



# 키넥트센서를 이용한 수화 인식시스템 설계

권범준, 지도교수 이덕규

서원대학교 소프트웨어전공

## 서론

수화는 현대사회에서 청각장애인들의 의사소통을 위해 가장 많이 사용되는 수단이다. 하지만 비장애인과 의사소통에는 원활하지 않으며 공공서에서도 필담이 주로 사용되는 실태이다. 이에 수화 사용자와 비사용자간 통역을 보다 간편하게 수행하기 위해 OpenNI에서 지원하는 Skeleton Tracking 을 키넥트 센서를 통해 구현하여 수화를 인식한 후 한국어로 변환하는 기능을 구현하고자 한다.

## 기존 기술

Skeleton Tracking은 인체의 골격을 이용하여 정보를 얻는 기술이다. 인체를 20개의 관절(Joint)과 19개의 뼈(Bone)의 형태로 인지하여 3차원 좌표상에서 센서 인지범위에 따라 각 골격과 센서간의 거리를 결정한다. 센서가 임의의 축(예시로 Z축)에 위치할 때 X축은 센서와의 거리를 각성분으로 측정하며 Y축은 고도성분으로 측정한다. 하나의 키넥트 센서는 2명의 골격만 Tracking 할 수 있다. 따라서 각 골격에 고유 ID를 부여한다. 수화를 인식하기 위해 좌우측 Hand, Elbow와 Spine 골격을 사용한다.

TensorFlow는 기계학습을 위한 라이브러리다. 구글이 오픈소스로 공개하여 일반 사용자도 접근하기 용이하다. 추상화를 통해 개발자가 간소화된 코드를 입력할 수 있다. 데이터를 학습할 때 GPU와 CPU를 모두 사용 가능하며 GPU 가속을 이용할 경우 대량 연산을 빠른 속도로 처리하여 결과를 산출하는 시간이 경감된다. 본 주제는 동작을 인식시키기 위해 많은 양의 영상 데이터의 학습을 필요로 하므로 고사양의 GPU와 TensorFlow를 이용해 학습에 필요한 시간을 줄일 수 있다.

## 제안

수화 인식을 위해 키넥트 센서를 사용한다.

키넥트 센서는 모션인식을 위해 별개의 센서를 필요로 하지 않고 다양한 상호작용 기능을 하나의 기기에서 구현 가능하다는 장점이 있다. 키넥트 센서는 한 개의 적외선 센서와 2개의 RGB 카메라 렌즈, 하단부와 양측면에 4개의 마이크로 구성되어 있다. 내부는 CPU, RGB 카메라, IR-Depth 센서, multi-array microphone 등의 센서를 가지고 있으며 PC에서 사용할 수 있고 OpenNI를 지원하여 C++ 개발환경을 조성할 수 있다. 키넥트 OpenNI 라이브러리에는 수화 인식에 필요한 골격인식 및 손 인식 기능이 탑재되어 있다. 키넥트 센서를 이용한 수화 인식 과정은 1. 동작 인식, 2. 영상데이터 산출, 3. 대상 동작의 기계학습(동일한 동작을 같은것으로 인식하기 위해 반복하는 과정), 4. 수화 모션데이터 산출로 이루어져 있다.

## 분석

### 영상 확인

키넥트 OpenNI 환경에서는 키넥트 센서를 이용해 인식된 영상정보를 최대 60 fps로 실시간 재생하거나 녹화하여 mp4 및 여러 동영상 확장자로 저장할 수 있다. 대상의 움직임은 키넥트 센서가 영상 데이터로 처리한다. 대상이 수화 'a'를 인식시킬 경우 센서를 통해 영상데이터 A가 생성된다.

## 인식과정

적외선 프로젝트가 적외선을 무수히 많은 점들로 투사하면, 오른쪽의 적외선 카메라가 대상에 반사 되어 돌아오는 것을 인식한다. 이를 통해 깊이 인식 및 모션인식이 가능하다. 가운데 렌즈는 RGB 카메라로 일반 카메라와 같은 RGB 영상을 인식한다. 키넥트 OpenNI 라이브러리에는 골격인식 및 손 인식 기능이 탑재되어 있다. 이를 이용하면 추측되는 골격이 세가지 상태로 결정된다. 'Tracked'은 관절의 연결과 위치가 정확히 인지된 상태이고, 'PositionOnly'는 관절의 위치 값은 알 수 있지만 정확히 어느 관절인지 인지하지 못한 상태이다. 마지막으로 'NotTracked'는 아무것도 인지되지 않은 상태이다.

이 중 인식을 위해서는 Tracked 상태의 골격이 필요하며, Tracked 된 경우 Tracking ID가 할당된다. 수화 인식에 필요한 Tracking ID는 총 5개로 HAND\_RIGHT, HAND\_LEFT, ELBOW\_RIGHT, ELBOW\_LEFT, SPINE이다. 다섯 개의 Tracking ID를 사용해 모션데이터를 추출한다.

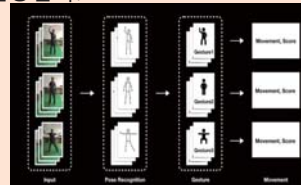
### 모션데이터 추출

OpenNI 개발환경에서 실시간 영상 및 인식된 데이터(Tracking ID)를 사용자가 지정한 폴더에 저장한다.

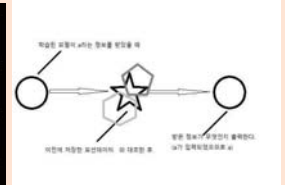
이를 C++ 환경에서 Load하여 새로 인식되는 정보와 교차비교할 수 있다.

앞서 생성된 영상데이터 A를 텐서플로우 라이트를 이용하여 학습시킨다. 학습이 완료된 영상데이터는 학습된 모션데이터  $\alpha$ 를 산출한다.

다른 수화를 인식시켜 학습한 수화를 올바르게 인식하였는지 검증한다.



pic.1 텐서플로우 라이트의 학습과정



pic.2 인식 및 결과 검증 도식

## 결론

키넥트 센서로 모션 캡처에 대해 적은 비용과 비교적 쉬운 방식으로 접근할 수 있다. 키넥트 센서는 수화에 필요한 골격을 인식하는데 어려움이 없다. 또한 OpenNI 개발환경이 지원되어 기존에 사용하는 C++ 환경에서 시스템 구축이 가능하다. 이 개발환경을 통해 수화를 인식하고 번역하는 프로그램을 개발할 수 있을 것으로 기대된다.

## 참고문헌

- 이현석; 김승필; 정완영. 키넥트의 모션 인식 기능을 이용한 수화번역 시스템 개발. 융합신호처리학회 논문지, 2013, 14.4: 235-242.
- SIN, Min-Ho, et al. Kinect 센서를 활용한 제스처 수화인식 시스템. Korea Information Processing Society Review, 2013, 20.3: 45-49.
- 김승희, et al. 키넥트를 이용한 수화 번역 시스템. 한국통신학회 학술대회논문집, 2016, 1077-1078.

