
프로젝트 주제 제안서



제 출 일	2023년 2월 20일 (월)
과 목 명	융합소프트웨어종합설계1
지도 교수	교수 이계식
학 과	컴퓨터공학과
학 번	2018250033, 2018250056
이 름	유규빈, 함상진

목차

1. 프로젝트 배경
2. 해결책
3. 기술 자료
4. 최종 결과물 및 목표
5. 예상되는 어려움
6. 활용 사례
7. 프로젝트 효과

1. 프로젝트 배경

시간이 지나면서 인터넷 방송의 발전으로 아프리카 TV, 트위치, 유튜브와 같은 인터넷 방송, 동영상 플랫폼의 이용자 수가 증가하고 이에 따라 인터넷 방송도 다양한 분야로 발전하고 하려는 사람 또한 많아졌다.

또, 메타버스에 대한 관심 증가로 VR 기술이 발전하였으며 인터넷 방송에서도 VR 기술이 다양한 콘텐츠에 사용된다. 그 중 하나로 버추얼 유튜버(V-Tuber)를 예로 들 수 있는데, 모니터 상의 실제 사람이 아닌 2D 혹은 3D 캐릭터를 이용해 방송하는 사람들을 말한다. 사람의 움직임을 VR 기기나 faceRig, Live2D 등의 프로그램을 이용하여 보여주는 데에 사용한다.

그러나 VR 기기는 착용해야만 움직임을 추적할 수 있고 가격이 비싸다는 단점이 있다. 또 장비 착용이 필요없는 faceRig 와 Live2D 는 표정과 간단한 머리의 움직임 정도만 추적 가능하며 그 외의 손동작, 몸의 움직임은 추적이 불가능하다.

2. 해결책

- 이미지를 분석하여 움직임을 추적해 사람의 자세를 추적해주는 컴퓨터 비전 기술을 사용한다.
- 딥러닝의 발전으로 Human pose detection, tracking 와 같이 성능이 좋아진 컴퓨터 비전 분야의 모델을 사용
- 위 모델을 이용해 VR 기기없이 영상 촬영만으로 사람의 움직임을 실시간으로 따라하는 캐릭터를 만들어 방송에 활용한다.

3. 기술 자료

자세 추정 모델

- Open Pose : 둘 이상의 대상을 추적하는 자세 추정 모델로 실시간에 좋은 성능을 보이는 모델
- Deep Cut : 이미지의 사람 수를 감지한 후 각 이미지의 관절 위치를 예측하는 모델로 Deep Cut 역시 둘 이상의 대상을 추적이 가능하다.

둘 이상의 대상 추적 모델

- Deep Sort : Object Tracking Frame Work(객체 추적 프레임 워크)로 SORT(Simple Online and Realtime Tracking)에 Deep Learning 을 접목시켜 성능향상을 이룬 Frame Work 이다.

모델 학습을 위한 데이터 셋

- Human3.6M : 약 360 만 장의 이미지를 가지고 있고 3D Pose Estimation 에서 가장 표준으로 사용된다.
- HumanEva-I : 약 4 만 장의 이미지를 가지고 있고 Human3.6M 이 표준으로 사용되기 전까지 HumanEva-I 가 주로 사용된다.

스켈레톤 모델 렌더링

- Unity

4. 최종 결과물 및 목표

- 기존에 존재하는 얼굴만 인식하는 프로그램과는 달리 VR 장비없이 두 명 이상의 상반신을 인식하는 프로그램
- 카메라로 입력받은 이미지를 딥러닝을 활용한 자세 추정 모델을 사용하여 사람의 자세에 대한 정보를 계산한다.
- 계산된 정보를 토대로 3D 모델을 실시간으로 렌더링한다.

5. 예상되는 어려움

- 과격한 움직임에 대한 트래킹
- 별다른 장비(키넥트) 없이 하나의 카메라로 깊이 정보를 얻는 방법
- 모델 학습에 필요한 데이터셋 준비

6. 구체적인 활용 사례

- 신체 부위를 노출하지 않고 시청자와 더 다양한 소통을 하고싶은 인터넷 방송
- 둘 이상의 사람에 대한 트래킹
- 전문적이지 않지만 간단한 모션 캡처
- VR 장비 없이 캐릭터와 카메라만 있다면 방송을 시작할 수 있는 접근성이 좋아진다.