

자세 추정을 위한 모션 캡처 데이터 복원

윤여수 · 박현준*

청주대학교

Restoring Motion Capture Data for Pose Estimation

Yeo-su Youn · Hyun-jun Park*

Cheongju University

E-mail : yyso1@cju.ac.kr / hyunjun@cju.ac.kr

요 약

자세 추정을 위한 모션 캡처 데이터 파일에는 주변 환경과 움직임의 정도에 따라 부정확한 데이터가 존재할 수 있으므로, 이를 보정하는 작업이 필요하다. 기존에는 직접 후처리 과정을 통해 부정확한 데이터를 복원하였으나, 최근에는 자동화된 방법으로 LSTM, R-CNN 등 다양한 종류의 신경망을 사용한다. 하지만 신경망 기반의 데이터 복원 방법들은 컴퓨터 자원을 많이 요구하므로, 본 논문에서는 신경망 기반의 방법보다 자원 사용량은 낮추면서 데이터 복원율은 유지하는 방법을 제안한다. 제안하는 방법은 자세 측정 데이터(c3d)를 활용하여 부정확한 자세 데이터를 자동으로 복원한다. 실험 결과, 데이터의 부정확한 정도에 따라 89%에서부터 99% 정도의 데이터 복원율을 보였다.

ABSTRACT

Motion capture data files for pose estimation may have inaccurate data depending on the surrounding environment and the degree of movement, so it is necessary to correct it. In the past, inaccurate data was restored with post-processing by people, but recently various kind of neural networks such as LSTM and R-CNN are used as automated method. However, since neural network-based data restoration methods require a lot of computing resource, this paper proposes a method that reduces computing resource and maintains data restoration rate compared to neural network-based method. The proposed method automatically restores inaccurate motion capture data by using posture measurement data (c3d). As a result of the experiment, data restoration rates ranged from 89% to 99% depending on the degree of inaccuracy of the data.

키워드

pose estimation, motion capture, c3d data, data restoration, human pose data

1. 서 론

모션 캡처(motion capture)란 사람이나 물체의 움직임을 기록하는 과정이다. 모션 캡처는 측정 방식에 따라 대상에 마커(센서)를 부착시키거나, 적외선을 이용하는 등의 방법으로 대상의 움직임을 디지털 형태로 기록하는 방식인 광학 시스템과 관성

센서와 자기 시스템 등을 이용하여 대상의 움직임을 측정하는 비광학 시스템으로 분류되어있다. 최근에는 대상에 마커를 붙이지 않고 움직임을 측정할 수 있는 마커리스 방식도 사용되고 있다[1].

모션 캡처는 동작, 관절의 움직임, 표정 등을 3D 데이터로 추출하여 빠르고 자연스럽게 대상의 움직임을 데이터화 할 수 있다. 모션 캡처는 비디오 게임, 애니메이션, 영화 제작, 보행 교정, 운동선수

* corresponding author

들의 자세 교정 등에 이용되고 있다.

모션 캡처가 대상의 움직임을 쉽고 빠르게 데이터화 할 수 있는 장점이 있지만, 단점도 존재한다. 측정 방식의 제한된 환경, 완벽하지 않은 데이터 값, 정밀한 측정을 위한 기기 비용 등의 문제점이 있다. 최근에는 제한된 환경과 기기 비용 등의 문제를 해결하기 위해 딥러닝을 이용한 마커 리스 방식이 주목받고 있지만[2, 3], 영화 제작이나 애니메이션 제작 같은 정밀한 모션 캡처 데이터가 필요한 상황에서는 광학식 방식을 이용한 측정 방식을 사용하고 있다. 하지만 광학식 방식을 사용하여 데이터를 획득하더라도 값이 완벽하지 않으므로, 전문가들이 이를 후보정하여 복원하는 과정을 거친다.

본 논문에서는 사람의 움직임을 측정하는 모션 캡처 데이터의 부정확한 데이터 값 복원 작업을 자동화하는 방법을 제안한다. 먼저 소실된 데이터가 있는지 판단 후 이를 복원한다. 복원된 데이터는 각 프레임의 마커마다 움직인 거릿값을 측정하여, 한 프레임에서 신체가 이동할 수 있는 최대 거릿값 이상이면 소실 처리 후 다시 복원하는 과정을 거친다. 결과적으로 원본 데이터를 변조하기 전과 변조 후 복원한 데이터의 차이를 통해 복원율이 유의미한 결과값을 보여주는 것을 확인하였다.

II. 모션 캡처 데이터 복원 방법

2.1 소실된 데이터 복원

먼저 모션 캡처 데이터에서 소실 처리된 데이터가 있는지 판단한다. 이때 소실 처리된 데이터는 모션 캡처 데이터를 측정할 때, 마커가 측정되지 않아 데이터에 빈 값, 혹은 0으로 들어간 값을 의미한다. 식 (1)은 소실 처리된 데이터 x 가 존재할 때, 데이터 x 의 이전 프레임값과 i 번째 이후 프레임의 소실되지 않은 값을 이용하여 속도 v 를 구하는 식이다.

$$v_{frame} = (x_{frame+i} - x_{frame-1}) / (i+1) \quad (1)$$

소실된 데이터가 이전 프레임의 데이터에서 i 번째 이후에 있는 데이터라면, 속도 v 에 i 를 곱해준 후 이전 프레임의 데이터에 더해준 값으로 복원한다. 그림 1은 소실된 데이터를 이전 프레임과 다음 프레임의 데이터를 이용하여 복원하는 방법의 예시를 보여준다. 초록색 원들은 부착된 마커들을 의미하고 점선으로 이루어진 원은 소실된 마커를 의미한다.



그림 1. 소실된 데이터 복원 방법

2.2 거리를 벗어난 데이터 복원

소실된 데이터를 복원한 데이터는 복원 데이터, 첫 번째 프레임부터 마지막의 이전 프레임까지의 데이터는 이전 데이터, 두 번째 프레임부터 마지막 프레임까지의 데이터는 다음 데이터라고 정한다. 신체가 한 프레임 동안 이동한 거릿값 데이터는 다음 데이터와 이전 데이터를 이용하여 구한다. 복원 데이터에서 거릿값의 크기가 임곗값 이상인 위치값을 구하기 위해서는 거릿값과 복원 데이터의 데이터 모양을 맞추어야 한다. 데이터 모양을 맞추기 위해서 첫 번째 프레임의 데이터 값이 항상 정상적이라는 가정하에 거릿값 데이터의 처음 부분에 0을 넣어준다. 모양을 맞춘 후에는 거릿값 데이터에서 임곗값 이상인 위치값들을 구한다. 위치값들은 복원 데이터에 대입하여 해당되는 복원 데이터의 값을 소실 처리한다. 이때, 임곗값은 모션 캡처 데이터에 저장된 좌표값들의 크기에 따라 바뀌며, 신체 질량 중심의 좌우 방향의 평균적인 속도인 1.15 m/s[4]에 24의 프레임 수를 나눈 값의 2배인 0.10 m/frame을 기준으로 ± 0.05 하여 설정한다. 소실 처리된 데이터는 다시 2.1의 방법으로 복원 과정을 거치고, 거릿값이 벗어난 위치값들이 없어질 동안 위 과정을 반복한다. 그림 2는 거리가 벗어나 원래 있어야 할 위치를 벗어난 데이터가 소실 처리되는 예시를 보여준다. 초록색 원은 마커를 의미하고 빨간색 점선으로 이루어진 원은 원래 있어야 할 위치를 보여준다.

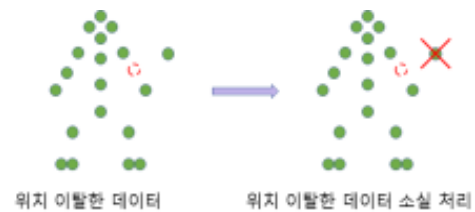


그림 2. 위치 이탈한 데이터 소실 처리

III. 실험 및 결과 분석

본 논문에서는 Intel i7-1065G7 1.30GHz CPU, 16GB RAM 사양 PC에서 Python을 사용하여 제안된 방법을 구현하였다.

실험에 사용된 데이터 파일은 www.c3d.org 사이트에서 제공하고 있는 c3d 파일이다[5]. 사용된 데

이터 파일 (a), (b), (c)에 대한 정보는 표 1과 같다.

표 1. 실험 데이터 파일 정보

파일	데이터 파일 이름	프레임 수	마커 수
(a)	Sample29/OptiTrack-IITSEC 2007.c3d	7931	34
(b)	Sample29/Facial-Sing.c3d	3000	38
(c)	Sample00/Innovative Sports Training/Static Pose.c3d	120	30

데이터 파일들은 본 논문에서 제안한 방식에 따라 복원 작업을 거친 후, 복원된 데이터를 소실 및 변경하였다. 소실은 데이터값을 0으로, 변경은 데이터값을 최솟값과 최댓값 사이의 랜덤한 값으로 조작하였다. 변조된 데이터들은 다시 복원 작업을 거쳤다. 복원율은 원본 데이터를 복원한 데이터에 대한, 변조된 데이터를 복원한 데이터의 비율이 5% 미만이면 정확한 값이라고 판단하여 계산하였다. 표 2는 데이터 변조와 복원을 위한 최솟값과 최댓값, 임계값에 대한 정보를, 표 3은 데이터의 변조율에 따른 복원율을 보여준다.

표 2. 실험 데이터에 대한 정보 및 임계값

파일	최솟값	최댓값	임계값
(a)	-1524.45	2018.60	100
(b)	-401.31	1156.71	100
(c)	-0.02	1.40	0.15

표 3. 변조율에 따른 데이터 복원율(%)

파일 \ 변조율	20%	40%	60%
(a)	98.20	96.31	90.50
(b)	96.63	94.59	89.03
(c)	99.27	97.47	92.04

실험 결과, 파일 (c)는 다른 파일들에 비해 데이터 수가 적어 상대적으로 높은 복원율을 보여주었으며, 파일 (b)는 원본 데이터의 부정확한 정도가 다른 파일들에 비해 높아 상대적으로 낮은 복원율을 보여주었다. 결과적으로, 데이터 수가 적을수록, 원본 데이터의 정확도가 높을수록 실험에 의한 복원율이 높다는 것을 확인하였다. 그림 3은 파일 (a)의 변조율이 60%였을 때 복원 전과 후의 영상을 비교한 예시를 보여준다.

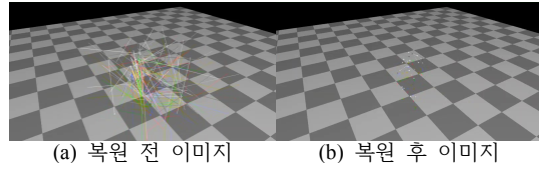


그림 3. 복원 전과 후의 이미지

IV. 결 론

본 논문에서는 간단한 수식을 사용하여 모션 캡처 데이터를 자동으로 복원하여 자세 추정을 하는 방법을 제안하였다. 제안한 방법에 이용한 데이터는 c3d 포맷의 모션 데이터를 사용하였으며, 복원 방법은 마커의 속도와 이동한 거리를 이용하여 부정확한 데이터를 복원하였다.

실험 결과, 데이터 변조율에 따라 정확도의 차이가 발생하였으며, 전체적으로 89% 이상의 복원율을 보여주었다. 광학식 방식으로 모션 캡처 데이터를 생성할 때는 부정확한 정도가 50% 이상 되지 않으므로, 결과적으로는 제안한 방식으로 데이터 복원 작업을 거칠 시에 94% 이상의 복원율을 보여준다고 할 수 있다.

Acknowledgements

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 지역지능화혁신인재양성(Grand ICT연구센터) 사업의 연구결과로 수행되었음”(IITP-2021-2020-0-01462)

References

- [1] Wikipedia. Motion Capture [Internet]. Available : https://en.wikipedia.org/wiki/Motion_capture.
- [2] S. U. Kim, H. Im, and J. Kim, “Deep Learning-Based Human Motion Denoising,” *Journal of IKEE*, Vol. 23, No. 4, pp. 1295-1301, 2019.
- [3] X. B. Peng, A. Kanazawa, J. Malik, P. Abbeel, and S. Levine, “Sfv: Reinforcement learning of physical skills from videos,” *ACM Transactions On Graphics (TOG)*, Vol. 37, No. 6, pp. 1-14, 2018.
- [4] N. K. Son and H. D. Kim, “Age effects on center of mass during obstacle crossing,” *Journal of Korean Society of Physical Medicine*, Vol. 9, No. 1, pp. 93-99, 2014.
- [5] C3D.ORG. C3D Files [Internet]. Available : <https://www.c3d.org/sampled.html>.