

딥러닝 기반 휴먼 모션 생성 모델 활용을 위한 데이터 전처리 자동화 모듈 개발

김상준, 신유정, 유동희*

부산가톨릭대학교, 부산가톨릭대학교, *부산가톨릭대학교

weareyoyo12@naver.com, yujung5042@naver.com, *dhyu@cup.ac.kr

Development of data preprocessing automation module to utilize deep learning-based human motion generation model

Kim Sang Jun, Shin Yoo Jung, Yu Dong hui*

Catholic University of Pusan, Catholic University of Pusan, *Catholic University of Pusan

요약

최근 디지털 전환이 가속됨에 따라 Human Motion Modeling에 대한 관심과 연구가 확대되고 있다. 본 연구에서는 현재 활발히 연구되고 도입되는 Human Motion Modeling의 주요 범주 중 Human Motion Prediction에 대해 살펴보고 그 대표적인 오픈소스 모델인 QuaterNet의 환경 구축 및 모션 생성을 위한 데이터 전처리 모듈 개발의 과정에 대해 다룬다.

I. 서론

최근 이른바 디지털 전환시대를 맞아 휴먼 모션 모델링에 대한 관심과 연구가 확대되고 있다. 휴먼 모션 모델링은 인간의 복잡한 동작을 인간의 관절에 해당하는 Joint들의 Translation과 Rotation의 변화로 일반화하여 모델링하는 기법이다. 현재 연구되는 휴먼 모션 모델링의 범주로는 Human Motion Prediction/Generation, Humanoid Motion Control, Cross-modal Motion Synthesis가 대표적이다[1].

본 연구에서는 현재 활발히 연구되고 도입되는 휴먼 모션 모델링의 주요 범주 중 Human Motion Prediction에 대해 살펴보고 그 대표적인 오픈소스 모델인 QuaterNet을 소개하고 QuaterNet 모델에 범용 데이터 셋을 활용하기 위한 환경 구축 및 학습 및 새로운 모션 생성을 위한 전처리 자동화 모듈 개발 과정을 다룬다.

II. 본론

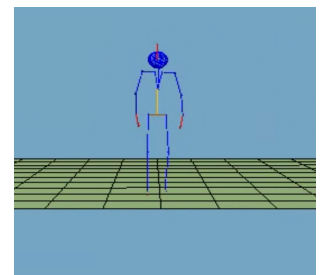
인간의 자연스러운 움직임을 컴퓨터상에서 생성하는 Human Motion Prediction은 인간의 움직임을 일종의 시계열로 보고 그 시퀀스를 분석하기 위해 일반적으로 RNN기반 딥러닝 모델을 사용하여 Seed Motion에 이어질 휴먼 모션을 예측 및 생성한다. 인간의 움직임은 확률 과정적 특징을 지니는데, 이로 인해 예측된 시퀀스가 자연스럽게 인지될 수 있는 범위라 하더라도 Ground Truth와는 다소 차이를 보일 수 있다. 때문에 Human Motion Prediction의 성능 평가는 Seed Motion으로부터 400ms 내외의 생성된 시퀀스 프레임들의 joint들의 오차의 평균으로 정량적으로 평가하는 Short-term(prediction)과 Seed Motion으로부터 1초와 그 보다 긴 시퀀스를 생성하여 그 자연스러움 정도를 사람이 정성적으로 평가하는 Long-term(generation)으로 구별하는 것이 일반적이다.

Human Motion Prediction에는 인간의 움직임을 static data인 Skeleton과 time-varying data인 Motion으로 일반화한 모션 캡처 데이터가 학습 및 예측에 사용된다. 모션 캡처 데이터의 형식에는 TRC, HTR, BVH 등이 있다.

이중 대표적인 BVH 포맷은 BioVision사에 의해 정의된 텍스트 파일 포맷으로, 모델의 골격구조의 계층정보와 초기 위치, 이동과 회전에 대한 움직임 정보가 포함된다. BVH는 크게 2개의 Section으로 구성되는데, HIERARCHY SECTION에서는 캡처한 Actor(Subject)의 골격구조가 정의되며, MOTION SECTION은 총 프레임 수와 프레임 타임 그리고 각 프레임에 할당된 joint별 position, rotation으로 캐릭터의 동작을 기술한다.

```
HIERARCHY///생체정보
ROOT Hip
{
  OFFSET 0.000000 0.000000 0.000000
  CHANNELS 6 Xposition Yposition Zposition Xrotation Yrotation
  JOINT Spine1
  {
    OFFSET -0.000000 28.528410 0.000000
    CHANNELS 3 Zrotation Xrotation Yrotation
    JOINT Neck
    {
      OFFSET -0.000000 23.181000 0.000000
      CHANNELS 3 Zrotation Xrotation Yrotation
      JOINT Head
      {
        OFFSET -0.000000 9.358000 1.724980
        CHANNELS 3 Zrotation Xrotation Yrotation
        End Site
        {
          OFFSET -0.000000 17.249779 0.000000
        }
      }
    }
  }
}
(생체)

MOTION///시간변화정보
Frame: 1284
Frame Time: 0.008333
151.8031302 89.615639 2.728137 1.186190 -1.171126 92.272568
```



BVH 파일 구조

BVH 형식으로 표현된

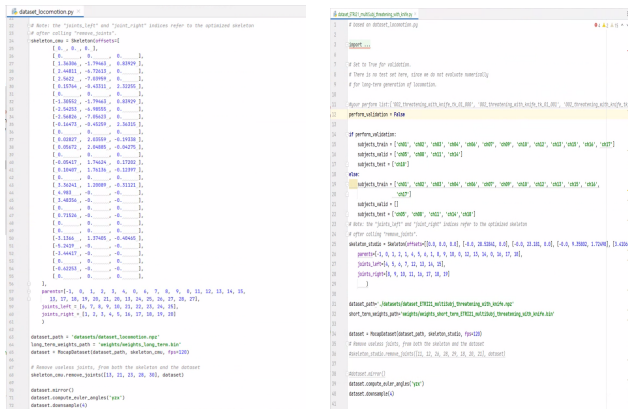
Skeleton의 예

[그림 1]

본 연구에서는 Human Motion Prediction을 위한 모델로 Facebook Research에서 2018년에 오픈소스로 공개한 QuaterNet[3][4]을 활용하였다. QuaterNet은 회전을 기존 오일러 각이 아닌 Quaternion으로 정의함으로써 Gimbal lock 현상으로 인한 오류를 회피하고, 손실함수로 Forward Kinematics에 따른 positional loss를 도입하여 관절 회전기반 손실함수와 관절 위치

기반 손실함수의 장점을 결합한 Human Motion Prediction 분야의 대표적인 모델이다. QuaterNet은 Short-term과 Long-term으로 구분되는 모션 생성을 위해 GRU 기반의 pose network를 기본으로 하되, Long-term의 경우 보행 모션 생성 시 캐릭터의 발과 이동 궤적의 매칭을 최적화 하는 pace network가 추가된다.

QuaterNet의 활용을 위해서 데이터 셋을 준비하는 과정은 크게 세 단계로 나눌 수 있는데, 첫 번째는 데이터 셋 확보 및 변환단계로, 확보된 데이터 셋을 모션 데이터의 액션명, quaternion으로 변환된 joint의 회전 값, Actor(Subject)구분, Root joint의 위치인 trajectories 정보를 추출하여 npz 포맷으로 변환한다. 두 번째는 앞서 준비된 데이터 셋(*.npz)에 적용될 Skeleton의 기본 정보 및 각 모션데이터에 대한 미러링, 다운 샘플링 여부 결정, 모델 학습 시 train data와 validation data의 구분하여 정의한다. 세 번째로는 pace network에서 궤적과 매칭 될 발의 joint를 지정한다.



기존 CMU dataset에 대해
정의된 데이터 셋 정보

기능 1, 2, 3이 적용되어 범용
데이터 셋에 대해 자동으로
정의된 데이터 셋 정보

[그림 2]

본 연구에서는 QuaterNet을 활용하여 Long-term generation을 수행하되, QuaterNet에서 사용된 기본 데이터 셋인 CMU 모션 캡처 데이터 셋이 아닌 다양한 모션 캡처 데이터도 활용할 수 있도록 위의 3가지 과정을 자동화 하였는데, 이를 위해 다음과 같은 기능을 추가한 데이터 전처리 모듈을 개발하였다.

1. 단일 take상 기록된 복수의 모션 시퀀스가 annotation상으로 구분한 경우, annotation된 프레임들만 추출하는 기능
2. 학습과 모션 생성에 사용될 BVH 파일의 Skelton의 offset, joint의 left/right 구분, joint들의 계층구조, fps 정보를 자동으로 파악하고 해당 내용에 대한 정의하는 기능
3. 모델 학습에 사용될 train data와 validation data를 자동으로 구별하여 데이터 셋을 형성하는 기능

본 연구에서 개발한 모듈을 통해 데이터 셋을 확보하고 자동으로 데이터 셋에 대한 정보를 추출 및 정의하여 학습 및 모션생성을 한 결과 정상적인 모델 기능사용을 확인 할 수 있었다.

III. 결론

본 논문에서는 딥러닝 기반 휴먼 모션 생성을 위해 오픈소스인 모델인 QuaterNet을 활용하되, QuaterNet의 기본데이터가 아닌 범용 모션 캡처 데이터를 활용할 수 있도록 모션 캡처 데이터의 주요 정보를 추출하고 이를 모델 학습시 사용할 수 있도록 하위 데이터셋을 구분하는 과정을 자동화 하는 모듈을 개발하였다.

향후 본 연구는 생성된 모션과 Seed Motion과의 연결성을 측정하고, 생성된 모션 정보의 가중치를 메타버스 플랫폼에 적용시키는 두 방향으로 진행될 예정이다.

참 고 문 헌

- [1] Zijie Ye, "Human motion modeling with deep learning: A survey", AI Open, 2022, pp35-39.
- [2] Julieta Martinez. et al. "On human motion prediction using recurrent neural networks", arXiv:1705.02445, 2017.
- [3] Dario Pavllo. et al."QuaterNet: A Quaternion-based Recurrent Model for Human Motion", British Machine Vision Conference(BMVC), 2018.
- [4] facebookresearch, "QuaterNet", (<https://github.com/facebookresearch/QuaterNet>).
- [5] "CMU graphics lab motion capture database", (<http://mocap.cs.cmu.edu>)