Frequency-domain audio features

Valerio Velardo

Join the community!

thesoundofai.slack.com

Previously...

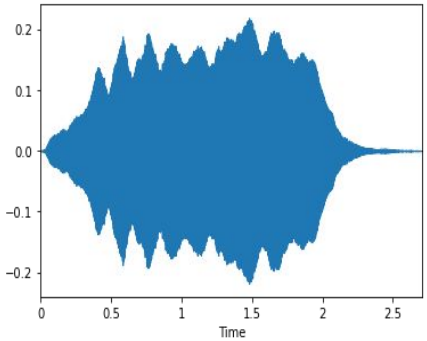
- Mel-Frequency Cepstral Coefficients

Frequency-domain features

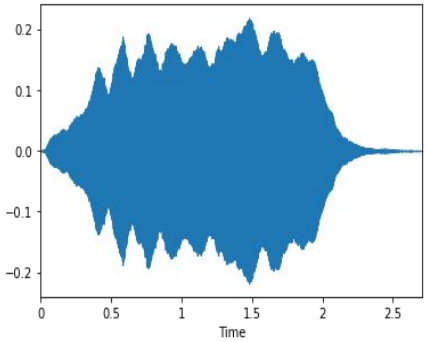
- Band energy ratio (BER)
- Spectral centroid (SC)
- Bandwidth (BW)
- ...

Extracting frequency-domain features

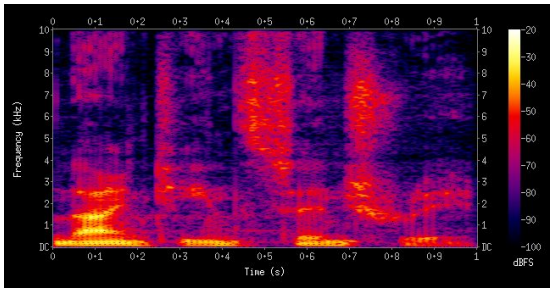
Extracting frequency-domain features



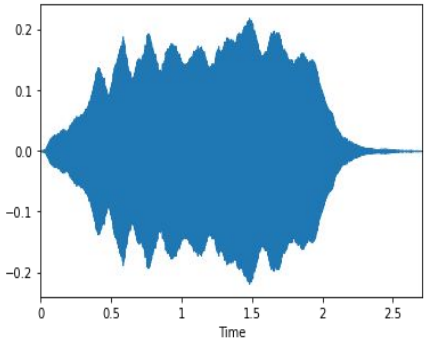
Extracting frequency-domain features



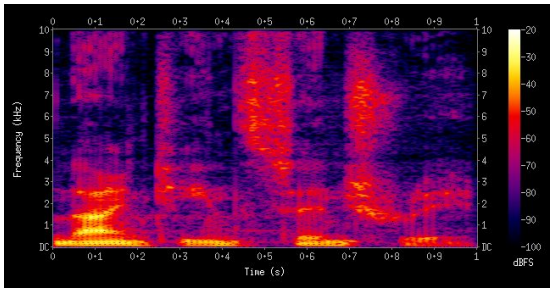
STFT



Extracting frequency-domain features



STFT



FEATURE COMPUTATION

Math conventions

- $m_t(n)$ -> Magnitude of signal at frequency bin *n* and frame *t*

Math conventions

- $m_t(n)$ -> Magnitude of signal at frequency bin *n* and frame *t*
- *N* -> # frequency bins

Band energy ratio

- Comparison of energy in the lower/higher frequency bands
- Measure of how dominant low frequencies are

Band energy ratio

$$BER_t = \frac{\sum_{n=1}^{F-1} m_t(n)^2}{\sum_{n=F}^{N} m_t(n)^2}$$

Band energy ratio

Power at *t, n*

$$BER_t = \frac{\sum_{n=1}^{F-1} m_t(n)^2}{\sum_{n=F}^{N} m_t(n)^2}$$

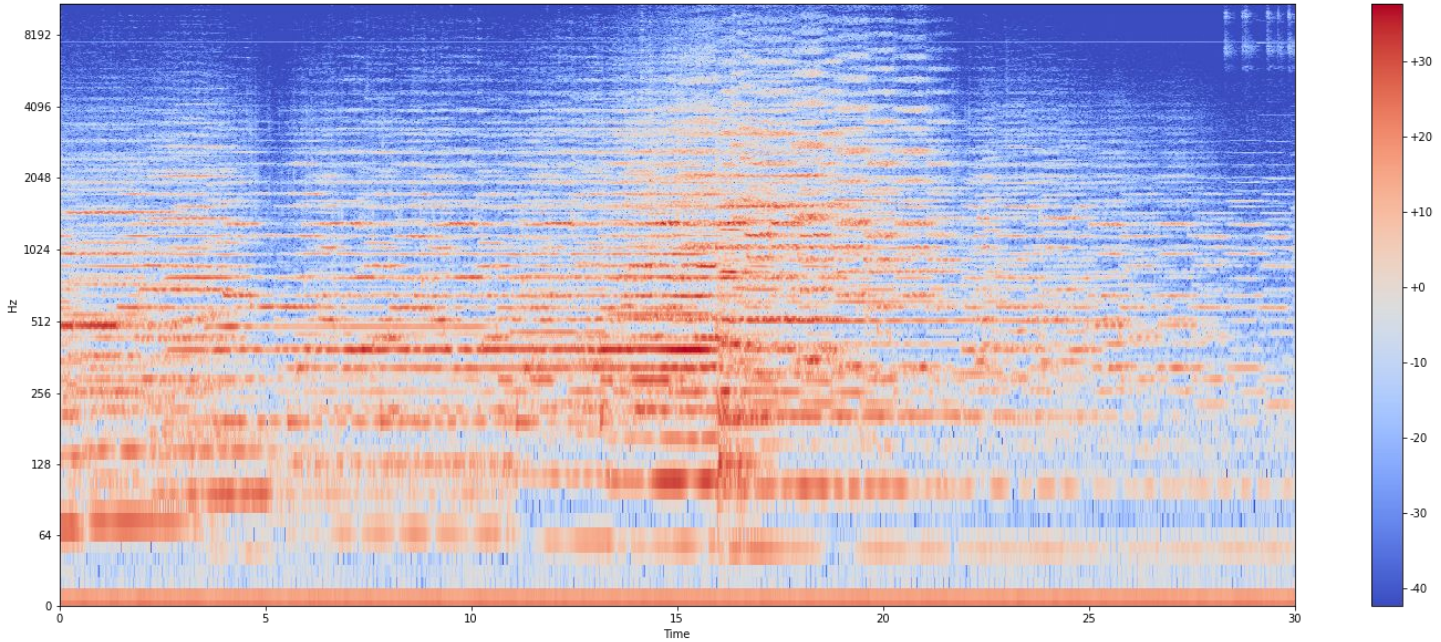
Band energy ratio

Power at *t, n*

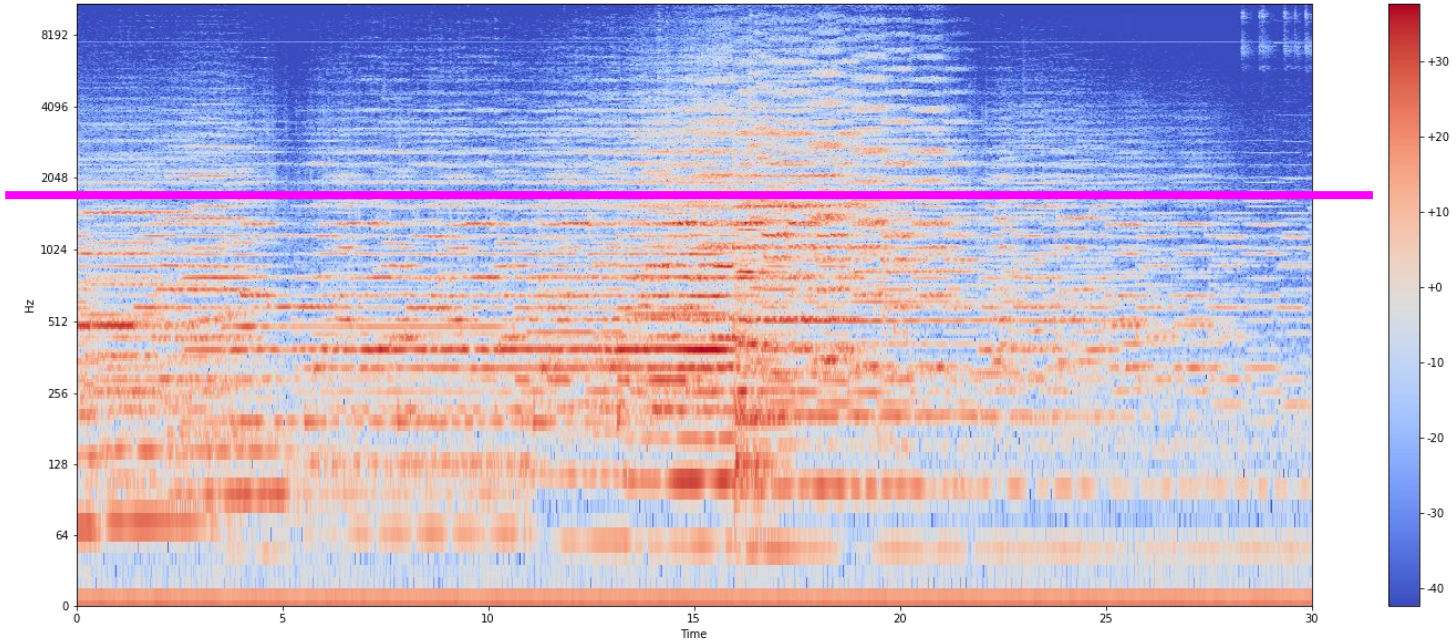
Split frequency

$$BER_t = \frac{\sum_{n=1}^{F-1} m_t(n)^2}{\sum_{n=F}^{N} m_t(n)^2}$$

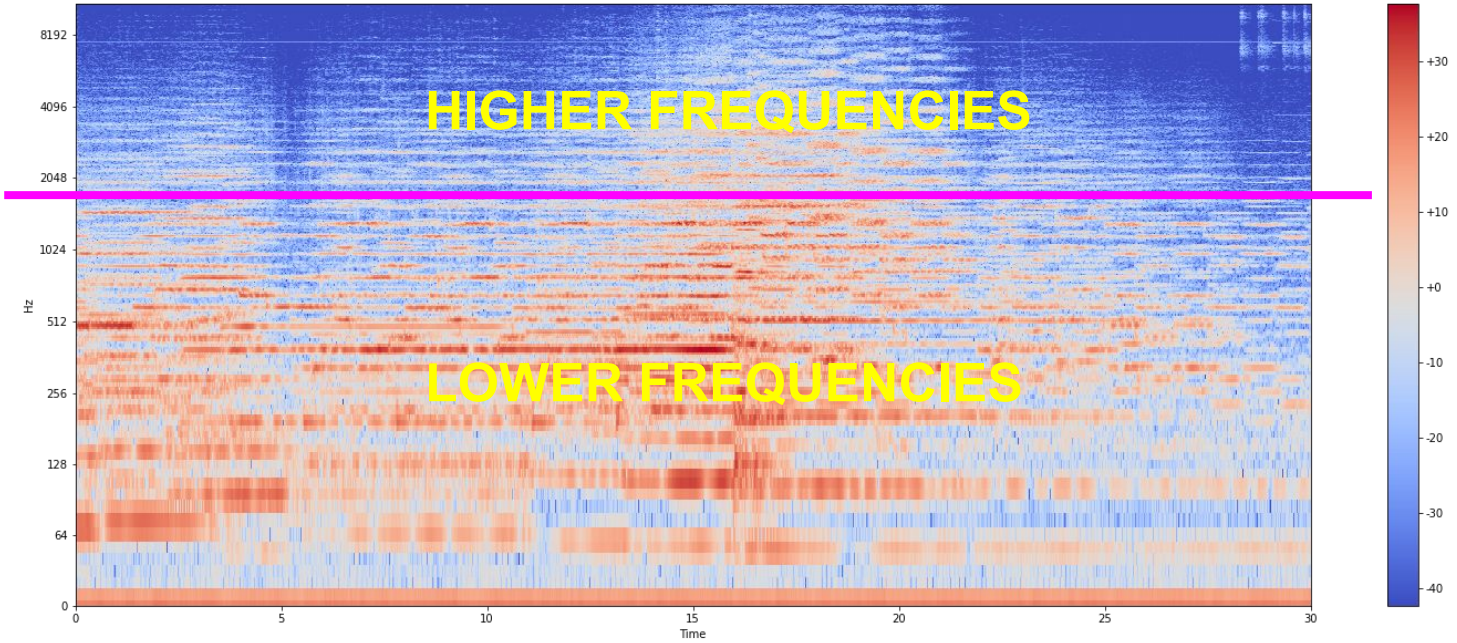
Band energy ratio



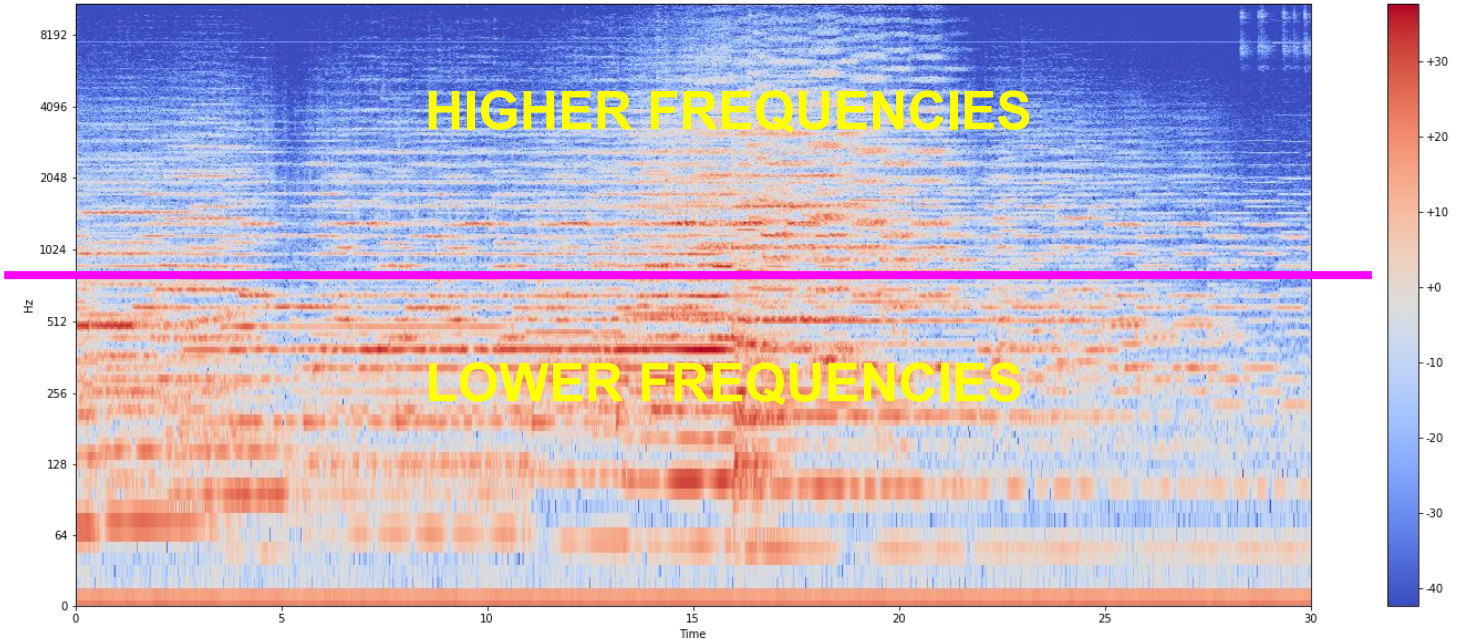
Band energy ratio



Band energy ratio



Band energy ratio



Band energy ratio

$$BER_t = \frac{\sum_{n=1}^{F-1} m_t(n)^2}{\sum_{n=F}^{N} m_t(n)^2}$$

Band energy ratio

Power in the lower frequency bands

$$BER_t = \frac{\sum_{n=1}^{F-1} m_t(n)^2}{\sum_{n=F}^{N} m_t(n)^2}$$

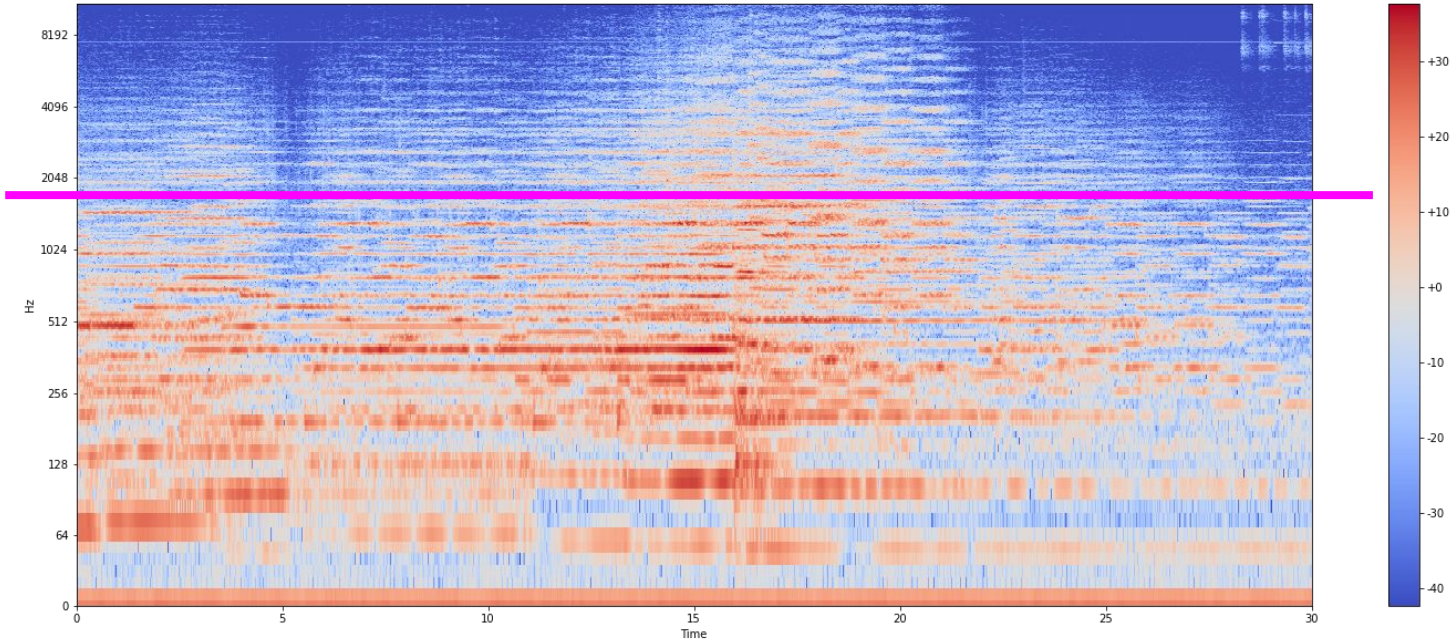
Band energy ratio

Power in the lower frequency bands

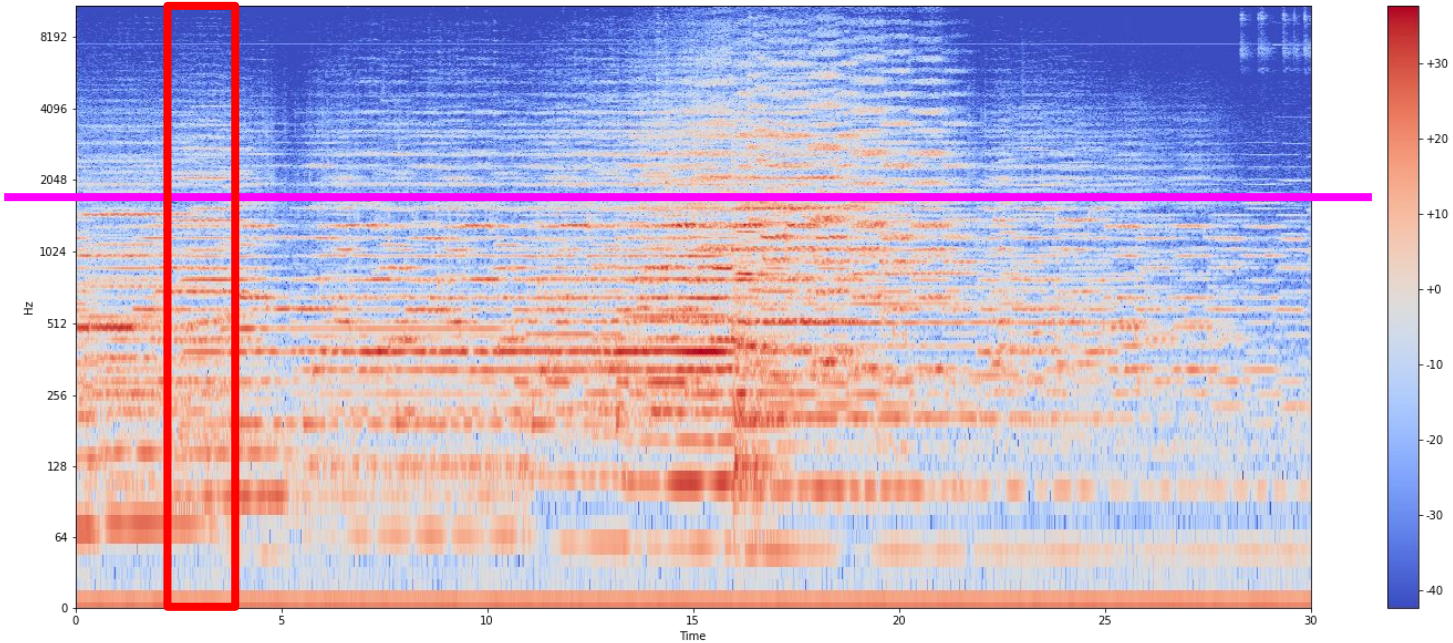
$$BER_t = \frac{\sum_{n=1}^{F-1} m_t(n)^2}{\sum_{n=F}^{N} m_t(n)^2}$$

Power in the higher frequency bands

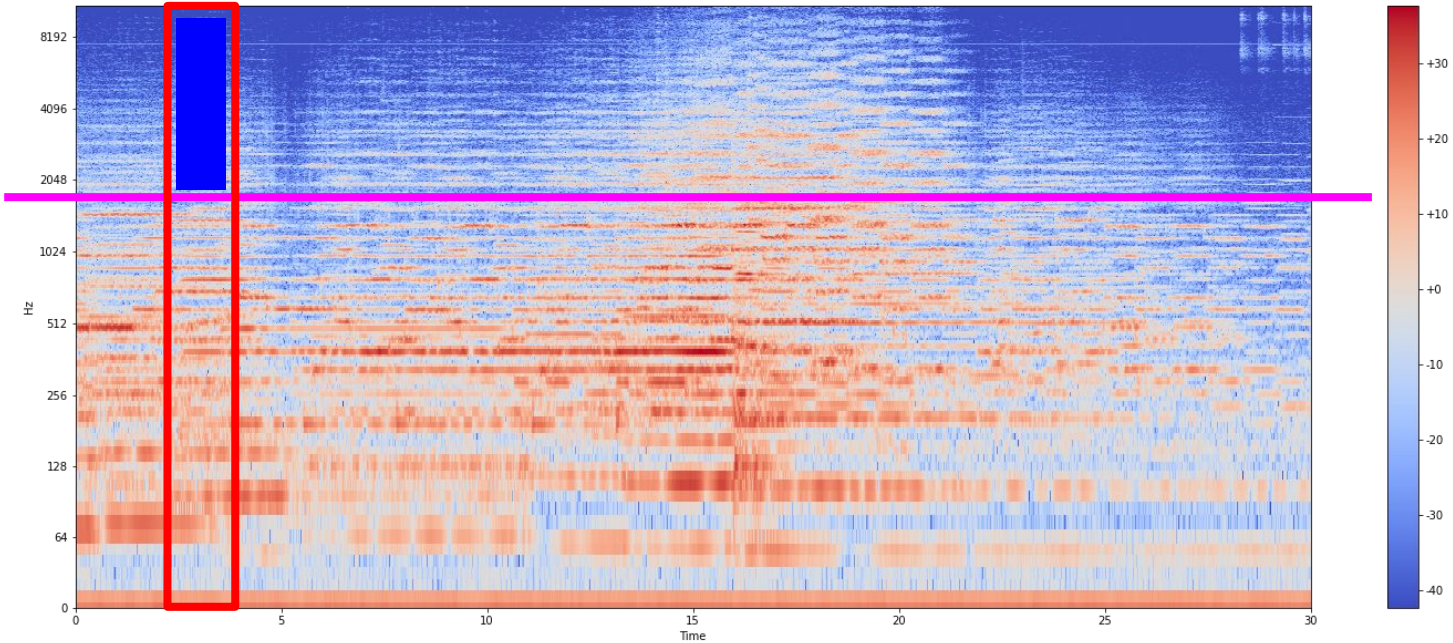
Band energy ratio



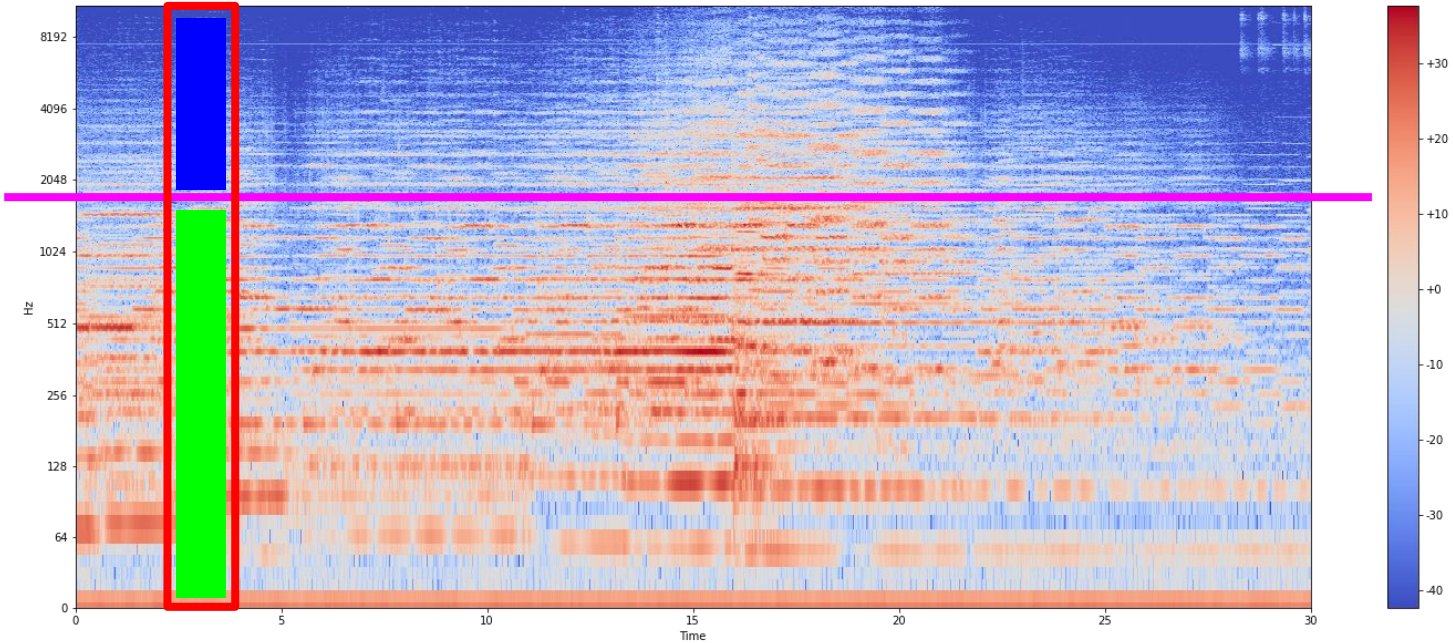
Band energy ratio



Band energy ratio



Band energy ratio



Band energy ratio applications

- Music / speech discrimination
- Music classification (e.g., music genre classification)

Spectral centroid

- Centre of gravity of magnitude spectrum
- Frequency band where most of the energy is concentrated
- Measure of “brightness” of sound

Spectral centroid

- Weighted mean of the frequencies

Spectral centroid

- Weighted mean of the frequencies

$$SC_t = \frac{\sum_{n=1}^{N} m_t(n) \cdot n}{\sum_{n=1}^{N} m_t(n)}$$

Spectral centroid

- Weighted mean of the frequencies

$$SC_t = \frac{\sum_{n=1}^{N} m_t(n) \cdot n}{\sum_{n=1}^{N} m_t(n)}$$

Frequency bin

Spectral centroid

- Weighted mean of the frequencies

Weight for *n*

$$SC_t = \frac{\sum_{n=1}^{N} m_t(n) \cdot n}{\sum_{n=1}^{N} m_t(n)}$$

Spectral centroid

- Weighted mean of the frequencies

$$SC_t = \frac{\sum_{n=1}^{N} m_t(n) \cdot n}{\sum_{n=1}^{N} m_t(n)}$$

Sum of weights

Spectral centroid applications

- Audio classification
- Music classification

Bandwidth

- Derived from spectral centroid
- Spectral range around the centroid
- Variance from the spectral centroid
- Describe perceived timbre

Bandwidth

- Weighted mean of the distances of frequency bands from SC

Bandwidth

- Weighted mean of the distances of frequency bands from SC

$$BW_t = \frac{\sum_{n=1}^{N} |n - SC_t| \cdot m_t(n)}{\sum_{n=1}^{N} m_t(n)}$$

Bandwidth

- Weighted mean of the distances of frequency bands from SC

$$BW_t = \frac{\sum_{n=1}^{N} |n - SC_t| \cdot m_t(n)}{\sum_{n=1}^{N} m_t(n)}$$

Weight for *n*

Bandwidth

- Weighted mean of the distances of frequency bands from SC

Distance of frequency band from spectral centroid

Weight for *n*

$$BW_t = \frac{\sum_{n=1}^{N} |n - SC_t| \cdot m_t(n)}{\sum_{n=1}^{N} m_t(n)}$$

Bandwidth

- Weighted mean of the distances of frequency bands from SC

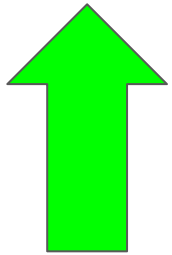
Distance of frequency band from spectral centroid

Weight for *n*

$$BW_t = \frac{\sum_{n=1}^{N} |n - SC_t| \cdot m_t(n)}{\sum_{n=1}^{N} m_t(n)}$$

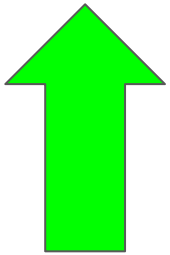
Sum of weights

Bandwidth

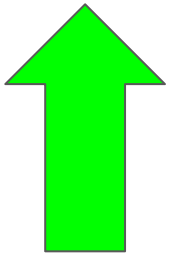


Energy spread across frequency bands

Bandwidth

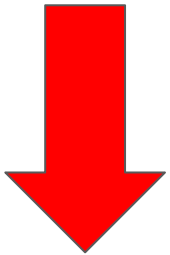


Energy spread across frequency bands



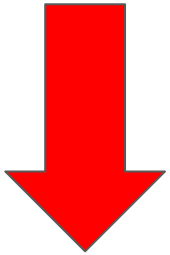
BW_t

Bandwidth

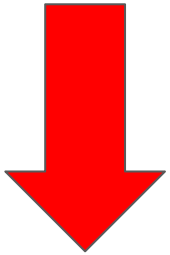


Energy spread across frequency bands

Bandwidth



Energy spread across frequency bands



BW_t


WHAT'S THE DIFFERENCE BETWEEN BANDWIDTH AND SPECTRAL SPREAD?
THEY'RE THE SAME
imgflip.com

Bandwidth applications

- Music processing (e.g., music genre classification)

What’s up next?

- Implement band energy ratio in Python (almost!) from scratch
- Visualise BER for music in different genres