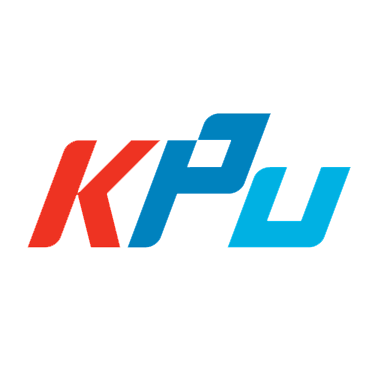
****

**스마트 스피커**

|  |  |
| --- | --- |
| **작성자** | 2016156028 임지섭 |
| **최종수정일** | 21.05.22 |

목차

**장 제목 입력(수준 1)1**

장 제목 입력(수준 2)2

장 제목 입력(수준 3)3

**장 제목 입력(수준 1)4**

장 제목 입력(수준 2)5

장 제목 입력(수준 3)6

1장

**2장 전처리**

음성 데이터는 굿모닝, 내일 날씨 알려줘, 출근길 교통 상황 어때 총 3개의 class를 사용하였고, 인당 250개 씩 녹음을 하였다. 녹음을 한 결과 아이폰과 안드로이드 폰의 녹음 품질이나 설정 등에 차이 때문에 아날로그 신호에서 디지털 신호로 변환된 값에 차이가 발생하는 것을 알 수 있었다. 그러므로 안드로이드로 녹음한 데이터를 기준으로 변환을 실시하였다.

특이사항: MediaRecoder의 samplingRate 8000Hz

사용될 주요 python 라이브러리

librosa == 0.8.0

matplotlib == 3.3.3

**1. MFCC**

1) 변환

(1) wav2melspectrum

[code]

librosa.feature.melspectrogram(audio, sr = sampling\_rate, n\_mels = 128)

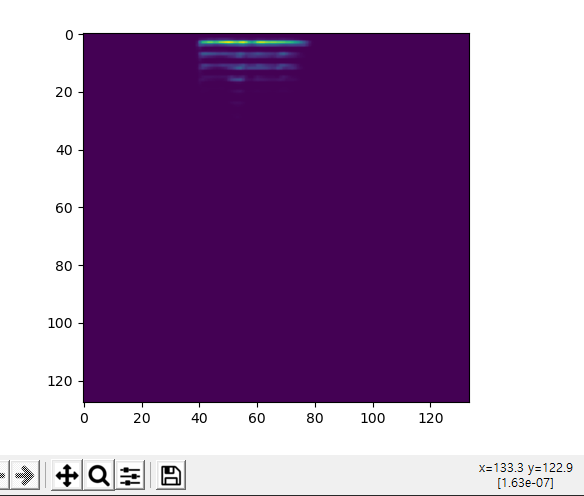


그림 [shape: 128x134]

(2) digital2DB

[code]

log\_S = librosa.power\_to\_db(S, ref=np.max)

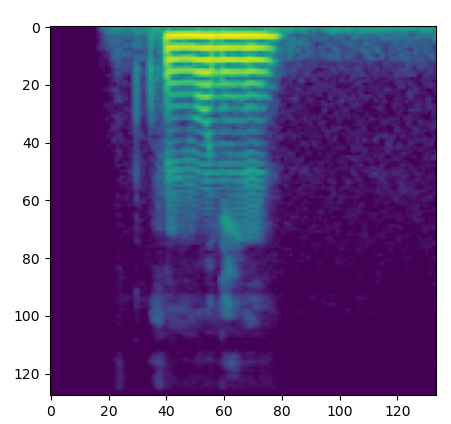


그림 [shape 전과 동일]

(3) MFCC

[code]

mfcc=librosa.feature.mfcc(S = log\_S, n\_mfcc=20)

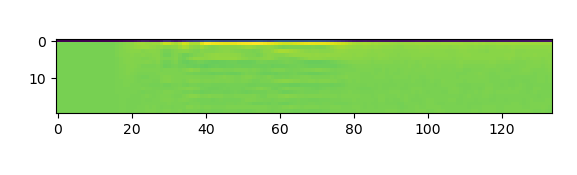


그림 3[shape: 20x128]

(4) delta화

[code]

delta2\_mfcc = librosa.feature.delta(mfcc, order=2)

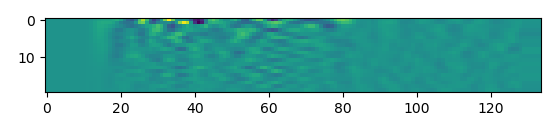


그림 4[shape: 20x128]

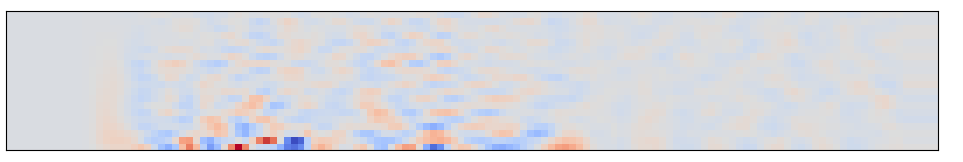


그림 5[specshow]

설명: Speech signal이 time-variant 한데, 우리가 frame으로 잘라서 데이터를 보고 있으므로 정보를 충분히 반영하지 못할 수 있다. 이러한 점을 보완하기 위해 각 frame 간의 차이를 구해 인풋 데이터의 time-variant한 특성을 반영해 feature로 사용하는 게 Deltas의 기본 아이디어

단점: 본래 signal보다 백색 소음이 증폭되는 문제가 있음