

손가락을 활용하여 증강현실과 같은 화면에서 그림을 그릴 수 있는 화상강의용 웹 플랫폼

A Web based Video Conference Platform that Draws in Augmented Reality by Fingers

안다숨(Dasom Ahn), 윤영기(YeongGi Yoon), 황채윤(Chaeyoon Hwang)

요약

현재 코로나 19로 인하여 화상 강의의 증가는 세계적으로 증가하는 추세를 보이고 있지만, 대면에 비교하여, 집중도 혹은 수업 태도 같은 결과를 보면 다소 부정적임을 알 수 있다. 이를 해결하기 위해 특수한 장비 없이 웹 브라우저 만으로, 화상강의를 도와줄 수 있는 웹 플랫폼을 제작해보았다. 손가락만으로 자유자재로 공중 혹은 현재 스크린에 자연스러운 필기가 가능하며, 이를 통하여 참여자 간의 자연스러운 소통 및 상호작용의 유도가 가능해진다. 복잡한 어플리케이션이나 프로그램을 다운받을 필요 없이, 웹으로 접속하여 실행만 하면 손쉽게 사용할 수 있도록 웹 환경에서 사용성을 주었다. 또한, 화면 공유시에 WebRTC의 Mesh 구조인 p2p를 적절하게 사용하여, 클라이언트 측의 부담을 줄였다.

키워드: 웹 플랫폼, OpenCV, WebRTC, ICE, 화상강의

Abstract

By effect of COVID19, increase of video conference has grown rapid, but the consequence isn't as good as it's growth regarding student's focus or participation. To solve this problem, we have developed a Web Platform that can help these kinds of video conference without needing any tools and only using a Web Browser. Using fingers, an instructor can write through the monitor like an augmented reality by moving around fingers in the air, and by doing so, it leads to interconnection between instructor and students. Without any downloading a program or an application, it just needs to connect via Web and launch. Also, while sharing a screen we used a WebRTC technique in mesh architecture to connect in p2p and reduce the burden of client.

Keyword: Web Platform, OpenCV, WebRTC, ICE, video conference

1. 서론

코로나19로 인해 여러 대학과 초, 중, 고등학교에서 원격 온라인 교육이 일상화되었다. 전 세계적으로도 화상강의의 증가율이 2.7배에 다다르고 있는데, 원격 수업으로 인한 단점도 발생하고 있다. 대표적으로, 교육자와 학생 간 의사소통 및 상호작용이 잘 이루어지지 않는 등 학습의 질이 하락하게 되었다. 또한, 원격 수업이 보편화되며 태블릿, 터치펜 등의 부가적인 장비가 필요하게 되었다. 특히 고가의 태블릿이 현재는 어린 학생들에게도 필수품이 되어버리는 상황까지 왔다. 한편 교육자의 입장에서 온라인 강의를 진행하는데 있어 태블릿, 터치펜, 칠판 등의 보조 장비가 필요한 것은 마찬가지이다.

이러한 화상영상의 증가 현상은 교육에서 뿐만이 아니다. 최근 들어, 사내에서 화상회의가 빈번하게 이루어지고 있고, 친구들 사이에서도 화상통화를 하는 경향이 늘어났다. 화상회의 및 통화는 상대의 얼굴을 볼 수 있다는 점에서 유용하지만, 즉각적인 기록이 어렵다는 단점이 있다.

따라서 상호 간 얼굴을 볼 수 있다는 장점을 유지하면서, 텍스트 공유도 유연하게 할 수 있는 'AR Note'를 개발하였다. 카메라 화면 위에서 다른 도구 없이 손가락만을 활용하여 필기를 할 수 있고, 손가락 개수를 조정하며 여러 기능을 활용할 수 있다. 이는 증강현실의 개념으로, 화면을 바라보며 즉각적으로 텍스트를 추가할 수 있다는 장점이 있고, 사물 위에 필기를 해야 하는 경우에도 유용하게 쓰일 수 있다. 더 나아가 온라인 학습으로 저조하게 된 학습률의 증진과, 화상영상 시 요구되는 고가의 보조 장비의 부담을 줄이는 효과가 기대된다. 본 논문에서는 'AR Note'의 사용방법, 구조, 핵심 모듈에 대해 설명할 것이다.

2. 관련 연구

2.1 Open-CV

Open-CV는 Open Source Computer Vision의 약자로, 영상 처리에 사용할 수 있는 오픈 소스 라이브러리이다. 실시간 이미지 프로세싱에 중점을 두어, 윈도우, 리눅스 등 다양한 OS에 실행 가능한 크로스 플랫폼이며, 오픈소스 BSD 라이선스에 위반되지 않는 한 무료로 사용 가능하다. Tensorflow, Torch / PyTorch 등 딥러닝 프레임워크를 지원한다.

프로젝트를 진행하기 위해, 손동작을 바로바로 인식 할 수 있는 모듈을 딥러닝으로 구현하려고 하였으나, 학습되는데 시간과 데이터가 부족하여 프로젝트를 진행하는데 있어, 이미 구글에서 학습이 된 모듈인 MediaPipe를 이용하여 프로젝트를 완성하였다.

손가락을 활용하여 증강현실과 같은 화면에서 그림을 그릴 수 있는 화상회의용 웹 플랫폼 개발

2.2 MediaPipe

MediaPipe는 구글에서 학습이 된 오픈소스 프레임워크로, 머신러닝 솔루션이며, 실시간 미디어 처리에 강세를 가지고 있다. 손 뿐만이 아닌, 얼굴인식, 포즈, 사물인식 등 사람뿐만이 아닌 사물에 대하여 까지도 다양하게 학습이 완료되어 있다. 또한, Android, iOS, C++, Python 등 다양한 환경에서 작업 할 수 있는 크로스 플랫폼이기에, 추후에 프로젝트의 확장 가능성을 생각하여 학습 모듈로 선택하게 되었다.

2.3 WebRTC

Web Real-Time Communication의 약자로, 별다른 소프트웨어 없이, 카메라, 마이크, 화면 등을 사용하여, 실시간 커뮤니케이션을 제공하는 기술이다. 즉, 잘 알고있는 화상통화 혹은 화상 공유를 구현할 수 있는 오픈소스이다. 비디오 및 음성 데이터를 P2P 방식으로 전송할 수 있게 지원한다. Latency가 낮고, 별 다른 소프트웨어 필요 없이 실시간 커뮤니케이션을 할 수 있기에 본 프로젝트에 사용하였다. 또한, JavaScript API로 제공되어, 웹 플랫폼에 사용하기에 적합하였다.

3. 사용 방법

3.1 사용 화면

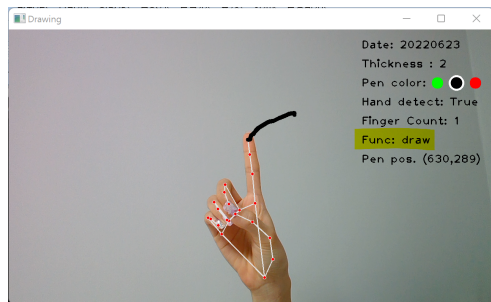


그림 1. 손가락으로 그리는 모습을 보여주는 파이썬 모듈

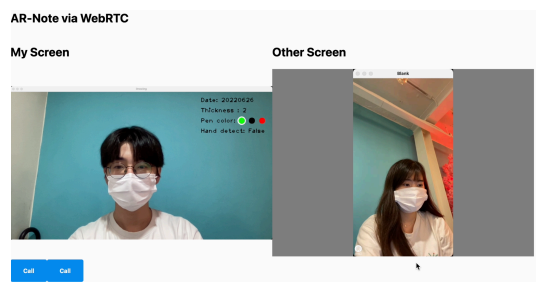


그림 2. WebRTC로 연결된 화면

Index	Name	Description
1	Date	현재 날짜 표기
2	Thickness	사용 중인 펜의 두께 표기
3	Pen color	사용 중인 펜 색상 표기
4	Hand detect	손이 인식되었는지 여부 표기
5	Finger Count	손가락 개수 표기
6	Func	현재 손가락에 대한 기능 표기
7	Pen pos.	펜의 좌표 표기

표 1. 사용화면에 대한 설명

3.2 손가락 개수에 따른 기능

Num	Function
0개	화면 스크린샷 촬영
1개	그리기
2개	그리기 멈춤 및 펜 이동
3개	펜 두께 변경
4개	펜 색상 변경
5개	그린 내용 전체 삭제

표 2. 손가락 개수에 따른 기능 설명

4. 설계

본 장에서는 애플리케이션 제작을 진행하며, 구축된 전체적인 시스템 구조와 데이터의 출처와 같은 지식을 설명한다. AR-Note의 전체적인 구조는 그림 5와 같으며, 크게 파이썬 모듈 내부와 WebRTC 환경이 있다.

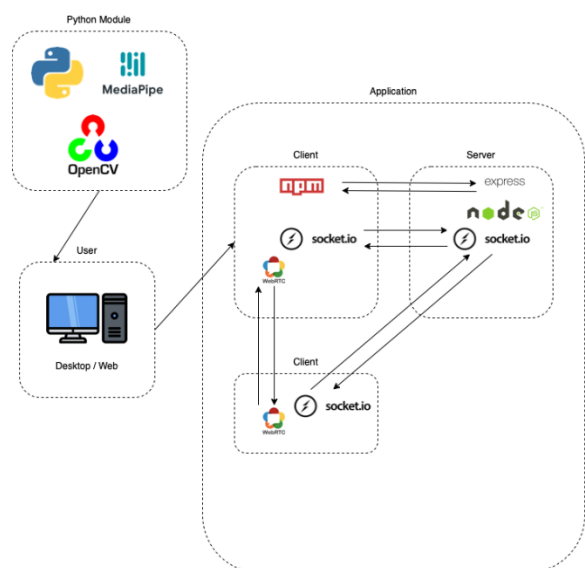


그림 5 AR-Note 시스템 구조

4.1 Python모듈

본 프로젝트에서는 손을 인식하여 증강현실에서 그림을 그릴 수 있는 부분을 파이썬 모듈로 만들어서 사용하고 있다. Open-CV 와 mediapipe를 사용하여, 모듈을 제작하였으며, 손가락의 개수 및 위치에 따라 어떠한 행동을 할 지 정하여 유저가 원하는 행동을 할 수 있다.

4.2 WebRTC

webRTC는 p2p기반의 서비스로, 처음 peer들을 연결할 때만 서버가 개입하고, 이후에는 서버 없이 단말 간에 데이터를 실시간으로 주고받을 수 있게 한다. 이전에 통신 기업들이 독점하던 기술이 현재는 구글이 오픈소스화하며 표준화되었다. 대표적 활용 사례로는 최근 붐이 일어났던 클럽하우스, 슬랙, 아자르, 카카오톡, 구글미트 등이 있다.

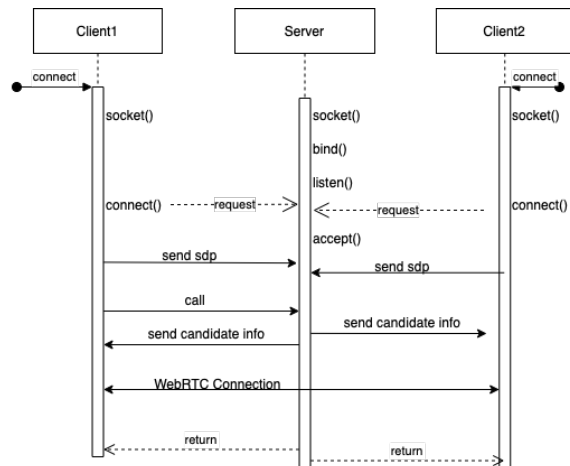


그림 3. webRTC 통신 과정

본 프로젝트에서 webRTC를 도입한 목적은, p2p 기반의 서비스가 실시간 화상회의 시스템에 적용되면 유리하기 때문이다. 화상회의가 시작되면, 회의에 참가한 여러 참여자가 실시간으로 데이터를 주고받게 된다. 이때 데이터의 형식이 오디오, 비디오이기 때문에 latency가 작을수록 유리한데, 이를 webRTC 기술이 실현시킬 수 있다. 이 외에도 webRTC 기술을 사용하면, 설치나 플러그인이 불필요하다는 장점도 가질 수 있다.

4.3 시스템 구조

AR-Note의 시스템 환경은 크게 Python 모듈과 WebRTC 모듈로 구성된다. Python 모듈에서는 Mediapipe의 Hands를 사용해 손가락을 구별하고, 펼쳐진 손가락의 개수가 계산된다. 3장에서 설명한 대로 개수별로 다른 기능을 구현하였고, 이후 OpenCV 라이브러리로 화면 구성과 필기 과정이

이뤄진다.

강의 진행자와 피진행자간 화면 통신은 WebRTC 모듈에서 구현된다. 서버 연결은 node.js를 사용하였으며, Socket IO로 서버와 클라이언트를 TCP 방식으로 연결해주었다. 통신에 참여할 참여자들이 모두 준비됐을 경우, WebRTC를 이용해 참여자 간의 데이터 전달만이 이루어진다. 이 모든 과정은 별 다른 설치 없이 진행되며, 서버를 따로 거치지 않기에 부담이 적고 속도가 빠르다.

5. 결론 및 향후 연구

Covid-19으로 인하여, 화상강의는 뗄 수 없는 존재가 되었다. 하지만, 이러한 화상강의를 도와주는 도구가 존재하지 않아, 많은 학생 및 강사들이 어려움을 겪고 있다. AR-Note를 활용하여, 강사들과 학생들간의 교류가 더 많아지고, 학습에 있어서 도움을 받을 수 있다. 하지만, 현재 AR-Note를 로컬에서 실행하여 화면공유를 하는 형태로 존재한다. 더 높은 접근성을 위하여, 서버에서 Python Module을 실행 후 화면만 받을 수 있게 넘겨준다면 더 수준 높은 플랫폼이 제공될 것이라 생각한다. 또한, 현재 WebRTC는 기존 라이브러리가 제공하는 ICE 및 Stun 서버 즉 일반적인 NAT를 활용한 통신만 가능하다. 즉 보안정책이 강하거나 사설망을 사용하는 경우, 이러한 Stun 서버로만 통신이 불가능하여 Turn 서버를 구현하여 모든 연결망을 p2p 방식으로 소통할 수 있게 만들 예정이다. 또한, 현재 사용하고 있는 Mesh구조인 경우에, Client가 4~5명 이상이 되면, Client의 부담이 너무 강해 제대로 활용이 어렵다. 이러한 부분은 개선하기 위해, 미디어 서버인 SFU 를 중개 서버로 두어, Client의 부담을 줄일 예정이다.

참고 문헌

- [1] 김지원, 박영신, 김경이 and 양길석. (2021). COVID-19에 따른 대학 온라인 수업에 대한 교수자와 학습자의 인식 및 경험 분석. 교육연구, 80, 33-58.
- [2] Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C., Yong, M.G., Lee, J., Chang, W., Hua, W., Georg, M., & Grundmann, M. (2019). MediaPipe: A Framework for Building Perception Pipelines.
- [3] Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C., & Grundmann, M. (2020). MediaPipe Hands: On-device Real-time Hand Tracking.
- [4] Cui, Jian & Lin, Zhuying. (2016). Research

손가락을 활용하여 증강현실과 같은 화면에서 그림을 그릴 수 있는 화상강의용 웹 플랫폼 개발

and Implementation of WebRTC Signaling via
WebSocket-based for Real-time Multimedia
Communications.