

드론으로 취득된 음성정보에서 구조요청 소리를 듣고 구분하기

연구 진행 방식
경희대학교 캡스톤 디자인 1

박준혁
이민혜
이수인

목차

- 데이터 형태
- 데이터 전처리 방법
- 클래스 구분 방법
- 발원방향 예측 방법
- 예상되는 문제점

데이터 형태

- 스테레오 타입의 데이터 50개 (wav 타입)

스테레오 타입 ?

다른 방향에 있는 2대의 마이크를 통해 녹음한 음향을
2대의 스피커로 재생하는 방식



데이터 전처리 방법

- 관련 논문(1)에 따르면 logarithmic scaling of the magnitude preprocessing 방법이 music tagging task를 수행하는 것에 있어 상당한 개선 효과가 있다고 한다.
- 위 논문 저자는 logarithmic scaling of the magnitude preprocessing 방법이 music tagging task 뿐만 아니라 ‘환경적 소리 기술’에도 충분히 적용될 수 있다고 말한다.
- Melspectrograms를 적용시키고, log 함수를 적용시킨다

멜 스펙트럼

멜 스펙트럼은 주파수의 단위를 다음 공식을 따라 멜 단위(mel unit)로 바꾼 스펙트럼을 말한다.

$$m = 2595 \log_{10} \left(1 + \frac{f}{700} \right)$$

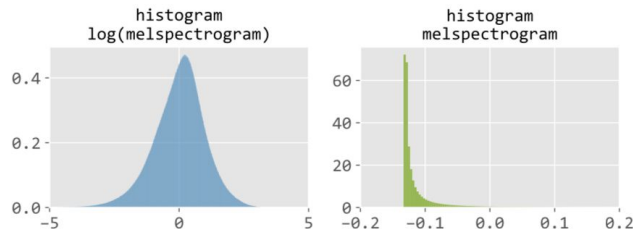


Fig. 6: histograms of the magnitude of melspectrogram time-frequency bins with (left) and without (right) logarithmic compression. The number of bins are 100 and both are normalised, i.e., $\sum_{i=1}^{100} 0.01 \times y_i = 1$. Log compression significantly affects the histogram, making the distribution Gaussian (left), otherwise extremely skewed (right). This is after standardisation and based on randomly selected 100 tracks from the training set.

(1) A COMPARISON OF AUDIO SIGNAL PREPROCESSING METHODS FOR DEEP NEURAL NETWORKS ON MUSIC TAGGING (2018, Queen Mary University of London, London, UK)
: deep neural networks를 이용한 music tagging 작업에서 audio preprocessing의 영향을 연구한 논문

클래스 구분에 사용할 모델 설계

취득된 구조요청 음성정보에서 남녀노소 클래스 구분하기

(= Audio Classification)

- CNN : 부분 부분의 특징을 추출하여서 이미지를 인식하고 구별함
- => 추후 데이터셋을 확인해보고 상황에 맞춰 수정 예정

발원방향 예측에 사용할 모델 설계

(= Sound Localization Problem)

- 데이터 셋이 스테레오 타입으로 주어지므로 발원방향의 추정이 가능
- 좌우 채널로 음원의 차이가 존재
- CNN + delay-line localization 방법 사용

delay-line localization : 양 채널로 동일한 사운드가 얼마만큼의 간격을 두고 입력되는지에 기반해 소리 발원의 방향을 추정하는 기법

양 채널의 **delay-line**과 그에 따른 방향을 학습시키고 추정하는 방식으로 진행

예상되는 문제점

1. 부족한 데이터셋

- 총 50-60개의 데이터셋이 존재하기 때문에 학습시키기에는 부족한 갯수라고 판단됨
- 이를 해결하기 위해 데이터 **oversampling**을 통해 데이터 갯수를 전체적으로 늘려 학습할 계획

2. overfitting

- 부족한 데이터셋과 이를 해결하기 위한 **oversampling** 때문에 **overfitting** 현상이 발생할 것으로 판단
- 적절하게 **train**과 **test**셋을 나누어 **overfitting**이 발생한지 확인하고 이에 따라 **dropout** 기법 등을 사용할 예정