

모방학습과 모션 블렌딩을 이용한 객체 상호작용 애니메이션

신정민¹, 한상원¹, 신병석¹

¹인하대학교 전기컴퓨터공학과

sin37minduke@gmail.com, starski6302@gmail.com, bsshin@inha.ac.kr

Object Interaction Animation Using Imitation Learning and Motion Blending

Jeong-Min-Shin¹, Sang-Won Han¹, Byeong-Seok Shin¹

¹Dept. of Electrical and Computer Engineering, Inha University

요 약

애니메이션은 주어진 키프레임(key frame)에 맞추어 움직이기 때문에, 다른 객체와 상호작용할 때 상대편 물체의 위치나 방향을 애니메이션에 맞추어 변환해야 한다. 이 논문에서는 모방학습으로 애니메이션을 학습하고, 모션 블렌딩(motion blending) 기법으로 객체 간 상호작용을 학습하여 새로운 애니메이션을 생성하는 방법을 제안한다. 에이전트(agent)는 오브젝트의 상태를 관측하고 주어진 모션들을 블렌딩하는 방법으로 다양한 행동을 취하고 목적에 대한 보상을 받는다. 에이전트가 행동하는 과정에서 모션 블렌딩 비율에 대한 가중치를 계산하는 함수를 설계하고, 생성되는 애니메이션이 사람이 취할 수 있는 동작에 가깝도록 회전 각도 clamping 함수와 보상 시스템을 설계하여 반영한다. 모방학습 기반 모션 블렌딩 기법은 객체의 변화에 상호작용하는 애니메이션을 기존 강화학습 기반 애니메이션 생성 기법보다 적은 학습량으로 생성할 수 있음을 확인했다.

1. 서론

애니메이션은 주로 3D 모델을 제작하는 디자이너들이 리깅(rigging)을 통해 모델을 제어할 수 있는 위치를 지정한 뒤, 이를 통해 모델의 움직임을 직접 조작하여 생성된다. 이렇게 생성된 애니메이션이 독립적으로 사용되는 경우에는 정교하고 정확한 움직임을 보여준다. 그러나 3D 모델이 주어진 환경 내에서 다양한 상태를 가지고 있는 여러 객체와 상호작용하는 경우, 디자이너에 의해 생성된 애니메이션을 직접 사용하면 오브젝트에 상태에 맞지 않는 동작을 실행하게 되어 오히려 부자연스러운 동작이 나타나게 된다.

이에 3D 모델이 객체의 변화에 대응하여 상호작용 애니메이션을 생성하는 다양한 방법이 제안되었으며, 최근에는 주어진 환경에서 에이전트가 직접 상태를 관측하고 그에 따라 행동하여 최적의 보상을 얻는 행동을 학습하도록 하는 기법인 강화학습을 애니메이션에 적용하는 방법이 제안되고 있다.

강화학습은 주어진 환경 안에서 유한한 시간 동안 에이전트가 행동을 선택하고, 이를 통해 얻은 보상의 합산을 최대화하도록 만드는 기계학습 기법이다 [1]. 강화학습은 주어진 문제를 형식화하는 과정을 필요로

하며, 이를 위해 주로 마르코프 의사결정과정을 사용한다. 이것은 주어진 환경과 문제를 에이전트의 상태, 행동, 보상, 상태 전이 확률의 요소로 정의한다. 에이전트는 정의된 요소를 통해 행동을 선택하고 그에 따른 보상에 따라 행동 정책을 수정하며 보상을 최대화하게 된다 [2].

모델의 애니메이션 생성을 위한 기계학습은 주로 스켈레톤 구조 및 관절 영역의 변형에 대해 학습한다. 이를 위해 제안된 방법 중 하나인 Deepmimic은 모션 캡처를 통해 모델이 스켈레톤 구조 및 관절 영역의 변형을 학습하고 모션을 생성하게 한다 [3]. 한편, 모델이 기존의 모션을 기반으로 학습하여 새로운 모션을 생성하는 경우 관절 기반 모델 대신 다른 모델을 사용하는 방법도 제안되었다. 일례로 근육을 중심으로 움직이는 물리 기반 모델의 이족보행을 구현하기 위해 관절 기반 모델을 대체한 근골격 모델을 사용하여 관절 토크 한계를 제공하고 사람이 수행할 수 있는 범위로 애니메이션을 제한하는 방법이 있다 [4-6]. 이와 같은 강화학습 기반의 애니메이션 생성 방법은 주어진 환경에 대해서 학습을 온전히 완료하지 않더라도 애니메이션을 환경에 최적화된 동작으로 생성할

수 있다는 장점이 있다 [3]. 하지만 모델의 행동에 영향을 미치는 관절 구조의 개수가 많을수록, 목표 애니메이션을 학습하기 위해 고려해야 하는 변수가 많아지면서 학습 비용이 크게 증가하는 문제가 있다.

이 논문에서는 필요한 학습의 횟수를 감소시키기 위해 모방학습과 모션 블렌딩 기법을 사용한 강화학습 방법 및 모션 블렌딩 함수를 제안한다. 동작의 기반이 되는 모션들과 모션 블렌딩을 사용하여 주어진 상황에 대응해서 반응할 수 있는 새로운 애니메이션을 생성할 때 모방 학습을 이용하여 적은 학습 횟수로 주어진 환경과 목적에 알맞은 애니메이션을 생성하였다. 논문에서 제안하는 모방학습 기법은 기존 애니메이션에 기반한 정보에 기반하여 각 애니메이션 별 반영 비율을 가중치로 주어 모션 블렌딩에 기반한 새로운 애니메이션을 생성한다. 제안된 모방학습 기반 모션 블렌딩 기법을 테스트하기 위해 물리 엔진에 의거해서 움직이는 오브젝트와 상호작용하는 환경을 구성하였으며, 오브젝트에 대해 목표 동작을 달성하기 위해 두 개의 관찰 항목과 목표 달성을 위한 관절 기반의 보상 시스템을 설정하였다. 제안된 모방학습 기반 모션 블렌딩 기법은 해당 보상 시스템 내에서 적은 학습 횟수로 목표를 달성하는 적절한 동작을 생성하였다.

2. 모방 학습 기반 모션 블렌딩 애니메이션 생성

이 논문에서는 애니메이션 데이터를 기반으로 학습하여, 3D 모델이 객체와의 상호작용에서 성공적으로 목표를 수행하면서 동시에 자연스러운 애니메이션 동작을 수행하도록 학습시킨다. 객체와 상호작용하는 환경에서 현재 프레임에서 오브젝트의 위치, 회전 각도 두 가지를 상태 변수로써 관측한다. 상태에 따라 에이전트가 취할 행동을 결정하며, 그 결과로 보상을 받게 된다. 에이전트가 객체와의 상호작용에서 모션 블렌딩의 결과가 목표하는 애니메이션에 가깝게 동작하도록 보상 시스템을 설계하고, 에이전트가 최적의 보상을 얻어내기 위해 다양한 행동을 시도하게 한다.

목표를 달성하기 위해 에이전트는 참조 모션들을 블렌딩하여 새로운 모션을 생성하게 된다. 에이전트는 입력으로 주어진 환경과 상태로부터 각 애니메이션마다 임의의 가중치 값을 결정하고, 이를 모든 애니메이션이 가지는 가중치의 합으로 나누어 산술평균을 구함으로써 각 애니메이션의 반영 비율을 구한다. 반영 비율은 모델의 각 관절 회전 각도에 적용되기 때문에 clamp 함수를 통해 정규화를 진행한다.

$$\theta_{t+1} = \text{clamp}\left(\frac{w_i}{s_w}, 0, 1\right) \times \theta_t$$

w_i 는 애니메이션의 i 번째 모션에 부여된 가중치를 의미하고, s_w 는 모든 모션에 부여된 가중치의 합이다. θ_t 과 θ_{t+1} 은 현재 프레임과 다음 프레임의 회전 각도이다. θ_{t+1} 를 이용해서 모델의 각 관절 별로 애니메이션 계층 구조를 고려한 변환의 회전 각도에 반영하여 다음 프레임에 모델이 취할 동작을 적용한다. 그러나 위의 모션 블렌딩 함수를 사용하는 경우, 모델의 일부 관절에서 실제로 자연스럽게 움직일 수 있는 회전 각도보다 적게 회전하는 경우가 발생하였다. 따라서 모델이 생성하고자 하는 애니메이션이 모션 블렌딩의 결과로 나올 수 있는 관절의 회전 각도 범위보다 더 넓은 회전 각도 범위를 필요로 하는 경우, 위 식을 변형한 새로운 모션 블렌딩 함수를 사용하였다.

$$\theta_{t+1} = \text{clamp}\left(\frac{w_i}{s_w}, 0, 1\right) \times \theta_t + \text{clamp}(w_{add}, c_{min}, c_{max})$$

이 모션 블렌딩 함수는 회전 각도 범위를 넓히기 위한 추가적인 값을 더한다. w_{add} 는 추가 가중치 값이며, 상수 c_{min} 과 c_{max} 는 추가할 회전 각도 범위의 최소값과 최대값으로 정의한다. 가중치 값에 일정한 범위로 제한된 회전 각도를 추가적인 값으로 적용함으로써, 모델이 기존의 모션 블렌딩 함수로 결정되는 회전 각도 범위보다 더 넓은 회전 범위를 가진다.

한편, 다음 프레임에 취할 동작을 모션 블렌딩 함수의 결과로 구하였을 때 다음 프레임 관절의 움직임이 이전 프레임에 비하여 크게 변화할 수 있으며, 이때 부자연스러운 움직임이 나타날 수 있다. 이를 해결하기 위해 한 프레임 당 애니메이션 회전 각도 변화가 일정 각도 이상을 넘지 않도록 회전 각도의 clamping 함수를 다음과 같이 정의하였다.

$$\begin{cases} t = 0, \theta_{t+1} \\ t \geq 1, \theta_{t+1} = \text{LerpUnclamped}(\theta_t, \theta_{t+1}, \frac{1}{Cangle}) \end{cases}$$

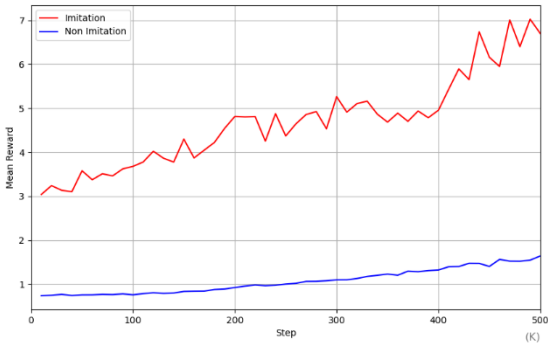
LerpUnclamped 함수는 현재 프레임으로부터 넘어가는 다음 프레임의 회전 각도가 급격하게 변하지 않도록 $Cangle$ 값의 비율만큼만 보간하여 다음 프레임의 회전 각도를 결정한다. 상수 $Cangle$ 은 기존에 모션 블렌딩 결과의 프레임이 넘어갈 때 각도를 한 프레임에 회전시킬 각도의 한계각으로 나눈 것이다. LerpUnclamped 함수를 통해 $Cangle$ 만큼 보간된 회전 각도를 다음 프레임 관절의 회전 각도에 대입하여 모델의 동작이 급격하게 변화하지 않고 연속적인 움직임을 나타낼 수 있도록 하였다.

3. 실험 결과

애니메이션 학습은 Unity 2022.3.7f1 환경에서 Unity 엔진의 플러그인인 ML-agents 0.30.0 버전을 사용하여

진행하였고, 강화학습의 알고리즘은 PPO(Proximal Policy Optimization)을 사용하였다. 모방학습에 기반한 애니메이션 생성은 20,000,000의 step의 학습을 진행하였고, 비교를 위해 모방학습을 적용하지 않은 강화학습 기반 애니메이션 생성 기법과의 동일한 조건과 보상 시스템으로 학습을 진행하였다.

모방 학습 기반 모션 블렌딩을 위해 5개의 모션을 사용하였고, 각각의 모션은 Root Transform Rotation을 모델의 골반이 바라보는 방향이 아닌 실제 이동하는 방향을 기준으로 설정하여 애니메이션 포즈에 반영하였다. 보상 시스템은 구체 오브젝트가 추락하지 않은 시간, 구체 오브젝트를 잡고 있는 휴머노이드 모델의 관절의 위치 및 방향을 사용하여 설정하였다.

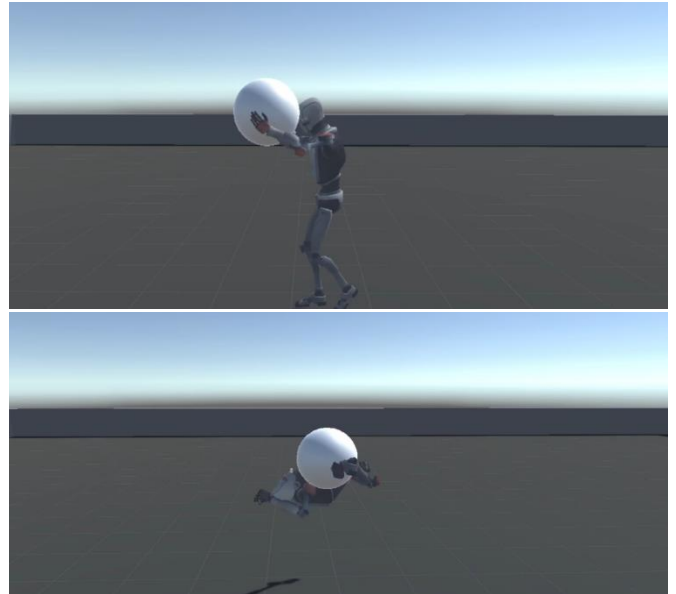


(그림 1) 애니메이션 모방학습 기반 모션 블렌딩을 적용한 경우(파란색)와 적용하지 않은 강화학습(빨간색)의 학습 회수당 보상 그래프

그림 1은 각각 모방학습 기반 모션 블렌딩 기법이 사용된 경우와 모방학습이 적용되지 않은 강화학습의 step이 진행되면서 에이전트가 얻은 보상을 10,000번의 step당 평균값으로 처음 학습부터 500,000step까지의 두 학습의 보상 결과값을 나타낸 그래프이다. 그림 1의 그래프는 모방학습 기반 모션 블렌딩 기법이 기존의 모방학습 및 모션 블렌딩이 적용되지 않은 강화학습 기법과 비교하여 적은 학습 횟수에서도 에이전트가 받는 보상이 크게 증가하는 것을 보여준다.

그림 3은 모방학습 기반 모션 블렌딩 기법의 결과로 생성된 애니메이션과, 모방학습 기반 모션 블렌딩이 적용되지 않은 기존 강화학습 기법의 결과로 생성된 애니메이션의 모습이다. 모방학습의 결과로 생성된 애니메이션은 양손으로 구체 오브젝트를 지지하여 안정적으로 사람의 걷는 모션을 모방하여 학습되었지만, 모방학습 기반 모션 블렌딩이 적용되지 않은 강화학습의 경우 구체를 떨어뜨리지 않는 보상을 달성하기 위해 부자연스러운 자세로 오브젝트를 지지하는 모습이 나타났다. 이는 모방학습 기반 모션 블렌딩

기법이 자연스러운 자세를 위해 복잡한 보상 시스템을 추가로 요구하지 않고 자연스러운 모션을 생성할 수 있음을 보여준다.



(그림 3) 모방학습 기반 모션 블렌딩 기법의 학습 결과 생성된 상호작용 애니메이션(위)과 모방학습 기반 모션 블렌딩 기법이 적용되지 않은 강화학습에 의해 생성된 상호작용 애니메이션(아래)

4. 결론

이 논문은 애니메이션 모방학습 기반 모션 블렌딩 기법을 통해 캐릭터의 애니메이션 학습으로 객체와 상호작용하는 애니메이션을 생성하는 강화학습 기법을 제안하였다. 주어진 모션들 각각에 설정한 가중치 값을 이용해서 모션 블렌딩 함수를 설계하고 정규화된 애니메이션 반영 비율을 구하였으며, 이를 학습 행동 변수로 설정하였다. 또한 애니메이션 회전 각도 clamping 함수를 정의하여 급격하고 부자연스러운 동작 변화가 방지하고 연속적인 움직임이 나타나도록 하였다. 모델의 모든 관절을 각각 학습 행동 변수로 설정한 애니메이션 학습과 비교하여 모션 블렌딩 기반 애니메이션 모방학습 기법이 더 적은 학습량으로 자연스러운 애니메이션을 생성하였다.

모방학습 기반 모션 블렌딩 기법을 통해 적은 학습 step으로도 일정 수준의 퀄리티를 보장하는 애니메이션을 빠르게 생성할 수 있었으나, 이 방법은 생성된 애니메이션이 모방학습에서 사용된 참조 모션에 의해 모델의 동작이 일정 범위 안으로 제한되어 움직일 수 있는 관절의 회전 각도 범위에 한계가 있다는 단점이 있다. 따라서 향후 연구를 통해 모방학습 기반 모션 블렌딩 기법의 단점을 개선하고 더 다양하고 자연스러운 모션의 생성이 가능할 것으로 예상된다.

Acknowledgement

이 성과는 2022 년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.NRF-2022R1A2B5B01001553). 이 논문은 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.RS-2022-00155915, 인공지능융합 혁신인재양성(인하대학교)).

참고문헌

- [1] Sutton, Richard S., and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
- [2] Bellman, Richard. "A Markovian decision process." *Journal of mathematics and mechanics* (1957): 679-684.
- [3] Peng, Xue Bin, et al. "Deepmimic: Example-guided deep reinforcement learning of physics-based character skills." *ACM Transactions On Graphics (TOG)* 37.4 (2018): 1-14.
- [4] Wang, Jack M., et al. "Optimizing locomotion controllers using biologically-based actuators and objectives." *ACM Transactions on Graphics (TOG)* 31.4 (2012): 1-11.
- [5] Lee, Seunghwan, et al. "Dexterous manipulation and control with volumetric muscles." *ACM Transactions on Graphics (TOG)* 37.4 (2018): 1-13.
- [6] Lee, Yoonsang, et al. "Locomotion control for many-muscle humanoids." *ACM Transactions on Graphics (TOG)* 33.6 (2014): 1-11.