

Codec MPE-LPC de Baixo Atraso

José Sindi Yamamoto e Fábio Violaro

Neste trabalho é apresentado um codificador-decodificador (Codec) de voz de baixo atraso baseado na técnica MPE-LPC ("Multipulse Excited-Linear Prediction Coding"). Primeiramente, é feita uma descrição geral do algoritmo de geração da excitação multipulso. Em seguida, é apresentada a técnica empregada para a redução do atraso de algoritmo de codificação e alguns detalhes de implementação como tabelas de quantização e valores de constantes utilizados. Finalmente, são apresentados alguns resultados de desempenho obtidos em termos de Razão Sinal/Ruído Segmental e Distância Cepstral.

1. INTRODUÇÃO

O algoritmo de codificação de voz MPE-LPC fornece uma boa qualidade de voz a taxas de bits entre 9,6 e 16 kbit/s. Entretanto, a extração de parâmetros tais como coeficientes do filtro de síntese LPC (preditor LPC), período de "pitch" e ganho do filtro de síntese a longo-prazo (preditor de "pitch") são baseados na formação de blocos do sinal de voz original segundo uma estrutura não-realimentada conhecida como "forward", resultando em um atraso de algoritmo de codificação da ordem de 25ms.

Tal magnitude de atraso é indesejável para aplicações em redes telefônicas, pois, de acordo com a recomendação CCITT G.131 implica, independentemente de outros atrasos inerentes aos sistemas de transmissão, na utilização de canceladores de eco que não são normalmente utilizados nas centrais das redes telefônicas, a não ser em casos especiais como ligações via satélite.

Com o intuito de contornar este problema, foi implementado neste trabalho um codec MPE-LPC utilizando uma estrutura realimentada conhecida como "backward", de modo a se obter um baixo atraso e boa qualidade à taxa de bits em torno de 16 kbit/s. Em tal estrutura, os diversos parâmetros de codificação são obtidos a partir do sinal sintetizado ou quantizado, o que além de permitir uma redução no atraso de algoritmo de codificação de 25ms para aproximadamente 5ms ou menos, evita a necessidade de transmissão de informações paralelas.

2. ALGORITMO DE EXCITAÇÃO MULTIPULSO

2.1. Modelo de Excitação Multipulso

Segundo a técnica utilizada em codificadores por Predição Linear (LPC) tradicionais, a geração do sinal de excitação é realizada considerando-se a existência de apenas dois tipos de sons: sonoros e não-sonoros. No caso de segmentos de sinais de voz sonoros, o sinal de excitação é constituído de apenas 01 (um) pulso por período fundamental e, no caso de sinais não-sonoros, o sinal de excitação consiste na geração de um ruído branco. Devido a este modelamento um tanto falho, os LPC tradicionais são eficientes para reduzir a taxa de bits a valores muito baixos, mas às custas de uma redução notável na qualidade da voz. Além disto, de um modo geral, a qualidade não é melhorada aumentando-se a taxa de bits.

Em 1982, foi proposta por Atal e Remde [1], a técnica de excitação *multipulso* e os codificadores de voz LPC com este tipo de excitação ficaram conhecidos como MPE-LPC. De acordo com esta nova técnica, a necessidade de discriminação do sinal de voz em sonoros e não-sonoros é eliminada, gerando um tipo de excitação com uma maior densidade de pulsos independente de os sons serem sonoros ou não-sonoros. A posição destes pulsos e a amplitude dos mesmos são determinadas utilizando-se um método de análise por síntese que procura minimizar o erro quadrático médio, o qual é ponderado de forma a refletir melhor a percepção auditiva humana.

Na **Figura 1** é mostrado o diagrama em blocos de um sintetizador LPC com excitação multipulso. O diagrama é bastante parecido com o sintetizador LPC tradicional, exceto pela ausência do gerador de pulsos, gerador de ruído e do comutador para sinais de voz classificados como sonoros ou não-sonoros.

A excitação para o bloco sintetizador LPC é produzida por um único gerador de excitação que produz uma seqüência de pulsos localizados nos instantes $m_1, m_2, \dots, m_i, \dots$, com amplitudes $A_1, A_2, \dots, A_i, \dots$, respectivamente. Se o número de pulsos é aumentado arbitrariamente a valores grandes, de modo que haja um pulso a cada instante de amostragem, é possível obter-se uma réplica do sinal de voz original (ao custo de uma alta taxa de bits). Entretanto, é possível sintetizar todos os tipos de sons de voz, incluindo os sons sonoros e não-sonoros, com distorção quase imperceptível subjetivamente, empregando apenas poucos pulsos por bloco (por exemplo, 8 pulsos a cada bloco de 5ms).

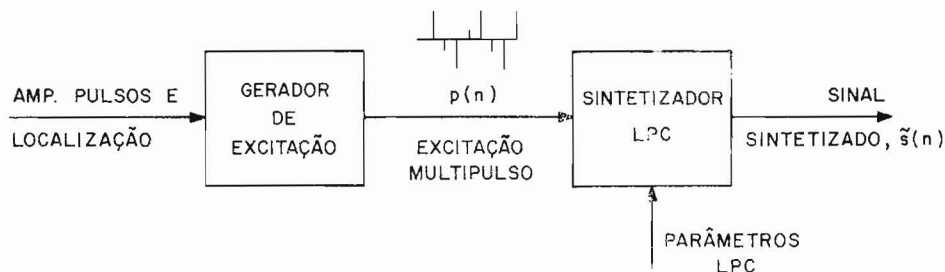


Figura 1. Diagrama em Blocos de um Sintetizador LPC com Excitação Multipulso.

2.2. Cálculo da Excitação Multipulso

Os procedimentos envolvidos num codec utilizando o algoritmo MPE-LPC terminam quando as posições m_i e amplitudes A_i dos pulsos de excitação são conhecidas em um dado intervalo de tempo. As amplitudes e posições dos pulsos são selecionadas de modo que o sinal de voz sintetizado satisfaça um critério de fidelidade relativo ao sinal original.

Muitas variações de métodos de cálculo das amplitudes e posições dos pulsos da excitação multipulso têm sido propostas na literatura [2-13]. Neste trabalho optou-se por uma implementação do método de Berouti et al. [2], já testada anteriormente num codec MPE-LPC à taxa de 13,65 kbit/s com estrutura "forward" [17,18]. Neste método, a análise multipulso pode ser representada segundo a **Figura 2**.

O fator μ utilizado nos dois filtros de síntese LPC, $H(z/\mu)$, pondera o resíduo de modo que a razão sinal sintetizado/ruído seja menor na região das formantes, pois para a audição humana o ruído na região das formantes é mascarado pelo sinal de voz.

O filtro de síntese LPC com ponderação excitado pela seqüência multipulso, $p(n)$, sintetiza o sinal de voz ponderado, $\tilde{s}_w(n)$. A diferença entre o sinal de referência ponderado, $d(n)$, e o sinal de voz sintetizado com ponderação, $\tilde{s}_w(n)$, deve ser minimizada durante o processo de geração da excitação multipulso, utilizando-se como medida o erro quadrático médio.

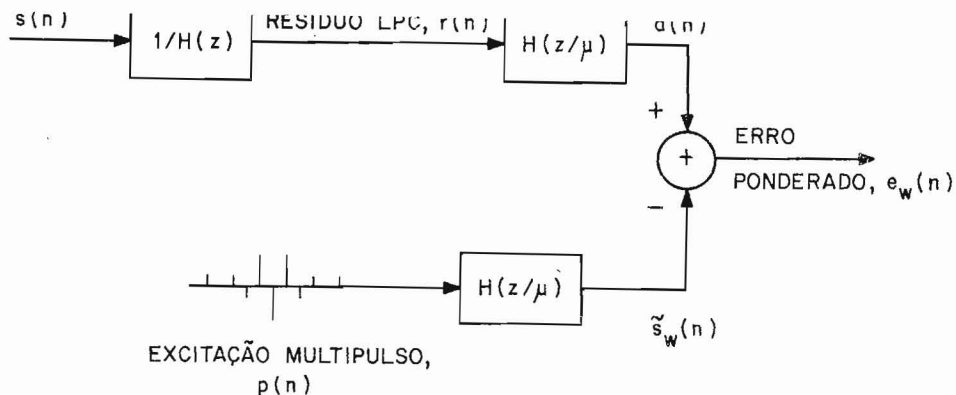


Figura 2. Modelo de Excitação Multipulso.

Em termos de implementação, o sinal de voz original é particionado em blocos de N amostras sendo alocados em cada bloco N_p pulsos. Sejam $h_v(n)$ e $h_w(n)$, a resposta causal ao impulso do filtro inverso, $1/H(z)$, e do filtro de síntese LPC com ponderação, $H(z/\mu)$, respectivamente. Então, o sinal de resíduo, $r(n)$, é dado por:

$$r(n) = \sum_{k=-\infty}^n s(k) h_v(n-k), \quad 0 \leq n < N \quad (1)$$

O sinal de referência com ponderação, $d(n)$, é obtido passando-se o sinal de resíduo, $r(n)$, pelo filtro de síntese LPC com ponderação, $H(z/\mu)$:

$$d(n) = \sum_{k=-\infty}^n r(k) h_w(n-k), \quad 0 \leq n < N \quad (2)$$

O sinal de voz sintetizado com ponderação, $\tilde{s}_w(n)$, é calculado passando-se a excitação multipulso, $p(n)$, pelo segundo filtro de síntese com ponderação, $H(z/\mu)$:

$$\tilde{s}_w(n) = \sum_{k=-\infty}^n p(k) h_w(n-k), \quad 0 \leq n < N \quad (3)$$

Nas expressões (1) a (3), o limite inferior da somatória foi tomado como $-\infty$ ao invés de zero, para levar em conta a memória do filtro devida aos blocos anteriores.

As amplitudes dos pulsos da excitação multipulso, A_i , e as suas posições correspondentes, m_i , são então obtidas através de um processo de minimização do erro quadrático médio ponderado dado por:

$$(e_w^2) = \sum_{n=0}^{N-1} |d(n) - \tilde{s}_w(n)|^2 \quad (4)$$

Após serem colocados N_p pulsos nas posições m_i , tem-se:

$$\langle e_w^2 \rangle = \sum_n |d'(n) - \sum_{i=1}^{N_p} A_i h_w(n - m_i)|^2, \quad (5)$$

onde A_i é a amplitude do i -ésimo pulso e $d'(n)$ corresponde a $d(n)$ após ter sido descontada a memória do filtro de síntese devida aos blocos anteriores (resposta à entrada nula). Diferenciando-se esta expressão em relação às amplitudes A_i , obtém-se o seguinte conjunto de equações simultâneas:

$$\begin{bmatrix} \Phi(m_1, m_1) & \Phi(m_1, m_2) & \dots & \Phi(m_1, m_{N_p}) \\ \Phi(m_2, m_1) & \Phi(m_2, m_2) & \dots & \Phi(m_2, m_{N_p}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi(m_{N_p}, m_1) & \Phi(m_{N_p}, m_2) & \dots & \Phi(m_{N_p}, m_{N_p}) \end{bmatrix} \times \begin{bmatrix} \hat{A}_1 \\ \hat{A}_2 \\ \vdots \\ \hat{A}_{N_p} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{N_p} \end{bmatrix} \quad (6)$$

onde:

$$\Phi(m_i, m_j) = \sum_{n=0}^{N-1} h_w(n - m_i) h_w(n - m_j), \quad 1 \leq i, j < N_p \quad (7)$$

$$\beta_i = \sum_{n=0}^{N-1} d'(n) h_w(n - m_i), \quad 1 \leq m < N_p \quad (8)$$

e \hat{A}_i , $i = 1, \dots, N_p$, são as incógnitas.

Na forma matricial obtém-se:

$$\underline{\Phi} \cdot \underline{\hat{A}} = \underline{\beta} \quad (9)$$

O método adotado para a resolução desta equação é o algoritmo de Cholesky baseado na decomposição triangular da matriz Φ [14].

Este algoritmo é aplicado no presente trabalho de uma maneira seqüencial à medida que a posição m_i de cada novo pulso é determinada. Pode-se demonstrar que a melhor posição para um único pulso é o valor de m que maximiza a relação [2].

$$\frac{\beta_i^2}{\Phi(m_i, m_i)} \quad (10)$$

Assim, a cada nova posição m_i do pulso determinada segundo este procedimento, resolvendo-se a Eq. 9 pelo método de Cholesky, determina-se o valor da amplitude \hat{A}_{lp} (lp é o número de pulsos até então alocados), e os valores das amplitudes \hat{A}_i , $i = 1, \dots, (lp - 1)$, são re-otimizados de modo a minimizar o erro quadrático médio ponderado dado pela Eq. 4.

Neste procedimento de determinação e re-otimização da amplitude dos pulsos, o efeito dos pulsos que foram sendo alocados deve ser removido da seqüência $\{d'(n)\}$ que será usada para determinar a posição ótima do próximo pulso:

$$d''(n) = d'(n) - \sum_{i=1}^{lp} \hat{A}_i h_w(n - m_i) \quad (11)$$

onde lp é o número de pulsos já alocados.

Substituindo-se o novo valor $d''(n)$ na Eq. 8, obtém-se uma fórmula para a atualização da função de correlação cruzada β_i :

$$\beta'_i = \beta_i - \sum_{i=1}^{lp} \hat{A}_i \Phi(m_i, m_i) \quad (12)$$

Desta maneira, para a determinação da posição do próximo pulso, isto é, do $(lp + 1)$ -ésimo pulso, a relação a ser maximizada é:

$$\frac{\beta_{lp+1}^2}{\Phi(m_{lp+1}, m_{lp+1})} \quad (13)$$

Para ser obter um melhor desempenho, o modelo de excitação multipulso utilizado neste trabalho inclui um filtro de síntese a longo-prazo, $H_p(z)$, antes do filtro de síntese LPC, como mostrado na **Figura 3**.

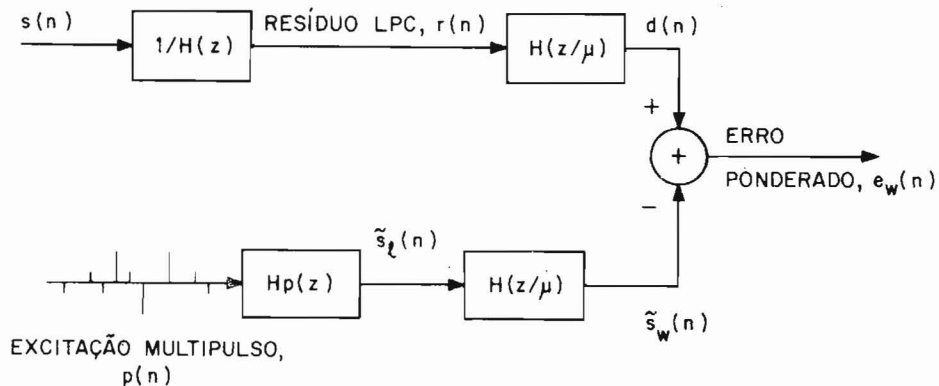


Figura 3. Modelo de Excitação Multipulso com Filtro de Síntese a Longo-Prazo.

3. Princípio de Funcionamento do Codec MPE-LPC de Baixo Atraso

3.1. Codificador

Na **Figura 4a**, tem-se o diagrama em blocos do codificador MPE-LPC de Baixo Atraso (MPE-LPC-BA). O funcionamento do codec é baseado na formação de blocos do sinal de voz original (i.e., opera de bloco em bloco) segundo uma estrutura "backward". Assim, inicialmente é feita a partição em blocos do sinal de voz original para a obtenção do sinal de referência ponderado, $d(n)$, do qual é subtraído o sinal sintetizado com ponderação, $\tilde{s}_w(n)$, obtendo-se o erro ponderado, $e_w(n)$. A partir do erro ponderado, $e_w(n)$, é obtido o sinal de excitação multipulso, $p(n)$, de acordo com o algoritmo descrito na seção 2.2.

O sinal multipulso, $p(n)$, excita um filtro de síntese a longo-prazo obtendo-se o sinal $\tilde{s}_l(n)$. O sinal $\tilde{s}_l(n)$ excita ao mesmo tempo um filtro de síntese LPC com ponderação e um outro sem ponderação. O sinal obtido através do filtro de síntese LPC com ponderação, $\tilde{s}_w(n)$, é usado para a geração do próximo

bloco de sinal de excitação multipulso, $p(n)$. O sinal obtido a partir do filtro de síntese LPC sem ponderação, $\tilde{s}(n)$, é utilizado para a extração dos coeficientes $a(k)$, $k = 1, \dots, p$ do filtro de síntese LPC de ordem p e para o cálculo do período de "pitch" e ganho do filtro de síntese a longo-prazo.

Na análise LPC, é utilizado o método de autocorrelação para a extração dos coeficientes do filtro de síntese LPC, através da formação de blocos sucessivos com superposição à esquerda e janela de Hamming do sinal sintetizado, $\tilde{s}(n)$. Com relação ao comprimento dos blocos, o comprimento de bloco, L , para a análise LPC está relacionado com o comprimento de bloco N , para a geração da excitação multipulso, pela seguinte igualdade:

$$L = \frac{3}{2} N \quad (14)$$

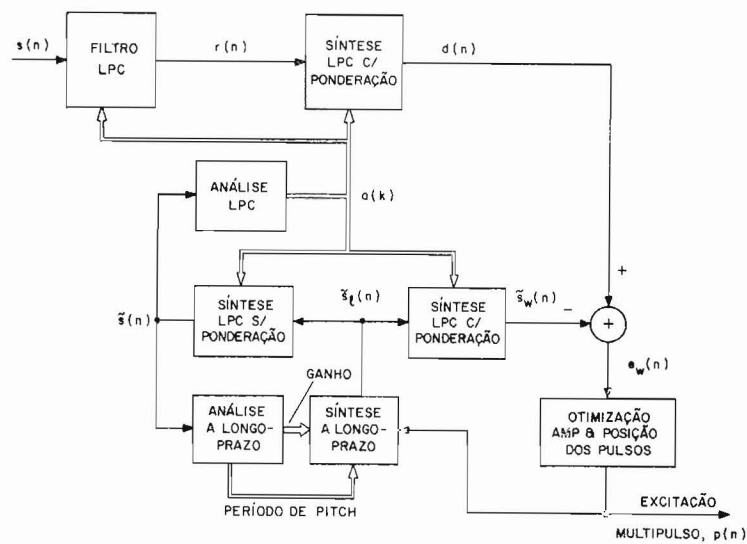
Analogamente, para a extração do período de "pitch" e ganho do filtro de síntese a longo-prazo, é feita a formação de blocos sucessivos com superposição e de comprimento maior que o período de "pitch" do sinal de voz sintetizado, $\tilde{s}(n)$. O algoritmo implementado para a detecção do período de "pitch" é baseado na utilização da função conhecida como AMDF ("Average Magnitude Difference Function") [15].

3.2. Decodificador

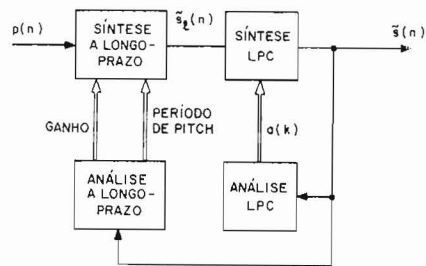
Um diagrama em blocos do decodificador do Codec MPE-LPC-BA está mostrado na **Figura 4b**. O sinal de excitação multipulso, $p(n)$, excita primeiramente o filtro de síntese a longo-prazo. O sinal obtido na saída do filtro de síntese a longo-prazo, $\tilde{s}_l(n)$, excita o filtro de síntese LPC, obtendo-se o sinal sintetizado, $\tilde{s}(n)$. Os coeficientes de ambos os filtros de síntese são atualizados do mesmo modo que no codificador.

4. Codificação da Posição dos Pulsos

Uma parte que merece bastante atenção num codificador baseado no modelo de excitação multipulso refere-se à representação dos N_p pulsos em um bloco de N amostras, pois existem no total $\binom{N}{N_p}$ possíveis padrões.



a) CODIFICADOR



b) DECODIFICADOR

Figura 4. Configuração MPE-LPC-BA

este trabalho é utilizado um algoritmo de codificação dos pulsos baseado em um esquema combinacional que pode ser descrito da seguinte maneira [16]:

Inicialização do algoritmo ($l = 0$);

Forma-se um vetor binário de comprimento N , cujos elementos são 0 ou 1 conforme a ausência ou presença de um pulso, respectivamente;

Percorre-se o vetor binário da posição N para a posição 1, à procura da posição n do próximo 1 no vetor e computa-se o índice l como sendo:

$$l = l + \binom{n-1}{m}, \quad (15)$$

de m é o número de 1's no vetor que ainda não foram encontrados rescido de 1.

Se os N_p elementos foram encontrados, então termina-se o algoritmo, e o valor de l representa o código do padrão associado com os N_p pulsos. Em caso contrário, volta-se ao passo 3.

O problema inverso, que é determinar a posição dos pulsos a partir do código l , é facilmente resolvido comparando-se o valor do código l com os mesmos valores combinacionais $\binom{n-1}{m}$.

O procedimento para a determinação do código l do padrão associado com N_p pulsos, bem como o procedimento inverso (i.e., a determinação da posição dos pulsos a partir do código l), utilizam valores numéricos representados pelo inteiro maior ou igual a $\log_2 \binom{N}{N_p}$ bits.

Quantização/Quantização Inversa da Amplitude dos Pulsos

As amplitudes dos pulsos da excitação multipulso, $p(n)$, são quantizadas de bloco em bloco conforme a técnica APCM (Adaptive Pulse Code Modulation). Em cada bloco, um conjunto de amostras A_i , ($i = 1, \dots, N_p$), é codificado do

seguinte modo: primeiramente, as amplitudes são comparadas entre si de modo a determinar a amplitude máxima em valor absoluto, AM:

$$AM = \max |A_i|, \quad 1 \leq i \leq N_p \quad (16)$$

O valor obtido para AM é, então, codificado com 06 (seis) bits conforme mostrado na **Tabela 1**. Esta tabela foi construída considerando que as amostras do sinal de entrada (PCM linearizado) são representadas com 13 bits (-4096 a 4095).

Níveis de decisão	Degrau de quantização	AM codificado (AM cod)	AM decodificado (AM dec)
0 15 16 31 32 47 . . 240 255	15	0 1 2 . . 15	15 31 47 . . 255
256 287 288 319 . . 736 767	31	16 17 . . 31	287 319 . . 767
768 831 832 895 . . 1728 ... 1791	63	32 33 . . 47	831 895 . . 1791
1792 ... 1919 1920 ... 2047 . . 3712 ... 4096	127	48 49 . . 63	1919 2047 . . 3839

Tabela 1. Trabalho de Quantização da Amplitude Máxima (AM)

Em seguida, as amplitudes A_i são normalizados em relação ao valor de AM decodificado, AM_{dec} :

$$AN_i = \frac{A_i}{AM_{dec}} \quad (17)$$

Finalmente, os valores de AN_i são quantizados com 3 bits conforme mostrado na **Tabela 2**. Assim, os valores das amplitudes decodificadas, A_{dec} , são obtidos da seguinte forma:

$$A_{dec_i} = AN_{dec_i} \times AM_{dec} \quad (18)$$

Níveis de decisão	AN_i codificado (AN_{cod})	AN_i decodificado (AN_{dec})
$[-1 \dots -0.75)$	0 = 000	-0.875
$[-0.75 \dots -0.50)$	1 = 001	-0.625
$[-0.50 \dots -0.25)$	2 = 010	-0.375
$[-0.25 \dots -0.0)$	3 = 011	-0.125
$[0.0 \dots 0.25)$	4 = 100	0.125
$[0.25 \dots 0.50)$	5 = 101	0.375
$[0.50 \dots 0.75)$	6 = 110	0.625
$[0.75 \dots 1.00)$	7 = 111	0.875

Tabela 2: Tabela de Quantização das Amplitudes Normalizadas (AN_i)

6. Resultados

O algoritmo MPE-LPC-BA foi implementado em linguagem de alto nível (Fortran) para a taxa de 16 kbit/s e avaliado usando-se medidas de Razão Sinal/Ruído Segmental (RSR seg) e Distância Cepstral (DC).

A razão sinal/ruído segmental é definida como:

$$RSR_{seg} = \frac{1}{N_b} \sum_{i=1}^{N_b} RSR(i), \quad (19)$$

onde: N_b = número total de blocos utilizados e $RSR(i)$ é a razão sinal/ruído no bloco i , definida como:

$$RSR(i) = 10 \log_{10} \frac{\sum_{j=M(i-1)+1}^{M \cdot i} x^2(j)}{\sum_{j=M(i-1)+1}^{M \cdot i} [x(j) - y(j)]^2} \quad (20)$$

onde: $x(j)$ e $y(j)$ são, respectivamente, o sinal de entrada e saída do codec, e M é o número de amostras por bloco. Neste trabalho foi utilizado $M = 128$.

A Distância Cepstral é definida como:

$$DC = \frac{10}{\ln 10} \left\{ [C_x(0) - C_y(0)]^2 + 2 \sum_{k=1}^P [C_x - C_y]^2 \right\}^{1/2} \quad (21)$$

onde: P é a ordem do filtro utilizado para o cálculo dos coeficientes cepstrais; C_x e C_y são os coeficientes cepstrais LPC obtidas para o sinal de voz original e sintetizado, respectivamente. Em paralelo a estas medidas, foram também realizados testes subjetivos informais para uma melhor avaliação da qualidade do sinal de voz sintetizado.

Nas **Tabelas 3 e 4** são apresentadas, respectivamente, a RSRseg e DC em função do atraso de algoritmo de codificação. Para efeito de comparação, na última coluna de cada tabela têm-se os resultados obtidos para um codec MPE-LPC utilizando uma estrutura "forward" com um atraso de algoritmo em torno de 25 e as seguintes características de alocação de bits por bloco de 160 amostras (20ms) [17,18]:

8 x Coeficientes LPC (Log-Área)	36 bits
1 x Ganho do preditor de longo-prazo	2 bits
Período de "Pitch"	7 bits
Amplitudes (8 pulsos/sub-bloco de 40 amostras e codificação APCM)	4X8X3 = 96 bits
Máxima amplitude do sub-bloco de 40 amostras	4X6 = 24 bits
Posição dos pulsos	27X4 = 108 bits
 TOTAL	 273 bits
TAXA DE BITS = 273 bits/bloco X 50 blocos (de 20ms cada)/s = 13.65 kbit/s.	

Com relação ao codec MPE-LPC-BA, os resultados foram obtidos utilizando-se um filtro de análise e síntese LPC de ordem $p = 12$, coeficiente de ponderação $\mu = 0.9$ e incluindo um preditor a longo-prazo de ordem 1.

Em termos de RS Rseg, observa-se um melhor desempenho do MPE-LPC-BA para um atraso em torno de 3,5ms. Entretanto, em termos de DC, que reflete melhor a qualidade subjetiva do codec, tem-se um melhor desempenho para um atraso em torno de 6,0ms, sendo a coerência deste resultado confirmada através de testes subjetivos informais. De fato, através de testes subjetivos informais tem-se constatado que o codec MPE-LPC-BA à taxa de 16 kbit/s e atraso de 6ms apresenta uma qualidade de sinal de voz equivalente ao Codec MPE-LPC à 13,65 kbit/s, o qual fornece uma qualidade de sinal de voz adequada para aplicação em redes telefônicas.

Comprimento do bloco (N)	Atraso (ms)	Nº de pulsos por bloco (N_p)	RSRseg (dB)	
			MPE-LPC-BA	MPE-LPC (13.65 kbit/s)
48	6,0	16	13,34	15,14
39	4,87	13	13,49	
28	3,50	9	14,06	
20	2,50	6	13,59	
14	1,75	4	12,64	

Tabela 3. Razão Sinal/Ruído Segmental x Atraso

Comprimento do bloco (N)	Atraso (ms)	Nº de pulsos por bloco (N_p)	DC (dB)	
			MPE-LPC-BA	MPE-LPC (13.65 kbit/s)
48	6,0	16	2,42	2,39
39	4,87	13	2,44	
28	3,50	9	2,64	
20	2,50	6	2,92	
14	1,75	4	3,36	

Tabela 4. Distância Cepstral x Atraso

7. Conclusões

Neste trabalho foi apresentado um algoritmo de codificação de voz MPE-LPC de baixo atraso para aplicações de transmissão de voz à taxa de 16 kbit/s sem utilização de cancelamento de eco.

O algoritmo é baseado na introdução de uma estrutura "backward" em blocos para a adaptação dos coeficientes dos preditores de curto-prazo e longo-prazo.

De acordo com os resultados obtidos, o algoritmo apresenta uma boa qualidade de voz à taxa de 16 kbit/s e atraso entre 3,5 e 6,0ms. Considerando-se que estes resultados foram obtidos sem muitos refinamentos no algoritmo, espera-se que seja possível manter esta qualidade mesmo a taxas menores que 16 kbit/s. Assim, técnicas de codificação de canal poderão ser utilizadas para aumentar a robustez contra erros no canal. Este ponto não foi analisado até o presente momento do trabalho, bem como refinamentos no algoritmo não foram investigados em detalhe.

Referências

- [1] B.S. Atal e J.R. Remde, "A new model of LPC excitation for producing natural-sounding speech at low bit rates", Proc. ICASSP, pag. 614-617, Paris, 1982.
- [2] M. Berouti, H. Garten, P. Kabal e P. Mermelstein, "Efficient computation and encoding of the multipulse excitation for LPC", Proc. ICASSP, pag. 10.1.1-10.1.4, San Diego, CA, 1984.
- [3] S. Ono, T. Araseki e K. Ozawa, "Improved pulse search algorithm for multi-pulse excited speech coder", Proc. IEEE Global Telecommun. Conf., pag. 287-291, 1984.
- [4] S. Singhal e B.S. Atal, "Improving performance of multi-pulse coders at low bit rates", Proc. ICASSP, pag. 1.3.1-1.3.4, San Diego, CA, 1984.
- [5] P. Kroon e E.F. Deprettere, "Experimental evaluation of different approaches to the multi-pulse coder", Proc. ICASSP, pag. 10.4.1-10.4.4, San Diego, CA, 1984.

- [6] J.P. Lefevre e O. Passien, "Efficient algorithms for obtaining multipulse excitation for LPC coders", Proc. ICASSP, pag. 957-960, Tampa, FL, 1985.
- [7] Y. Wabe, S. Tanaka, K. Ozawa e T. Araseki, "A multi-pulse LPC codec using digital signal processors". Proc. ICASSP, pag. 1429-1432, Tampa, FL, 1985.
- [8] E.F. Deprettere e P. Kroon, "Regular excitation reduction for effective and efficient LP-coding of speech", Proc. ICASSP, pag. 965-968, Tampa, FL, 1985.
- [9] H. Koyama e A. Gersho, "Fully vector-quantized multipulse LPC at 4800 bps", Proc. ICASSP, pag. 445-448, Tokyo, Japão, 1986.
- [10] S. Singhal, "Reducing computation in optimal amplitude multipulse coders", Proc. ICASSP, pag. 2363-2366, Tokyo, Japão, 1986.
- [11] P. Kroon, E. F. Deprettere e R. Sluyter, "Regular-pulse excitation – a novel approach to effective and efficient multipulse coding of speech", IEEE Trans. Acoust. Speech, Signal Processing, vol. ASSP-34, pag. 1054-1063, 1986.
- [12] K. Ozawa, S. Ono e T. Araseki, "A study on pulse search algorithm for multipulse excited speech coder realization", IEEE Journal on Selected Areas in Communications, vol. SAC-4, nº 1, Janeiro de 1986.
- [13] S. Singhal e B.S. Atal, "Amplitude optimization and pitch prediction in multipulse coders", IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, nº 3, Março de 1989.
- [14] L.R. Rabiner e R.W. Schafer, "Digital Processing of Speech Signals", pag. 407-411, Prentice-Hall, Englewood Cliffs, N.Y., (1978).
- [15] L.R. Rabiner e R.W. Schafer, "Digital Processing of Speech Signals", pag. 149-150, Prentice-Hall, Englewood Cliffs, N.Y., (1978).
- [16] R. Montagna e M. Omologo, "Some Results on Multipulse Linear Predictive Coding", Proc. IEEE Global Telecommunications Conference, pag. 802-806, Houston, 1986.
- [17] J.S. Yamamoto, "Codec de Referência a 8-16 kbit/s sem quantização", Relatório Técnico PDS-27, CPqD/TELEBRÁS, Abril de 1988.
- [18] E. Martino, "Half-Rate codec", Relatório Técnico nº 1 da LOU/DBP-TB/001/88, Fevereiro de 1989.

JOSÉ SINDI YAMAMOTO, graduou-se em Engenharia Elétrica em 1979 pela Universidade de Brasília — UnB, e obteve o grau de Mestre em Ciências em Engenharia Elétrica em 1982 pela Universidade Estadual de Campinas UNICAMP. Tem atuado desde 1980 em projetos de P&D em transmissão digital, inicialmente na UNICAMP e, a partir de 1984, no Centro de Pesquisa e Desenvolvimento da TELEBRÁS —CPqD. De 1985 a 1986 desempenhou atividades na área de codificação de voz, como pesquisador visitante, no Centro Studi e Laboratori Telecomunicazioni SpA CSELT, Turin, Itália. Atualmente, coordena as atividades de P&D do CPqD/TELEBRÁS na área de processamento digital de voz.

FÁBIO VIOLARO, graduou-se em Engenharia Elétrica em 1973, obteve o grau de Mestre em 1975 e o grau de Doutor em 1980, todos pela Faculdade de Engenharia Elétrica da Universidade de Campinas UNICAMP. É professor do Departamento de Comunicações desta faculdade e sua área de interesse se concentra em processamento digital de voz: codificação, reconhecimento e síntese.