

GLOBAL TELECOM

글로벌텔레콤이 꿈꾸는 IT WORLD

디지털 컨버전스 시대를 선도하는

IT 서비스 전문 기업

GlobalTelecom

(주)글로벌텔레콤은 급변하는 IT환경에 부응하여 시장의 기술 흐름을 한걸음 앞서 준비하여 보다 미래 지향적인 가치를 창출 할 수 있도록 최선을 다하고 있습니다

26년간 축적된 기술력

26년간의 풍부한 경험을 토대로 축적된 유무선 네트워크 기술력을 근간으로 Optical Network, 정보 보안 및 물리 보안 솔루션, 망분리 솔루션, 미디어 솔루션 등에 특화된 기술력을 갖추고 있으며, 이를 토대로 종합적이고도 광범위한 서비스를 제공하고 있습니다.

표준 프로토콜 기반 사용자 그룹별 Mobility

다양한 업계의 고객 레퍼런스

정부 및 공공기관, 미디어, 금융, 증권, 통신, 게임 및 포털, 모빌리티, 학교, 병원, 일반 기업체 등 다양한 업계의 고객 레퍼런스를 보유하고 있으며, 이를 통하여 더욱 더 전문적인 서비스를 제공하고 있습니다.

전문적인 서비스 제공

각 분야 솔루션 제조사와의 파트너십

Cisco, Extreme, Citrix, Arista, Paloalto Networks, FireEye, Ateme 등 다양한 분야의 솔루션 제조사와 파트너십을 체결하였으며, 이러한 제조사와의 기술협력을 통하여 사용자 환경과 요구에 최적화된 서비스를 제공하고 있습니다.

제조사와의 기술협력

인정받은 정보통신공사업 시공역량

당사는 정보통신공사업 면허를 소지하고 있는 우량 중소기업체로서 정보통신공사협회에서 공시한 시공능력평가 순위를 통해 그 능력을 인정받고 있습니다.

전국 10위 / 11,071 업체
서울 8위 / 2,395 업체

회계 안정성 확보

공인 기술신용평가등급이 A0 로써 고객과의 비즈니스 거래에 있어서 안정된 회사의 이미지로 좋은 평가를 받고 있습니다.

회사 신인도 상승



서울 본사

주소 : 서울시 영등포구 영등포로 150 생각공장 당산 A동 1304~1310호

전화 : 02.3483.0900

www.globaltelecom.co.kr

2023 IIBC

2023 IIBC CONFERENCE

www.iibc.kr

종합학술대회 논문집

제21권 제1호

일시 | 2023년 11월 16[목]~18[토]

장소 | 한국과학기술회관(KOFST), 제주메종글래드호텔

주관 및 주최 | (사)한국인터넷방송통신학회(IIBC), (사)국제문화기술진흥원(IPACT), 지식의 숲(KF)

후원 | 과학기술정보통신부, 한국연구재단, 한국과학기술단체총연합회, 한국과학기술회관(KOFST)

협찬 | (주)글로벌텔레콤, 한국정보기술(주), 한국ICT기술협회 부설 한국ICT인재개발원

종합학술대회 논문집

제 21 권 제 1 호

IIBC

한국인터넷방송통신학회

IIBC (사)한국인터넷방송통신학회
The Institute of Internet, Broadcasting and Communication

증강현실 기기 상에 음성인식(Speech to Text)을 통한 청각 장애인 지원 앱 개발 사례

Best Practices on supporting the Hearing Impaired Person through Speech to Text (STT) Mechanism on AR Device

임지성*, 이동규**, 박찬솔***, 김영철****

Ji-Sung Lim*, Dong-Kyu Lee**, Chan-Sol Park***, R. Young Chul Kim****

jsleem0400@g.hongik.ac.kr*, dksky99@mail.hongik.ac.kr**, c2193102@g.hongik.ac.kr***, bob@hongik.ac.kr****

요약

현재 청각 장애는 선천적 혹은 사고나 노화로 소리를 들을 수 있는 능력이 상당히 떨어진 상태이기에 소리로 얻을 수 있는 정보를 얻지 못하는 상태를 의미한다. 본 논문에서는 소리 정보를 시각적인 측면에서 정보를 얻을 수 있는 방법을 제안한다. 즉 증강현실 디바이스를 통한 소리 시각화이다. 이를 위해 두 개의 모듈을 사용할 모듈이 필요하다. 첫 번째는 일상생활 속 원활한 대화를 위한 STT를 이용한 음성 시각화이고, 두 번째는 데시벨에 따른 위험 감지 시각화를 통한 위험 예방이다. 이를 통해 소리 정보를 습득할 수 없었을 때에 비해 시각적 정보를 받아들일 수 있다. 이를 통해 앞으로 청각 장애인들의 일상생활이 좀 더 나아질 것을 기대한다.

키워드 : 청각장애인, 홀로렌즈, STT, 소리 시각화

I. 서론

본 논문은 2023년 홍익대학교 소프트웨어융합학과 종합 설계 과목에 결과물이다. 청각 장애인 소리를 들을 수 있는 능력이 상당히 떨어져 있거나 전혀 들리지 않는 상태의 장애인이다. 한쪽 귀에서만 발생할 수도 있고 양쪽 귀 모두에서 발생할 수도 있다. 원인으로서는 유전, 감염, 노화, 뇌막염, 교통사고 등이 있다. 청각 장애인들은 일상생활에서의 청각적 정보를 얻을 수 없기에 사람들과의 원활한 의사소통과 큰 소리를 통한 위험 감지가 불가능하다. 실제로 도로교통공사에서 측정한 자료의 경우 정상적인 사람이 음향기기 등으로 귀를 막았을 경우 소리 정보를 습득하지 못해 사고 위험이 늘어났다[1]. 본 논문에서는 소리 정보를 시각적인 측면에서 정보를 얻을 수 있도록 AR 디바이스를 통한 소리 시각화를 진행하였다. 이를 통해 시각적으로 표현된 소리 정보를 가지고 청각 장애인

인들의 일상생활에서 기존보다 좀 더 원활한 생활을 할 수 있을 것이다.

II. 관련 연구

2.1. 기존 청각 보조 도구

첫 번째 연구는 수화 동작을 스마트폰과 연동하여 음성과 텍스트로 통역해주는 수화 통역 장갑과 상황 인지 보조 IoT 장치이다[2]. 해당 연구의 시스템 설계는 AWS IoT Button과 Arduino UNO R3 보드를 기반으로 진행되었다. 구현된 장갑은 손가락 구부림 정도를 측정할 수 있는 Flex Senso, 손동작의 백터값을 추적하는 가속도 센서로부터 Arduino에서 선형 데이터를 구성한다. 구성된 데이터는 블루투스 통신으로 스마트폰에 전송된다. 전송된 데이터는 어플리케이션의 수화 동작의 특징을 담은 선형데이터와 비교하여 해당 동작이 수화인지 지화인지 판별한다. 데이터베이스의 값과 일치할 경우, TTS엔진을 이용하여 해당 단어를 스마트폰 스피커를 통해 음성으로 출력한다.

상황 인지 보조 IoT 기기는 AWS IoT Button은 연결 가능

*홍익대학교 소프트웨어융합학과 학부생
**홍익대학교 소프트웨어융합학과 학부생
***홍익대학교 소프트웨어융합학과 석사 과정
****홍익대학교 소프트웨어융합학과 교수

한 AP에 네트워크 통신을 위한 통신 인증을 설정한다. 해당 AP를 통하여 AWS Lamda를 호출한다. 호출된 AWS Lamda는 이벤트에 응답하여 코드를 실행한다. AWS Lamda에서 어플리케이션으로 클라우드 메세징 서비스 Firebase Cloud Messaging을 이용한 코드로 푸시 메시지를 보내면 사용자의 장갑에 부착된 LED 센서를 이용해 사용자에게 시각적 정보를 전달한다.

다음 연구는 딥 러닝을 이용한 소리 방향 및 종류 식별 시스템이다[3]. 해당 연구는 소리 종류 식별 모듈 구현을 위해 python 2.7버전과 TensorFlow를 사용하였다. 소리 특징 추출 및 입력 데이터는 UrbanSound8K에 Python 라이브러리인 LibROSA를 이용하여 특성을 추출하였다. 딥 러닝 신경망 구성으로는 신경망은 완전 연결 신경망으로 3개의 은닉 계층에 각각 뉴런을 300개, 200개, 100개 설정했다. 은닉, 출력 계층의 가중치 및 편향의 초기화는 평균이 0이고 표준 편차가 입력 데이터의 크기(193)의 제곱근의 역수인 정규분포를 따르는 임의의 값으로 설정한다. 기존 방식은 활성화 함수로 Sigmoid, optimizer로 Gradient Descent optimizer, learning rate은 0.1을 사용하였지만 제안 방식은 활성화 함수로 ReLU 함수, optimizer로 Adam optimizer, learning rate은 0.0001을 사용했다. 비용의 흔들림이 적고 빠르게 0으로 수렴해 나가는 제안 방식을 선택하였고, 훨씬 빠르고 정확도가 기존 방식보다 약 7% 개선 되었다.

마지막 연구는 청각 장애인을 위한 소리 감시 어플리케이션이다[4]. 해당 연구는 소리 감시를 위해 아두이노를 이용하여 청각장애인의 특정 소리를 감지하여 사용자에게 알려주는 생활보조기구를 개발하였다. 사용자가 어플리케이션의 실외 환경 모드를 실행시키면 백그라운드 환경에서 소리 신호를 실시간으로 감지하고 분석한다. 분석한 데이터는 변환 과정을 거쳐 double 형 값으로 변환한다. 변환된 오디오 데이터는 데시벨을 구하는 계산을 통해 값을 얻는다. 갑작스러운 큰 소리가 나는 경우를 판별하기 위해 실시간으로 들어오고 있는 소리 신호 값과 누적된 평균값의 차가 15dB 이상일 경우 사용자에게 주의를 준다.

2.2 Kospeech

Kospeech는 Python을 기반으로 하는 모듈식 및 확장 가능한 End-to-End 방식의 한국어 자동 음성 인식(ASR) 툴킷이다.[5] 한국어 음성인식을 위해 문자 하위 단어, 문자소에 대한 데이터 셋 처리 방법을 구성하였다. 특징 추출 방법으로는 Recurrent Neural Network(RNN) 등 순환 신경망 유형, LongShort Term Memory(LSTM), Gate Recurrent Unit(GRU) 등 다양한 옵션을 제공할 하는 툴킷이다. Kospeech의 모델 아키텍처는 Deep Speech 2, Listen Attend

and Spell, Transformer, Joint CTC-Attention LAS의 네 가지 모델로 구현된다.

III. AR 디바이스 기반의 청각 장애인 보조 시스템 설계

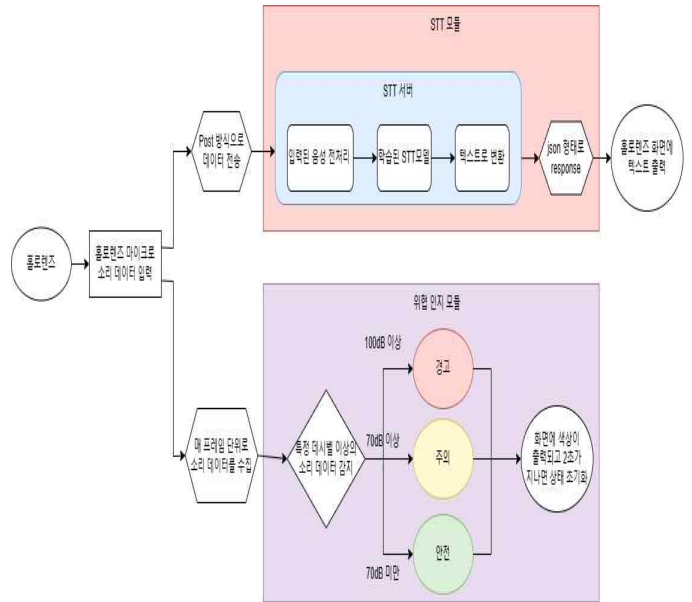


그림 1 AR 디바이스 기반의 청각장애인 보조 시스템 순서도

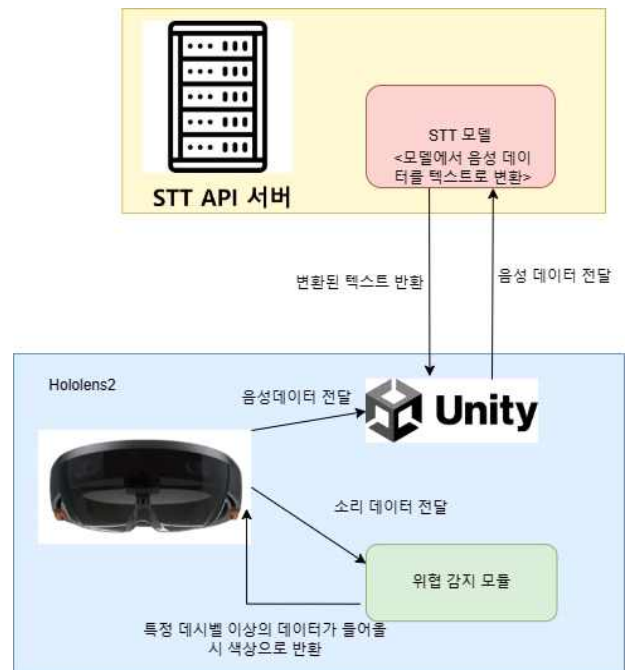


그림 2 전체 시스템 구조도

그림1은 본 논문에서 제안하는 AR 디바이스 기반의 청각 장애인 보조 시스템의 순서도이다. 이를 보면 AR 디바이스에

서 받아들인 소리 정보가 각각의 모듈에서 어떻게 진행되는지 확인할 수 있다. 그림2는 전체 시스템 구조도이다. 청각장애인 보조 시스템은 두 가지의 모듈로 구성되어 있다. 홀로렌즈의 마이크를 통해 소리 데이터를 수집을 하고 수집된 데이터를 각각의 모듈로 전달한다.

STT 모듈의 경우 AR 디바이스 상에서 구현되려면 STT 모델이 너무 무겁기에 서버를 구축하고 서버 안에서 딥 러닝을 할 수 있도록 구현하였다. 데이터를 서버로 전달하고 전달 받은 데이터는 STT모델에 맞게 데이터 전처리를 진행한다. 전처리가 완료된 데이터는 학습된 STT 모델에 들어가게 되며 음성 데이터는 텍스트 형식으로 변환된다. 변환된 텍스트는 AR 디바이스 화면에 출력하게 된다.

위험 인지 모듈의 경우 실시간으로 소리 데이터를 받게 되며 받은 데이터에서 데시벨을 측정, 측정된 데시벨이 기준(100dB이상 위험, 70dB 이상 경고, 70dB미만 안전)에 따라 구별하고 데시벨에 맞는 색상(위험: 빨강, 경고: 노랑, 안전: 초록)을 화면에 출력한다. 화면에 출력되고 2초 이상 해당 데시벨의 소리 데이터가 입력이 없으면 화면 상태를 초기화하도록 구현하였다.

IV. 적용사례

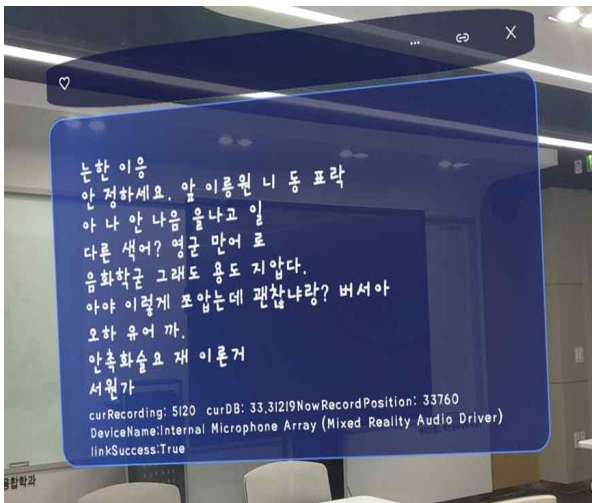


그림 3. STT 모듈 적용 결과

적용 사례로써 청각 장애인 보조 시스템을 홀로렌즈2 기기에 적용했다. 그림3은 STT모듈을 홀로렌즈에 적용하여 홀로렌즈 화면상에 텍스트를 출력하였다. 홀로렌즈에서 수집한 음성 데이터를 STT API 서버로 보내고 서버 내부에 있는 STT모델을 통해 텍스트로 변환하고 이를 홀로렌즈 화면상에 출력하였다.

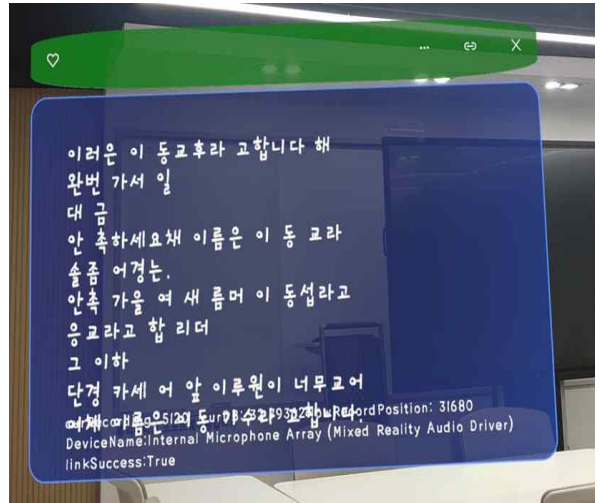


그림 4 위험 감지 모듈 적용 결과

그림4는 홀로렌즈의 마이크를 통해 들어온 소리 데이터를 위험 감지 모듈에서 데시벨 측정한다. 측정된 데시벨의 결과 값을 홀로렌즈 화면상에 색상으로 출력했다. 촬영 당시의 데시벨이 70dB미만이었기에 화면상에서 출력되는 색상은 초록색으로 표현되었다.

V. 결론 및 향후 연구

본 연구에서는 AR 디바이스 기반의 청각장애인 보조 시스템 구현을 제안했다. 이를 홀로렌즈와 Kospeech STT 모델을 API 서버에 적용하여 입력된 음성 데이터를 텍스트 형식으로 화면상에 출력하였다. 또한 소리를 통해 인식할 수 있는 위험을 색으로 시각화하여 청각 장애인이 반응 할 수 있도록 했다. 이를 통해 일상생활 속 타인과의 의사소통에 대해서 수화를 모르더라도 원활한 대화가 가능하다. 또한 소리를 통해 인지할 수 있는 위험을 도구를 통해 시각으로 인지할 수 있다. 추후 연구로써, 현재 개발한 시스템은 프로토타입으로서 현재 다소 떨어지는 STT 모델의 정확도를 개선해 나아갈 것이다. 데시벨 측정 모듈에서는 위의 바에서만 적용되는 색상을 화면 전체적으로 나타나게 하여 가시성을 갖출 예정이다. 또한 소리 분류 모델을 적용시켜 위험이 되는 소리가 무슨 종류의 위험인지 구분할 수 있도록 발전시킬 계획이다.

참고 문헌

- [1] 강수철, 류준범 "보행 중 음향기기 사용이 교통안전에 미치는 영향 연구" 2015.
- [2] 구성민, 장인국, 손영성 "청각장애인을 위한 오픈 소스 하드웨어 기반의 수화통역 장갑 & 상황 인지 보조 IoT 장치" 정보과학회 컴퓨팅의 실제 논문지. 2018-04 24(4):04-209

- [3] 김세영, 김현웅, 박찬호, 정목동 “청각 장애인을 위한 딥러닝 기반 소리 방향 및 종류 식별 시스템”한국정보과학회 학술발표논문집. 2017-06 2017(06):1896-1898
- [4] 이규진, 손현정, 최현석, 이상호 “청각장애인을 위한 관심소리 인식 및 알람서비스”한국통신학회 학술대회논문집. 2016-11 2016(11):152-153
- [5] 김수환, 배세영, 원채형 “KoSpeech: Open-Source Toolkit for End-to-End Korean Speech Recognition”In: Software Impacts. (Software Impacts, February 2021, 7)