
프로젝트 주제 제안서



국립한경대학교
HANKYONG NATIONAL UNIV.

제 출 일	2023년 3월 1일 (수)
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

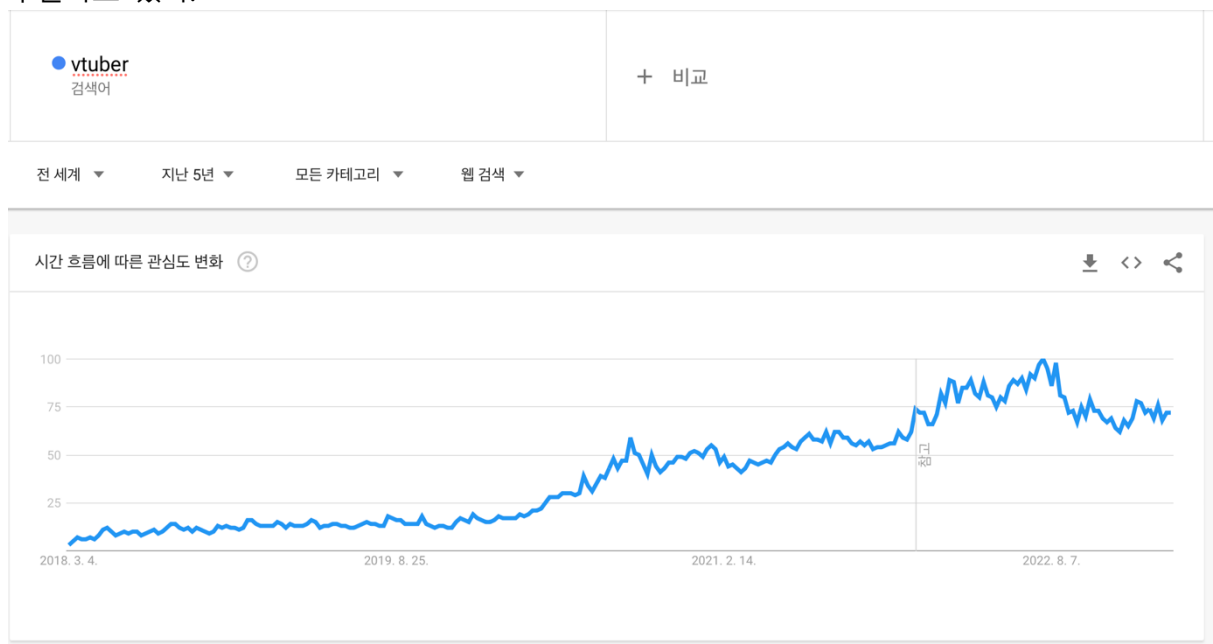
6. 활용 사례

7. 참고 문헌

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(Virtual Youtuber)를 예로 들 수 있는데, 버추얼 유튜버(버튜버)란 카메라나 특수 장비를 통해서 그 사람의 행동이나 표정을 대신 표현해주는 캐릭터가 등장해 방송을 진행하는 인터넷 방송인을 말한다. 버튜버라는게 다소 생소할 수 있는 장르일 수 있지만 구글 트렌드 분석에 따르면 코로나 19가 본격적으로 확산한 2020년 상반기부터 버튜버에 대한 관심도가 빠르게 증가하는 추세를 보였다. 업계에선 현재 활동 중인 버튜버만 1만 6000명이 넘는 것으로 추산하고 있다.



또, 아래 사진을 보면 알 수 있듯이 전세계 유튜버 일간 슈퍼챗 순위를 버튜버가 상위권을 차지하고 있다.

슈퍼챗 순위					2023.02.28.(화)	
순위	채널	슈퍼챗 수입	슈퍼챗 개수	영상		
1 ▲ 368	Laplus ch. ラプラス・ダー... #ホロライブ #Vtuber	+₩11,571,397	+456		▶ 32.4만	▶ 56.8만
2 ▲ 8	Hakos Baelz Ch. hololiv... #ホロライブ #hololive	+₩9,300,616	+280		▶ 13.8만	
3 ▲ 70	Kanata Ch. 天音かなた #ホロライブ #hololive	+₩9,217,293	+524		▶ 4	▶ 14.1만
4 ▼ 3	Pastor Jerry Eze #Streams of ... #church	+₩6,681,858	+825		▶ 30.7만	▶ 33.6만
5 NEW	Overly Sarcastic Produ... #Summary #Funny	+₩5,232,157	+78		▶ 3만	▶ 8.1만

버튜버는 표정이나 자세를 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 사용한다.
- 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 모델이다.
- 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

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

5. 예상되는 어려움

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

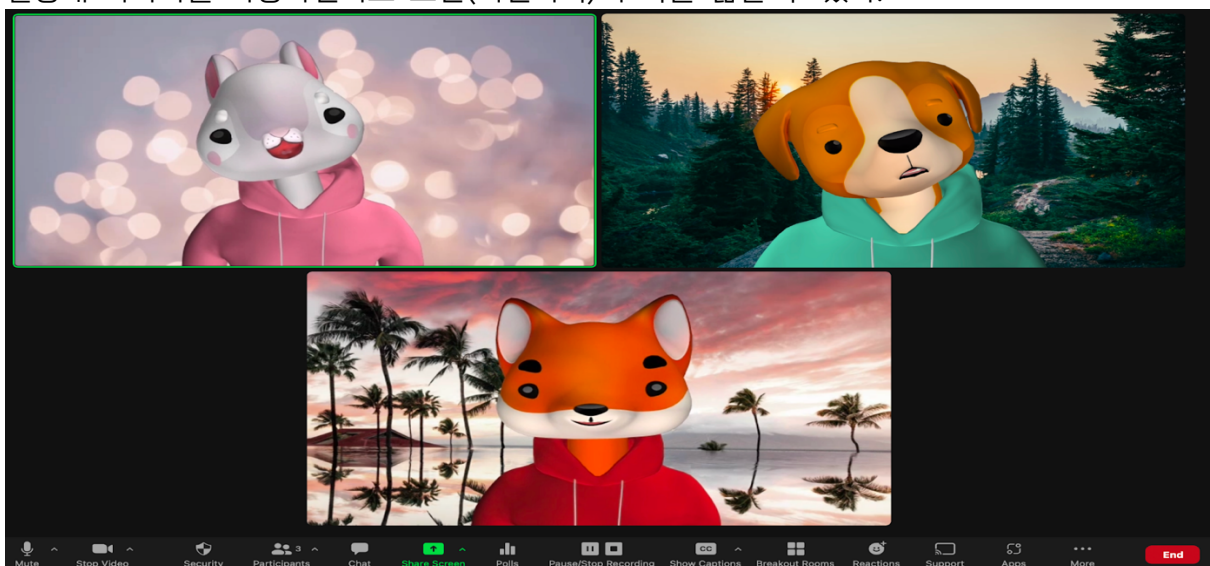
6. 활용 사례

자세를 인식해 계산한 정보를 토대로 3D 모델을 실시간으로 렌더링하는 기술을 처음에 언급한 버추얼 유튜버(V-Tuber)에 활용하려고 한다.

예를 들어, 아래 보이는 강서구청에서 운영하는 버튜버 "새로미"도 현재 몸통의 움직임은 인식하지만 팔 동작이 인식 안 되는 제한이 있다. 이때 우리가 개발할 상반신 인식 프로그램을 활용하여 더 다양하고 자유로운 비언어적 표현을 사용할 수 있게 만들 수 있다.



또, 줌에서도 공식적으로 아바타 모드를 업데이트한 만큼 화상회의에서도 아바타를 활용하는 경우가 많아졌다. 이러한 추세에 따라 화상회의할 때도 상반신 인식 프로그램을 활용해 아바타를 사용하면서도 표현(비언어적)의 폭을 넓힐 수 있다.



7. 참고 문헌

1. 구글 트렌드

(<https://trends.google.co.kr/trends/explore?date=today%205-y&q=vtuber>)

2. 전세계 유튜버 일간 슈퍼챗 순위

(<https://playboard.co/youtube-ranking/most-superchatted-all-channels-in-worldwide-daily>)

3. Human Pose Estimation 이란 ? (2022)

(<https://supermemi.tistory.com/entry/Human-Pose-Estimation-이란-2022>)

4. Human Pose estimation 관련 데이터셋

(<https://ctkim.tistory.com/entry/Human-Pose-estimation-관련-데이터-셋>)

5. i 강서 TV_강서구청 방송국 - 📻전국 지자체 최초!! 공무원 버튜브! 🇰🇷강서구🔥가 시작합니다 [V]loGS 1화

(<https://www.youtube.com/watch?v=K2LoboS8ePc>)

6. 줌 아바타

(<https://support.zoom.us/hc/en-us/articles/4642184011917-Using-Avatars-in-meetings-and-webinars>)