

욜로-오픈포즈 환경에서 사용자의 행동 인식을 위한 신체의 관절 좌표 추출에 관한 연구

이철민, 이동명

동명대학교 컴퓨터공학과

captaink7428@gmail.com, dmlee@tu.ac.kr

A Study on Extraction of Skeleton Coordinates of User Body for Object Behavior Recognition in Yolo-Open Pose Environment

Cheol Min Lee, Dong Myung Lee

Tongmyong Univ.

요약

본 논문에서는 이와 같은 오픈포즈(Open Pose)와 율로(You Only Look Once, YOLO)의 특징 중 장점만을 이용하여 사용자의 행동 인식을 위한 신체의 관절(Skeleton) 좌표를 추출하였다. 신체의 관절 좌표를 추출하기 위해서는 율로와 오픈포즈의 체계적인 실행 환경 구축이 필수적이며, 추출 프로세스 구조는 그래픽 드라이버 설치, 율로 실행 환경 설정, 오픈포즈 실행 환경 설정 그리고 신체의 관절 좌표 추출의 4단계로 구성된다. 주어진 관절 포인트 원손 및 오른손 관절에 대해 8회 실험 결과, 대체적으로 신체의 관절 좌표 추출 성능이 우수한 것으로 판단된다.

I. 서 론

오픈포즈(Open Pose)는 영상을 통해 객체의 관절(Skeleton)을 이용하여 자세를 인식하는 행동 감지 라이브러리(Library)이다. 그러나 오픈포즈는 객체의 자세만 인식할 뿐 객체의 행동(자세)의 유형을 인식하지는 않는다. 반면, 율로(You Only Look Once, YOLO)는 영상인식 전용의 학습 기반 플랫폼으로써 해당 객체의 유형을 인식하는데 사용된다. 그러나 이동 객체의 특정 신체에 대한 관절 좌표를 제공하지는 않는다.

본 논문에서는 이와 같은 오픈포즈와 율로의 특징 중 장점만을 이용하여 사용자의 행동 인식을 위한 신체의 관절 좌표를 추출하였다.

II. 관련연구

2.1 오픈포즈 기반 손동작 음계결정 알고리즘

본 연구와 관련된 연구인 오픈포즈 기반 음계결정 알고리즘[1]은 손동작 인식 알고리즘과 음계 출력 알고리즘으로 구성된다. 손동작 인식 알고리즘은 손모양을 추출하고 손의 중심 좌표를 추출하는데, 손모양은 완전 연결망(fully connected network)를 사용하여 21개의 포인트 좌표를 추출하여 추정된다.

손모양의 중심 좌표는 해당 좌표의 평균값으로 표현된다. 음계 출력 알고리즘은 영상 음계 지정과정과 음계 데이터를 저장하는 과정으로 구성되는데, 오른손 중심 좌표의 위치에 따라 정해진 영상좌표에 음계를 출력한다. 이 연구에서는 특정 신체의 관절 포인트 좌표를 이용하여 음계 데이터를 출력한다.

2.2 YOLO 기반 사용자 행동인식 및 위치추정

[2, 3]은 YOLO를 기반으로 사용자 행동 인식을 위한 실시간 모델을 제안하였다. 이 연구의 주된 특징은 적은 수의 프레임을 사용하여 사용자의 동작을 인식했다는 점이며, 심지어 단일 프레임만으로도 사용자의 동작을 인식한다는 점이다.

III. 신체의 관절 좌표 추출 기능 설계

3.1 프로세스 구조 및 기능 설계

신체의 관절 좌표를 추출하기 위해서는 율로와 오픈포즈의 체계적인 실행 환경 구축이 필수적이며, 추출 프로세스 구조는 그림 1과 같이 4단계로 구성된다.

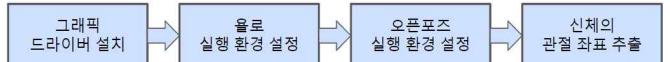


그림 1. 신체의 관절 좌표 추출 프로세스 구조

3.2 기능 설계

그림 1의 4가지 단계에 대한 기능을 다음과 같이 설계하였다. 첫째, 딥러닝(Deep Learning)을 병렬 처리하여 빠른 연산을 하기 위해서 버전 8 이상의 CUDA(Compute Unified Device Architecture)를 설치하고 그래픽 드라이버를 설치한다. 둘째, 율로 사용에 필수적인 파이썬(Python) 라이브러리 모두 설치한다. 그리고 코코기반의 학습 데이터셋(Data Set), 그리고 가중치 파일을 설치하고 경로를 설정한다.

셋째, 오픈포즈 사용에 필요한 비주얼 스튜디오 엔터프라이즈(Visual Studio Enterprise)를 설치하고 C++관련 SW 도구를 추가로 다운로드 받는다. CMake의 경로는 오픈포즈 파일과 오픈포즈 빌더 통해 바이너리 파일을 저장할 위치 경로를 설정하고 오픈포즈를 빌드한 후 비주얼 스튜디오에서 솔루션 빌드한다.

넷째, 율로와 오픈포즈를 파이썬 환경에서 사용하기 위해 빌드된 폴더, 학습 데이터 파일 그리고 가중치 파일(weight)의 3가지의 파일에 대한 경로를 설정하고, 신체 관절 좌표를 출력하기 위한 소스코드를 설계한다. 그 결과, 율로는 사용자 주변에 있는 사물의 좌표를, 오픈포즈는 신체의 관절 좌표 값을 각각 출력한다.

IV. 실험 및 결과분석

4.1 실험 환경

실험에는 엔비디아 그래픽 카드 2070 Ti를 장착한 데스크톱과 일반적인 웹캠을 사용하였고, OpenCV, 오픈포즈, 욜로 등의 라이브러리와 가상환경을 이용하기 위해 아나콘다(Anaconda)를 설치하였다. 개발 기능을 구현에 사용된 언어는 파이썬이다.

4.2 실험 방법

4.2.1 신체의 관절 포인트 추출

상기의 실험 환경을 모두 설정 한 후, 오픈포즈를 사용하여 감지한 신체의 관절 포인트를 그림 2와 같이 추출하였다[4]. 그 다음, 출력된 관절 포인트는 신체의 자세별로 파일에 저장하여 학습 데이터로 활용하여야 하지만 이 기능은 본 연구의 범위에서 제외하였다.

4.2.2 객체의 유형 인식

객체의 유형 인식은 이미 구현된 욜로 기반 오픈소스 코드를 이용하였다. 그림 3은 이를 이용해서 인식된 객체를 나타낸다. 이 코드는 추출된 객체의 좌표를 신체의 관절 좌표로 매칭하는 역할을 한다.



그림 2. 신체의 관절 포인트 추출



그림 3. 욜로를 이용한 객체 인식

4.2.3 신체의 관절 좌표 추출

신체의 관절 좌표 추출은 사용자의 몸에서 관절 포인트 번호를 그림 4와 같이 정의하여 이루진다. 일반적으로 신체 관절 좌표의 추출에 있어서 신체의 모든 관절을 대상으로 하는 경우는 확률적으로 드물기 때문에 본 논문에서는 신체의 특정 관절을 대상으로 좌표를 추출하고자 한다.

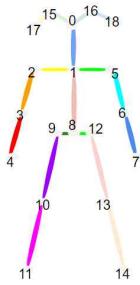


그림 4. 신체의 관절 포인트 번호 정의

4.3 실험 결과

본 본문에서는 사용자 양 손의 손목을 대상으로 신체의 관절 좌표가 정상적으로 추출되는지를 실험하였다. 그림 4에서 정의된 사용자의 관절 포인트 번호 중 4, 7번의 포인트가 관절 좌표로 추출된 결과는 그림 5와 같다. 사용자의 손목의 좌표는 그림 5에서 보는 바와 같이 왼쪽 손목은 빨간색이, 오른쪽의 손목은 노란색이 각각 표시되었다.

상기의 실험 결과가 어느정도 정확하게 관절 포인트와 관절 좌표간에 매칭된 것인지 확인하기 위하여 기본적인 실험을 실시하였다. 여기에서 충분한 실험에 의한 매칭 결과를 도출하지는 못 했으나 주어진 관절 포인트(왼손 및 오른손 관절)에 대해 8회 실험을 실시하였고 그 결과는 <표 1>과 같다. <표 1>에 의하면 왼쪽 손목과 오른쪽 손목의 최대 오차는

(1.30779, 1.30789), (1.32176, 3.86643)으로 나타났다. 이 결과는 충분한 실험에 의한 결과는 아니지만 대체적으로 신체의 관절 좌표 추출 성능이 우수한 것으로 판단된다.



그림 5. 손목 관절 좌표 추출 결과

<표 1> 신체의 관절 좌표 추출

순번	왼쪽 손목(x, y)	오른쪽 손목(x, y)
1	(146.65508 315.0602)	(183.24052 392.04724)
2	(145.38124 315.03842)	(183.25052 392.056)
3	(146.66048 315.0615)	(183.24524 390.78247)
4	(145.37921 315.01877)	(183.2368 392.04468)
5	(145.39413 315.03683)	(183.1841 392.05664)
6	(146.65999 316.32666)	(181.93993 394.64447)
7	(145.35269 316.31754)	(181.92876 394.6489)
8	(145.36653 315.06143)	(183.21915 392.08566)
최대 오차	(1.30779, 1.30789)	(1.32176, 3.86643)

V. 결론

본 논문에서는 욜로-오픈포즈 환경에서 사용자의 행동 인식을 위한 신체 좌표를 추출하였다. 이를 위한 신체 좌표 추출 프로세스는 그래픽 드라이버를 설치, 욜로 실행환경 설정, 오픈포즈 실행환경 설정, 그리고 신체 좌표 추출의 4단계로 구성된다. 주어진 관절 포인트 왼손 및 오른손 관절에 대해 8회 실험 결과, 대체적으로 신체의 관절 좌표 추출 성능이 우수한 것으로 판단된다. 추후, 사용자 및 특정 사물에 대한 학습을 통한 행동 인식 알고리즘을 추가적으로 제시할 예정이다.

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2019R1F1A1062670).

참 고 문 헌

- [1] 이철민, 이동명, “Openpose를 이용한 딥러닝 기반 손 모양 인식 알고리즘,” 제9회 2019 ICT융합기술서비스 학술대회 논문집, 한국공학 교육학회, 제9권, 제1호, pp.29-32, 2019. 8. 22.
- [2] Shubham Shinde, Ashwin Kotharia, Vikram Guptab “YOLO based Human Action Recognition and Localization,” 컴퓨터 미디어, pp.1-2, 2018.
- [3] Redmon, Joseph and Farhadi, Ali, “Yolo library,” <https://github.com/eriklindernoren/PyTorch-YOLOv3>.
- [4] G. Hidalgo, Z. Cao, T. Simon, S. E. Wei, H. Joo, and Y. Sheikh, “Open Pose library,” <https://github.com/CMU-Perceptual-Computing-Lab/openpose>.