Audio Data

Kaggle study
2019. 8. 4

Applications

Music Generation

OpenAI **MuseNet** https://openai.com/blog/musenet/

Speech Synthesis

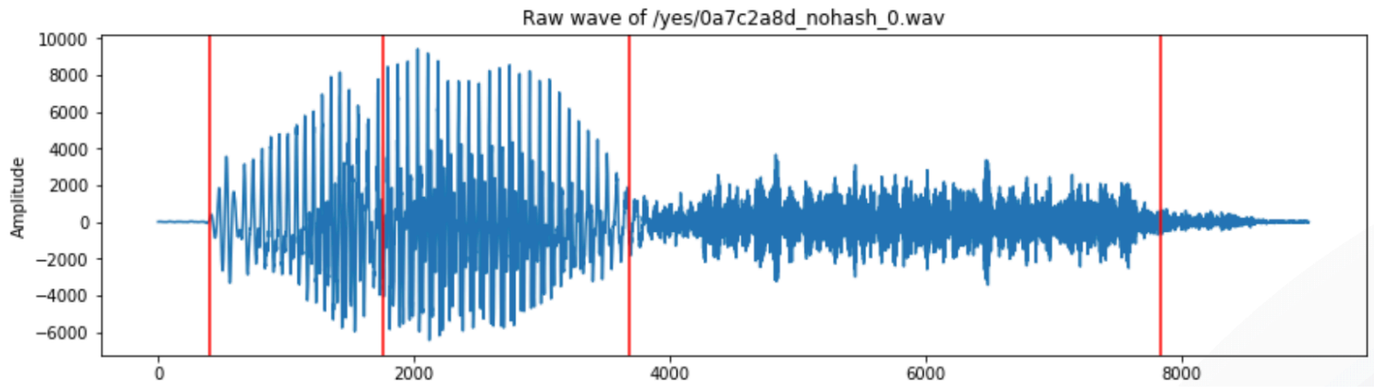
Google Tacotron https://google.github.io/tacotron/publications/tacotron2/index.html

Synthesizer

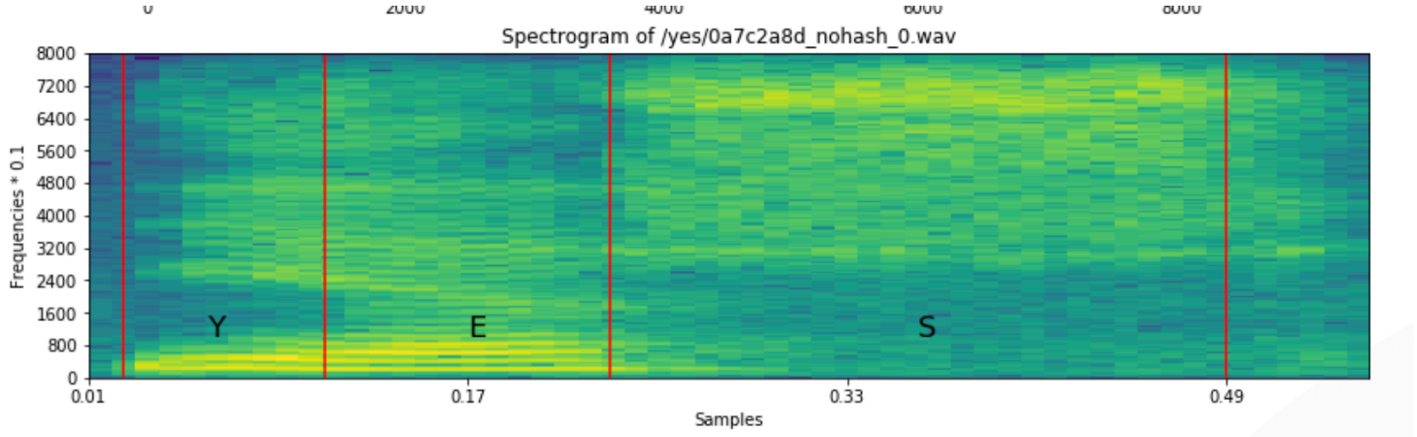
Synthesizer V https://www.youtube.com/watch?v=dddiw3eqwlU

Data Representation

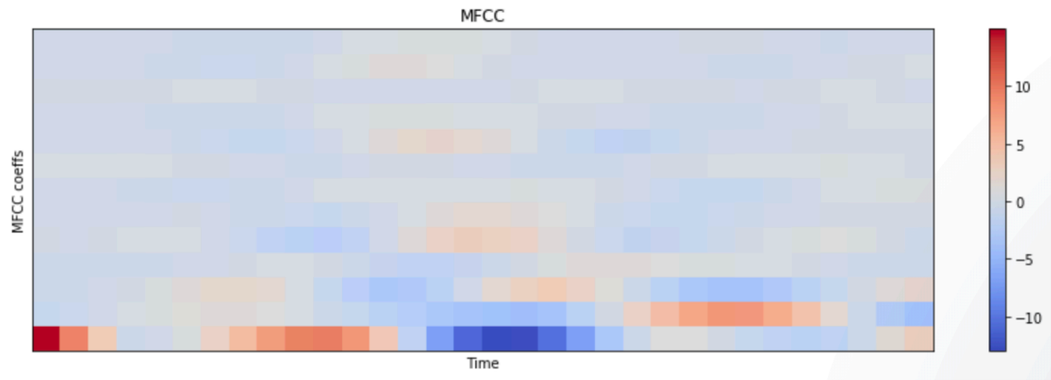
Raw data



Spectrogram



MFCC



Libraries

from scipy.io import wavfile

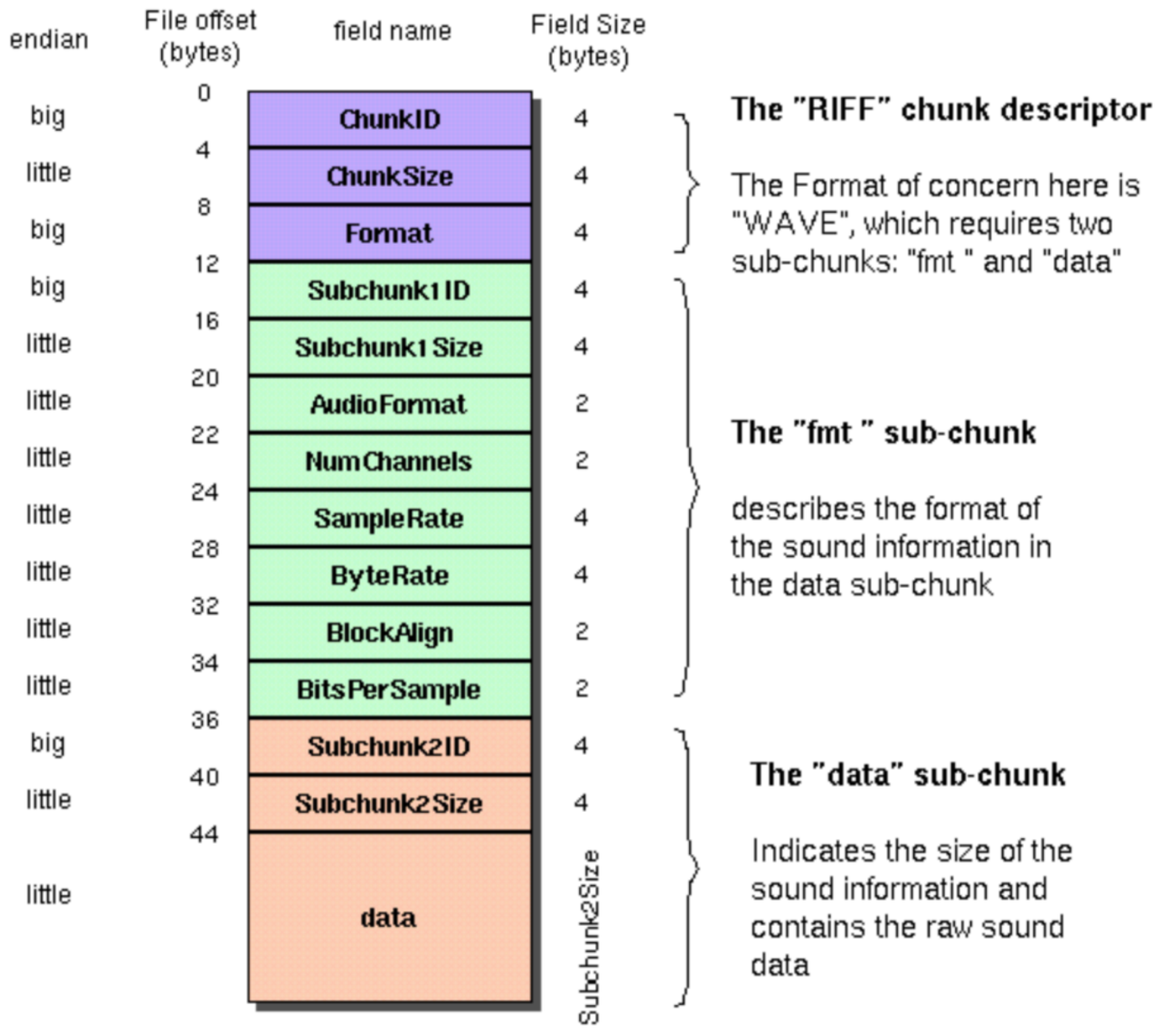
import librosa

Import Ipython.display

from scipy.fftpack import fat

Wav File format

The Canonical WAVE file format



Visualization and Data Preprocessing

https://www.kaggle.com/davids1992/speech-representation-and-data-exploration/notebook