

Giovanni Paolo Meloni (10295603)

Ciência da Computação

Disciplina: Processamento de Imagens

Prof. Moacir Ponti

SCC-0251

Comparação de Espectrogramas

Universidade de São Paulo

ICMC - São Carlos

28/06/201

Objetivo

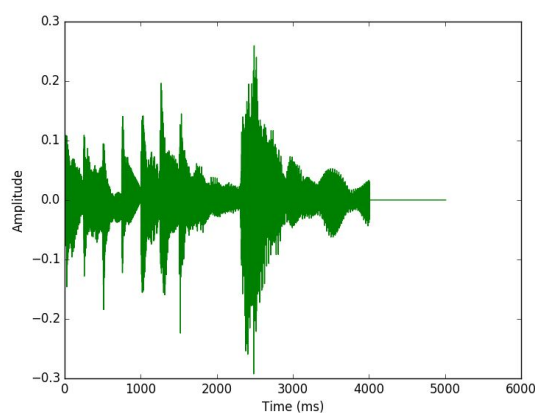
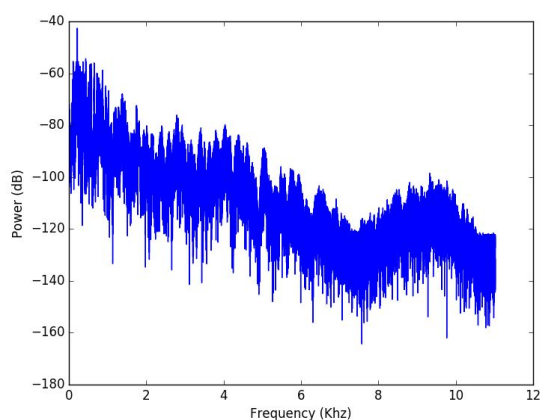
O trabalho tem como objetivo, a partir de um arquivo de áudio .wav, fazer seu espectrograma e comparar com outros para poder diferenciar o áudio entre um instrumento acústico, de corda ou elétrico.

Foram usados 4 arquivos de áudio (.wav) para que exista uma base de comparação. Os arquivos foram:

- [Classical Guitar](#)
- [Harpsichord](#) (Acoustic String)
- [Cello](#)
- [Electric Guitar, Single Coil Pickup](#)

Todos eles obtidos do **Centro para Pesquisa Computacional em Música e Acústica (CCRMA)** da Universidade de Stanford [\[1\]](#).

Classical Guitar:



Resolução

Para que o problema fosse resolvido primeiro foi necessário criar um vetor com os exemplos que seriam usados na comparação, para tanto existe um vetor que armazena os *arrays* de frequência desses áudios.

O usuário então deve entrar com o nome do arquivo de áudio (portanto esse deve estar dentro da pasta de execução do programa). O áudio será lido e uma amostragem de frequência e som serão retornados.

O som deverá ser convertido para valores reais (*float*) entre -1 e 1 e será necessário uma amostragem de pontos desse áudio (para termos uma quantidade discreta de pontos) além da duração desse som. Todas essas informações são necessárias para a criação das imagens como exemplificadas acima.

Com o vetor de comparação e o espectrograma do áudio de entrada em mãos, é preciso usar alguma lógica de comparação entre espectrogramas para que se tenha um resultado aceitável no fim. Para esse programa foi usado a distância Euclidiana entre esses dois *arrays* (espectrogramas). Aquele com menor valor na diferença Euclidiana era o que mais se assemelhava ao áudio de entrada.

Conclusão

Esse problema mostrou-se mais complicado do que o esperado. Para alguns áudios, seu espectrograma estava estourado o que leva à uma comparação errônea dos arquivos de áudio. Para arquivos que não estouraram, quando foi computado seu espectrograma o tamanho era muito distante de outros, o que impossibilita uma comparação. Ainda que o *array* fosse segmentado, não foi encontrado uma boa lógica para onde e nem quantas vezes segmentá-lo, portanto o resultado da comparação não era confiável.

Os resultados obtidos não foram bons nem perto do que se era esperado. De todos os áudios foi obtido sua amplitude para visualização mas nem todos obtiveram um espectrograma utilizável.