

Implementação de um Sistema de Codificação RELP para Fins Didáticos

Carlos Neto, Carlos Silva, Francisco Vaz

Resumo- Neste trabalho foi desenvolvido um sistema de codificação RELP para voz. Este tipo de codificador de média complexidade é bem conhecido e o seu estudo permite ao aluno trabalhar com diferentes tipos de processamento que também são usados noutros codificadores mais eficazes e modernos. Houve a preocupação de desenvolver uma interface com o utilizador muito simples permitindo uma alteração dos parâmetros do codificador e uma fácil visualização dos sinais nas diferentes fases da codificação.

O sistema foi desenvolvido usando Matlab em MS-Windows.

Abstract- In this work we implemented a RELP speech coder. This type of coder is well known and presents medium complexity. It is useful to study because it includes different signal processing techniques that are used in more modern and efficient speech coders. Special attention was given to the user's interface enabling a simple way to change the coding parameters and an easy display of the signals in different stages of the coding process.

The system was implemented using Matlab running on MS-Windows

I. INTRODUÇÃO*

A. Codificação por predição linear

A codificação (ou compressão) da voz é uma área do processamento de sinal que tem como objectivo obter representações compactas da voz para ter uma mais eficaz transmissão e/ou armazenamento [1-2]. Os codificadores conseguem baixar o ritmo de transmissão à custa da degradação da qualidade do sinal, o que implica o estabelecimento de um compromisso entre a qualidade e o ritmo de transmissão pretendidos.

O sistema RELP (*residual excited linear prediction*), que aqui se apresenta é um codificador de fonte. Neste tipo de codificadores o tracto vocal é modelado por um conjunto de filtros digitais variáveis no tempo. No RELP o modelo usado tem apenas pólos, é um modelo de predição linear (*linear prediction coding - LPC*) [3].

O modelo LPC admite que a voz é um sinal que pode ser modelado por um sistema como o representado na figura 1. Podem existir dois tipos de excitação (uma excitação periódica para sons vozeados e ruído branco para sons não vozeados).

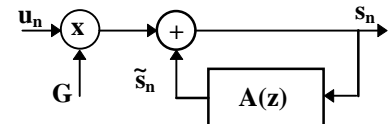


Figura 1 : O modelo LPC de produção de voz

O sinal de saída deste modelo é a soma de uma combinação linear das saídas anteriores (uma predição do sinal) com um termo dependente da excitação.

$$s_n = \sum_{k=1}^p a_k s_{n-k} + G u_n \quad (1)$$

Este modelo fica caracterizado pelos parâmetros -ganho G e coeficientes a_k - e pelo conhecimento do tipo de excitação.

Pode também interpretar-se este processo de geração do sinal como sendo uma predição do sinal com o sistema linear A(z)

$$\tilde{s}_k = \sum_{k=1}^p a_k s_{n-k} \quad (2)$$

Este valor é diferente do verdadeiro, existindo um erro de predição :

$$e_n = \tilde{s}_n - s_n \quad (3)$$

Os parâmetros a_k podem ser calculados minimizando o valor quadrático médio do erro de predição [1] obtendo-se o seguinte sistema de equações (equações de Yule-Walker ou normais)

$$\sum_{k=1}^p a_k R(i-k) = -R(i), 1 \leq i \leq p \quad (4)$$

e a equação

$$E_p = R(0) + \sum_{k=1}^p a_k R(k) \quad (5)$$

em que R(i) é a função de autocorrelação do sinal.

As características do sistema (a matriz de correlação $\|R(i-k)\|$ é uma matriz de Toeplitz) permitem a aplicação de um algoritmo rápido para a sua resolução, conhecido por algoritmo de Durbin ou de Levinson [1][6]. Este algoritmo é iterativo, isto é, calcula o modelo de ordem n a partir do modelo de ordem n-1, o que permite obter todos os coeficientes de todos os modelos desde a ordem 1 até p. São particularmente importantes os coeficientes $a_{i,i}$ -

* Trabalho realizado no âmbito da disciplina de Projecto

coeficientes de ordem i do modelo de ordem i . Estes coeficientes são habitualmente designados por coeficientes de reflexão, de correlação parcial ou PARCOR e simbolicamente representados por $k_i = a_{i,i}$.

B. O codificador RELP

O modelo RELP é considerado como uma técnica de codificação de complexidade média alta [5] [10], com qualidade telefónica, permitindo débitos binários entre os 9.6Kbit/s e os 4.8Kbit/s. O seu grau de complexidade facilita a sua implementação em tempo real.

Um sistema de codificação é constituído por um emissor-codificador e um receptor-descodificador. Nos codificadores baseados no modelo de predição linear, o emissor, após a segmentação do sinal, faz uma análise do sinal calculando os parâmetros do modelo de cada segmento e codifica-os em seguida. No receptor, após a descodificação dos parâmetros, uma predição do sinal pode ser obtida de acordo com a equação (2). No RELP codifica-se e transmite-se também o erro de predição $e(n)$, possibilitando que o sinal original seja recuperado (equação (4)), a menos de erros introduzidos pela codificação.

A economia de bits na transmissão é conseguida à custa das características espectrais do erro. Normalmente é suficiente enviar um conteúdo de frequências entre os 50Hz e os 800Hz permitindo efectuar uma decimação deste sinal antes da sua codificação. No receptor, usando métodos de interpolação, regenera-se o sinal de erro que será usado para excitar o filtro de síntese. O sinal de erro decimado deve conter a frequência fundamental, e se possível alguns harmónicos, para deste modo se obter uma boa reconstituição no receptor.

A codificação do erro pode ser feita no domínio do tempo ou da frequência. É mais frequente a codificação no domínio do tempo porque é mais rápida e permite uma codificação mais eficaz pelo que optámos por ela. A codificação consiste numa limitação da banda abaixo da frequência $B=W/L$ (W é a largura de banda original) e numa decimação à frequência de Nyquist de $2B$ que é feita retendo uma amostra em cada L amostras e ignorando as restantes. No fim faz-se a quantificação das amostras retidas. No receptor é feita a descodificação das amostras seguindo-se um interpolação e filtragem passa baixo.

Para que o sinal de erro reconstruído se aproxime mais do original é necessário fazer uma regeneração das altas frequências, o que constitui a parte mais delicada de todo o codificador RELP. C.K.Un e J.R.Lee [4] classificam em três categorias as técnicas utilizadas para regenerar as altas frequências: distorção não linear, duplicação espectral e implantação do *pitch*. Mais tarde foram propostos outros métodos de regeneração das altas frequências, mas optámos pela utilização do *Perturbed Spectral Folding*, que é um método de duplicação espectral em que se soma ruído a algumas amostras diferentes de zero. Na duplicação do espectro o sinal de banda base é duplicado para as bandas

superiores de forma a preencher as bandas vazias com réplicas da banda base. Este método pode ser implementado fazendo uma sobre-amostragem do sinal de banda base, caso este tenha sido previamente sub-amostrado. O processo é extremamente simples e não requer filtragens, ajustamento de ganhos ou distorções não lineares. Possui ainda a vantagem de evitar ruídos tonais, resultantes de periodicidades introduzidas pelo método simples de duplicação espectral.

Na figura seguinte apresenta-se um esquema simplificado do emissor e do receptor RELP:

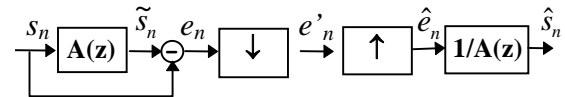


Figura 2 : Emissor e receptor RELP.

II. IMPLEMENTAÇÃO DO EMISSOR DO RELP

A figura 3 mostra um diagrama de blocos do emissor do codificador RELP. Cada bloco será analisado em detalhe nas seguintes fases de codificação.

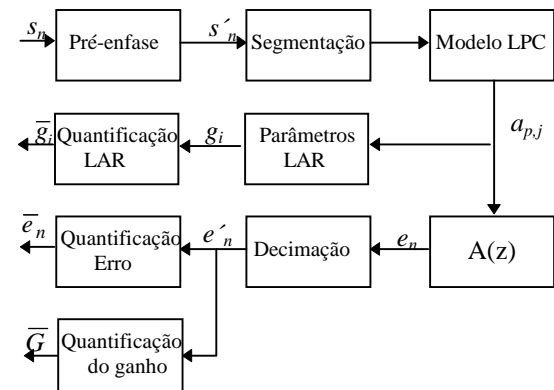


Figura 3 : Diagrama de blocos dos emissor RELP.

A. Pré-ênfase.

A envolvente do espectro de potência da voz humana é, em geral, claramente decrescente. É aconselhável efectuar uma filtragem de acentuação das altas frequências, operação designada por pré-ênfase e que reduz a gama dinâmica do espectro dos sons vozeados, realçando as componentes de alta frequência e melhorando a quantificação do sinal. Esta filtragem é feita com um filtro com um único zero real, cuja a função de transferência é $H(z) = 1 - \alpha z^{-1}$. O parâmetro α deve ter um valor próximo da unidade, usando-se correntemente o valor 0.95.

B. Segmentação

Um sinal de voz é um sinal não estacionário, no entanto se for feita uma divisão do sinal em segmentos suficientemente pequenos, podemos considerar estes como sendo quasi

estacionários. A janela de segmentação não deve ser muito pequena, devendo incluir pelo menos dois períodos da frequência fundamental para sons vozeados [10]. O *pitch* é no máximo 10ms [8], o que implica o uso de janelas na ordem dos 20 a 30 ms. É conveniente a utilização de janelas com transições suaves, como são por exemplo os casos das janelas de Hanning ou de Hamming. É também preferível efectuar a segmentação com sobreposição para evitar ruídos síncronos com o ritmo de segmentação. [10].

C. Determinação e Quantificação do modelo LPC

Para fazer o cálculo dos coeficientes LPC utilizou-se o algoritmo de Durbin que permite obter os parâmetros a_k e os coeficientes de reflexão k_i . Os a_k serão guardados para posterior cálculo do erro de predição e os k_i irão ser usados indirectamente para a codificação e transmissão.

A escolha da forma de transmissão dos parâmetros deve ser feita de modo a garantir:

- uma boa representação do sinal,
- a estabilidade do filtro após a quantificação,
- uma gama dinâmica pequena para economia de bits na quantificação.

Se não for feita quantificação, a informação quer dos coeficientes a_k quer dos de reflexão k_i é a mesma, o mesmo já não sucedendo após a quantificação porque os erros de quantificação introduzidos têm efeitos diferentes.

O método de cálculo utilizado assegura só por si a estabilidade do filtro mas ao quantificarmos a_k a estabilidade deixa de estar garantida devida os erros introduzidos pela quantificação. Só a utilização de uma precisão elevada (cerca de oito a dez bits) garante essa mesma estabilidade [10].

A estabilidade do filtro gerado a partir dos parâmetros k_i quantificados é mais fácil de verificar uma vez que é suficiente que estejam no intervalo]-1,1[. Este facto justifica a escolha dos k_i para serem quantificados. No entanto a sua utilização é também desaconselhada por outras razões.

Define-se sensibilidade espectral ao parâmetro k_i como sendo:

$$\frac{\partial \mathcal{S}}{\partial k_i} = \lim_{\Delta k_i \rightarrow 0} \left[\frac{1}{\Delta k_i} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \log \frac{P(k_i, w)}{P(k_i + \Delta k_i, w)} dw \right] \right] \quad (6)$$

onde $P(., w) = |H(e^{-jw})|^2 = 1 / |A(e^{-jw})|^2$ é o espectro de potência calculado com o modelo $H(z)$.

A figura seguinte adaptada de [7] mostra a evolução desta sensibilidade para os parâmetros k_i e para um modelo de doze pólos extraídos dum segmento de 20 ms de um sinal de voz amostrado a 10kHz. Cada curva representa a sensibilidade espectral de um dos doze coeficientes de reflexão no intervalo]-1,1[enquanto os outros onze coeficientes são mantidos constantes. A curva é típica tanto

de sons vozeados como não vozeados, assim como de oradores femininos como masculinos. Da figura torna-se claro que é necessário um maior número de bits de quantificação para os valores de k_i próximos de 1.

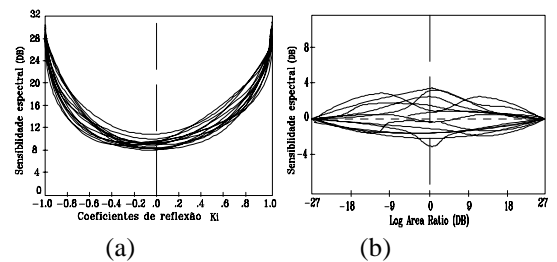


Figura 4 : Sensibilidade espectral
a) Coeficientes de reflexão. b) Parâmetros LAR.

O gráfico b da figura mostra a sensibilidade aos parâmetros LAR- *log area ratio*- definidos como

$$g_i = \log \frac{1 + k_i}{1 - k_i}, \quad i = 1, \dots, p \quad (7)$$

É clara a menor sensibilidade e uma maior uniformidade de comportamento dos parâmetros g_i em relação aos k_i .

A quantificação dos parâmetros *log-area-ratio* foi feita segundo o método apresentado em [7] e designado por *Optimum bit allocation*.

A distribuição dos bits de quantificação é feita minimizando o máximo desvio espectral devido à quantificação. O desvio total ΔS devido a variações de Δg_i , em que g_i é o parâmetro a quantificar, com $1 \leq i \leq p$, é dado por:

$$\Delta S = \sum_{i=1}^p \left| \frac{\partial \mathcal{S}}{\partial g_i} \Delta g_i \right| \quad (8)$$

Dado o número total de bits de quantificação M temos que proceder à sua distribuição pelos p parâmetros, M_i bits para cada g_i . O número total de níveis de quantificação é $N = 2^M$ e para cada parâmetro usam-se N_i . Entre estes parâmetros existem as seguintes relações:

$$\sum_{i=1}^p M_i = \sum_{i=1}^p \log_2 N_i = M \quad (9)$$

$$N_i = 2^{M_i}, 1 \leq i \leq p \quad (10)$$

O passo de quantificação para g_i é definido como sendo:

$$\delta_i = \frac{g_i - \underline{g}_i}{N_i} \quad (11)$$

onde \underline{g}_i e \overline{g}_i são os limites superiores e inferiores.

Usando uma aritmética de arredondamento por defeito, o máximo erro de quantificação cometido é metade do nível de quantificação:

$$|\Delta g_i|_{\max} = \frac{1}{2} \delta_i \quad (12)$$

Então o máximo desvio é após substituição :

$$(\Delta S)_{\max} = \sum_{i=1}^p \left| \frac{\partial \mathcal{S}}{\partial g_i} \right| \frac{g_i - \underline{g}_i}{2N_i} \quad (13)$$

Se fizermos

$$K_i = \left| \frac{\partial \mathcal{S}}{\partial g_i} \right| \frac{g_i - \underline{g}_i}{2N_i} \quad (11)$$

e por substituição de (11) na expressão de (10)

$$(\Delta S)_{\max} = \sum_{i=1}^p \frac{K_i}{N_i} \quad (14)$$

Pretendemos minimizar $(\Delta S)_{\max}$ em relação a N_i com a condição (6).

A minimização dá os seguintes resultados [7]:

$$N_1 = K_1 \left[N \left(\prod_{i=1}^p K_i \right)^{-1} \right]^{1/p} \quad (15)$$

$$N_i = \frac{K_i}{K_1} N_1, 2 \leq 2i \leq p \quad (16)$$

As expressões (14-16) anterior dizem-nos que a contribuição relativa dos diferentes parâmetros para o máximo desvio espectral deverá ser a mesma. Como já foi indicado os parâmetros *log-area-ratio* possuem uma sensibilidade espectral $\frac{\partial \mathcal{S}}{\partial g_i}$ que é aproximadamente

constante e semelhante para todos os g_i . Podemos então tirar a conclusão de que o passo de quantificação δ_i deve ser o mesmo para todos os parâmetros *log-area-ratio* e determinado por:

$$\delta = \left[2^{-M} \prod_{i=1}^p (g_i - \underline{g}_i) \right]^{1/p} \quad (17)$$

Sabendo a gama de variação dos parâmetros *log-area-ratio*, é possível distribuir um conjunto de bits de uma forma óptima. Neste trabalho a gama de variação dos parâmetros foi determinada empiricamente.

D.Determinação do sinal de erro

De acordo com a figura 2, o sinal de erro é obtido subtraindo ao sinal original, o sinal de voz processado pelo filtro $A(z)$.

Como já foi explicado na introdução, o sinal assim obtido contém informação redundante que pode ser eliminada fazendo uma decimação. A decimação reduz o número de amostras do sinal de erro o que permite reduzir de forma significativa o ritmo de transmissão.

O bloco decimador é constituído por um filtro passa baixo com largura de banda menor que a largura de banda do sinal original a que se segue uma sub-amostragem para a nova frequência de Nyquist. É necessário que este novo sinal de erro de largura B , contenha informação do *pitch* e alguns

formantes dos sons vozeados, e no caso de sons não vozeados que se assemelhe a ruído de banda limitada.

E.Quantificação dos parâmetros e do erro

Genericamente para fazer a quantificação o utilizador tem que inicialmente definir os valores máximos para cada um dos parâmetros *log area*, para os sinais de erro e ganho, para posterior normalização dos seus valores entre $]-1,1[$ para o sinal de erro e $[0,1]$ para o ganho. Tem ainda que atribuir um conjunto de bits de quantificação a cada um dos parâmetros anteriores. O número de bits atribuídos determina a precisão com que é representado cada parâmetro.

Conhecido o número de bits atribuído a cada parâmetro é possível calcular a resolução máxima permitida para a gama dinâmica a quantificar. Dividindo o valor do parâmetro a quantificar pela resolução e arredondando para o inteiro mais próximo de zero obtém-se um valor positivo ou negativo consoante o sinal do parâmetro quantificado. Este valor é transmitido (na nossa implementação em Matlab, escrito num ficheiro *.mat*) e recebido pelas rotinas do receptor (leitura do ficheiro *.mat* referido). As rotinas de recepção fazem a operação inversa: o valor recebido é desnormalizado atendendo à resolução e gama dinâmica definidas. Este valor é igual ao original a menos do erro cometido devido à quantificação. Estudou-se inicialmente a hipótese de não quantificar directamente o sinal de erro, mas sim a diferença entre as amostras. No entanto, após uma cuidadosa análise constatou-se que se obtinham piores resultados do que com uma quantificação simples e uniforme. Tal facto, deve-se às transições abruptas que o sinal de erro apresenta, alternando por vezes de altos valores positivos para altos valores negativos o que faz com que o sinal diferença apresente uma grande gama dinâmica e seja difícil de quantificar com boa resolução. Note-se que estas transições abruptas transportam informação importante acerca do *pitch*.

D.Determinação e quantificação do ganho

O ganho do filtro foi estimado de uma forma extremamente simples. Como o ganho deste tem que estar directamente relacionado com a energia do sinal, então o sinal de erro também deve reflectir essa mesma energia. Desta forma o ganho do filtro seria unitário, mas como o sinal de erro é normalizado no intervalo $]-1,1[$, o ganho é simplesmente o valor máximo do módulo do erro antes da normalização. Foi feita quantificação uniforme do ganho por não se verificar a necessidade nem ser habitual usar outro tipo de quantificação.

III. IMPLEMENTAÇÃO DO RECEPTOR DO RELP

A.Descodificação do erro, ganho e parâmetros *log-area-ratio*.

Estes parâmetros são codificados como uma sequência de número inteiros, que representam conceptualmente uma

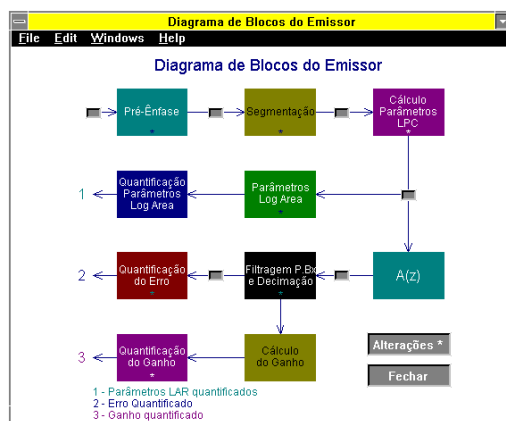
IV. SISTEMA IMPLEMENTADO

O trabalho foi implementado sobre uma plataforma MatLab (versão 4.2 para MS-Windows), possibilitando ao utilizador testar as potencialidades do codificador RELP, permitindo-lhe alterar os valores dos parâmetros programados nos diagramas de blocos do emissor e receptor.

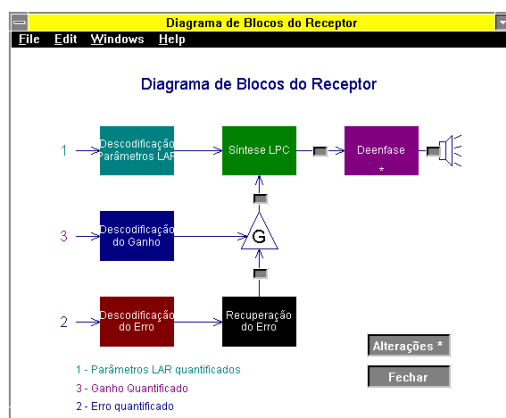
O sistema lê apenas ficheiros em formato .wav e considera que a frequência de amostragem é de 8k amostras/s.

Ao correr o programa, o utilizador fica perante sucessivos menus que lhe permitem explorar as várias facilidades implementadas. O primeiro menu permite a escolha das seguintes opções

- Introdução
- Ajuda
- Ler sinal
- Diagrama de blocos do emissor
- Diagrama de blocos do receptor
- Ouvir sinal original
- Ouvir sinal recuperado
- Inicializar parâmetros



a)



b)

Figura 7 : Menus gráficos para : a) emissor b) receptor.

Os menus das opções Diagrama de blocos do emissor/receptor estão representados na figura 7.

O utilizador pode seleccionar o bloco de processamento que pretende modificar (apenas os blocos assinalados por *). Após selecção um novo menu estará presente no ecrã, permitindo a alteração dos parâmetros do bloco de processamento seleccionado. A figura 8 apresenta como exemplo a definição dos parâmetros da janela de segmentação.

O utilizador pode também optar por visualizar o sinal em processamento. Para tal deverá seleccionar os pontos de visualização entre os blocos

Os seguintes parâmetros são susceptíveis de alteração:

- Número de parâmetros LPC.
- Largura de Banda base do erro: 400HZ, 600Hz, 800Hz, 1000Hz e 2000Hz.
- Número de bits atribuídos ao sinal de erro, aos parâmetros *Log-area-ratio* e ao ganho.
- Valores máximos de cada parâmetro *Log-Area-Ratio*, valor máximo do erro normalizado e valor máximo do ganho.
- Tipo de janela de segmentação (rectangular ou Hanning), duração da janela e tempo de sobreposição entre janelas .
- Posição do pólo do filtro de pré-ênfase, a posição do zero no filtro de deênfase ou ainda a desactivação de um e/ou outro filtros.

Sempre que mude um parâmetro que altere o ritmo de transmissão, o utilizador terá de imediato o novo ritmo, que é actualizado automaticamente e apresentado na janela em que está a trabalhar.

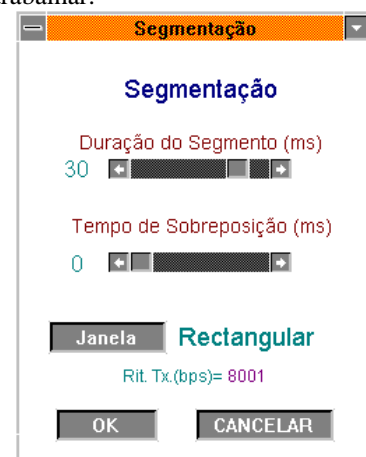


Figura 8 - Menu para definir os parâmetros da janela de segmentação.

Foi ainda implementada uma medida objectiva da relação sinal ruído segmental (SNR). A expressão seguinte representa a relação sinal ruído para cada segmento:

$$SNR = \frac{\sum_{i=1}^m (x_R^{(i)} - x_O^{(i)})^2}{\sum_{i=1}^m x_O^{(i)2}} \quad (19)$$

em que x_O é o sinal original e x_R o sinal recuperado. A relação sinal ruído do sinal recuperado é feita fazendo a média do SNR dos vários segmentos que o constituem.

O programa foi feito admitindo que o computador possui placa de som. Assim é possível ouvir o sinal original e o sinal reconstruído, podendo desta forma fazer-se uma avaliação qualitativa do desempenho do sistema com as alterações feitas aos parâmetros.

A opção "inicializar parâmetros" permite um regresso aos valores de defeito escolhidos para os parâmetros permitindo um ritmo de transmissão de 8000 bit/s.

V. RESULTADOS E CONCLUSÕES

O nosso objectivo foi tentar obter a melhor qualidade possível com um ritmo de transmissão de 8000 bit/s. A qualidade conseguida, tomando como medida a relação sinal ruído segmental é de cerca de 10 dB, variando com a duração da frase. Na tabela seguinte apresentam-se os valores obtidos para duas frases: Frase 1 "O sonho comanda a vida" e Frase 2 "Viseu é a capital da Beira Alta e a cidade mais bonita de Portugal"

Ritmo tx (bps).	8001	10080	12000
Frase 1	10.6	11.0	11.5
Frase 2	7.9	8.1	10.0

Tabela 1

Na audição de sinais reconstruídos pelo sistema, nota-se um realce das componentes de baixa frequência e uma atenuação das altas frequências. O sucedido deve-se ao facto da deficiente recuperação de altas frequências do sinal de erro. A figura 8 mostra esse efeito.

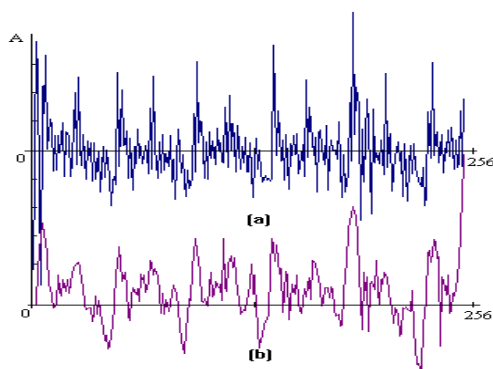


Figura 8 : Sinal de erro da palavra "olá" . a) original b)recuperado.

Fizemos testes informais de audição da frase: "O sonho comanda a vida". Foi dada a oportunidade a oito pessoas de ouvir a frase recuperada. Seis dos ouvintes identificaram a frase com facilidade. Os dois restantes não identificaram a palavra sonho na primeira audição. Mais uma vez estamos

perante um efeito de redução das altas frequências contidas no som fricativo "s", que quase desaparece na reconstrução, acrescido de um ênfase das baixas frequências maioritárias no resto da palavra.

Para terminar apresentamos na figura 9 a palavra "olá" na sua versão original e reconstruída após uma codificação a 8kbit/s. A sua audição não suscitou dificuldades de compreensão embora se verifique um aumento significativo do ruído de fundo.

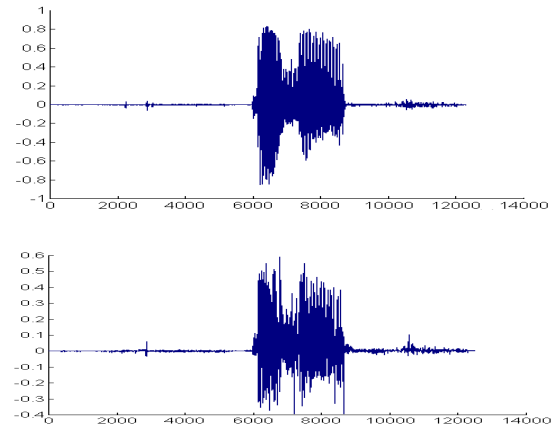


Figura 9 : (a) Sinal original. (b) Sinal reconstruído.

O grande inconveniente deste programa tem a ver com o facto do MatLab se basear numa linguagem interpretada. Cada vez que se pretende ver os efeitos da alteração de um parâmetro, quando se está a processar um sinal com alguns segundos, o programa pode levar alguns minutos a fornecer o resultado. Mas não esqueçamos que o objectivo foi desenvolver um programa para fins didáticos, facilmente utilizável, e este objectivo supomos que foi de facto concretizado.

REFERENCES

- [1] A.M Kondoz " Digital speech coding for low bit rate communications systems, Wiley, 1994.
- [2] L. Rabiner and B.H Juang, " Fundamentals of Speech recognition ", Prentice Hall, 1993.
- [3] H.K Un and D.T. Magill, " The Residual-Excited Linear Prediction Vocoder with Transmission Rate Below 9.6 kbits/s, IEEE Transactions on communications vol . com-23, nº 12, pp. 1466-1474, December 1975.
- [4] C.K. Un and J.R Lee "On spectral Flattening Techniques in Residual-Excited Linear Prediction Vocoding ", IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 216-219, 1982.
- [5] V.R. Viswannathan, A. L. Higgins, and W.H. Russell, " Design of a Baseband LPC Coder For Speech Transmission Over 9.6 kbit/s Noisy Channels, IEEE Transactions on Communicatons, vol. com-30 , nº4, pp. 663-673, April 1982.
- [6] J. Makoul, " Linear Prediction: A Tutorial Review ", Proc. IEEE, vol. 63, pp. 561-580, April 1975.

- [7] R. Viswanathan and J. Makoul, "Quantization Properties of Transmission Parameters in Linear Predictive Systems", IEEE Transactions on communications vol. ASSP-23, n° 3, pp. 309-324, June 1975.
- [8] T.V. Ananthapadmanabaha and B. Yegnanarayana "Epoch Extraction from Linear Prediction Residual for Identification of Closed Glottis Interval", IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-27, n°4 pp. 309-318, August 1979.
- [9] D.Y. Wong C.C. Hsiao, and J.D. Markel, "Spectral Mismatch Due to Preenphasis in LPC Analysis/Synthesis", IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-28, n°2 pp. 263-264, April 1980.
- [10] F. M. Gomes de Sousa, "Codificação de Fala para Transmissão em Canal Telefónico e Via Rádio", Universidade Técnica de Lisboa, Instituto Superior Técnico, Lisboa, junho de 1993.