

PSI-2533 Modelagem em Processamento de Sinais

Exercício 1 do Módulo de Modelagem em Processamento de Voz

12 de agosto de 2015

1. Um sinal de voz é amostrado na frequência $F_a = 44100$ Hz. Usa-se uma janela de Hamming de comprimento L para calcular a transformada de Fourier de curto prazo (STFT) do sinal. A STFT é amostrada no tempo com período P e em frequência em $N = 1024$ pontos.

- a) Pode-se demonstrar que o lóbulo principal da janela de Hamming tem uma largura simétrica bilateral completa de aproximadamente $8\pi/L$ rad/amostra. Como deve ser escolhido L se desejamos que a largura bilateral completa do lóbulo principal corresponda a aproximadamente 200 Hz de frequência analógica?
- b) Calcule a duração da janela em ms.
- c) Como devemos escolher P se desejamos uma transformada discreta de Fourier (TDF) a cada 10 ms? De quanto é a sobreposição entre janelas adjacentes?
- d) Qual é o espaçamento em Hz dos pontos amostrados no domínio da frequência?

2. Use o sinal de voz disponível no stoa para este Exercício 1. Cuidado! Este é um arquivo binário, que apesar da extensão `wav` não está no formato wave da Microsoft. No Matlab, ele pode ser lido com o comando

```
[s Fa] = audioread('BR-00432.balsen6.wav', 'double');
```

com frequência de amostragem $F_a = 44100$ Hz. Verifique a leitura correta da frequência de amostragem.

Você deverá obter um espectrograma de faixa larga dos primeiros 1,5 s deste sinal. Use uma janela de Hamming $w(n)$ com 50% de sobreposição entre aplicações adjacentes da janela. Recomenda-se comprimento $L = 300$ para a janela, mas pode ser outro valor que achar mais adequado. Experimente e exponha seus motivos para a escolha de L .

- a) Qual é o valor de P neste caso (vide Questão 1)?

- b) Calcule a duração da janela em ms.
- c) Qual é o espaçamento em Hz dos pontos amostrados no domínio da frequência?
- d) Utilize adequadamente a função `spectrogram` do Matlab para obter seu espectrograma em tons de cinza. Recomenda-se a pré-ênfase prévia do sinal usando, para realce das altas frequências, o filtro de pré-ênfase

$$H_E(z) = 1 - \mu z^{-1}$$

com coeficiente $0 \leq \mu \leq 1$. Ajuste o valor que lhe pareça mais adequado e forneça-o em seu relatório. Sugere-se $\mu = 0,98$, mas outro valor pode ser mais adequado dependendo de sua configuração e técnicas de visualização.

Obs.: Em octave, e antigamente em Matlab, o comando chama-se `specgram`.

- e) A partir do espectrograma, determine as frequências F_1 , F_2 e F_3 das três primeiras formantes por volta de 350 ms. Explique o cálculo da sua estimativa.
- f) Determine a frequência de corte de -3 dB do filtro passa-baixas correspondente à janela $w(n)$ a partir da magnitude da resposta em frequência $W(e^{j\omega})$. Forneça um gráfico mostrando sua determinação.
- g) A partir do espectrograma, qual é a largura de faixa do sinal de voz na sua opinião? Argumente.

3. Use o mesmo sinal de voz da Questão 2. Você deverá obter um espectrograma de faixa estreita dos primeiros 1,5 s deste sinal. Use uma janela de Hamming $w(n)$ com 50% de sobreposição entre aplicações adjacentes da janela. Recomenda-se comprimento $L = 1024$ para a janela, mas pode ser outro valor que achar mais adequado. Experimente e exponha seus motivos para a escolha de L .

- a) Qual é o valor de P neste caso (vide Questão 1)?
- b) Calcule a duração da janela em ms.
- c) Qual é o espaçamento em Hz dos pontos amostrados no domínio da frequência?
- d) Utilize adequadamente a função `spectrogram` do Matlab para obter seu espectrograma em tons de cinza. Recomenda-se a pré-ênfase prévia do sinal usando, para realce das altas frequências, o filtro de pré-ênfase

$$H_E(z) = 1 - \mu z^{-1}$$

com coeficiente $0 \leq \mu \leq 1$. Ajuste o valor que lhe pareça mais adequado e forneça-o em seu relatório. Pode usar o mesmo valor da Questão 2 se não tiver objeção.

- e) A partir do espectrograma, determine a frequência fundamental F_0 da locutora por volta de 350 ms. Explique o cálculo da sua estimativa.

- f) Determine a frequência de corte de -3 dB do filtro passa-baixas correspondente à janela $w(n)$ a partir da magnitude da resposta em frequência $W(e^{j\omega})$. Forneça um gráfico mostrando sua determinação.
- g) A partir do espectrograma, qual é a largura de faixa do sinal de voz na sua opinião? Argumente.