

인공지능 강화학습을 통한 일륜전동차 제어

천시현¹, 김병진²

¹폴리텍대학 청주캠퍼스 전기에너지과 학부생

²폴리텍대학 청주캠퍼스 전기에너지과 교수

2380948@naver.com, furthest@kopo.ac.kr

Single Wheel Vehicle controlled Reinforcement

Shi-Hyeon Cheon¹, Byoung-Jin Kim²

^{1,2}DeElectric Energy Department, Cheongju campus Korea Polytechnic

외발자전거는 서커스 곡예의 소재로 사용될 정도로 일반인이 타기 어렵다. 하지만, 구조가 단순하고 부피가 작고 가벼워 개인용 이동 수단으로 장점이 많다. 바퀴가 하나이므로 자유도가 높아서 좁거나 곡선반지름이 작은 길에서도 유리하다. 최근 전기자전거뿐만 아니라 전동보드, 전동휠 등 다양한 형태의 전력 구동형 개인 이동 수단들이 많이 등장하고 이용되고 있는데 스스로 균형을 유지하여 안전하고 쉽게 탈 수 있는 외발 전동차를 개발을 위해 인공지능 강화학습을 적용한 자동제어장치를 시뮬레이션을 통해 구현하고자 한다.

1. 서론

국내에서도 로봇을 위한 외륜 구동장치의 제어를 주제로 한 연구가 있었다. 이현탁, 박희재[1]는 외바퀴 구동장치가 피치 방향과 롤 방향이 서로 동역학적으로 디-커플 되어 있다는 가정하에 양축을 독립적으로 모델링하고 선형화하여 여기에 기반을 둔 최적제어기 LQR 제어기를 제안하였다. 정지수[2]는 역진자-균형제어 원리를 이용하여 피치 방향 제어하고, CMG 방식을 이용하여 롤 방향 제어하여 성능 개선을 꾀한다. 김필교와 4인[3]은 외바퀴 구동 로봇의 최적 설계에 중점을 두었다. 동력학 기반으로 시뮬레이션하고 바퀴 내부에 배터리를 장착하고 높은 플라이휠 회전속도를 채택하여 구조적으로 안정성 확보를 실험하였다.

일륜 전동차를 개인 이동 수단으로 사용하기 위한 균형 제어의 목표와 방법이 조금 다르다. 바퀴가 굴러 진행하고 있을 때 피치 방향 균형은 자이로(Gyro) 효과에 의해 유지되고, 정지하여 균형을 잃을 때는 운전자의 의지이므로 제어 대상이 아니다. 이 논문에서는 인공지능에 의해 롤 방향의 균형을 잡고 운전자의 운전 신호에 따라 효과적으로 속도 제어가 가능한 시스템을 시뮬레이션 결과와 함께 논하고자 한다.

인공지능의 신경망 학습 가운데 우리가 적용할 강화학습은 환경을 학습하는 기법이라 할 수 있다. 기계가(인공지능이) 컴퓨터 게임을 한다고 하면, 그 인공

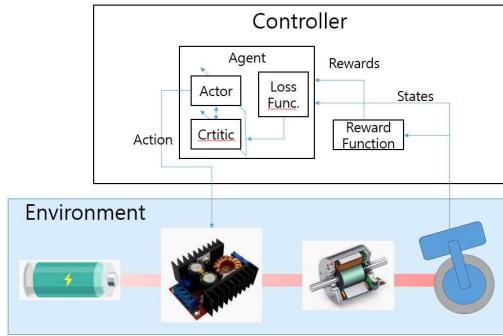
지능을 액터(Actor)라 하고 좌우로 이동 또는 발사 등의 액션(Action)을 취하게 되는데 그때마다 (게임 속) 환경의 상태가 변화한다. 인공지능은 그 환경을 읽고 다음 액션을 선택하는데 점수가(보상이) 최대가 되도록 한다. 초기 상태에서는 학습이 충분치 않아 점수가 형편없이 낮지만, 에피소드가(한 판의 게임이) 반복될수록 점차 학습이 완료되고 결국 해당 게임의 전문가가 된다. 사람이 훈련을 통해 외발자전거를 학습하듯이 인공지능도 강화학습을 통해 학습할 수 있다. 인공지능이 실존하는 외발자전거를 학습하기보다는 시뮬레이션을 통해 학습하는 것이 비용적 측면에서 훨씬 유리하다. 비행 조종사가 시뮬레이션을 통해 훈련하는 것도 같은 이유이다.

2. 일륜전동차의 기계적 모델링

바퀴가 굴러가는 방향을 롤 축, 그 왼쪽을 피치 축이라 할 때, 바퀴는 초속 1m로 굴러가고 있다면 자이로 효과로 인해 스스로 피치 축 방향으로 균형을 잡고 있다. 바퀴의 속도가 충분하지 않아 좌우 균형을 잃게 되어도 운전자의 자발적 의지로 보고 이를 상관하지 않도록 한다. 바퀴 내부에 모터가 내장되어 있고 바퀴 축에 연결된 지지대 위에는 그 모터를 제어 및 구동에 필요한 여러 장치가 달려있고 그 위에 운전자가 탈 수 있는 안장이 설치된 구조이다. 안장을 포함한 지지대 전체를 안장 지지대라고 하면 이것이 중력에 의해 기울어지지 않도록 제어하고자 한다. 운전자가 의도적으로 상체를 앞쪽으로 기울이면 인공지능 제어기는 바퀴의 전진 방향으로 동력을 가하고, 뒤

쪽으로 기울이면 후진 방향으로 동력을 가하여 균형과 속도 제어를 동시에 해결하려 한다. 이러한 제어상태에서는 절대 모터의 기계적 최대속도를 내어서는 안 된다. 최대속도에서는 운전자가 앞쪽으로 균형을 잃게 되면 이를 바로 잡아 줄 추가동력을 낼 수 없기 때문이다.

2. 인공지능 강화학습을 통한 제어기



(그림 1) 강화학습 기반 제어기

Controller 속에 Agent는 인공지능이다. Python의 tensorflow 모듈을 사용하였으며, Actor가 취할 수 있는 Action은 [-1, 0, 1]인데, 모터에 전달하는 에너지를 의미하고 음수는 역방향을 나타낸다. Action은 초당 50회 주기적으로 생성하는데 이것이 모터 드라이버에 PWM 파형으로 전달된다. 드라이버는 모터를 구동하고 이에 따라 바퀴의 속도, 가속도, 압장 지지대의 중심축으로부터의 각도 및 각속도 등이 변환된다. 즉 환경의 상태는 [각도, 각속도, 속도, 가속도] 4개의 값으로 구성된 벡터(텐서)이다.

각도가 작을수록 높은 점수를 보상한다. 이 과정을 한 step으로 하고 step마다 action, states, rewards를 Agent에 전달하면 이 값을 누적하면서 episode가 끝날 때까지 기다린다. 한 episode에서 최대 step 수를 제한하는데, 여기서는 1,000회로 한정하였다. episode가 종료되는 추가적인 조건으로는 reward가 충분할 때, 여기서는 475점(한 에피소드 동안 균형을 유지하고 있는 경우 받을 수 있는 리워드), 또는, 상태가 특정 값을 벗어날 때, 여기서는 각도가 30도 이상의 경우(균형을 잃는 경우)이다. episode 뒤에 Agent는 action 이 얼마나 유효했는지를 평가한다. 이 과정이 네트워크의 역전파로 진행된다. 기계학습은 episode 종료 시점에 이루어진다. 에피소드는 최대 1000번 진행한다.

Actor-Critic 기법은 Q-learning, DQN 같은 전통적 강화 학습보다 진보된 기법으로 Actor와 Critic 두 개의 신경망을 두어 보다 안정적으로 빠르게 학습할 수 있다. Actor는 상태에 따라 어떤 액션이 유리한지 선택하는 역할을 담당한다. Critic은 Actor가 취한 action이 얼마나 유리한지를 평가한다. 네트워크는 심층구조이고 여기서는 6층 구

조이다.

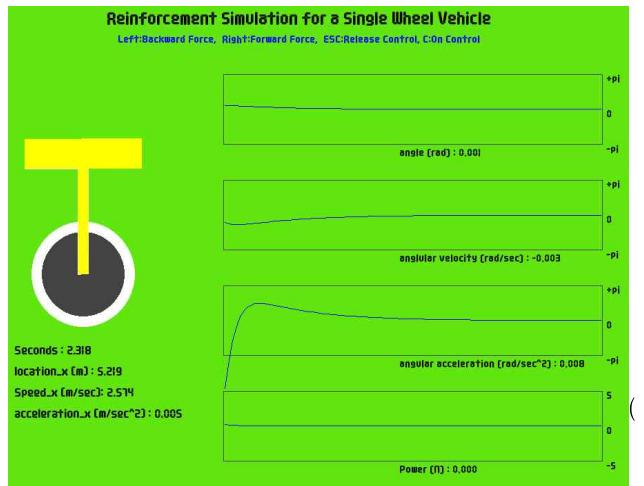


그림 2) 시뮬레이션 결과

학습은 575회 episode 만에 완료되었으면 평균 reward는 480점이다. 학습을 빠르게 진행하기 위해 학습할 때는 실시간 렌더링을 하지 않고 학습이 종료된 이후 렌더링을 진행하였다. 그림 3은 학습 이후 시뮬레이션 렌더링 화면이다. 초기 각은 0.4[rad] 초기 속도는 1 [m/sec]이다. 이후 2.318초까지의 상황을 그래프로 보여지고 있다. 시뮬레이션 렌더링은 실시간으로 표현되며 필요에 따라 제어기를 제거해 보거나, 외란을 주어 제어의 효과와 상태를 확인해 볼 수 있다.

3. 결론

본 논문에서는 일륜전동차의 자율 균형 제어기를 강화 학습기반 인공지능망으로 활용될 수 있고 그 효과를 시뮬레이션해 보였다. 일반적인 피드백제어기와는 달리 훈련과 학습이라는 과정이 필요하며 이를 위해 일륜전동차를 시뮬레이션할 수 있도록 하였으며 이 과정에서 인공지능제어기는 비선형성 환경과 예측 불가 외란에 잘 대처할 수 있도록 학습한다.

본 논문은 과학기술정보통신부 정보통신창의인재양성사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.

참고문헌

- [1] 이현탁, 박희재, “외륜 이동로봇의 균형제어 알고리즘”, Journal of the Korean Society of Manufacturing Technology Engineers, 26:1, 144-149, 2017
- [2] 정지수, 박희재, “Control Moment Gyro 원리를 이용한 외륜로봇의 균형제어”, Journal of the Korean Society of Manufacturing Technology Engineers, 27:3, 228-234, 2018
- [3] 김필교, 박준형, 하민수, 정슬, “외바퀴 구동 Gyrobo의 제작 및 밸런싱 제어 구현”, Journal of Institute of Control, Robotics and Systems, 19(6), 5021-507, 2013