

자가 접촉을 고려한 사람 자세 및 체형 예측 시스템에 관한 연구

최문경, 허정우, 이경준, 이상훈
연세대학교

{bryan1302, gjwjddn9, naive2kj90, slee}@yonsei.ac.kr

A Study on the Pose and Shape Regression System Considering Self-Contact

Choi Moon Kyeong, Huh Jung Woo, Lee Kyung June, Lee Sang Hoon
Yonsei Univ.

요약

한 사람의 서로 떨어져 있는 두 신체 부위 간의 접촉은 일상적인 상황에서 빈번하게 발생한다. 그러나 이미지에서 사람의 체형과 자세를 예측하는 기존 모델은 자가 접촉이 있는 이미지에 대해서는 성능이 감소하는 경향이 있다. 본 논문은 사람 자세 및 체형 예측 모델이 예측한 사람의 접촉 영역이 실제 접촉 영역과 일치하도록 회귀자를 학습시키는 새로운 방법을 제안한다. 이는 자가 접촉이 발생하더라도 정확한 자세와 형태를 예측할 수 있도록 한다. 또한, 접촉 영역을 이산적으로 표현하는 대신 연속적으로 나타내기 위해 '접촉 지도'를 제안한다. 접촉 지도를 통해 각 신체 부위에 대한 접촉 여부 뿐만 아니라 접촉과 가까운 정도를 나타내어 보다 풍부한 특징을 추출할 수 있다. 실험을 통해 접촉 지도를 사용해 학습된 사람 자세 및 체형 예측 모델이 더 정확히 사람 자세와 체형을 예측할 수 있다는 것을 보였다.

I. 서론

한 사람의 서로 떨어져 있는 두 신체 부위 간의 접촉은 일상적인 상황에서 빈번하게 발생한다. [5] 예를 들어, 손으로 얼굴을 만지거나 다리를 꼬는 등의 행동이 이에 해당한다. 그러나 기존의 이미지에서 사람의 체형과 자세를 예측하는 모델은 자가 접촉이 있는 이미지에 대해서는 성능이 감소한다는 문제가 존재한다. 이는 기존 모델에서 사용된 손실 함수가 한 지점과 다른 지점 사이의 상대적 거리를 고려하지 않았기 때문으로 해석된다. 이는 결과적으로 사람 자세 및 체형 예측 모델이 자가 접촉이 발생하는 두 신체 부위가 서로 관통하거나 분리되는 체형과 자세를 예측하도록 만들었다.



그림 1 자가 접촉이 있는 이미지

이러한 문제를 해결하기 위해 본 논문에서는 사람 자세 및 체형 예측 모델이 예측한 사람의 접촉 영역이 실제 접촉 영역과 일치하도록 예측 모델을 학습시키는 새로운 방법을 제안한다. 예측 모델은 예측한 사람의 접촉 영역과 실제 접촉 영역 사이 차이를 최소화하도록

학습된다. 이는 예측 모델이 자가 접촉이 발생하더라도 정확한 자세와 형태를 예측할 수 있도록 한다.

또한, 접촉 영역을 이산적으로 표현하는 대신 연속적으로 나타내기 위해 '접촉 지도'를 제안한다. 접촉 지도를 통해 각 신체 부위에 대한 접촉 여부 뿐만 아니라 접촉과 가까운 정도를 나타내어 보다 풍부한 특징을 추출할 수 있다. 또한 접촉 지도를 통해 접촉 영역을 3 차원이 아닌 2 차원으로 나타내어 연산량을 감소시킨다.

제안하는 방법으로 학습된 우리 모델을 기존 사람 자세 및 체형 예측 모델과 예측 정확도를 비교하였다. 우리 모델은 두 평가지표 모두에서 더 낮은 오류율을 얻었다. 이를 통해 제안하는 방법으로 학습된 사람 자세 및 체형 예측 모델이 더 정확히 사람 자세와 체형을 예측할 수 있다는 것을 보였다.

II. 본론

본 논문에서 제안하는 접촉 지도와 유사하게, 손과 물체 사이 상호작용에 관한 논문[1]에서 접촉 지도가 사용되었다. 이는 모든 물체의 포인트에 대해 가장 가까운 손 포인트 사이 거리를 2 차원 지도로 나타내었다. 그러나 사람은 손, 물체와 달리 하나의 메쉬로 이루어져 있어 같은 방법으로 사람의 자가 접촉을 나타내기 어렵다. 이를 위해 TUCH[2]에서 제안한 접촉의 정의를 변형하여 접촉과 가까운 정도를 정의하였다. TUCH 는 특정한 자세의 사람 메쉬에서 유클리드 거리가 충분히 가까운 두 정점이 측지선 거리 (표면을 따라 이동할 때의 거리)가 충분히 먼 경우를 접촉으로 정의하였다.

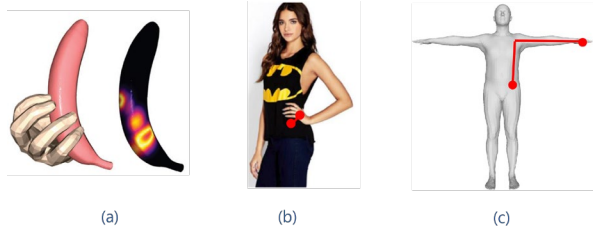


그림 2 (a) 손과 물체 사이 접촉 지도 (b) 두 정점 사이 유클리드 거리 (c) 두 정점 사이 측지선 거리

이러한 정의를 변형하여, 한 사람의 정점에서부터 측지선 거리가 충분히 큰 정점 중 유클리드 거리가 가장 가까운 정점까지의 거리를 접촉과 가까운 정도라고 정의하였다. 주어진 사람 메쉬의 모든 정점에 대해 접촉과 가까운 정도를 계산한 뒤 사람 메쉬 UV 맵 상에 나타내었고, 이를 접촉 지도라고 정의하였다. 그림 3 은 주어진 사람의 메쉬, 해당 메쉬의 접촉 지도, 접촉 지도를 텍스처로 입힌 사람 메쉬를 나타내고 있다. 접촉 지도에서 색깔이 어두울수록 접촉에 가깝다는 것을 의미한다.

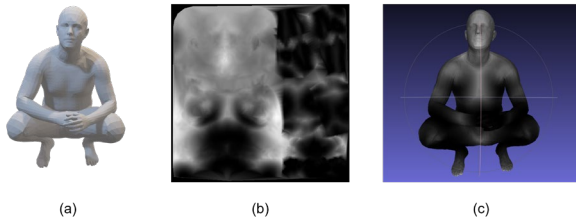


그림 3 (a) 주어진 사람 메쉬 (b) 해당 메쉬의 접촉 지도 (c) 접촉 지도를 텍스처로 입힌 사람 메쉬

접촉 지도를 사용해 사람 자세 및 체형 예측 모델을 학습시키는 전체 구조는 다음과 같다. 먼저 주어진 이미지로부터 사람의 자세와 체형을 예측한다. 그 다음 접촉 지도 디코더를 사용해 사람의 자세와 체형으로부터 접촉 지도를 예측한다. 마지막으로 예측한 접촉 지도와 정답 접촉 지도의 차이를 손실 함수로 사용, 차이를 최소화한다. 학습 과정에서 접촉 지도 디코더가 사람의 자세와 체형으로부터 접촉 지도를 예측할 수 있도록 AMASS 데이터셋[6]을 사용해 미리 학습시켰다.

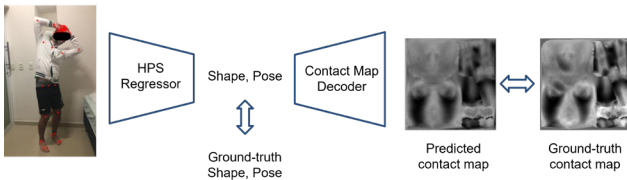


그림 4 제안하는 사람 자세 및 체형 예측 모델 학습 구조

예측한 접촉 지도와 정답 접촉 지도의 차이를 손실 함수로 사용하여 학습된 우리 모델과 기존 사람 자세 및 체형 예측 모델 SPIN[3]의 예측 정확도를 비교하였다. 테스트 데이터셋으로는 51,000 개의 프레임과 프레임 속 사람의 체형과 자세를 포함하고 있는 3DPW[4] 데이터셋을 사용하였다. 평가 지표로는 Mean Per-Joint Position Error (MPJPE), Procrustes-aligned version (PA-MPJPE)를 사용하였다. 우리 모델은 두 평가지표 모두에서 더 낮은 오류율을 얻었다. 이를 통해 접촉 지도를 사용해 학습된 사람 자세 및 체형 예측 모델이

더 정확히 사람 자세와 체형을 예측할 수 있다는 것을 보였다.

	MPJPE (mm)	PA-MPJPE (mm)
SPIN	96.9	59.2
Ours	81.5	53.4

III. 결론

본 논문에서는 사람 자세 및 체형 예측 모델이 예측한 사람의 접촉 영역이 실제 접촉 영역과 일치하도록 회귀자를 학습시키는 새로운 방법을 제안하였다. 또한, 접촉 영역을 이산적으로 표현하는 대신 연속적으로 나타내기 위해 '접촉 지도'를 제안하였다. 접촉 지도를 통해 각 신체 부위에 대한 접촉 여부 뿐만 아니라 접촉과 가까운 정도를 나타내었다. 실험을 통해 접촉 지도를 사용해 학습된 사람 자세 및 체형 예측 모델이 더 정확히 사람 자세와 체형을 예측할 수 있다는 것을 보였다.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1A2C3011697) and the Yonsei Signature Research Cluster Program of 2023 (2023-22-0008).

참 고 문 헌

- [1] Jiang, Hanwen, et al. "Hand-object contact consistency reasoning for human grasps generation." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.
- [2] Muller, Lea, et al. "On self-contact and human pose." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
- [3] Kolotouros, Nikos, et al. "Learning to reconstruct 3D human pose and shape via model-fitting in the loop." Proceedings of the IEEE/CVF international conference on computer vision. 2019.
- [4] Von Marcard, Timo, et al. "Recovering accurate 3d human pose in the wild using imus and a moving camera." Proceedings of the European conference on computer vision (ECCV). 2018.
- [5] Kwok, Yen Lee Angela, Jan Gralton, and Mary-Louise McLaws. "Face touching: a frequent habit that has implications for hand hygiene." American journal of infection control 43.2 (2015): 112-114.
- [6] Mahmood, Naureen, et al. "AMASS: Archive of motion capture as surface shapes." Proceedings of the IEEE/CVF international conference on computer vision. 2019.