

INVESTIGAÇÃO DOS COEFICIENTES CEPSTRAIS DA FREQUÊNCIA MEL PARA EXTRAÇÃO DE CARACTERÍSTICAS DE GÊNEROS MUSICAIS

Rafael Alves Bonfim de Queiroz¹, João Fernando Marar², Cleber Morio Okida³
^{1,2,3} UNESP, Faculdade de Ciências, Departamento de Computação, Lab. SACI
bonfim@fc.unesp.br; feramarar@fc.unesp.br

1. Introdução

Os gêneros musicais são rótulos categóricos utilizados para classificar e caracterizar músicas digitais e para organizar grandes coleções disponíveis na Web [2]. Esses rótulos são relacionados com a instrumentalização utilizada, estrutura rítmica e conteúdo harmônico da música. Como o gênero musical é um conceito relativamente subjetivo, até a indústria musical é muitas vezes contraditória ao atribuir gêneros musicais [3]. A classificação automática de gêneros musicais poderá auxiliar ou substituir o usuário humano nesse processo, assim como prover um componente importante para um sistema de recuperação de informações para músicas.

Dentro deste cenário, investigamos os coeficientes cepstrais da frequência mel (MFCCs) para extração de características de músicas.

Os MFCCs são coeficientes característicos do som, obtidos de um tipo de representação cepstral. O termo cepstral origina de cepstrum, que é o resultado da Transformada de Fourier de um espectro decibel. A escala mel é uma escala perceptual de pitch, que é a frequência fundamental do som percebida. O nome mel originou-se da palavra melodia para indicar que a escala é baseada em comparação de pitch.

2. Fundamentos dos MFCC

Os MFCCs apresentam um bom desempenho em extração de características presente em sinais da fala humana [2]. Pois, conseguem representar os aspectos perceptualmente relevantes do espectro do som de curta duração. O modelo matemático para calcular os MFCCs [1] é definido por :

$$c(n) = \sum_{k=1}^K \log|S(k)| \cdot \cos\left[n\left(k - \frac{1}{2}\right)\frac{\pi}{K}\right] \quad (1)$$

para $0 \leq n \leq P$, onde $c(n)$ é o n -ésimo coeficiente mel cepstral, P é o número de coeficientes mel cepstrais extraídos, K é o número de filtros digitais utilizados e $S(k)$ é o sinal de saída do banco de filtros digitais. Seguem as etapas do algoritmo implementado para se calcular a Equação 1: **1) Pré-ênfase** : amplifica o espectro do sinal do som; **2) Janela de Hamming**: minimiza o efeito de borda provocado com a segmentação da amostra de dados, que ocorre devido às descontinuidades no início e no fim de cada bloco segmentado; **3) Transformada discreta de Fourier (DFT)**: determina as componentes de frequência

predominantes em um dado segmento de som, resultando os valores de suas amplitudes; **4) Banco de filtros na escala mel**: transforma os valores da amplitude da DFT para as saídas do banco de filtros na escala mel, obtemos nesta etapa o $S(k)$; **5) Obtenção dos coeficientes cepstrais**: calcula o logaritmo decimal do $S(k)$ e depois resolve a transformada cosseno definida na Equação 1.

3. Resultados

Na Figura 1 é apresentado os cinco coeficientes cepstrais de um trecho de 1,5 s de uma música do gênero Canto Gregoriano e outra do gênero Rock. Tais coeficientes foram obtidos com o algoritmo descrito na seção anterior. No eixo das ordenadas, a amplitude (valor) dos coeficientes é adimensional.

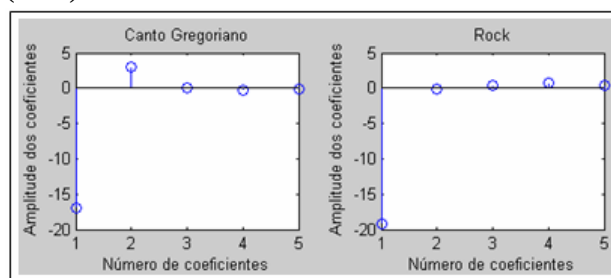


Figura 1 – Cálculo dos MFCC

4. Conclusão

Foi estudado e implementado o algoritmo para calcular os MFCC de trechos de músicas, pois se pretende em trabalhos futuros construir um classificador baseado em Redes Neurais que utilizará os MFCCs como padrões de treinamento e teste da Rede Neural. Na Figura 1, percebe-se que os cinco MFCCs do gênero Canto Gregoriano e Rock possuem amplitudes diferentes, assim podemos considerá-los como padrões distintos entre os dois gêneros investigado.

5. Referências

- [1] A. Petry et. al. , Reconhecimento Automático de Pessoas pela Voz Através de Técnicas de Processamento Digital de Sinais, XI SEMAC, 2000.
- [2] G. Tzanetakis and P. Cook, Musical genre classification of audio signals, *IEEE Transactions on Speech and Audio Processing*, 10(5):293–302, 2002
- [3] Aucouturier, J. J. and Pachet, F., Representing musical genre: A state of the art. *Journal of New Music Research*, 32 (1):83–93, 2003

¹ Aluno de IC do CNPq