

# 심층 강화 학습과 모션 매칭을 이용한 캐릭터 애니메이션 생성

이정민<sup>0</sup>, 권태수, 이윤상  
한양대학교 컴퓨터소프트웨어학과  
{j0064423, taesoo, yoonsanglee}@hanyang.ac.kr

## Generating Character Animation using Deep Reinforcement Learning and Motion Matching

Jeongmin Lee<sup>0</sup>, Taesoo Kwon, Yoonsang Lee  
Department of Computer Science, Hanyang University

### 요약

캐릭터 애니메이션을 생성할 수 있는 방법으로는 전통적인 키프레임 애니메이션뿐만 아니라 모션캡처데이터를 이용한 모션편집 및 최적화, 머신 러닝 등 다양한 방법들이 있다. 본 연구에서는 현재 게임 업계에서 가장 널리 쓰이고 있는 방법 중 하나인 모션 매칭[1]과 심층 강화학습을 함께 이용해서, 심층강화학습으로 학습된 정책이 모션 매칭의 쿼리를 출력하여 자연스러운 모션을 생성하는 방법을 제안한다. 이를 통해 캐릭터가 달성해야 할 목표(goal)를 사용자가 지정하는 방식으로 직관적으로 캐릭터를 제어하는 것이 가능하며, 제어되는 캐릭터는 학습 도중 경험하지 못한 상황에도 유연하게 대처할 수 있다.

### 1. 서론

인터랙티브하게 제어 가능한 캐릭터 모션 생성은 컴퓨터그래픽스 학계에서 지속적으로 연구되어 온 주제이다. 자연스러운 결과 모션을 위해 모션캡처데이터를 이용하여 그래프 구성 및 탐색, 최적화 등의 방법들이 사용되어 왔으며, 최근 들어서는 인공지능 기술의 발전에 힘입어 지도학습, 비지도학습, 강화학습을 망라한 다양한 심층학습 기법을 이용하여 모션을 생성하는 연구들이 활발하게 진행되고 있다[2].

그 중 심층강화학습은 그 동안 주로 물리 시뮬레이션되는 캐릭터의 제어를 위해 사용되어 왔으며, 안정적으로 자연스러운 동작을 수행하며 목표를 달성하고 학습 도중에 경험하지 못한 상황에 대해서도 대응하는 동작을 만들어내는 등 좋은 결과를 보여왔다[3]. 물리 시뮬레이션 없이 운동학적으로 기술된 모션데이터만을 이용해 모션을 생성하는 연구에 있어서는 심층강화학습이 많이 사용된 편은 아니나, 최근 들어 이 문제에 있어서도 심층강화학습의 사용이 점차 늘고 있다[4][5].

\* 구두발표논문

\* 본 논문은 요약논문 (Extended Abstract)으로서, 본 논문의 원본 논문은 현재 타 학술대회 (논문지)에 제출 준비 중임.

\* 본 연구는 정부 (과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행되었음 (NRF-2019R1C1C1006778, NRF-2020R1A2C1012847).

본 논문은 게임 업계에서 캐릭터 애니메이션 생성을 위해 가장 널리 쓰이고 있는 방법 중 하나인 모션 매칭 기법과 심층강화학습을 함께 활용해서, 심층강화학습으로 학습된 정책이 모션 매칭의 쿼리를 출력하여 자연스러운 모션을 생성하는 방법을 제안한다. 모션 매칭[1]은 현재 캐릭터의 상태와 앞으로 취해야 할 동작에 대한 정보를 특성 (feature)로 추약하고 일정 간격으로 이를 전처리 과정에서 생성한 특성 데이터베이스에 쿼리로 보내, 저장된 모션 데이터 중 가장 가까운 프레임을 찾아 재생을 이어가는 방법이다. 본 논문은 모션 매칭 쿼리 중에 앞으로 취해야 할 동작에 해당하는 부분의 정보를 심층 강화학습으로 학습된 정책을 통해 출력하도록 하여 캐릭터가 지능적으로 다음 순간의 동작을 찾아 나가며 사용자가 지정하는 목표를 달성하도록 하는 방법을 제안한다.

본 연구에서 제안하는 방법은 전신 모션의 생성을 심층 강화학습으로 학습된 정책을 통해 하는 것이 아니라, 모션 매칭의 쿼리만 정책이 출력하고 실제로 전신 모션은 모션 매칭 기법에 의해 생성되기 때문에, 경량화된 네트워크 구조로 표현되는 정책을 사용할 수 있다. 또한 저차원으로 표현된 포즈의 특성 값을 곧바로 학습하기 때문에 학습 시간 역시 [5]와 같은 기존의 연구들에 비해 짧다. 단순히 목표 위치를 찾아가도록 하는 경우, 일반적인 데스크탑 정도의 사양에서 10분 안에 수렴할 수 있다.

### 2. 모션 매칭

모션 매칭[1]은 전처리 단계에서 모션 데이터베이스의 각 프레임마다 특징적인 요소를 추출해 각 프레임에 대한 특성값을 특성 데이터베이스로 저장한다. 런타임에서는 일정 시간 간격마다 현재 캐릭터의 상태와 도달하고자 하는 목표를 조합한 새로운 특성값을 생성하여 이를 특성 데이터베이스에 쿼리로 보내 이 쿼리와 가장 가까운 특성값을 가지는 동작 프레임을 찾아나가는 식으로 자연스러운 캐릭터 애니메이션을 생성하게 된다. 이 연구에서 활용된 데이터베이스는 기존의 연구에서 사용된 평지를 걸어다니는 동작에 한정된 데이터베이스로[2], 다음 27차원의 특성값을 사용한다.

$$\text{feature}_i = \{p_i^{lfoot}, v_i^{lfoot}, p_i^{rfoot}, v_i^{rfoot}, v_i^{hip}, t_i, d_i\} \in \mathbb{R}^{27}$$

각 요소는 양쪽 발의 위치 (p)와 속도 (v), 캐릭터 루트의 속도 (h), 그리고 현재 프레임으로부터 동일 시간 간격으로 선택된 3개의 미래 프레임에서 가지는 루트의 2차원 위치 (t)와 2차원 속도 (d) 값을 포함한다. 이 연구에서 캐릭터 프레임은 캐릭터가 바라보는 정면 방향, 위를 가리키는 벡터 (Y축), 그리고 이 둘 모두에 수직인 다른 벡터를 세 축으로 하는 프레임으로 정의되며 위의 모든 특성 요소들은 이 캐릭터 프레임에 대해 표현된다. p, v, h는 캐릭터의 현재 상태에 해당하며, t, d는 미래의 정보로, t와 d를 새롭게 조합하면 캐릭터의 미래 동작을 원하는 대로 조작할 수 있다.

### 3. 심층 강화학습

본 연구에서는 앞서 언급한 캐릭터의 미래 3개 프레임에서의 2차원 위치 (t, 6차원)와 방향 (d, 6차원)을 액션으로 출력하도록 심층 강화 학습을 통해 정책을 학습한다. 출력된 액션은 모션 매칭에서 다음 프레임 탐색을 하기 위한 쿼리로 구성된다. 학습된 정책의 출력은 12차원의 벡터이다.

정책의 입력이 되는 상태 (state)는 모션 매칭 특성의 다른 요소들인 현재 캐릭터의 상태 (p,v,h, 15차원)와 달성하고자 하는 목표에 따른 추가적 요소를 이어 붙여 사용한다.

만약 캐릭터가 지면 상의 특정 위치로 이동하도록 하는 것이 목표라면, 현재 캐릭터의 루트 프레임에 대한 목표 위치 (2차원)만을 더한 17차원의 관찰값을 사용하고, 만약 캐릭터가 산재한 장애물들을 피해 목표 위치에 도달하고자 한다면 주변 위치에 대한 정보가 더 필요하다. 이 연구에서는 DeepLoco[6]에서 제안한 네트워크 구조를 차용해 주변 환경에 대한 정보를 압축했다. 캐릭터의 현재 프레임을 기준으로 16\*16개의 정사각형 센서 내로 캐릭터가 처한 환경의 대략적인 정보를 압축할 수 있다.

보상은 캐릭터의 현재 위치와 목표하고자 하는 위치가 멀리 떨어질수록 큰 음수 보상을 부여하며, 최대 1000 스텝 동안 목표를 달성하지 못하면 에피소드는 강제로 종료된다. 또 캐릭터의 주변에 장애물이 있는 경우 캐릭터가 장애물과 부딪칠 때 매우 큰 음수 보상과 함께 에피소드를 강제 종료하도록 했다.

### 4. 실험 결과

강화학습은 rllib 라이브러리의 PPO Trainer를 사용했으며, 16개의 CPU core를 사용했다. 캐릭터만 존재하는 환경에서 목표 위치를 찾아가는 때는 2층 구조의 간단한 (32\*32) 완전연결 계층을 사용했으며 학습은 10분 내로

수렴했다. 캐릭터가 피해가야 할 장애물이 다수 존재하는 경우 캐릭터의 현재 위치를 중심으로 40m\*40m범위 내에 200개의 길이 10m내의 랜덤한 길이의 육면체 장애물을 설정했으며 각 장애물은 0.5m/s보다 느린 속도로 움직이도록 했다. 캐릭터는 주변 6m\*6m내에서 16\*16개의 센서로 주변의 정보를 압축하며, 이 센서의 정보와 캐릭터의 현재 상태를 입력으로 학습에 필요한 정보를 압축한다. 캐릭터는 주어진 정보를 활용해 최종적으로 목표 위치에 도달하기까지 필요한 행동을 취할 수 있게 학습된다.

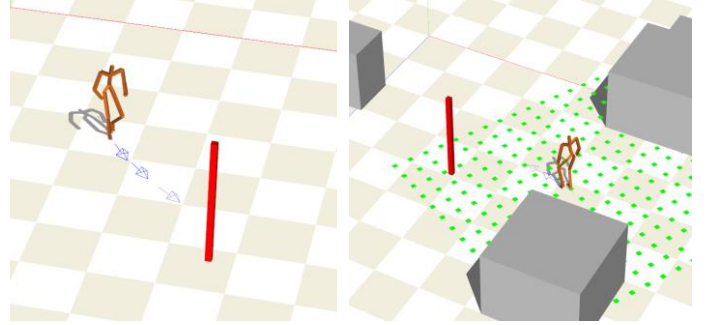


Figure 1: Experimental results.

목표 위치만 있는 경우, 캐릭터는 목표 위치의 방향으로 크게 몸을 돌려 뛰다가 근처에 오면 걸어서 정확한 위치를 찾아가며, 장애물이 있는 경우 장애물을 피해서 목표 위치에 닿을 수 있다.

### 5. 결론

이 연구에서는 게임 산업에서 캐릭터 동작 생성에 널리 쓰이는 방법인 모션 매칭을 이용해서 보다 쉽고 간단한 구조의 강화학습으로 목표 위치에 도달하는 캐릭터 동작을 생성한다.

### 참고문헌

- [1] S. Clavet, "Motion matching and the road to next-gen animation", 2016
- [2] Daniel Holden, Taku Komura, and Jun Saito, "Phase-Functioned Neural Networks for Character Control", in SIGGRAPH 2017, 2017
- [3] Xue bin Peng, Pieter Abbeel, Sergey Levine, and Michiel van de Panne, "DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills", in SIGGRAPH 2018, 2018
- [4] Hung Yu Ling, Fabio Zinno, George Cheng, and Michiel van de Panne, "Character Controllers using Motion VAEs", in SIGGRAPH 2020, 2020
- [5] Kyungmin Cho, Chaelin Kim, Jungjin Park, Joonkyu Park, and Yunyong Noh, "Motion Recommendation for Online Character Control", in SIGGRAPH ASIA 2021, 2021
- [6] Xue Bin Peng, Glen Berseth, Kangkang Yin, and Michiel van de Panne, "DeepLoco: Dynamic Locomotion Skills Using Hierarchical Deep Reinforcement Learning", in SIGGRAPH 2017, 2017