내었다. 주구조물의 질량, 강성 및 감쇠를 각각  $m_1, k_1, c_1$ 으로 나타내었고, 스마트 TMD의 질량 및 강성을  $m_2, k_2$ 로 나타내었다. 이 해석모델은 다음과 같이 동적평형방정식으로 표현할 수 있다.

$$M\ddot{x} + C\dot{x} + Kx = Df_c + Ef_e \quad (3)$$

여기서  $x = [x_1 \ x_2]^T$ 는 스마트 TMD와 주구조물의 상대변위,  $f_e = -M\ddot{x}_g$ 는 지진하중,  $M = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}$ ,  $C = \begin{bmatrix} 0 & 0 \\ 0 & c_2 \end{bmatrix}$ ,  $K = \begin{bmatrix} k_1 & -k_1 \\ -k_1 & k_1 + k_2 \end{bmatrix}$ 는 각각 스마트 TMD와 주구조물의 질량, 감쇠 및 강성행렬,  $D = [1 \ -1]^T$ 와  $E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^T$ 는 MR 감쇠기 제어력과 지진하중의 작



<Fig. 2> Example structure model

용 위치를 나타내는 위치행렬,  $f_c$ 는 MR 감쇠기 제어력을 나타낸다. 수치해석을 통한 제어알고리즘 개발을 위하여 이러한 운동방정식을 다음의 상태공간방정식으로 나타내었다.

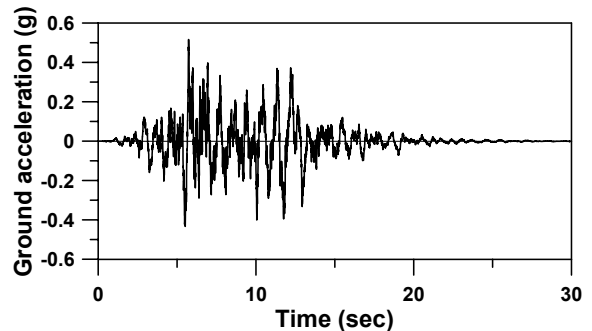
$$\dot{z} = Az + Bf_c + Hf_e \quad (4)$$

여기서  $A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix}$ 는 시스템행렬,  $B = \begin{bmatrix} 0 \\ M^{-1}D \end{bmatrix}$ 와  $H = \begin{bmatrix} 0 \\ M^{-1}E \end{bmatrix}$ 는 위치행렬이며,  $z = [x_1 \ x_2 \ \dot{x}_1 \ \dot{x}_2]^T$ 는 상태벡터이다.

스마트 TMD의 지진응답 제어 성능을 평가하기 위해서 인공지진하중을 사용하였다. 인공지진하중을 생성하기 위하여 선행 연구<sup>9)</sup>에서 제안된 강진지역 지진의 주파수특성을 표현하는 필터를 사용하였고, 이를 식 (5)에 나타내었다.

$$F(s) = \frac{4\zeta_g \omega_g s}{s^2 + \zeta_g \omega_g s + \omega_g^2} \quad (5)$$

여기서,  $\omega_g = 2\pi \text{ rad/s}$ 이며  $\zeta_g = 0.3$ 의 값을 사용하였다. 인공지진의 최대지반가속도(Peak Ground Acceleration, PGA)는 0.5g로 하였고, 길이 30초, 0.005초의 시간 간격을 갖도록 하였다. 생성한 인공지진의 지반가속도 시간이력을 <Fig. 3>에 나타내었다. 인공지진하중을 생성하기 위해 백색잡음을 위 필터에 통과시킨 지진하중 시간이력에 포락선(Envelope)을 적용시켜 일반적인 지진하중의 형태를 갖도록 하였다.



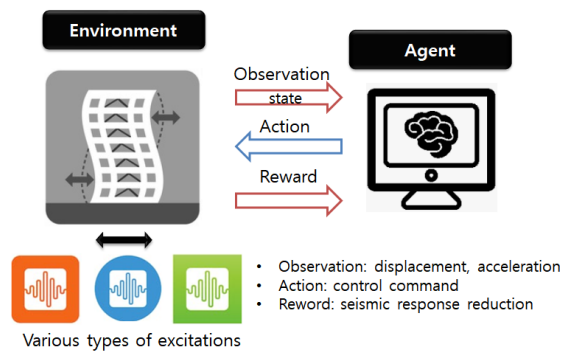
<Fig. 3> Ground acceleration time history of artificial earthquake

### 3. DDPG를 활용한 제어알고리즘 개발

본 연구에서 스마트 TMD의 제어알고리즘 개발을 위해서 사용하고자 하는 DDPG는 강화학습 알고리즘의 하나이다. 강화학습은 기계학습(Machine learning)의 일종으로서 개발자가 일일이 코드로 명시하지 않은 동작을 컴퓨터가 데이터로부터 학습하여 실행할 수 있도록 하는 알고리즘을 개발한다. 강화학습은 <Fig. 4>에 나타난 바와 같이 환경(Environment)과 에이전트(Agent)의 상호작용을 통해서 알고리즘을 개발한다.

학습을 하는 에이전트는 환경에서 제공해주는 정보를 관측(Observation)하고 이를 가공한 상태(State)를 기반으로 의사결정을 하여 환경에 대해 행동(Action)을 취한다. 본 연구에서는 스마트 TMD를 제어하는 명령전압을 에이전트가 취할 행동으로 정의하였다. 에이전트는 특정 상태에서 특정 행동을 할 때 보상(Reward)을 받게 된다. 보상은 에이전트가 취한 행동에 대해서 좋고 나쁨의 평가를 수치적으로 받는 것을 의미한다. 보상의 적절한 설계는 강화학습의 효율성을 결정하는 중요한 요소 중 하나이다.

DDPG 알고리즘은 기존의 DPG(Deterministic Policy Gradient) 알고리즘에 심층 인공신경망(Deep Neural Network)을 적용하여 개선한 것이다. DPG 알고리즘은 행동의 출력 방식으로 연속적인 행동을 선택할 수 있는 특징이 있으며, 이를 이용한 DDPG 또한 연속제어에 활용할 수 있다<sup>19)</sup>. 기존의 정책 기반 강화학습 알고리즘에서는 가능한 행동에 대한 선택지가 있었으며, 정책은 가능한 행동을 취할 확률을 출력하였다. 이는 DQN의 행동 출력 방식과 같고, 스마트 TMD의 제어명령에 적용하기 위해서는 연속적인 제어명령을 적절한 간격(Step)으로



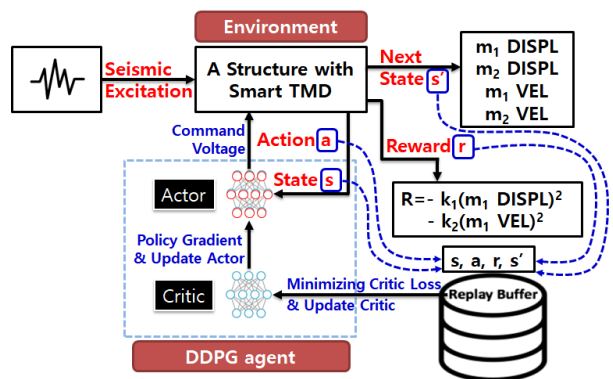
<Fig. 4> Concept of reinforcement learning

이산화시켜서 가능한 행동에 대한 선택지를 제공한다.

그러나 DDPG 알고리즘은 Deterministic이라는 이름에서 알 수 있듯이 결정론적이므로 행동 선택지가 존재하여 그 행동들에 대한 확률분포를 출력하는 것이 아니라 연속적인 값 중에서 1가지 행동에 해당하는 값을 결정하여 출력한다. 즉, 이산화시킨 제어명령이 아니라 실수 범위에서 연속적인 행동값을 출력할 수 있으므로 DDPG는 DQN에 비하여 더욱 세밀한 제어가 가능하다. DDPG의 신경망은 상태를 입력받아 행동을 출력하는 액터(Actor) 모델 및 상태와 행동을 입력받아 큐 함수값을 예측하는 크리티크(Critic) 모델로 구성된 액터-크리티크(Actor-critic) 기반의 강화학습 알고리즘이다.

본 연구에서 사용된 DDPG 알고리즘의 구성을 <Fig. 5>에 나타내었다. 그림에서 보는 바와 같이 액터는 상태(a)를 입력받아 스마트 TMD를 제어할 명령전압을 행동(a)로 출력한다. 이때 상태는 앞 절에서 설명한 상태공간 방정식의 상태벡터 성분으로 구성된다. 즉, 주구조물과 스마트 TMD의 변위 및 속도를 의미한다. 행동에 의해서 환경은 다음 스텝의 상태(s')를 출력하고 상태를 이용하여 보상(r)을 계산한다. 이렇게 스텝마다 계산된 s, a, r, s' 값을 강화학습에 그대로 사용하는 것이 아니라 재생 버퍼(Replay buffer)에 저장한 후 랜덤한 샘플링 과정을 통해서 학습을 수행한다.

DDPG에서는 액터와 크리티크를 표현하는 네트워크와 학습의 목표인 정답을 출력하는 타겟 네트워크를 따로 설정하여 학습의 안정성을 증가시킨다. DQN에서도 DDPG와 마찬가지로 학습의 효율성을 위해서 타겟 네트워크 개념을 사용하며, 특정한 주기마다 학습 중인 네트워크를 타겟 네트워크로 한 번에 복사해서 업데이트 한



<Fig. 5> Structure of DDPG

다. 이때 학습의 효과는 높아질 수 있지만 때때로 발산하는 경우도 생긴다. 이를 방지하기 위해서 DDPG에서는 소프트 타겟 업데이트(Soft target update) 개념을 사용하여 네트워크의 변수들을 그대로 복사하지 않고 식 (6)과 (7)에 나타난 바와 같이 타겟 네트워크의 변수 값(Weights)이 학습된 네트워크를 천천히 추적하도록 한다.

$$\theta_{critic}^{target} \leftarrow \tau\theta_{critic} + (1 - \tau)\theta_{critic}^{target} \quad (6)$$

$$\theta_{actor}^{target} \leftarrow \tau\theta_{actor} + (1 - \tau)\theta_{actor}^{target} \quad (7)$$

여기서  $\theta$ 는 액터 및 크리틱 네트워크를 구성하는 Weights,  $\tau$ 는 업데이트 비율을 의미하고  $\tau$ 값의 비율만큼 타겟 네트워크의 Weights를 학습된 네트워크에서 업데이트 한다. 소프트 타겟 업데이트를 사용하면 급격한 네트워크의 변화를 방지하여 학습에 의해 안정적으로 네트워크가 수렴하도록 도와준다.

DQN과 같은 이산적인 행동 환경을 대상으로 하는 알고리즘에서는  $\epsilon$ -greedy 기법에 따라서 랜덤한 행동을 통해 탐험(Exploration)을 수행한다. 그러나 DDPG와 같은 연속적인 행동 환경은 수행할 수 있는 행동의 경우의 수가 매우 많아서 행동의 선택지를 만들어 랜덤한 선택을 하는 것이 불가능하다. 이에 연속적인 행동 환경에서의 탐험은 이산적인 행동 환경에서 사용했던  $\epsilon$ -greedy 기법을 그대로 적용하기 어렵다. 따라서 선택지를 만들지 않고 실수 범위에서 랜덤하게 행동을 선택하는 방법을 통해서 탐험을 수행해야 한다. 이를 위해서 본 연구에서 사용하는 DDPG에서는 랜덤하게 평균으로 회귀하면서 노이즈를 생성하는 OU(Ornstien Uhlenbeck) 노이즈 기법<sup>20)</sup>을 적용하였다. OU 노이즈는 다음의 식 (8)에 의해서 계산된다.

$$dx_t = \theta(\mu - x_t)dt + \sigma dW_t \quad (8)$$

여기서  $dx_t$ 는 랜덤한 실수값의 변화량,  $\theta$ 는 노이즈 값이 평균으로 얼마나 빨리 회귀할지를 조절하는 평균 회귀속도,  $\mu$ 는 OU 프로세스를 통해서 회귀하는 평균값,  $\sigma$ 는 노이즈 값이 얼마나 크게 변동하면서 평균으로 회귀할지 조절하는 OU 프로세스의 변동성,  $dt$ 는 시간의

변화량,  $dW_t$ 는 위너 확률(Wiener stochastic) 프로세스의 변화량을 의미한다. 위너 확률 프로세스의 변화량은 평균이 0이고 분산이  $dt$ 인 정규분포를 따른다.

DDPG의 학습 성능을 결정하는 많은 요소 중 보상을 어떻게 결정하느냐가 매우 중요한 요소로 알려져 있다. 본 연구에서는 식 (9)에 나타난 계산식을 이용해서 DDPG 에이전트의 행동에 대한 보상을 얻는다.

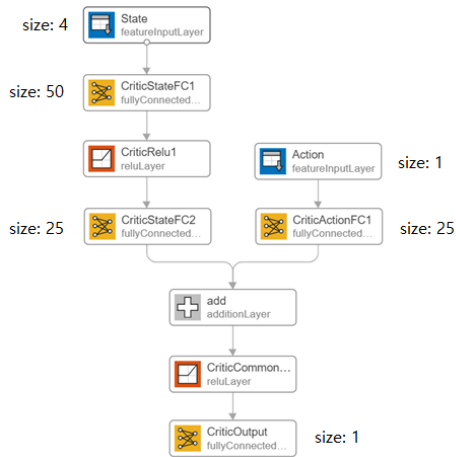
$$r = -(k_1 d_1^2 + k_2 v_1^2) \quad (9)$$

여기서  $d_1$ 과  $v_1$ 은 환경의 상태를 통해서 얻을 수 있는 구조물의 변위와 속도를 의미하고,  $k_1$ ,  $k_2$ 는 변위 및 속도에 대한 가중치이다. 이 식을 통해 알 수 있듯이 변위 및 속도응답을 기반으로 계산되는 보상은 값이 커지면 에이전트가 취한 행동이 나쁜 행동임을 의미하므로 음수로 취하여 벌칙의 의미를 나타낸다. 구조물의 변위 및 속도응답에 제곱을 취하여 응답이 클수록 강하게 벌칙을 줄 수 있도록 하였다. 변위응답과 속도응답 중 어떤 응답을 행동의 판단 기준으로 더 중요하게 고려할지는 가중치를 통해서 결정된다. 본 연구에서 대상으로 하는 예제구조물은 변위보다 속도응답이 더 크게 발생하므로 두 응답을 비슷한 수준으로 고려하기 위해서  $k_1$ ,  $k_2$ 의 값을 각각 300과 1로 설정하였다.

#### 4. DDPG 기반 제어알고리즘의 성능 검토

전술한 바와 같이 DDPG 에이전트는 액터와 크리틱으로 구성되고, 이들의 네트워크 구성 방법이 매우 중요하다. 입력된 상태와 행동의 가치를 평가하는 크리틱 네트워크의 구성을 <Fig. 6>에 나타내었다. 그림에 각 레이어의 크기를 함께 나타내었고, 상태와 행동을 함께 입력받은 두 네트워크가 합쳐져서 하나의 가치값을 출력한다.

액터 네트워크는 4개의 크기를 가진 상태를 입력받아 1개 크기의 행동(명령전압, 0-1V)을 출력한다. 액터 네트워크는 3개의 노드를 가진 1개의 완전연결(Fully connected) 레이어로 구성되고 두 레이어의 출력 시 tanh 활성화함수를 적용하였다. DDPG 에이전트를 학습시킬 때 사용한 하이퍼파라미터를 <Table 1>에 나타내었다. 표에서 보는 바와 같이 10,000 에피소드를 진행하



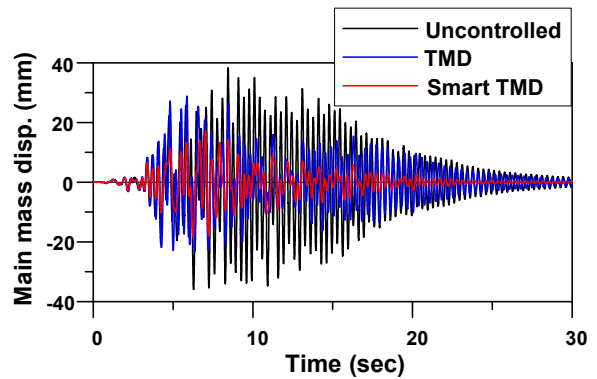
〈Fig. 6〉 Configuration of critic network

〈Table 1〉 Hyperparameter for DDPG

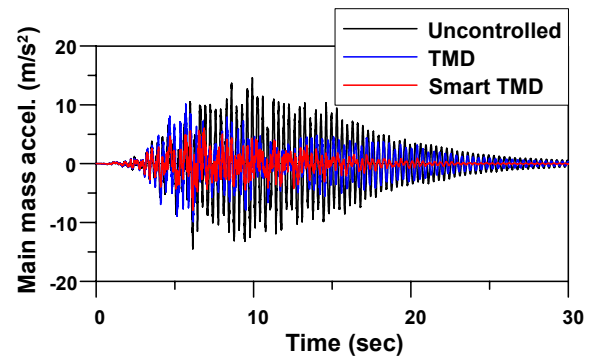
Item	Value
Learning rate (Actor)	0.0001
Learning rate (Critic)	0.001
Target smooth factor	0.001
Discount factor	0.99
Mini batch size	128
Noise variance	0.3
Noise variance decay rate	0.00001
Optimizer	Adam
Max. episode	10,000

여 DDPG 에이전트를 학습하였다. 일반적으로 액터-크리틱 기법에서 크리틱의 학습률(Learning rate)은 액터의 학습률보다 크게 설정하므로 액터와 크리틱에 대해 각각  $1 \times 10^{-3}$ 과  $1 \times 10^{-4}$ 의 학습율을 적용하였다. DDPG 에이전트의 탐색을 위해서 사용되는 노이즈의 하이퍼파라미터 중 변동성을 나타내는  $\sigma$ 는 0.3, 평균회귀속도( $\theta$ )는  $1 \times 10^{-5}$ 의 값을 사용하였다.

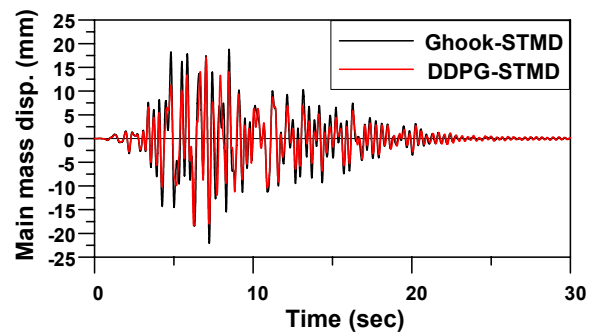
학습한 DDPG 에이전트 모델을 스마트 TMD의 제어 알고리즘으로 이용하여 제어한 예제구조물의 지진응답 저감 성능을 검토하였다. 이를 위해서 예제구조물을 제어하지 않은 경우와 최적의 수동 TMD를 사용하여 제어한 경우를 함께 비교하였다. 본 연구의 대상인 대공간구조물의 안전성과 연관이 있는 변위응답과 지붕에 설치된 장치의 낙하 문제와 관련이 있는 가속도응답에 대한 시간이력을 〈Figs. 7, 8〉에 나타내었다. 또한 DDPG 제어 알고리즘의 성능을 비교하기 위해서 준능동 제어장치



〈Fig. 7〉 Displacement time histories



〈Fig. 8〉 Acceleration time histories



〈Fig. 9〉 Comparison of DDPG and groundhook

를 위해 널리 사용되고 있는 Groundhook 제어알고리즘을 사용하여 스마트 TMD를 제어한 주구조물의 변위응답을 〈Fig. 9〉에 나타내었다.

수동 TMD를 사용하면 변위응답과 가속도응답 모두 제어하지 않은 경우에 비하여 대폭 줄일 수 있었다. 수동 TMD를 사용한 최대 변위응답 저감율은 약 25%, 최대 가속도응답 저감율은 약 30%로 가속도응답 저감에 더 효과적인 것을 알 수 있다. DDPG로 제어한 스마트 TMD를 사용한 경우에는 최대 변위응답 저감율은 약 52%, 최대 가속도응답 저감율은 약 46%로 수동 TMD보다 제어 성능이 크게 향상되었으며 변위응답에 대한 제어 성능이 더 뛰어난 것을 알 수 있다. 수동 TMD에

비하여 스마트 TMD는 변위응답 및 가속도응답을 각각 약 36% 및 22% 더 줄일 수 있었다. DDPG 제어알고리즘과 Groundhook 제어알고리즘을 비교한 <Fig. 9>를 통해서 최대변위는 약 16%, 최대가속도는 약 10% DDPG 제어알고리즘이 더 우수한 제어 성능을 나타내는 것을 확인하였다.

## 5. 결론

본 연구에서는 스마트 TMD의 효율적인 제어를 위한 알고리즘을 개발하기 위해서 강화학습 알고리즘인 DDPG를 사용하였으며 효용성 및 적응성을 검토하였다. 이를 위하여 단자유도 예제구조물과 MR 감쇠기로 구성된 스마트 TMD 실험체를 이용하여 강화학습의 환경을 구성하였다. 주구조물과 스마트 TMD의 변위 및 속도를 DDPG 에이전트로 전달되는 상태로 결정하였고, 스마트 TMD의 제어를 위한 명령전압을 행동으로 정의하였다. 수치 해석 결과, DDPG를 통해서 스마트 TMD를 우수하게 제어할 수 있는 알고리즘을 개발할 수 있는 것을 확인하였다. 강화학습을 통한 건축구조물의 동적응답 제어는 국내외적으로 초기연구단계로서 추후 더 다양한 강화학습 알고리즘의 활용성 및 적응성에 대한 추가적인 연구가 필요하다고 판단된다.

### 감사의 글

본 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임. (No. NRF-2019R1A2C1002385)

### References

1. Spencer Jr., B. F., & Nagarajaiah, S., "State of the art of structural control", *Journal of Structural Engineering*, Vol.129, No.7, pp.845~856, 2003, doi: 10.1061/(ASCE)0733-9445(2003)129:7(845)
2. Kim, H. S., & Kang, J. W., "Vibration Control Performance Evaluation of Hybrid Mid-Story Isolation System for a Tall Building", *Journal of Korean Association for Spatial Structures*, Vol.18, No.3 pp.37~44, 2018, doi: 10.9712/KASS.2018.18.3.37
3. Korkmaz, S., "A review of active structural control: challenges for engineering informatics", *Computers & Structures*, Vol.89, No.23-24, pp.2113~2132, 2011, doi: 10.1016/j.compstruc.2011.07.010
4. Ok, S. Y., Park, K. S., Song, J. H., & Koh, H. M., "Multi-Objective Integrated Optimal Design of Hybrid Structure-Damper System Satisfying Target Reliability", *Journal of the Earthquake Engineering Society of Korea*, Vol.12, No.2, pp.9~22, 2008, doi: 10.5000/EESK.2008.12.2.009
5. Pastia, C., & Luca, S. G., "Vibration Control of a Frame Structure Using Semi-Active Tuned Mass Damper", *Bulletin of the Polytechnic Institute of Jassy, CONSTRUCTIONS. ARCHITECTURE Section*, Vol.59, No.4, pp.31~40, 2013, Retrieved from <http://www.bipcons.ce.tuiasi.ro/Archive/392.pdf>
6. Nagarajaiah, S., & Narasimhan, S., "Smart base-isolated benchmark building. Part II: phase I sample controllers for linear isolation systems", *Structural Control Health Monitoring*, Vol.13, No.2-3, pp.589~604, 2006, doi: 10.1002/stc.100
7. Kim, H. S., & Kang, J. W., "Seismic Response Control of Retractable-roof Spatial Structure Using Smart TMD", *Journal of Korean Association for Spatial Structures*, Vol.16, No.4, pp.91~100, 2016, doi: 10.9712/KASS.2016.16.4.091
8. Kim, H. S., & Kang, J. W., "Optimal Design of Smart Outrigger Damper for Multiple Control of Wind and Seismic Responses", *Journal of Korean Association for Spatial Structures*, Vol.16, No.3, pp.79~88, 2016,

- doi: 10.9712/KASS.2016.16.3.079
9. Bitaraf, M., Ozbulut, O. E., Hurlebaus, S., & Barroso, L., "Application of semi-active control strategies for seismic protection of buildings with MR dampers", *Engineering Structures*, Vol.32, No.10, pp.3040~3047, 2010, doi: 10.1016/j.engstruct.2010.05.023
  10. Koo, J. H. (2003). Using Magneto-Rheological Dampers in Semiactive Tuned Vibration Absorbers to Control Structural Vibrations (Doctoral dissertation). Virginia Polytechnic Institute and State University, USA.
  11. Dyke, S. J., Spencer Jr., B. F., Sain, M. K., & Carlson, J. D., "Modeling and control of magnetorheological dampers for seismic response reduction", *Smart Materials and Structures*, Vol.5, No.5, pp.565~575, 1996
  12. Bathaei, A., Zahrai, S. M., & Ramezani, M., "Semi-active seismic control of an 11-DOF building model with TMD+MR damper using type-1 and -2 fuzzy algorithms", *Journal of Vibration and Control*, Vol.24, No.13, pp.2938~2953, 2018
  13. Kim, H. S., & Kang, J. W., "Multi-objective Integrated Optimization of Diagrid Structure-smart Control Device", *Journal of the Computational Structural Engineering Institute of Korea*, Vol.26, No.1, pp.69~77, 2013
  14. Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo, A., & De Felice, F., "Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions", *Sustainability*, Vol.12, No.2, 2020, doi: 10.3390/su12020492
  15. Buşoniu, L., de Bruin, T., Tolić, D., Kober, J., & Palunko, I., "Reinforcement learning for control: Performance, stability, and deep approximators", *Annual Reviews in Control*, Vol.46, pp.8~28, 2018, doi: 10.1016/j.arcontrol.2018.09.005
  16. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... Hassabis, D., "Human-level control through deep reinforcement learning", *Nature*, Vol.518, pp.529~533, 2015, doi: 10.1038/nature14236
  17. Warburton, G. B., "Optimum absorber parameters for various combinations of response and excitation parameters", *Earthquake Engineering and Structural Dynamics*, Vol.10, No.3, pp.381~401, 1982, doi: 10.1002/eqe.4290100304
  18. Sueoka, T., Torii, S., & Tsuneki, Y. (2004). The Application of Response Control Design Using Middle-Story Isolation System to High-Rise Building. Proceedings of the 13th World Conference on Earthquake Engineering, Canada, Retrieved from [https://www.iitk.ac.in/nicee/wcee/article/13\\_3457.pdf](https://www.iitk.ac.in/nicee/wcee/article/13_3457.pdf)
  19. Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., ... Wierstra, D. (2016). Continuous control with deep reinforcement learning. Proceedings of the 4th International Conference on Learning Representations, Puerto Rico, Retrieved from <https://arxiv.org/pdf/1509.02971.pdf>
  20. Uhlenbeck, G. E., & Ornstein, L. S., "On the Theory of the Brownian Motion", *Physical Review*, Vol.36, No.5, pp.823, 1930, doi: 10.1103/PhysRev.36.823
- Received : April 19, 2021
  - Revised : April 26, 2021
  - Accepted : April 26, 2021