



Dados não estruturados

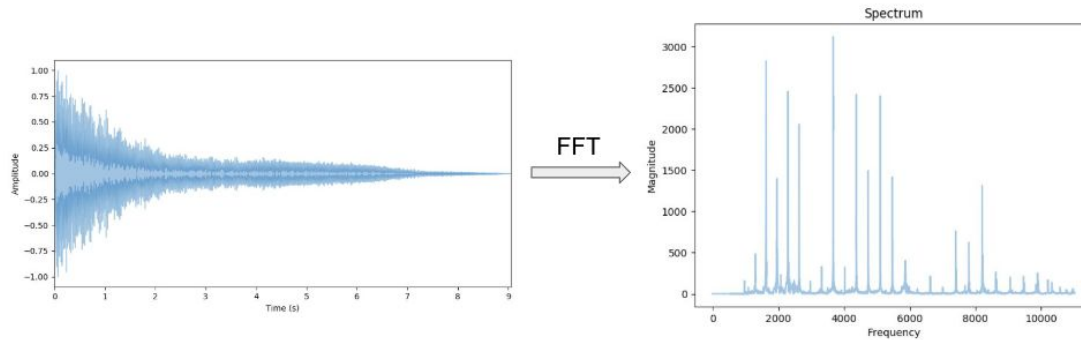
PDS - Aula 11



Objetivos

- Espectrograma
- Extração de features do áudio

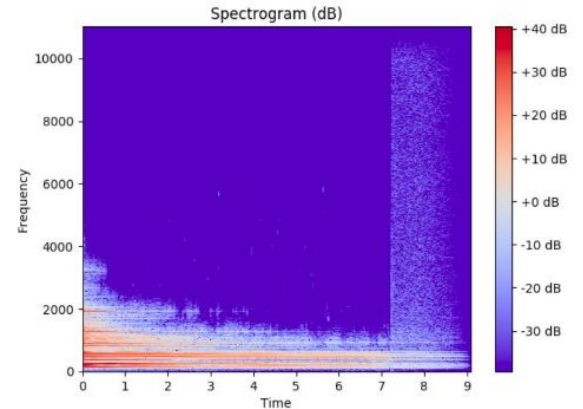
Transformada de Fourier



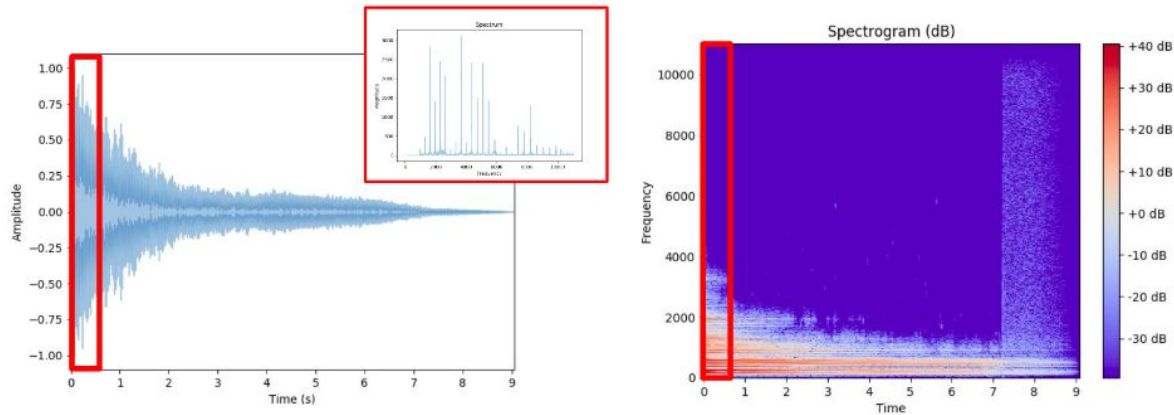
- Convertemos do domínio do tempo para o domínio de frequência
- Mas não temos mais informação de tempo

STFT (Short Time Fourier Transform)

- Computa FFT para diferentes intervalos
- Preserva informação de tempo
- Espectrograma (relação de tempo + frequência + magnitude)

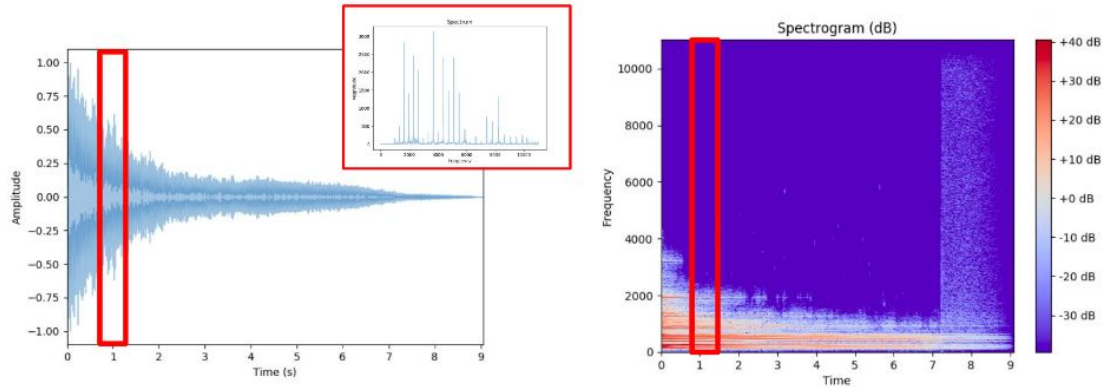


STFT (Short Time Fourier Transform)

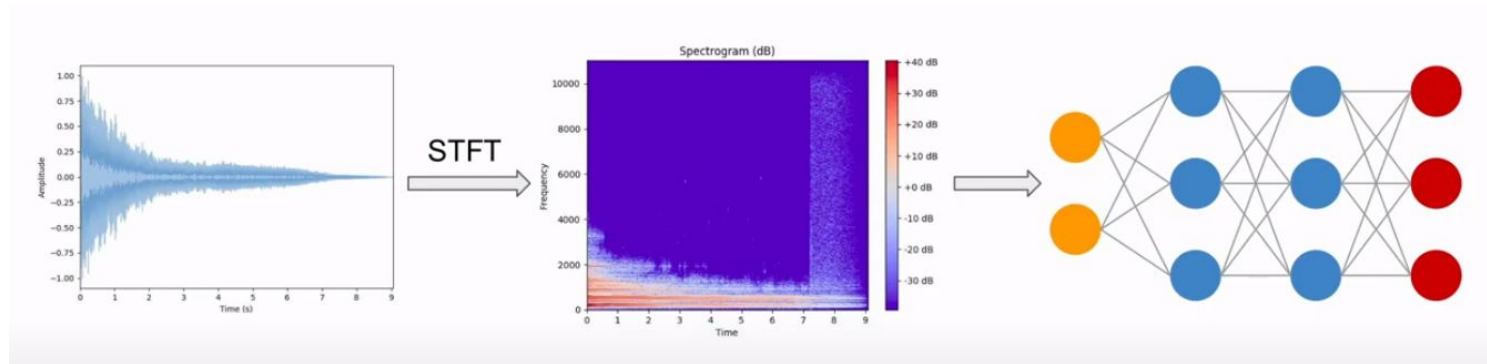


Spectrograms: an Introduction: <https://www.youtube.com/watch?v= FatxGN3vAM>

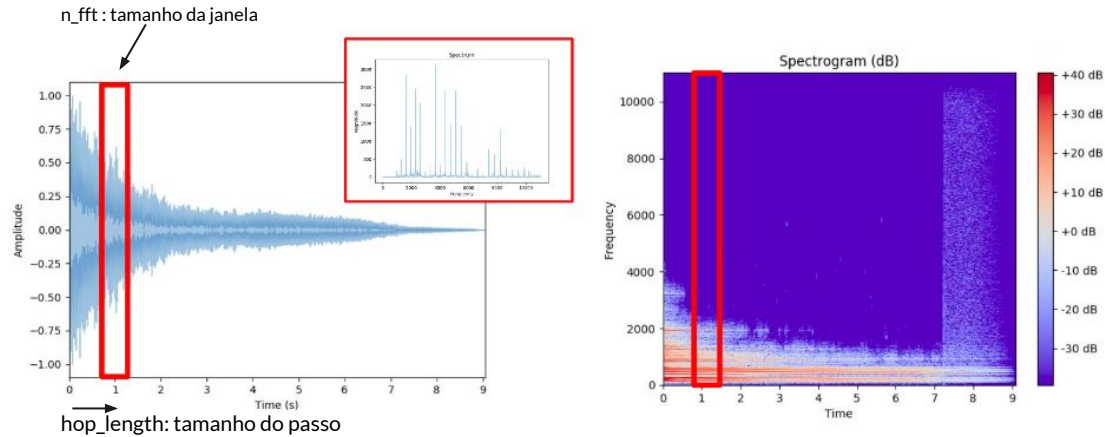
STFT (Short Time Fourier Transform)



Pré processamento de áudio



STFT (Short Time Fourier Transform)





Mão na massa

1) Sintetizar os seguintes áudios e tocar com o Áudio:

- a) C2 - 65 Hz
- b) C4 - 262 Hz
- c) G6 - 1568 Hz
- d) A6 - 1760 Hz

Usando:

Amplitude = 10

Frequência de amostragem = 44kHz

Fase = 0

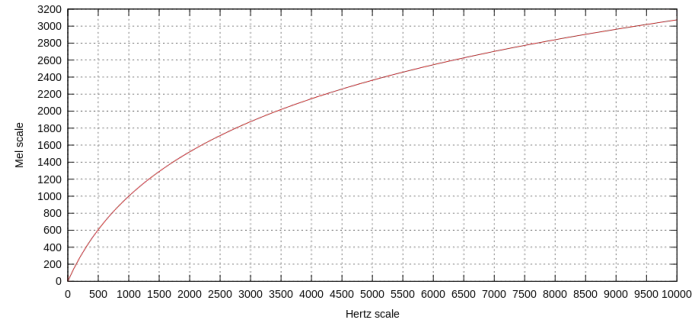


Humanos percebem numa frequência log

Escala de Mel

- Escala de Log
- 1000 Hz = 1000 Mel

$$m = 2595 \cdot \log\left(1 + \frac{f}{500}\right)$$





Aplicações

- Reconhecimento de voz
- Classificação de gênero de música
- Classificação de instrumento

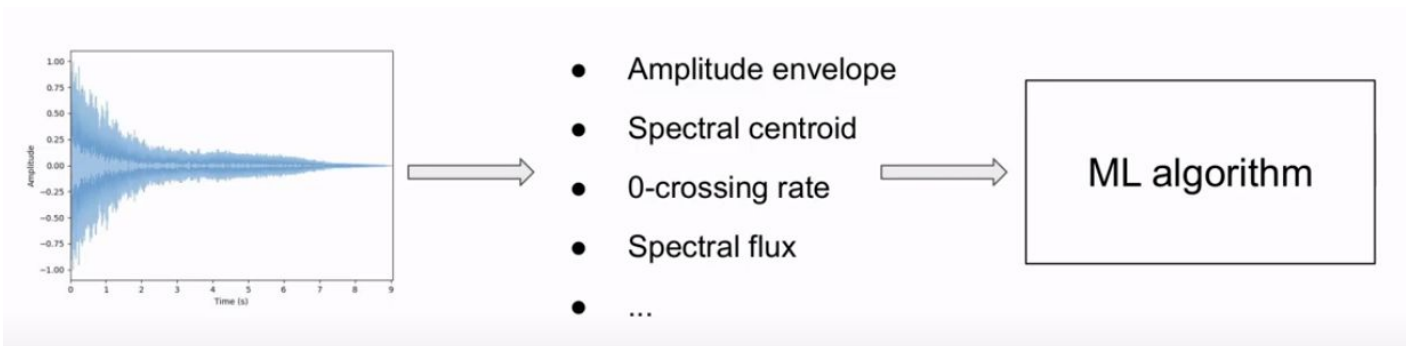


Espectrograma de Mel

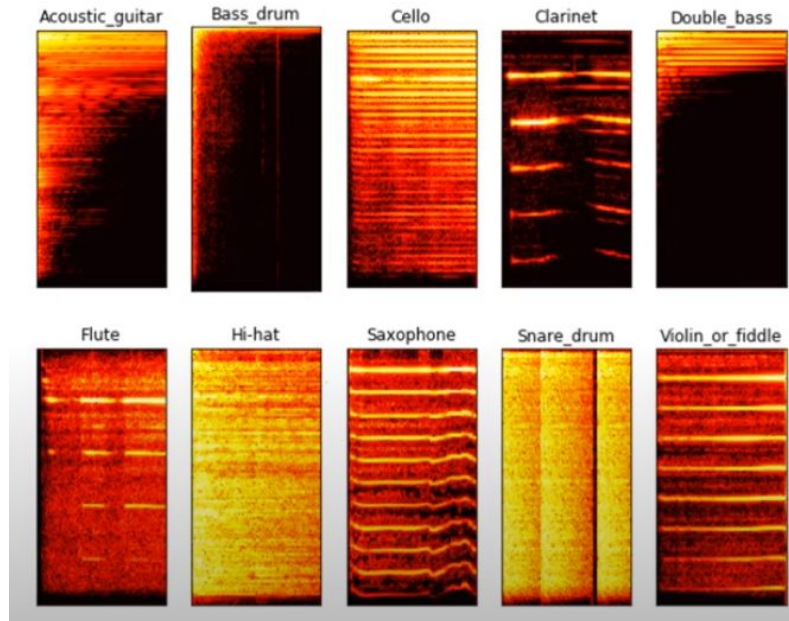
- 1) Calcular STFT
- 2) Converter amplitude para db
- 3) Converter escala de frequência para escala de mel

Formas tradicionais para extração de feature

- Domínio do tempo (0-crossing rate, amplitude envelop, etc)
- Domínio frequência (Spectral Centroid, Spectral Rolloff, Spectral Bandwidth)



Espectrograma para instrumentos musicais





Mão na massa

- 1) Com o arquivo de áudio fornecido, siga os seguintes passos:
 - a) Importar librosa
 - b) Carregar áudio (`librosa.load`) limitando para somente para os primeiros 5 segundos com método e `sample_rate = 44100` hertz
 - c) Plotar no domínio do tempo (`librosa.display.waveplot`)
 - d) Aplicar stft com a função `librosa.stft` com passo (`hop_length`) de 512 frames e *tamanho da janela* de 2084 frames
 - e) Calcular a magnitude da saída da operação do exercício anterior usando `np.abs`
 - f) Plotar o espectrograma x=tempo, y=frequência, cor = amplitude, com o método `librosa.display.specshow`
 - g) Plotar o espectrograma x=tempo, y=frequência, cor = db (amplitude_to_db), com o método `librosa.display.specshow`



Mão na massa

*Desafio: Gerar os espectrogramas para cada um dos gêneros musicais no dataset fornecido, da maneira que acharem melhor (ou seja, vocês vão escolher, offset, duration, sample_rate, n_fft, hop_length, escala, etc), mas com o **objetivo distinguir a partir da imagem do espectrograma a qual gênero musical ele pertence**,*

Entrega: 1 plot de espectrograma por gênero (1 rock, 1 pop, 1 metal, 1 jazz e 1 classical)

- a) *Baixar o mini dataset <http://opihi.cs.uvic.ca/sound/mini-genres.tar.bz2> que contém uma amostra de diferentes gêneros de música*
- b) *Carregar os áudio com um limite de 10s por arquivo*
- c) *Fazer o gráfico de espectrograma para pelo menos um arquivo de áudio por gênero musical*

***bonus:** montar o script que "ande" pelas subpastas e faça o espectrograma de forma automática*



Extração de feature: Mão na massa

Com o sinal vamos aplicar as seguintes funções para extração de feature:

- Spectral Centroid (centro de massa): `librosa.feature.spectral_centroid(x, sr=sr)`
- Spectral Rolloff : `librosa.feature.spectral_rolloff(x, sr=sr)`
- Spectral Bandwidth: `librosa.feature.spectral_bandwidth`
- Zero Crossing: `librosa.zero_crossings`
- Chroma feature (energia por nota): `librosa.feature.chroma_stft`