



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE SOBRAL
ENGENHARIA DA COMPUTAÇÃO E ENGENHARIA ELÉTRICA

Processamento Digital de Sinais (SBL0085)

Prof.: C. Alexandre Rolim Fernandes

Trabalho AP4 – MFCC de Sinais de Voz

- Trabalho Individual
- Apenas simulações, sem trabalho escrito
- Os códigos devem estar bem organizados e comentados, para que seja possível entendê-los e corrigi-los. Códigos que estejam desorganizados ou sem os devidos comentários explicativos terão penalização na nota.
- Fazer todas as questões em **um só arquivo**.
- O seu código deve **gerar automaticamente todos os gráficos e resultados** solicitados.
- Enviar no SIGAA o **código cujo nome do arquivo deve ser igual ao seu nome**.
- Não enviar código em PDF.
- Prazo e forma de entrega: dia 22/01/26 às 23h59, no SIGAA.

Coeficientes Cepstrais de Frequência Mel (MFCC) de Sinais de Voz

Foram fornecidos 2 arquivos de áudio, no formato .wav, correspondentes às palavras “sim” e “não”. O objetivo desta prática é calcular os Coeficientes Cepstrais de Frequência Mel (MFCC) destes áudios para saber estes coeficientes podem funcionar como parâmetro para distinguir estes sinais. Em outras palavras, o propósito é determinar se os MFCC são eficientes ferramentas para caracterizar estes sinais de áudios.

Observações:

- **Você de escrever uma justificativa para os resultados obtidos em todas as questões**, em forma de comentário no código.

- Se você for fazer eu código em Python, mandar o arquivo .ipynb já compilado.

- 1) Carregue os 2 sinais de áudio e gere os gráficos destes sinais.
- 2) Filtre os sinais com um filtro de Pré-ênfase usando $\alpha=0,95$. Além disso, você deve uniformizar o tamanho dos sinais, de forma que os 2 sinais tenham a mesma duração (truncando ou acrescentando zeros ao final). Gere os gráficos dos sinais filtrados. Não usar função pronta para esta filtragem.
- 3) Remova a média dos sinais filtrados e, em seguida, divida estes sinais em quadros (*frames*) de 10ms cada. Os sinais foram gerados com uma frequência de amostragem $F_s = 44,1 \text{ KHz}$. Gere o gráfico de 1 *frame* de cada um dos 2 sinais.
- 4) Aplique a janela de Hamming em todos os *frames* obtidos e gere o gráfico de 1 *frame* janelado de cada um dos 2 sinais (os mesmos *frames* da Questão 3). É permitido usar função pronta para a obtenção da janela de Hamming.
- 5) Calcule o módulo ao quadrado da Transformada Discreta de Fourier (DFT) dos *frames* obtidos na Questão 4. Visto que o módulo ao quadrado da DFT é simétrico em relação à frequência zero, você deve eliminar as redundâncias da DFT descartando as frequências negativas. Gere o gráfico do módulo ao quadrado da DFT de 1 *frame* de cada um dos 2 sinais (os mesmos *frames* da Questão 3). O eixo x deste gráficos deve corresponder às frequências entre 0 e $F_s/2$. É permitido usar função pronta para a DFT.
- 6) Aplique os filtros triangulares de escala Mel nas DFTs calculadas na Questão 5. Use 40 filtros triangulares em cada DFT de cada *frame*. Gere o gráfico do módulo ao quadrado da DFT filtrada de 1 filtro de 1 *frame* de cada um dos 2 sinais (os mesmos *frames* da Questão 3 - escolhendo 1 dos 40 filtros). É permitido usar função pronta para os filtros triangulares de escala Mel.

- 7) Calcule a energia de cada banda Mel gerada na Questão 6. Gere o gráfico da energia das bandas Mel em função dos 40 índices Mel para 1 *frame* de cada um dos 2 sinais (para os mesmos *frames* da Questão 3).
- 8) Calcule o Logaritmo das energias Mel geradas na Questão 6. Gere o gráfico do log da energia das bandas Mel em função dos 40 índices Mel para 1 *frame* de cada um dos 2 sinais (para os mesmos *frames* da Questão 3).
- 9) Calcule a Transformada Discreta do Cosseno (DCT) dos log da energia gerado na Questão 8. Pegue apenas os 16 primeiros parâmetros da DCT e ignore o restante. Gere o gráfico do MFCC para 1 *frame* de cada um dos 2 sinais (para os mesmos *frames* da Questão 3). É permitido usar função pronta para a DCT.
- 10) Para cada um dos 2 sinais, concatene (junte) os MFCC de todos os *frames* (calculados na Questão 9) em um só vetor, formando um grande vetor com todos os MFCC de um mesmo sinal. A ideia principal deste trabalho é saber se este vetor com todos os MFCC é uma ferramenta útil para caracterizar o sinal de áudio. Gere o gráfico deste vetor como todos os MFCC para os 2 sinais. Conclua se os MFCC podem funcionar como parâmetro para distinguir estes 2 sinais de áudio fornecidos.