Check for updates

Journal of KIIT. Vol. 20, No. 10, pp. 119-125, Oct. 31, 2022. pISSN 1598-8619, eISSN 2093-7571
http://dx.doi.org/10.14801/jkiit.2022.20.10.119

순환신경망을 이용한 동작 데이터의 발 접촉 검출 방법

강창구*, 김미경**

Detecting Foot Contact of Motion Data using Recurrent Neural Network

Changgu Kang*, Meekyoung Kim**

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1F1A1060381)

요 약

환경과 상호작용 행동분석에 관한 연구는 다양한 분야에서 이루어지고 있다. 상호작용은 환경과의 접촉 행위에서 시작되며 접촉은 중요한 맥락정보로 인식된다. 이러한 중요성에도 불구하고 환경의 복잡도와 사람 또는 사물의 가림 현상으로 정확한 접촉 인식이 어려우며, 대부분의 연구에서는 접촉 검출을 위해 환경 데이터를 요구한다. 본 논문에서는 이러한 제약사항들을 극복하고 정확한 행동 인식을 위한 중요한 접촉정보인 발 접촉 검출 방법을 제안하였다. 제안된 방법은 동작 데이터의 연속성에 적합한 순환신경망을 이용하며, 동작 데이터를 가공한 공간과 시간적 정보를 학습을 위한 특징으로 사용한다. 끝으로 우리는 수집된 동작 데이터를 이용하여 모델 학습을 수행하며, 학습된 모델의 테스트 결과 약 95% 정확도를 보여준다.

Abstract

Research on the analysis of interaction behavior in the environment is being conducted in various fields. Interaction begins with the act of contact with the environment, and contact is recognized as important contextual information. Despite this importance, accurate contact recognition is difficult due to the complexity of the environment and the occlusion of people or objects, and most studies require environment data for contact detection. In this paper, we propose a foot contact detection method, which is important contact information for accurate behavior recognition, overcoming these constraints. The proposed method utilizes a recursive neural network suitable for the continuity of motion data, and uses the spatial and temporal information processed by the motion data as a feature for learning. Finally, we perform model training using the collected motion data, and the test result of the trained model shows about 95% accuracy.

Keywords

behavior recognition, foot contact detection, recursive neural network, interaction

* 경상국립대학교 컴퓨터과학부 교수
- ORCID: https://orcid.org/0000-0003-4060-6835
** 대구대학교 컴퓨터정보공학부 교수(교신저자)
- ORCID: https://orcid.org/0000-0002-3750-0346

· Received: Aug. 05, 2022, Revised: Aug. 24, 2022, Accepted: Aug. 27, 2022
· Corresponding Author: Meekyoung Kim
Dept. of Computer & Information Engineering, Daegu University, Korea
Tel.: +82-53-850-6586, Email: koms1701@daegu.ac.kr

Ⅰ. 서 론

환경에서 사람들의 상호작용 행동 분석에 관한 연구는 컴퓨터 비전, 컴퓨터 그래픽스, 인지공학, 그리고 HCI 등 다양한 분야에서 연구가 진행되고 있다[1]-[6]. 이러한 행동 분석을 위해서는 기본적으로 행위자에 대한 동작 데이터가 필요하며, 이러한 데이터를 획득하기 위하여 깊이 카메라 또는 다양한 물리적 센서 장비를 이용하게 된다. 이렇게 획득된 동작 데이터와 행위자의 움직임을 파악하기 위한 맥락정보(Contextual Information)들을 획득하게 되고 이러한 맥락정보의 해석으로 사람들의 행동 분석이 이루어진다[2][3].

맥락정보에서 가장 중요하며 많이 사용되는 정보는 행위자의 접촉정보이다[4]. 환경과의 접촉정보를 통하여 행위자가 하는 행동 해석이 가능하게 된다. 예를 들어 한 사람이 상부 다리, 엉덩이, 그리고 등의 연속된 접촉이 발생하면 “앉고 있다”라는 행동으로 해석될 수 있으며, 엉덩이, 다리, 등, 머리가 연속적으로 접촉이 발생하면 “눕고 있다”라고 해석될 수 있다. 이러한 행동 분석을 위한 접촉정보의 중요성에도 불구하고 대부분의 행동 인식기술에서 상세한 접촉정보 획득에는 어려움이 있다. 실험실 환경에서는 물리적인 센서를 착용하여 행위자의 부분별 접촉 여부를 인식할 수 있지만, 카메라 장비를 사용하는 일반적인 환경에서는 사람 또는 물체의 가림형상으로 접촉 여부를 정확히 인식하는 것은 어렵다. 또한 기존 연구에서 환경정보, 물리적 장치 등을 이용하여 접촉여부를 인식하였다[5]-[7]. 사람의 몸은 관절로 구분되어 다양한 부분으로 나뉘어 있다. 그중에서도 발 접촉정보는 운동학적 제약조건을 제공함으로써 인간의 행동 인식 정확도를 향상하기 위한 중요한 요소이다[8].

우리는 이러한 제약사항을 극복하고 정확한 행동 인식에 중요한 접촉정보인 발 접촉 검출 방법을 제안한다. 동작 데이터는 연속성을 가지고 있으므로, 제안된 방법은 이러한 특징에 적합한 순환신경망을 이용한다. 전처리과정에서는 동작 데이터를 가공하여 공간적과 시간적 특징이 포함된 특징 벡터로 만들며, 학습 과정에서는 순환신경망 중 하나인 LSTM 모델을 이용하여 접촉 검출을 위한 최적화된 가중치를 찾게 된다. 최종적으로 테스트 과정을 거쳐 학습된 모델을 평가한다.

본 논문의 구성은 다음과 같다. 2장은 상호작용 행동 분석 및 접촉 검출에 관한 연구동향에 대해 기술한다. 3장에서는 데이터수집, 특징 구성, 그리고 학습기 구축에 대해서 상세히 기술한다. 4장에서는 제안된 방법의 성능을 평가하고 실험결과를 설명한다. 마지막으로 5장에서는 결론에 대해 기술하고 한계점과 추후연구에 대하여 기술한다.

Ⅱ. 연구 동향

많은 분야에서 상호작용 행동 분석과 관련된 다양한 주제의 연구들이 진행됐다. 우선 환경과의 상호작용 행동을 분석하기 위해서는 사용자의 자세 정보와 환경정보가 필요하다. 하지만 환경에서 이러한 정보 획득은 상호작용하는 물체에 대한 가림형상으로 인해 완전한 정보 획득이 어렵다. 이러한 문제를 해결하기 위해 Hassan은 상호작용 환경에서 캡쳐된 불완전한 사람의 자세 데이터를 3D 기하학 정보를 이용하여 적합한 자세 추정방법을 제안하였다[9].

다음으로 상호작용은 사람이 환경 공간에서 수행되며 이러한 환경에서 상호작용 영역 표현방법과 관련된 연구도 수행되었다. Savva는 3차원 깊이 지도로 표현되는 행동 지도(Action map)를 이용하여 환경에서 상호 작용 확률 정보를 표현하였으며, 주어진 동작이 실행 가능한 공간을 찾는 방법을 제안하였다[10]. 상호작용에 있어서 가장 중요한 정보는 접촉정보이며 깊이 지도를 통한 표현은 이러한 접촉정보에 대한 반영이 부족하다. 이러한 문제점을 해결하기 위해서 Gupta는 3D 장면 데이터에 사람-장면 접촉 관계를 명확히 표현하는 방법을 제안하였다[4].

접촉정보는 사람의 행동과 관련된 중요한 정보이다. Kang은 캐릭터 표면의 접촉 가능한 후보 점들을 할당하고 주어진 환경에 안정적인 자세를 유지하기 위한 접촉점을 성별하고 적합한 자세를 생성하는 연구를 진행하였다[11]. 발 접촉 검출을 위한 다양한 연구들도 진행되었다. Kim은 발 접촉 검출을 위해 물리적 센서를 이용하였으며, Ma는 역학정

보를 이용하며, 동작 데이터와 행위자의 무게 정보를 요구한다[5][8].

이와 같은 대부분의 연구들은 접촉 검출을 위해 추가적 정보 또는 장치들이 필요하였다. 제안된 방법은 일반적인 형식의 동작 데이터만을 이용하여 발 접촉 검출을 수행한다.

III. 데이터 수집 및 학습기 구축

우리는 동작데이터로부터 발 접촉 검출을 위한 순환신경망 모델을 사용하며, 이와 같은 모델을 사용하기 위하여 먼저 데이터 수집과 전처리 과정이 필요하다. 본 장에서는 이러한 데이터 수집과 전처리 과정에 대해서 기술하고, 마지막으로 사용된 학습모델에 대해 설명한다.

3.1 데이터 수집

학습데이터는 동작 데이터와 동작 데이터의 각 프레임당 접촉정보로 구성된다. 하지만 대부분의 동작 캡쳐장비로 저장된 데이터들은 동작 데이터만 있으며 접촉 여부는 확인할 수 없다. 그래서 우리는 동작 데이터의 각 프레임별 접촉 상태를 직접 레이블링 작업을 수행하였다. 레이블링 작업은 각 동작 데이터의 프레임당 진행되기 때문에 많은 시간이 필요한 작업이다. 이를 프레임 구간별 레이블링이 가능하도록 레이블링 프로그램을 개발하였다.

그림 1은 프로그램을 이용한 접촉 레이블링 작업 과정을 보여준다. 재생된 동작에서 골격의 파란색으로 그려진 부분은 현재 접촉되어 있음을 나타낸다. 프로그램에는 동작 자동 재생, 프레임 앞 뒤 이동, 구간별 접촉 설정 등의 기능이 구현되어 있다. 접촉은 동작 데이터의 뼈대 구조에 맞게 각 부분별 접촉 레이블링 작업과 추가적으로 무릎, 팔꿈치에 대한 접촉 레이블링 작업이 가능하도록 하였으며, 모두 18개의 접촉 레이블링 작업을 할 수 있다. 접촉 유무에 따라 0과 1로서 표현되며, 결과 파일은 동작 데이터 파일명과 동일한 CSV 형식의 파일로 저장된다.

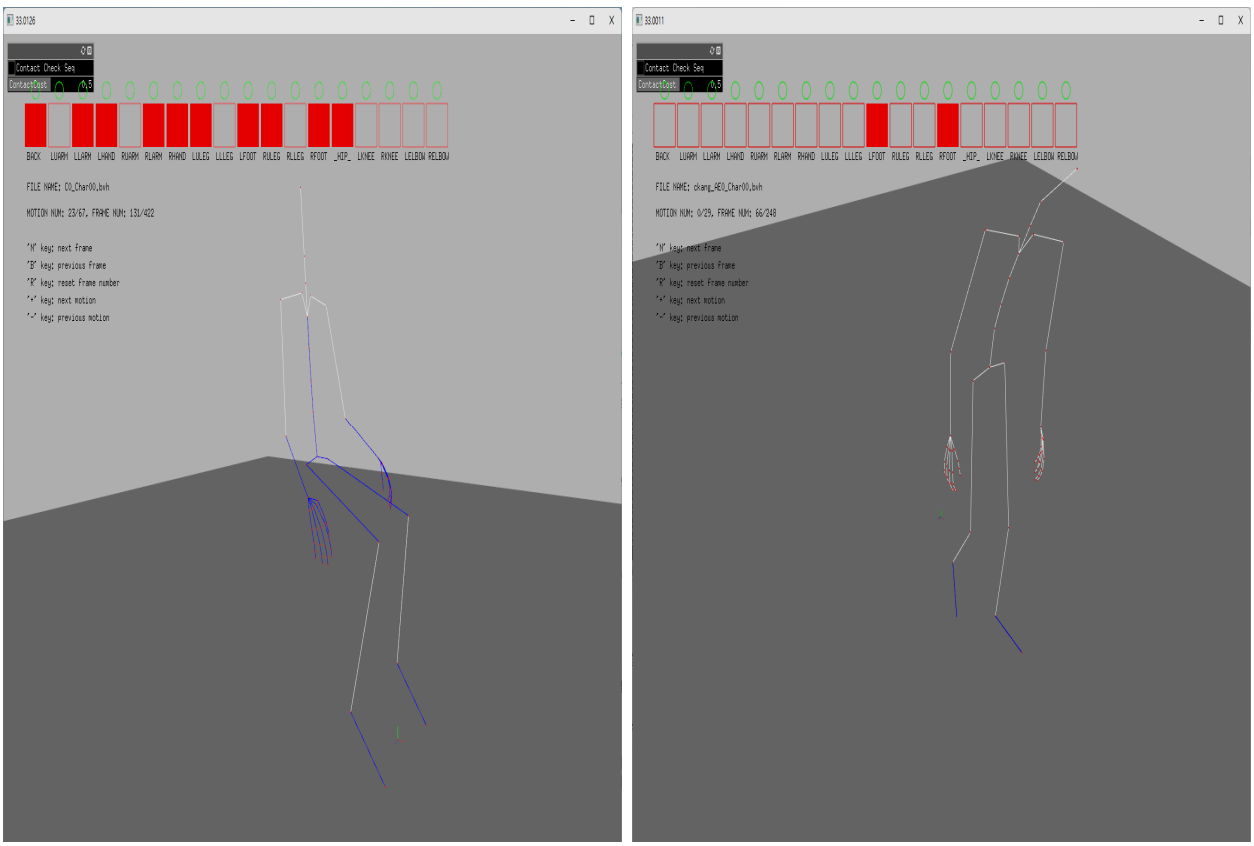


그림 1. 프로그램을 이용한 접촉 레이블링 작업
Fig. 1. Labeling a contact using a program

3.2 특징 구성

그림 2는 캐릭터의 골격 구조와 골격 구조 중에서 특징으로 사용된 부분을 보여준다. 전체 골격 구조에서 점선 테두리로 그려진 12개(Left hand, Left ForeArm, Left, Arm, Right Arm, Right Fore Arm, Right Hand, Left Up Leg, Left Leg, Left Foot, Right upLeg, Right Leg, Right Foot)의 부분으로부터 공간적 특징과 시간적 특징을 각각 만들어낸다. 동작 캡쳐장비를 이용하여 저장된 데이터는 각 프레임당 각 관절의 절대적인 각도로 저장되어 있다. 우리는 원본 데이터에 전처리를 통하여 공간적 특징과 시간적 특징을 각각 만들고 식 (1)과 같이 하나의 특징 백터로 결합한다.

$$feature = concat(f_{sptial}, f_{sequencial}) \quad (1)$$

$$f_{spatial} = \frac{Distance_{ralative}}{Distance_{geodesic}} \quad (2)$$

공간적 특징은 식 (2)와 같이 엉덩이(Hips)로부터 특징으로 사용되는 부분의 그래프 지오데식(Graph geodesic)거리와 상대거리의 비로서 정의하였다.

그림 3은 뼈대에서 공간적 특징을 계산하기 위한 그래프 지오데식 거리와 상대거리 측정 방법을 보여준다. 엉덩이(Hips)에는 왼손(Left Hand)까지의 지오데식 거리는 점 A, B, C를 통과하여 계산된 실선으로 그려진 거리를 말한다. 그리고 상대거리는 엉덩이로부터 왼손까지의 직선 거리를 의미한다. 그림에서는 점선으로 그려진 거리를 말한다.

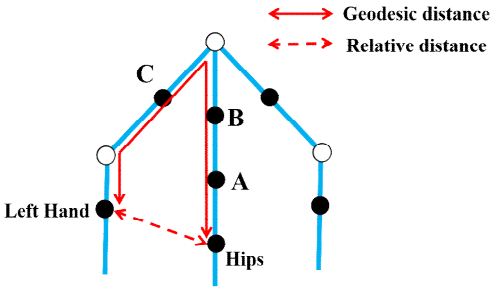


그림 3. 공간적 특징
Fig. 3. Spatial feature

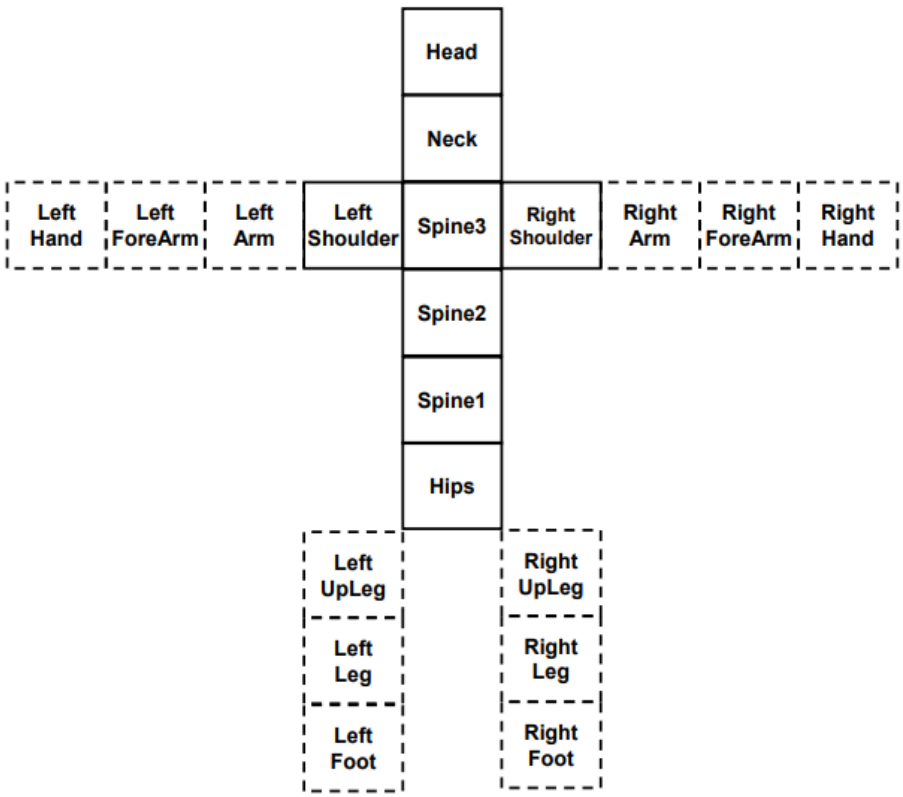


그림 2. 골격 구조와 특징 부분
Fig. 2. Skeleton architecture and feature parts

시간적 특징으로는 3축에 대한 각속도(Angular velocity)를 이용한다. 골격에서 특징으로 사용되는 부분은 각 부분당 공간적 특징 1개와 시간적 특징 3개의 차원이 4인 특징 벡터를 만든다. 최초 캡쳐된 동작 데이터의 관절구조는 각 손가락 관절이 세분화되어 있지만, 이를 좌, 우 손(Left hand, Right hand)으로 부분을 단순화한다. 최종적으로 각 프레임당 48 차원의 특징벡터가 만들어진다.

3.3 학습기 구축

동작데이터는 연속성을 가지고 있기 때문에 이에 적합한 학습모델인 순환신경망을 이용하며, 전처리 과정을 통하여 만들어진 특징벡터들 입력으로 순환신경망 모델 학습을 수행한다. 순환신경망 모델은 LSTM(Long Short-Term Memory)[12]을 사용하게 되며, LSTM 모델에서 출력된 결과는 2개 층의 DNN (Deep Neural Networks) 이용하게 되며 최종 출력인 양 발의 접촉값들을 출력하게 된다. 그림 4는 학습을 위해 구축된 전처리와 학습모델을 보여준다.

표 1은 구축된 학습모델의 구성요소별 출력크기를 보여준다. LSTM의 시퀀스 길이는 32로 사용하며 전처리 과정에서 시퀀스 길이에 맞게 학습데이터 크기를 만들어낸다. LSTM의 출력크기는 48×1, DNN의 첫 번째 층인 DNN1은 16×1, 두 번째 층인 DNN2는 2×1이다. 적합을 막기 위해 드롭아웃(Dropout) 계수는 0.4, 학습률은 후보군 중에 가장 손실값이 값이 작은 0.001을 설정하였다. 배치사이즈는 16, 32, 64을 비교하였을 때 손실값의 차이는 적으나 학습 속도를 고려하여 32로 설정하였다.

표 2는 학습률과 배치사이즈에 따른 손실값을 보여준다. 손실함수는 최소제곱오차, 최적화 알고리즘으로 아담(Adam)[13]을 각각 이용한다.

표 1. 구성요소별 출력크기
Table 1. Output size for each component

Component	Output size
Pre-processing	32x48
LSTM	48x1
DNN1	16x1
DNN2	2x1

표 2. 학습률과 배치사이즈에 따른 손실값 비교
Table 2. Loss according to learning rate and batch size

Learning rate	Loss	Batch size	Loss
0.0001	0.037	16	0.016
0.001	0.015	32	0.017
0.01	0.022	64	0.019

Ⅳ. 실험 및 분석

우리는 모델 학습을 위해 캡쳐 된 동작 데이터들을 이용하여 학습을 진행하였다. 동작 데이터는 “앞으로 걷기”, “좌(우)회전 걷기”, “걷다가 180도 회전 후 걷기” 등 다양한 형태의 데이터 81개와 “앉기”, “등받이 의자에 앉아서 기대기”, “손을 사용하며 앉기” 등 앉기와 관련된 데이터 39개를 수집하였다. 120개의 동작데이터 중 90개는 학습데이터로 사용하며, 그림 5는 학습과정에서 손실값의 변화를 나타낸다.

30개의 동작 데이터를 이용하여 테스트를 수행하였다. 발 접촉 유무에 따라 1 또는 0의 값으로 표현된 동작 데이터의 테스트 결과 정확도는 약 95%로 나타났다. 대부분의 에러구간의 발이 디디는 시점과 떼는 시점에서 발생하였다. 이유는 발을 디디는 시점과 떼는 시점의 경우 접촉되는 범위와 정도에 따라 접촉여부를 0과 1로 이진화된 구분이 어렵기 때문이다.

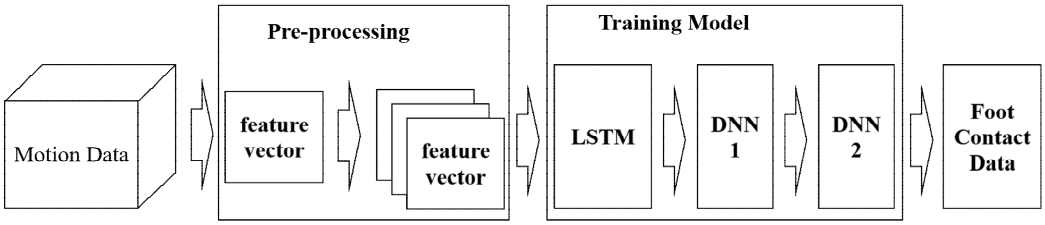


그림 4. 전처리와 학습모델
Fig. 4. Pre-processing and training model

발 접촉/미접촉이 반복적으로 나타나는 걷기 형태의 동작에서는 정확도가 약 92% 낮아졌으며, 앉기 형태와 같은 발 접촉/미접촉 반복 횟수가 적은 동작에서는 약 98%의 높은 정확도를 보였다.

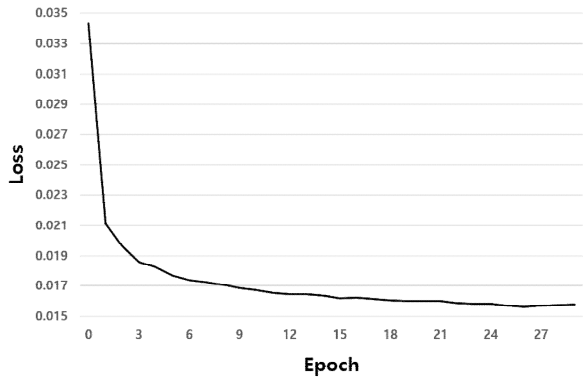
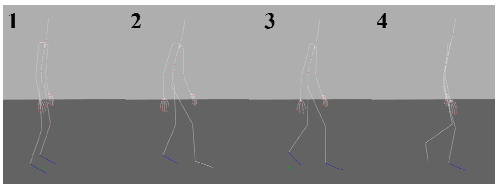
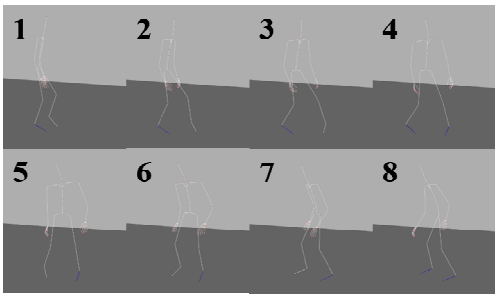


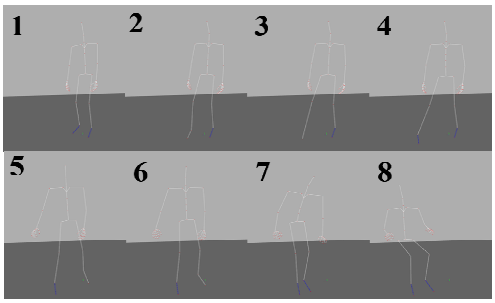
그림 5. 학습에 따른 손실변화
Fig. 5. Training loss



(a) 걷기
(a) Walking



(b) 걷다가 180도 회전 후 걷기
(b) Walking and turning 180 degrees



(c) 의자에 앉기
(c) Sitting on a chair

그림 6. 발 접촉상태 검출 결과 모션
Fig. 6. Motion for the result of foot contact state detection

그림 6은 모델에의 검출된 접촉 데이터들을 이용하여 재생한 결과이다. (a)는 “걷기” 동작일 때의 결과로서 양발, 왼발, 양발, 오른발 순으로 걷기 동작의 접촉 순서를 보여준다. (b)는 “걷다가 180도 회전 후 걷기” 동작으로 회전시 회전축인 오른발이 회전하는 동안 접촉이 잘 되고 있음을 확인할 수 있다. 그리고 (c) “의자에 앉기” 동작의 경우 앉기 동작의 경우 발이 지면에 붙어 있는 상태에서 다리가 옆으로 이동하는 등 복잡한 형태의 이동이 발생한다. 이와 같은 동작에서도 접촉이 잘 되고 있음을 확인할 수 있다.

Ⅴ. 결론 및 향후 과제

환경과 상호작용 행동분석에 관한 연구는 다양한 분야에서 이루어지고 있다. 상호작용은 환경과의 접촉 행위에서 시작되며 접촉은 중요한 맥락정보로 인식된다. 사람의 몸은 관절로 구분된 다양한 부분으로 나뉘며 그 부분들의 접촉여부에 따라 상호작용 행동의 형태가 구분된다. 그중에 발 접촉정보는 운동학적 제약조건을 제공함으로써 인간의 행동 인식 정확도를 향상하기 위한 중요한 요소이다.

본 논문에서는 이러한 제약사항들을 극복하고 정확한 행동 인식을 위한 중요한 접촉정보인 발 접촉 검출 방법을 제안하였다. 먼저 학습데이터 수집을 위해 발 접촉 검출 방법은 순환신경망을 이용하였으며, 캡쳐된 동작 데이터를 이용하여 학습과 테스트를 수행하였다. 제안된 방법은 발 접촉을 검출하는 방법으로서 사람의 몸에 다양한 부분에 대한 접촉 검출이 불가능하다. 이러한 한계점을 극복하기 위하여 행동 데이터만을 이용하여 손, 팔, 다리에 대한 접촉 검출 방법 연구와 접촉기반 행동 인식 연구를 향후 과제로 한다.

References

[1] Jong Gwan Lim, "Hand Motion Segmentation for IMU Applications based on Multidimensional Evaluation", The Journal of Korean Institute of Information Technology, Vol. 18, No. 10, pp. 79-85. Oct. 2020. https://doi.org/10.14801/jkiit.2020.18.10.79

[2] Yao Bangpeng and Li Fei-Fei, "Modeling mutual context of object and human pose in human-object interaction activities", 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, Jun. 2010. https://doi.org/10.1109/CVPR.2010.5540235.

[3] Gupta Abhinav, et al., "Understanding videos, constructing plots learning a visually grounded storyline model from annotated videos", 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, Jun. 2009. https://doi.org/10.1109/CVPR.2009.5206492.

[4] Siwei Zhang, et al., "PLACE: Proximity learning of articulation and contact in 3D environments", 2020 International Conference on 3D Vision (3DV) IEEE, Fukuoka, Japan, pp. 642-651, Nov. 2020. https://doi.org/10.48550/ arXiv.2008.05570.

[5] Myeongkyu Kim and Donghun Lee, "Development of an IMU-based foot-ground contact detection (FGCD) algorithm", Ergonomics, Vol. 60, No. 3, pp. 384-403, Mar. 2017. https://doi.org/10.1080/00140139.2016.1174314.

[6] Pappas Ion PI, et al., "A Reliable Gait Phase Detection System", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 9, No. 2, pp. 113-125, Jun. 2001. https://doi.org/10.1109/7333.928571.

[7] Sunwoo Park, Sungjae Kang, Seonhong Hwang, and Youngho Kim, "Comparison of Gait Phase Detection Systems for Automatic FES Walking", International Technical Conference on Circuits Systems, Computers and Communications, pp. 1172-1173, 2007.

[8] H. Ma, W. Yan, Z. Yang, and H. Liu, "Real-Time Foot-Ground Contact Detection for Inertial Motion Capture Based on an Adaptive Weighted Naive Bayes Model", IEEE Access, Vol. 7, pp. 130312-130326, Sep. 2019. https://doi.org/10.1109/ACCESS.2019.2939839.

[9] Hassan Mohamed, et al., "Resolving 3D human pose ambiguities with 3D scene constraints", Proc. of the IEEE/CVF international conference on computer vision, Seoul, Korea, pp. 2282-2292, Nov. 2019. https://doi.org/10.1109/ICCV.2019.00237.

[10] Savva Manolis, et al., "SceneGrok: Inferring action maps in 3D environments", ACM transactions on graphics(TOG), Vol. 33, No. 6, pp 1-10, Nov. 2014. https://doi.org/10.1145/2661229.2661230.

[11] Changgu Kang and Sung-Hee Lee, "Environment-adaptive contact poses for virtual characters", Computer Graphics Forum, Vol. 33. No. 7, Oct. 2014. https://doi.org/10.1111/cgf.12468.

[12] Sepp Hochreiter and Jürgen Schmidhuber, "Long short-term memory", Neural computation, Vol. 9, No. 8, pp 1735-1780, Nov. 1997. https://doi.org/10.1162/neco.1997.9.8.1735.

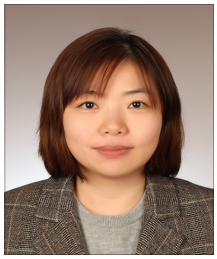
[13] Diederik P. Kingma and Jimmy Ba, "Adam: A method for stochastic optimization", 3rd International Conference for Learning Representations, San Diego, Dec. 2014. https://doi.org/10.48550/arXiv.1412.6980.

저자소개

강 창 구 (Changgu Kang)

2010년 2월 : 광주과학기술원 정보기전공학부(공학석사)
2017년 8월 : 광주과학기술원 전기전자컴퓨터공학부(공학박사)
2018년 3월 ~ 현재 : 경상국립대학교 컴퓨터과학부 부교수
관심분야 : 컴퓨터 그래픽스, 증강현실, 인공지능

김 미 경 (Meekyoung Kim)



2010년 2월 : 한국과학기술원 수리과학부(이학석사)
2017년 8월 : 한국과학기술원 문화기술대학원(공학박사)
2019년 3월 ~ 현재 : 대구대학교 컴퓨터정보공학부 조교수
관심분야 : 컴퓨터 그래픽스, 컴퓨터 애니메이션/시뮬레이션, 인공지능