

Survey

🕒 작성일시	@2022년 7월 19일 오전 11:41
📁 유형	
📌 상태	진행 중
📌 에픽	
📌 우선순위	
📌 작업	
📅 타임라인	
👤 엔지니어	
☰ 속성	

대표적인 데이터셋

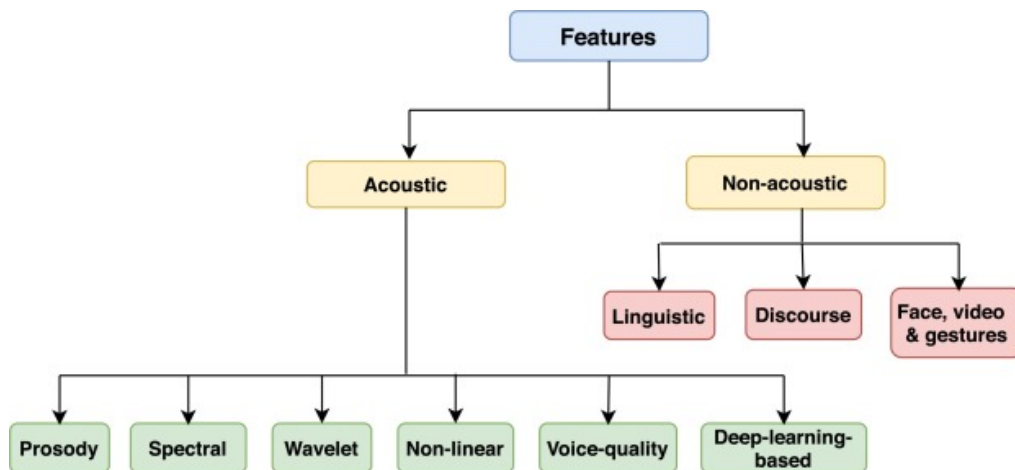
Database	Type	# Recordings	# Speakers	Sampling rate (KHz)	Emotions
IEMOCAP [32]	Acted (Multi-modal)	10309	10	16	Anger, Happiness, Boredom, Fear, Neutral, Surprise, Excitation Frustration, and others.
AFEW [29]	Acted (Multi-modal)	1426	330	44.1	Anger, Disgust, Fear, Sadness, Happiness, Neutral, and Surprise.
RECOLA [31]	Natural (Multi-modal)	7 hours	35	44.1	Arousal degree (1-5), Valence degree (1-5).
CHEAVD [30]	Natural (Multi-modal)	2600	238	44.1	26 Non-prototypical emotions + 6 basic emotions and Sadness.
FAU-Aibo [39]	Natural	18216	51	16	Anger, Emphatic, Neutral, Joy, and Rest.
SUSAS [40]	Natural	16000	7	8	Low stress, Middle stress, High stress, and Neutral.

01. Emotion Modeling

- 분류 및 회귀 문제로 모델링
- 감정은 연속적인 활동
 - 감정의 수준은 연속적인 변수
- 7가지 기본 이산 감정
 - 분노, 행복, 중립, 슬픔, 좌절, 혐오, 지루함

02. Features

- 음성 신호에서 특징 추출
- 특징 추출 전에 전처리 필요
 - 일반적으로 감정적 발화는 발화 수준에서 레이블링 되지만 레이블 된 감정은 전체 발화에 존재하지 않음.
 - 따라서, 고유한 두드러진 세그먼트를 찾기 위한 전처리 필요
 - 제거되는 음성 신호에 배경 잡음 존재
- 특징은 음향 및 비음향으로 분류됨.



02-01. Acoustic Features

- low-level descriptors (LLD) features : 20-30ms 크기의 프레임으로 분할한 다음 특징 추출
- utterance length(발화 길이)는 데이터셋마다 상이함.
 - 각 데이터의 평균, 중앙값, 최대값, 최소값, 왜도 등을 이용하여 고정시킴.
- acoustic features
 - prosodic
 - 프레임 지속 시간이 일반적으로 30-100ms로 길다.

- 낮은 각성 감정과 높은 각성 감정(예: 행복과 슬픔)을 잘 구별하지만, 슬픔과 두려움 같은 유사 감정은 잘 구별하지 못함.

- spectral

- 프레임 지속 시간이 10-30ms
- 음성 신호의 서로 다른 주파수 대역의 에너지 추출
- cepstral feature : MFCC, LPCC, PLP 계수
 - 역 스펙트럼 변환에 의해 계산
 - 로그 스펙트럼의 변화를 포착하고 스펙트럼의 변화 포착

- Fourier


- **Speech emotion recognition using Fourier parameters (2015)**
- 추출된 특징 평가
 - vector-quantization classifier
 - hidden Markov classifier
- CASIA, EESDB dataset에서 MFCC와 푸리에를 비교하면 정확도가 향상됨.
- 두 가지를 결합한 경우에도 향상

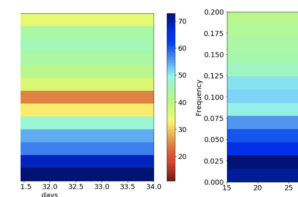
- Gabor-spectrogram

- **Spectral features based on local hu moments of Gabor spectrograms for speech emotion recognition (2016)**
- 1. 스펙트럼의 로그 에너지 계산
- 2. 에너지 스펙트럼을 Gabor 웨이블릿으로 컨볼루션하여 Gabor-spectrogram 계산
- 3. Gabor local Hu 계산
 - a. 블록 Hu 전략과 특징을 역상관시키기 위한 discrete cosine transform의 후속 적용을 위해 계산
- 4. PCA를 사용하여 중복 제거

Time series analysis: Obtaining the spectrogram using the Gabor transform

The practice of time series analysis deals with fluctuations which vary in time. The ubiquitous by now Fourier Transform, has provided us with a convenient tool to model such functions. Several variations of the Fourier Transform algorithm are

 <https://tng-daryl.medium.com/time-series-analysis-obtaining-the-spectrogram-using-the-gabor-transform-ff62c7a9ecb4>



- wavelet-based
- voice-quality-based
- nonlinear
- deep-learning-based