

5º Trabalho Laboratorial

Codificação LPC: análise e re-síntese do sinal. Manipulação intonacional do sinal de erro artificial

Objetivos

Usando o método da predição linear (LPC) e o ambiente Matlab, realizar a codificação deslizante do sinal de áudio de forma a evidenciar o modelo LPC deslizante e o sinal de erro, em função da ordem do modelo LPC e do tipo de segmento do sinal de áudio (fala vozeada, não-vozeada ou mista), explorando também a utilização da pré-ênfase.

Apreciar subjectivamente (timbre e periodicidade) e objectivamente (espectro e periodicidade) o sinal de erro e o grau da sua brancura.

Re-sintetizar um excerto do sinal com uma duração de aproximadamente 1 segundo em duas modalidades, a primeira com o sinal de erro original para produzir uma réplica exacta do sinal original e a segunda com o sinal de erro original substituído por um sinal de erro artificial com idêntica distribuição temporal da energia, constituído por uma combinação linear de trem de impulsos de frequência constante para uma entoação monótona e de ruído branco para produzir os segmentos vozeados, de excitação mista e não-vozeados.

Usar a parte: "E disto percebe"

Material/equipamento necessário

- *PC desktop, Matlab, sinal fornecido e auscultadores.*

Descrição do trabalho a realizar

1- Análise:

1.1 - Para o ficheiro "ASRF20.wav" fornecido, escrever um script em Matlab para realizar o cálculo dos coeficientes LPC ($p=12$), de forma deslizante, com segmentos janelados de duração fixa (Hamming, 30 ms, sobreposição 2/3) e determinação do sinal de erro. O sinal deve ser previamente re-amostrado para $f_s=16\text{kHz}$.

1.2 - Apreciar o sinal de erro obtido através da audição crítica e da medição das suas características de forma deslizante. Explorar a influência de p nas características do sinal de erro como no slide 34 dos apontamentos da aula teórica.

Nota 1: - Poderá ser usada inicialmente a função LPC do Matlab mas também deve ser desenvolvido o próprio algoritmo LPC (na versão AUTO) a partir da definição constante nos apontamentos, para comparar resultados; a informação poderá ser eventualmente complementada pela ajuda do Matlab, disponível em: <https://www.mathworks.com/help/signal/ref/lpc.html>, por exemplo.

Nota 2: - Considerar depois uma variante do processo com a utilização prévia de um filtro de pré-ênfase no sinal (por exemplo de 1ª ordem e frequência de corte de 50Hz,

como $y[n]=x[n]-0.98*x[n-1]$, para 16 kHz) e comparar resultados ao nível dos coeficientes LPC obtidos e do sinal de erro.

2- Re-Síntese:

2.1 - Re-sintetizar o sinal na forma original com o sinal de erro original e comparar as formas de onda original e re-sintetizada. Aprecie as diferenças que poderão ser causadas pela utilização da pre-ênfase, seguida da correspondente de-ênfase no final.

2.2 - Re-síntese com combinação de um tom fundamental monótono e uma excitação de ruído branco as partes de um segmento de aproximadamente 1 segundo de duração em função da seu tipo de excitação

Depois de realizar a anotação de fonemas em PRAAT, realizar a codificação LPC, como em 1 e, para os segmentos obtidos na segmentação resultante da anotação, re-sintetizar com sinal de erro artificial monótono (trem de impulsos de frequência constante ou ruído branco ou combinação dos dois) com ganho apropriado.

3 - Relatório: Fazer os comentários relevantes, relatar incluindo os algoritmos. O resultado será um breve relatório, os programas MATLAB com gráficos e ficheiros “.wav”.

4 - Anexo:

O modelo auto-regressivo utilizado na codificação/descodificação LPC do sinal permite a consideração imediata destas operações no domínio da frequência com os filtros de análise e de síntese correspondentes, como a seguir se mostra.

A regressão de produção de $s(n)$ (FIR) é a seguinte:

$$s(n) = \sum_{k=1}^p a_k s(n-k) + e(n) \quad (1)$$

e a transformada z vale:

$$S(z) = [\sum_{k=1}^p a_k z^{-k}]S(z) + E(z) = \widetilde{S(z)} + E(z) \quad (2)$$

que pode re-escrever-se como:

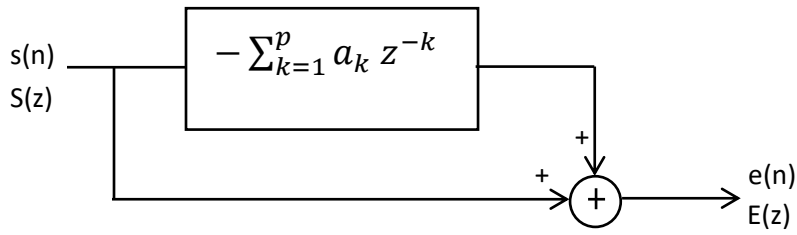
$$S(z)A(z) = E(z) \quad (3)$$

onde:

$$A(z) = 1 - \sum_{k=1}^p a_k z^{-k} \quad (4)$$

e, teremos assim a situação de análise, realizada pelo filtro $A(z)$, que por essa razão se denomina filtro de análise (do tipo FIR).

O diagrama seguinte ilustra a operação de análise LPC:

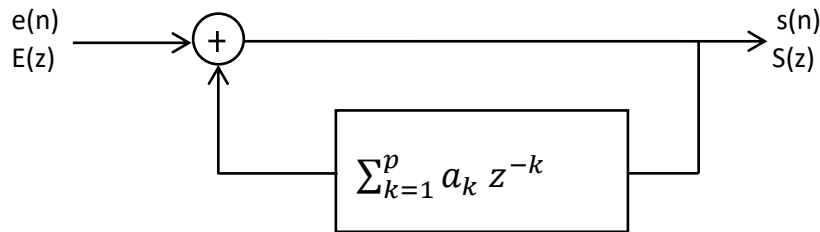


Que implementa a equação: $E(z) = [1 - \sum_{k=1}^p a_k z^{-k}]S(z)$ (5)

Explicitando agora $S(z)$ em função de $E(z)$ na eq. (3), obtém-se a seguinte equação de síntese:

$$S(z) = \frac{E(z)}{A(z)} = \frac{E(z)}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (6)$$

Onde podemos identificar o filtro de síntese $1/A(z)$ (do tipo IIR). O diagrama seguinte ilustra a operação de síntese LPC:



Onde o filtro de síntese será : $\frac{1}{A(z)} = \frac{1}{1 - \sum_{k=1}^p a_k z^{-k}}$ (7)

E a codificação/descodificação LPC pode ser visualizada como operação de filtragem.

A codificação corresponde a uma filtragem FIR e a descodificação corresponde a uma filtragem IIR.

Caso se deseje utilizar em vez do sinal de erro original, $e(n)$, um outro sinal $u(n)$ normalizado com energia unitária, utilizar-se-á um ganho, G , de forma que $e(n)$ seja substituído por $G u(n)$ e resulta que $G^2 = E_p$, a energia do erro de predição, donde pode calcular-se o valor de G :

$$G^2 = R(0) + \sum_{k=1}^p a_k R(k) \quad (8)$$

Em que $R(0)$ e $R(k)$, são os coeficientes de auto-correlação com deslocamentos 0 e k .