

3D 모델 기반의 3D Pose Estimation의 성능 향상 알고리즘

*이솔 *박정탁 *박병서 *서영호

*광운대학교

*solls1124@kw.ac.kr *littlepine97@kw.ac.kr *bspark@kw.ac.kr *yhseo@kw.ac.kr

Performance Enhancement Algorithm of 3D Pose Estimation
based on 3D Model

*Sol Lee *Jung-Tak Park *Byung-Seo Park *Young-Ho Seo

*Kwangwoon University

요약

본 논문에서는 Openpose의 신뢰도를 이용해 3D pose estimation의 정확도를 높이는 방법을 제안한다. 모델의 앞뒤양옆 네 방향에서 pose estimation의 진행하기 위해 3D 모델에 AABB(Axis Aligned Bound Box)를 생성한 다음, box의 네 옆면으로 모델을 투영시킨다. 각 면에 투사된 2D image에 대해 Openpose 2D pose estimation의 진행한다. 네 면에서 생성한 2D 스켈레톤들의 평균을 통해 3D 상의 교차점을 획득한다. Openpose에서 제공하는 신뢰도(confidence)를 이용하여 잘못 나온 2D 관절을 제외하는 것으로 더 정확한 pose estimation의 수행하였다. 실험적인 방법을 통해 신뢰도 0.45 이상의 값을 가지는 joint만을 사용해 3D 교차점을 구함으로써 3D pose estimation의 정확도를 높였다.

1. 서론

Pose estimation이란 컴퓨터 비전에서 사람의 신체 관절이 어떻게 구성되어있는지 위치를 추정하는 문제이다. 결과 데이터의 종류에 따라 2D pose estimation과 3D pose estimation로 나눌 수 있으며, 본 논문에서는 2D pose estimation의 결과를 정합해 3D로 만든다. 2D pose estimation은 Openpose를 이용하는데, Openpose는 이미지에서 사람의 신체 위치뿐만 아니라 해당 위치값이 얼마나 신뢰할 수 있는 가인 신뢰도를 함께 제공한다. 가려진 신체 부위에 대해서는 추측을 하지 않거나 추측을 하더라도 매우 낮은 신뢰도 값을 제공하며, 가려지지 않은 부위에 대해서는 높은 신뢰도 값을 제공한다. 신뢰도 값을 반영하지 않고 Openpose에서 제공한 모든 값을 전부 사용할 경우, 잘못 나온 값이 계산에 사용돼 수 cm 이상의 차이를 만든다. 또한, 신뢰도 값의 특성을 사용하지 않으면 이미지의 방향을 수동으로 구분하는 번거로움이 있다. 신뢰도 값을 사용할 경우, 잘못 나온 점을 계산에서 제외시킬 수 있기 때문에 더 정확도가 높은 pose estimation이 가능하다. 나아가 여러 방향에서 신뢰도의 특성이 다르기 때문에, 이를 이용해 자동으로 앞뒤양옆 이미지 분류가 가능하다.

2. 제안한 알고리즘

본 논문에서는 카메라 시스템을 통해 얻은 RGB, Depth 이미지 데이터를 사용한다. 해당 데이터에 대해 Openpose에서 제공하는 신뢰도

를 분석한 결과 신뢰도 0.45를 기준으로 정확도에 차이를 보였다. 따라서 신뢰도의 기준을 0.45로 실험적으로 결정했다. 해당 신뢰도의 기준은 사용한 3D 모델, 포즈, Openpose model의 트레이닝 횟수 등에 따라 다를 수 있다. 본 논문에서 제안하는 3차원 스켈레톤 추출 방식은 [1]의 알고리즘을 따른다. 본 논문에서는 [1]의 방법에서 조인트를 결정하는 알고리즘을 개선한다. 제안한 알고리즘은 그림 1과 같다.

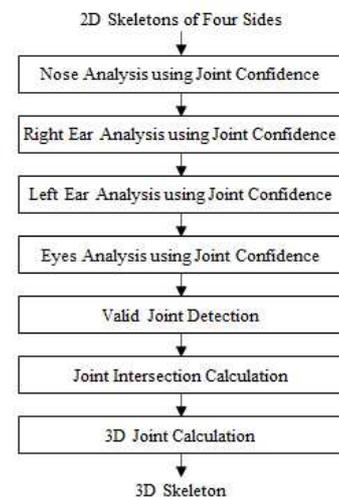


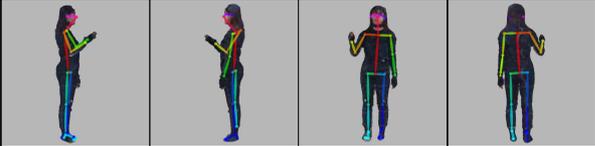
그림 1. 제안한 조인트 결정 알고리즘

먼저 3D model에 AABB(Axis Aligned Bound Box)를 생성한 다음, box의 네 옆면으로 모델을 투영시킨다. 각 면에 투사된 multi-view image에 대해 Openpose를 이용하여 2D pose estimation로 진행한다. 관절의 위치와 함께 제공받은 신뢰도를 이용하여 잘못 나온 2D 관절을 제외함으로써 좀 더 정확한 pose estimation로 가능하다. 신뢰도가 0.45보다 낮은 관절을 제외한 뒤, 남은 관절들만을 가지고 3D 교차점을 구한다. 그렇게 구한 교차점이 3D 모델 내부에 위치하지 않을 경우, 후처리를 통해 3D 관절이 3D 모델 내부로 이동시킨다.

3. 실험 결과

실험에는 한 팔을 든 채 가만히 서있는 자세로 촬영한 10 프레임의 3D 데이터를 사용했다. 해당 데이터의 한 프레임에 대해 네 방향에서 2D pose estimation한 결과를 그림 2에 나타내었다. 정면 영상을 보면 한 팔은 들고 다른 팔은 내리고 있다. 데이터의 오른쪽에서 본 영상에서는 몸통에 반대쪽 팔이 가려져 양 팔을 다 들고있다고 잘못 추측하였다.

정면에서 나온 신뢰도 값은 평균 76.92%이며, 그에 비해 옆면에서 잘못 추측한 팔꿈치(joint 6)와 손목(joint 7)의 신뢰도 값은 각각 35.13%, 41.33%로 상대적으로 매우 낮다. 10 프레임의 데이터에 대해 잘못 추측한 관절의 신뢰도의 최대값을 신뢰도의 기준으로 삼았다.



joint	X	Y	confidence									
0	508.08	123.39	83.79%	494.65	120.70	85.30%	529.98	123.29	85.32%	0	0	0
1	426.34	221.54	70.70%	579.10	227.00	75.59%	532.65	243.41	82.21%	469.93	229.78	70.17%
2	423.47	226.95	72.52%	609.14	210.72	59.92%	445.49	237.95	77.56%	554.60	227.02	52.23%
3	469.93	355.35	80.59%	513.60	358.07	45.39%	379.87	360.91	87.23%	628.33	358.09	60.78%
4	565.47	292.52	67.46%	442.65	300.78	70.90%	358.17	289.84	88.46%	641.83	281.62	50.73%
5	442.62	213.36	54.56%	554.47	235.29	67.38%	614.64	246.19	74.81%	385.43	232.52	40.24%
6	502.71	319.82	35.13%	535.61	385.44	61.66%	628.29	371.72	71.25%	374.46	377.20	26.12%
7	579.09	273.38	41.33%	497.33	516.46	70.79%	644.64	508.18	79.72%	360.87	491.75	19.76%
8	453.64	491.81	53.20%	562.78	486.37	54.94%	513.78	502.70	63.57%	486.34	494.58	48.12%
9	445.38	494.51	56.19%	576.43	480.91	40.48%	464.53	499.97	59.43%	426.42	500.03	41.60%
10	437.25	704.67	65.00%	570.94	691.04	45.16%	467.15	707.37	75.73%	431.77	704.75	60.09%
11	434.52	906.78	65.88%	581.94	852.12	5.45%	467.17	884.82	80.61%	434.53	879.41	54.58%
12	469.96	491.84	46.78%	554.65	491.78	56.25%	573.69	505.46	59.31%	540.92	491.80	45.48%
13	450.85	693.83	48.72%	557.24	701.95	68.43%	562.78	707.39	72.66%	538.26	707.35	52.23%
14	442.66	895.75	18.54%	570.98	906.75	74.92%	560.08	860.36	73.86%	535.52	884.82	28.16%
15	491.74	115.09	91.00%	0	0	0	510.90	104.23	83.31%	0	0	0
16	0	0	0	510.85	101.52	88.27%	551.81	109.66	80.53%	0	0	0
17	445.49	117.89	90.87%	0	0	0	489.00	120.56	85.82%	510.97	120.54	30.27%
18	0	0	0	554.64	112.34	80.11%	579.23	126.01	80.11%	418.06	120.64	31.92%

그림 2. multi-view에서의 관절의 위치와 신뢰도 값

또한 그림 2의 정면에서는 모든 관절에 대해 높은 신뢰도 값을 보였고, 뒷면에서는 귀(joint 17, 18)의 위치는 낮은 신뢰도로 추측하지만, 눈(joint 15, 16)의 위치는 특정하지 못했다. 옆면에서는 모델의 오른쪽에서 영상을 촬영한 경우에는 모델의 왼쪽이 나오지 않기 때문에 왼쪽 눈(joint 16)과 귀(joint 18)의 위치를 특정하지 못해 신뢰도가 0이 나왔고, 반대쪽에서 찍은 경우에도 동일하게 적용되었다. 이러한 방향에 따른 신뢰도의 특성을 이용해 자동으로 면을 분리하고 3D 교차점을 구했다.

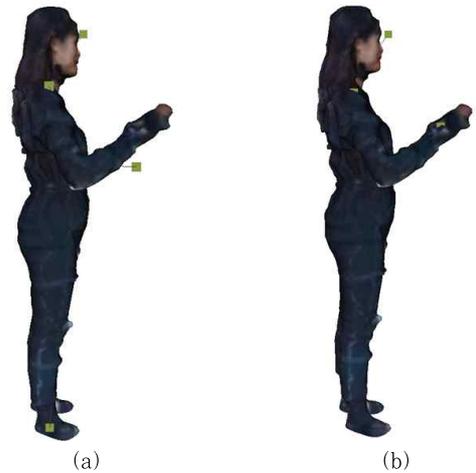


그림 3. 신뢰도를 반영하기 전(a)과 후(b)의 3D 조인트 결과

그림 3은 신뢰도가 0.45 미만인 2D skeleton을 제외하기 전과 후의 3D 조인트이다. 들고있는 팔과 내린 팔의 좌우를 정확히 구분하지 못하여 잘못 추측한 경우에는 신뢰도가 크게 떨어진다. 이때 잘못 추측한 관절을 제외하지 않고 3D 교차점을 구한 경우에는 그림 3(a)와 같이 원래의 3D 모델과 크게 차이가 날 수 있다. 반면 잘못 추측한 관절을 제외한 경우, 그림 3(b)와 같이 좀 더 정확한 3D 교차점을 구할 수 있다. 그림 4는 신뢰도가 0.45 미만인 관절을 제외하기 전과 후의 10프레임 데이터에 대한 3D 조인트이다.

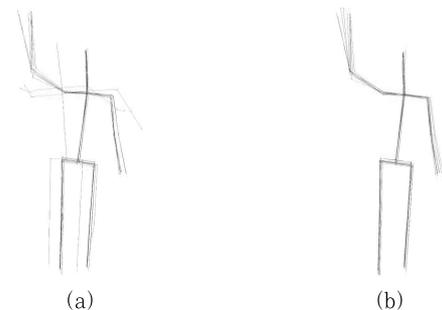


그림 4. 신뢰도를 반영하기 전(a)과 후(b)의 10프레임의 3D 조인트 결과

4. 결론

실험 데이터에 대해 Openpose가 제공하는 신뢰도를 분석해 신뢰도의 기준을 정할 수 있다. 일정 신뢰도 이상의 관절 값만을 계산에 사용함으로써 3D pose estimation의 정확도를 향상시켰다.

Acknowledgement

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학CT연구센터지원사업의 연구결과로 수행되었음 (IITP-2021-0-01846)

Reference

[1] 김경진, 박병서, 김동욱, and 서영호. “다시점 카메라 시스템을 이용한 고정밀 스켈레톤 추출 기법.” 한국방송미디어공학회 학술대회논문집(2020): 297-299