신체 부착 마커 기반 실시간 3 차원 인간 자세 복원에 대한 연구

김정수, 최성화, 이성민, 장민규, 박예승, 이상훈
연세대학교

(integer528, csh0772, lseong721, jmg1002, pys940617, slee)@yonsei.ac.kr

A Study on reconstructing real-time 3D human body pose based on body attached markers

Kim Jungsu, Choi Seonghwa, Lee Seongmin, Jang Mingyu, Park Yeseung, Lee Sanghoon
Yonsei Univ.

요 약

오늘날 인간 자세 복원에 대한 연구는 크게 발전하고 있으며, 다양한 산업에서 위 분야에 대한 요구사항 또한 많아지고 있다. 기존 다양한 게임 및 그래픽 산업에서 3 차원 사람 자세를 복원할 때는 사람에게 전신 마커를 부착한 슈트를 착용하게 한 후 이에 대한 데이터를 취득하였다. 하지만, 이러한 슈트 착용은 실험 진행자가 데이터 측정을 함에 있어 움직임에 제약을 주며, 측정에 더 많은 비용을 야기하게 한다. 이에, 본 논문은 7 개의 신체 정보만 주어졌을 때도 이를 통해 사람의 온전한 자세를 복원할 수 있는 방법을 제시한다. 제시한 딥러닝 모델을 통하여 학습 및 복원을 진행했을 때, SMPL(Skinned Multi Person Linear Model)을 이용하여 실시간으로 자세 복원이 이루어지도록 모델 구성을 하였으며, 머리, 등, 배, 오른손, 왼손, 오른발, 왼발에 있는 데이터만으로 전신의 자세를 정확하게 복원하는 결과를 얻었다.

Ⅰ. 서 론

3 차원 인간 자세 복원 기법은 게임, 그래픽 등 다양한 분야에 많은 수요가 있다. 다양한 그래픽 기반 콘텐츠 및 가상현실, 증강현실에서 인간은 항상 콘텐츠의 중심에 있으며, 그들의 행동을 재현하는 것이 어색하거나 재현의 품질이 낮을수록 콘텐츠 자체의 가치가 떨어지게 될 수 있기 때문이다.

이러한 이유로 3D 인간 자세 복원을 위해 다양한 연구가 진행되고 있으며, 크게 나누어 사람에게 3D 마커를 부착하여 관련 데이터를 얻는 연구와 일반 이미지에서 사람의 자세를 복원하는 연구로 크게 나눌 수 있다. 이 중 3D 마커를 부착하여 자세 복원을 하는 것은 높은 정확도 및 실시간성을 보장하는 방법이며, Optitrack 과 Vicon 이 대표적으로 이러한 기술을 갖고 있다[1, 2]. 하지만 많은 수의 마커를 사용자의 신체에 부착할 경우 사용자에게 움직임 제약을 통한 원활한 데이터 취득에 방해가 될 수 있다.

이에 본 논문에서는 머리, 등, 배, 양손, 양발을 통한 총 7 개의 신체 정보만 갖고 있을 때도 신체의 모든 자세 3 차원 인간 모델 템플릿인 SMPL(Skinned Multi Person Linear Model)을 이용하여 복원할 수 있는 딥러닝 모델을 제안한다.

Ⅱ. 본론

본 연구에서는 머리, 등, 배, 양손, 양발로 하여 총 7 개의 신체 정보에서 전신 자세를 복원한다. 또한, 이 정보는 기존 산업에서 사용되는 마커와 동일하게 6 DoF(Degree of Freedom)의 정보를 활용한다. 이를 활용하여 전신의 자세를 SMPL 형태로 복원하는 기법을 제안한다[3]. 이 중 신체 정보를 훈련하기 위해 구성한 데이터셋은 그림 1 의 형태로 제작 및 구성되었다.

그림 1 AMASS DB 를 이용한 데이터셋 구성

데이터셋 구성의 경우 그림 1 의 형태로 구성을 하였다. AMASS Database 의 ACCAD 서브 데이터셋을 이용하여 데이터베이스의 인간 자세에 대한 SMPL 데이터를 얻은 후, 머리, 등, 배, 양손, 양발로 하여 총 7 개의 신체 정보에 해당하는 위치 값을 추출하였다[4]. 이를 통해 7 개 신체 정보에 대한 3 차원 위치 값과 3 차원 회전 값을 입력 정보로 추출하여 데이터셋으로 구성하였다. 이를 통해 총 7x6 형태의 값을 입력 정보로 구성하였다.

이러한 데이터 구성 이후에는 전반적인 네트워크 구성을 통한 딥러닝 학습 및 SMPL 템플릿을 이용한

자세 복원을 진행하였다. 자세한 네트워크 구조를 그림 2 를 통해 표현하였다.

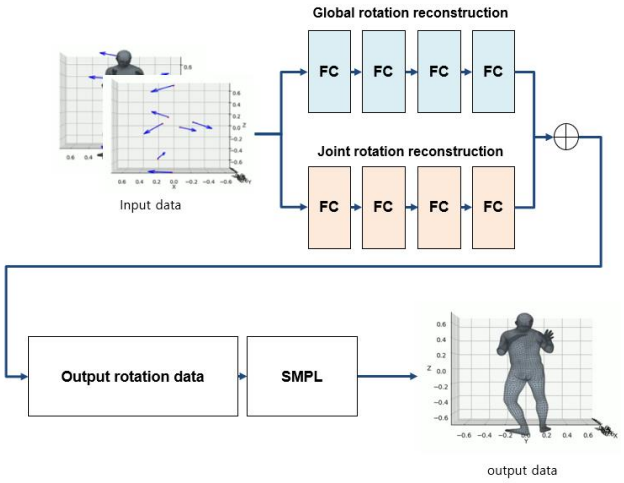

그림 2 자세 복원 네트워크 구조

그림 2 는 자세 복원 네트워크의 전반적인 구조이다. SMPL 템플릿의 경우 자세 복원을 위해 1x3 형태의 global rotation 파라미터를 통해 사람 전체의 회전 상태를 정의하고, 23x3 형태의 joint rotation 파라미터를 통해 각 관절의 회전 상태를 정의한다. 위 네트워크 또한 SMPL 템플릿에 맞춰 네트워크 구성을 진행하였다. 위 구조의 경우 자세의 global rotation 파라미터를 복원하기 위한 네트워크와 joint rotation 파라미터 정보를 얻기 위한 네트워크를 따로 구성하였다.

상단에서는 global rotation 복원을 위해 42 개의 입력 정보를 받고, 이를 Fully Connected Layer(FC)를 4 개를 거치게 하였다. Fully Connected Layer 의 경우 4 번을 거치게 하며, 각각의 네트워크가 출력을 입력으로 받고, 다시 출력을 하는 형태로 구성을 하여 입력 정보의 특징이 더욱 강화되어 나타날 수 있도록 보정하는 역할을 진행하였다. Global rotation 복원의 경우 각 joint rotation 복원보다 객관적, 주관적 평가 시 더 크게 영향을 끼치므로 따로 브랜치를 분리하여 학습을 진행하였다.

하단에서는 joint rotation 복원을 위해 42 개의 입력 정보를 받고, global rotation 파라미터 복원과 동일하게 진행하였다. 이를 통해 4 개의 Fully Connected Layer 를 지나게 한 후, 23x3 형태로 joint rotation 값을 얻게 하였다. 이를 통해 그 후, 두 값을 서로 더하여 24x3 형태로 값을 얻을 수 있도록 하였다. 이후, SMPL 템플릿을 이용하여 위 회전 값을 사람 자세로 변환하는 과정을 거쳐 사람 자세를 복원하였다.

각각의 모듈, 즉 joint rotation 부분과 global rotation 부분은 서로 연관되어 있지 않으며 독립적으로 수행된다. 이에 각각 얻은 결과값에 대해 Fully Connected Layer 를 통한 추가적인 과정을 거치지 않고, 단순히 두 행렬을 더하여 24x3 형태로 나타나게 하였다. 이를 통해 SMPL 자세 복원을 위한 회전행렬의 24x3 형태를 얻었으며, 이를 SMPL 템플릿에 넣어 자세 복원을 진행하였다.

위와 같이 모델 학습 및 결과를 얻은 후, 그림 3 을 통해 복원된 자세를 살펴보는 과정을 거쳤다. SMPL 형태로 자세 복원 후, 제대로 된 자세 복원을 더 자세히 확인하기 위해 오리지널 데이터와 복원된 데이터의 표면을 구성하는 과정을 거쳤다. 복원된 자세를 살핀 결과, 그림 3 과 같이 대부분의 자세에서 정상적으로 복원이 이루어진다. 전체 자세의 복원과 각 관절의 복원 모두 적절한 형태로 이루어지는 것을 확인하였으며, 복원에 걸리는 시간 또한 실시간으로 이루어지는 것을 확인하였다.

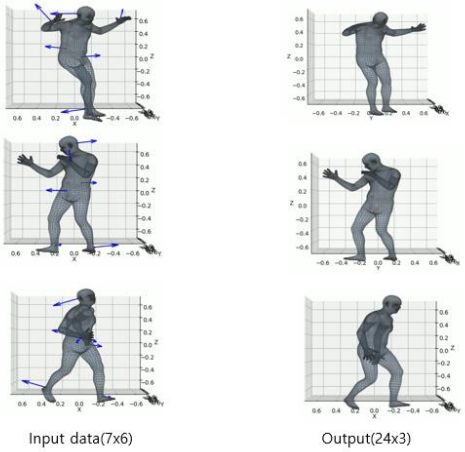

그림 3 인간 자세 복원 예시

Ⅲ. 결론

본논문에서는 제한된 신체 정보를 토대로 인간의 자세를 복원하는 연구를 다루었다. 이러한 기술을 통해 데이터 취득자는 데이터를 얻을 때 더 적은 마커를 이용해 사람 자세를 복원할 수 있고, 이에 더 자유로운 자세가 가능하며 움직임의 제약 또한 적어지게 될 것으로 기대한다.

ACKNOWLEDGMENT

이 성과는 2020 년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2020R1A2C3011697)

참 고 문 헌

[1] Optitrack, (https://optitrack.com/)

[2] Vicon (https://www.vicon.com/)

[3] Loper, Matthew, et al. "SMPL: A skinned multi-person linear model." ACM transactions on graphics (TOG) 34.6 (2015): 248

[4] Mahmood et al. " AMASS: Archive of Motion Capture as Surface Shapes" International Conference on Computer Vision, 2019