

청각 장애인을 위한 스마트 글라스, SoundView

국민대학교 소프트웨어융합대학
소프트웨어학부 고가을, 김예린, 류성호, 정승우

1. 프로젝트 추진 배경 및 주요 기능

청각장애인은 위기 상황에 대한 정보의 습득이 즉각적이지 않아 빠르게 대처할 수 없게 되고 이는 심각한 인명피해로 이어질 수 있다[1]. 또한 재난 상황만이 아니라, 일상 생활의 전반적인 영역에서도, 의사소통의 단절로 인해, 음성으로 전달되는 정보들에 대한 소외가 발생되며, 교육에서조차도 기회 제한이 발생된다[2].

따라서 우리는 세 가지 주요 기능을 중심으로, 청각장애인이 일상생활에서 불편함을 느낄 만한 요소들을 해소하고자 한다.

- 1) 소리의 크기 안내 기능을 제공한다.
- 2) 일상 생활에서 인지가 필요하다고 판단되는 소리에 대한 알림 기능을 제공한다.
- 3) 상대가 말하는 소리를 텍스트로 바꾸어 투명 디스플레이에 띄워주는 기능을 제공한다.

2. 시스템 설계

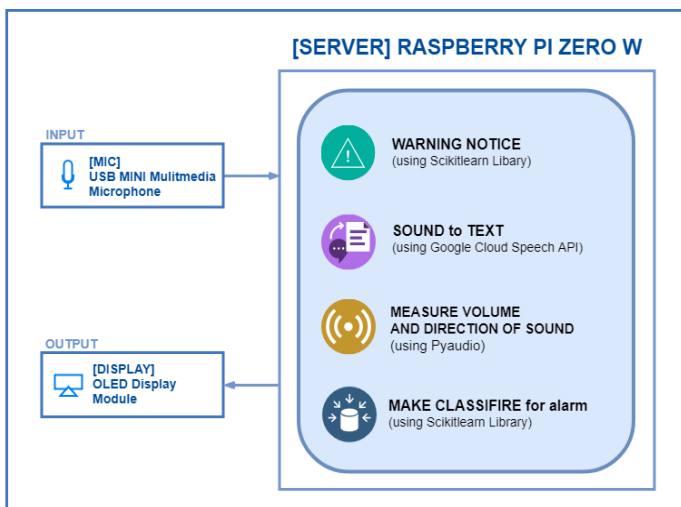


그림 1 시스템 구조도

2.1 소리의 크기

소리의 크기는 위험 상황을 판단할 수 있는 우선적인 방법이다. 일반적으로 소리의 크기로 대표되는 진폭(db)을 제공하고자 한다. 소리의 크기(db)는 아날로그 신호를 분석 가능한 Audio Level 로 실시간 변환하여 측정한다. Audio Level 로 변환된 디지털 신호의 크기는 OLED 화면을 통해 사용자에게 제공된다. 실시간 소리의 크기는 python 라이브러리를 사용해 수치화 하고, 이를 사용자 에게 잘 보여지도록 Audio Level Meter 중 peak 값을 사용한 변환 방법을 사용한다 (그림 2 참조).

실시간으로 측정된 일정 구간의 Audio Level 의 평균값은 데이터의 최대, 최소값의 차를 이용해서 구한다. 이를 통해 최종 변환된 Audio Level(dBu)를 실시간으로 사용자에게 보여지도록 한다.

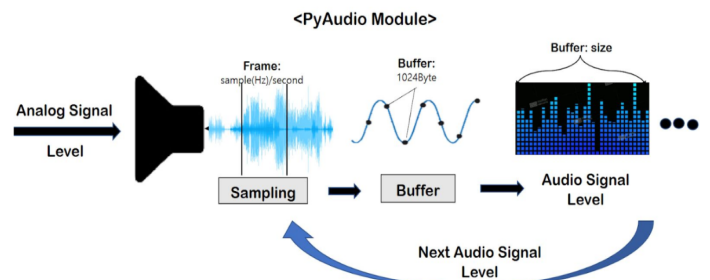


그림 2 PyAudio 모듈 작동의 흐름도

2.2 소리 알림

소리 알림을 위해서, 실시간 소리를 일정한 길이로 잘라 wav 형식으로 저장하고, MFCC 알고리즘을 이용하여 특징데이터를 추출해 npy 파일로 저장한다. 웹 상에서 수집한 다양한 소리 데이터에 위와 같은 과정을 거쳐 다수의 특징 데이터 집단을 생성하고, 로지스틱 회귀 분석을 통해 일정 수준이상의 분류 성능을 내는 분류기를 구현한다. 이후 실시간으로 들어 오는 소리에도 위와 같은 특징 추출 과정을 거쳐 분류기에서 그 소리의 유형을 판단한다.

분류 성능 확인을 위해 시각화한 혼돈 행렬에서는 왼쪽 상단에서 오른쪽 하단까지의 대각선이 짙은 남색이며, 나머지 영역은 밝은 색일수록 분류 성능이 좋다는 것을 의미한다. 아래 <그림 3>의 [행렬 1]에서, 소리 데이터 특징 추출을 위한 FFT 기반 분류기는 높은 성능을 기대하기는 어렵다. 좀 더 정확한 분류를 위해 멜 주파수 캡스트럴 계수인 MFCC 를 사용하고, FFT 를 이용한 [행렬 1]과 비교하여 MFCC 를 이용한 [행렬 2]가 분류 성능을 사용함으로 성능이 향상될 수 있다.

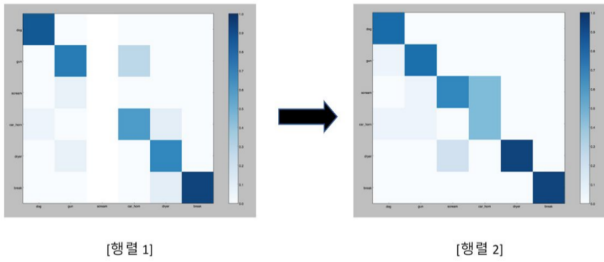


그림 3 분류기 별 예측도 구분

2.3 텍스트 변환

소리를 텍스트로 바꾸는 음성 인식 API 는 구글에서 제공하는 Google Cloud Speech API[3] 를 사용하였다. 인식되는 음성 언어에 대해서 여러 의미를 가진 문장으로 인식이 될 수 있지만 상위의 대안을 가진 문장 하나만 출력을 한다. 음성을 텍스트로 바꾼 결과값을 실시간으로 투명 디스플레이에 띄워주는 기능을 제공하기 위해서 반복할 때마다 음절 단위로 바뀐 텍스트를 투명 디스플레이에 실시간으로 업데이트한다.

3. 시스템 설계

3.1 개발 환경

소프트웨어는 Ubuntu 기반의 환경에서 Python 개발도구와 Google 사의 Google speech api 라이브러리들을 활용하였다. 스마트 안경의 제작을 위해서는 Raspberry Pi [4] 보드에서, 아두이노 모듈 및 Edutige Microphone 을 사용 하였으며, 출력을 위해서는 0.96 Inch OLED Display Module 을 사용 하였다.

3.2 하드웨어 구조도

스마트 글라스 구조는 <그림 4> 와 같이 구성 한다.

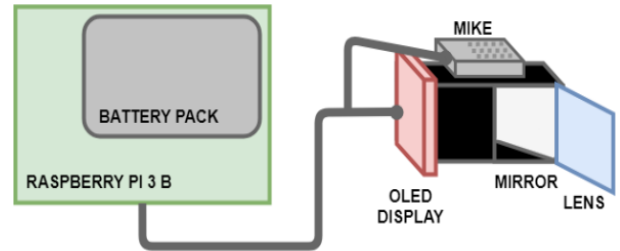


그림 4 스마트 글라스 구조도

아래 <그림 5>는 OLED 디스플레이의 광원이 사용자의 눈에 보이기까지의 과정을 시각화한 것이다.

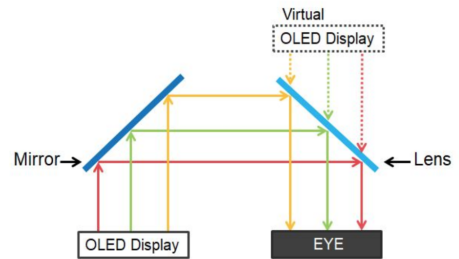


그림 5 OLED 광원이 반사각으로 투시 되는 과정



그림 6 OLED 모듈의 텍스트 출력

참고문헌

- [1] 한국 장애인 고용공단 고용개발원, “한눈에 보는 2017 장애인 통계”, 조사통계 2017-01 호.
- [2] 권순황, “교육과정 변천에 따른 청각장애 교육과정 적합성 고찰”, 한국청각.언어장애 교육연구 vol.7, no.2, pp1-21. 2016.
- [3] Cloud speech api, <https://cloud.google.com/speech/>
- [4] 라즈베리 파이, <https://www.raspberrypi.org/>