

# Comportamento de Codificadores Diferenciais de Voz com Preditores Fixos em Presença de Erros no Canal

Abraham Alcaim

O desempenho objetivo (medido em termos da relação sinal-ruído segmentada média) de codificadores diferenciais de voz (delta e ADPCM) operando a uma taxa de transmissão de 32 kbit/s em canais ruidosos é apresentado. Os codificadores utilizam um preditor linear fixo de 1ª ou 2ª ordem e um quantizador ajustável ao longo do tempo. Para os sistemas delta são considerados os seguintes algoritmos de adaptação do quantizador: 1) Jayant, 2) Castellino e Scagliola, e 3) CVSD. Além disso, é analisado o emprego de um mecanismo de dissipação do efeito de erros em sistemas delta com adaptação instantânea. No caso dos sistemas ADPCM os seguintes algoritmos de adaptação do quantizador são analisados: 1) Jayant, 2) Jayant com mecanismo de dissipação do efeito de erros, 3) energia segmentar sem transmissão explícita do parâmetro de ajuste do quantizador, e 4) algoritmo de adaptação silábica. Tanto os codificadores delta como os codificadores ADPCM que utilizam algoritmos de adaptação instantânea necessitam de um mecanismo de dissipação do efeito de erros para operar em canais ruidosos. Nesse caso, esses codificadores superam em desempenho aqueles que usam adaptação silábica.

## 1. Introdução

Sistemas de codificação diferencial, delta e ADPCM ("Adaptive Differential Pulse Code Modulation"), utilizando preditores fixos e quantizadores ajustáveis ao longo do tempo são os mais simples para codificação digital de voz a 32 kbit/s. O desempenho de vários métodos de codificação diferencial foi avaliado em um estudo recente [1]. No presente trabalho o comportamento desses métodos é analisado em presença de erros no canal de comunicações. Além disso e com particular interesse são analisados aqui codificadores delta e ADPCM menos sensíveis a erros no canal que aqueles considerados no trabalho acima referido.

O diagrama em blocos de um sistema ADPCM é mostrado na Fig. 1. Para cada amostra de voz  $s_i$  o quantizador produz uma saída  $z_i$  de acordo com as seguintes equações

$$\text{Erro de predição: } e_i = s_i - \hat{s}_i \quad (1)$$

O autor é Professor do Centro de Estudos em Telecomunicações da PUC/RIO, 22453, Rio de Janeiro, RJ.

$$\text{Sinal estimado: } \hat{s}_i = \sum_{j=1}^p a_j \cdot r_{i-j} \quad (2)$$

$$\text{Sinal reconstruído localmente: } r_i = \hat{s}_i + z_i \quad (3)$$

Saída do quantizador:

$$z_i = \begin{cases} \sigma_i y_j, & 0 \leq \sigma_i x_j \leq e_i < \sigma_i x_{j+1} \\ -\sigma_i y_j, & -\sigma_i x_{j+1} \leq e_i < -\sigma_i x_j \leq 0 \end{cases} \quad (4)$$

onde os valores  $\{x_j; j=1, \dots, N\}$  e  $\{y_j; j=1, \dots, N\}$  que definem o quantizador são aqueles que minimizam o erro médio quadrático de quantização para uma dada distribuição  $p(e)$  de entrada do quantizador [2] - [4].

A cada instante  $i$  o algoritmo de adaptação fornece uma estimativa  $\sigma_i$  do desvio padrão da entrada  $e_i$  do quantizador, cujos níveis são adaptados por essa estimativa conforme indicado em (4).

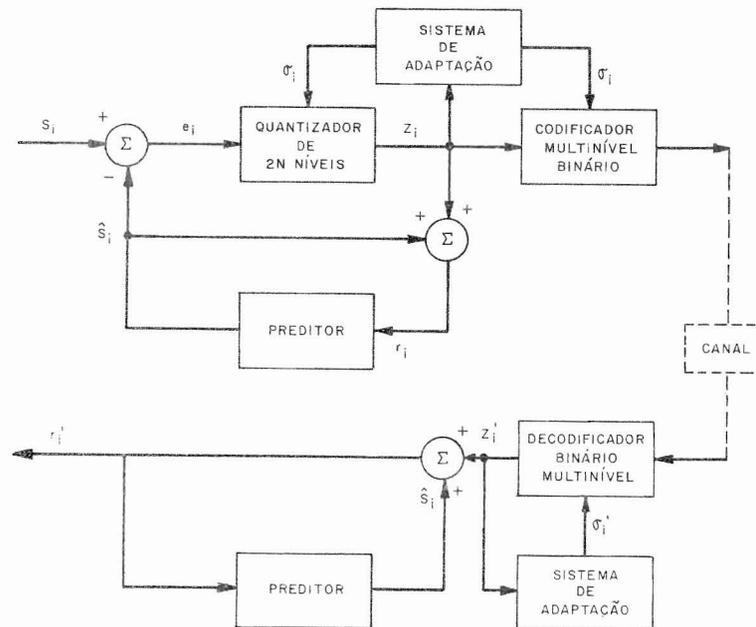


Figura 1. Sistema ADPCM.

O codificador delta é basicamente um ADPCM que usa apenas dois níveis de quantização e uma taxa de amostragem bem mais elevada que a taxa de Nyquist, para compensar de certa forma a quantização grosseira de 1 bit por amostra. O diagrama em blocos do sistema delta é mostrado na Fig. 2. Para cada amostra  $s_i$  do sinal de voz o bit de informação  $b_i$  é gerado através de (1) e (2), juntamente com

$$\text{Sinal reconstruído localmente: } r_i = \hat{s}_i + d_i \quad (5)$$

$$\text{Passo: } d_i = b_i \cdot |d_i| \quad (6)$$

$$\text{Bit de informação: } b_i = \begin{cases} +1, & e_i \geq 0 \\ -1, & e_i < 0 \end{cases} \quad (7)$$

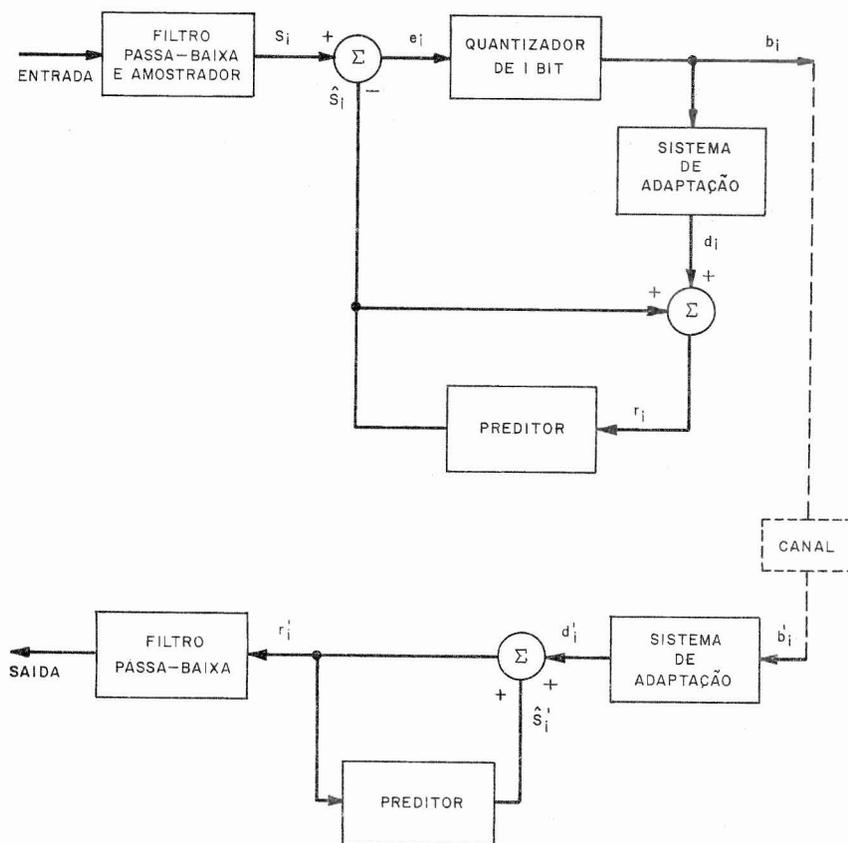


Figura 2. Sistema delta adaptativo.

A magnitude do passo a cada instante,  $d_i$ , depende da lógica utilizada pelo sistema de adaptação.

Devido ao ruído no canal de comunicações a seqüência de bits recebida em geral difere daquela transmitida. A conseqüência disso é obviamente um sinal decodificado  $r'_i$  diferente daquele reconstruído localmente no codificador ( $r_i$ ).

A medida de desempenho utilizada neste trabalho é a relação sinal-ruído segmentada média (RSS) [1] devido à sua forte correlação com medidas subjetivas [5]. Os desempenhos foram obtidos através de simulação dos codificadores, do canal ruidoso e dos decodificadores em um computador digital CYBER 170/835 e foram utilizadas três sentenças faladas por um homem adulto [1]. Os sinais de voz foram filtrados em 3300 Hz, amostrados a uma taxa de 8000 amostras/s para os sistemas ADPCM e de 32000 amostras/s para os sistemas delta e em seguida digitalizados em 12 bits/amostra.

Neste trabalho foram considerados preditores fixos de 1ª e 2ª ordem, cujos coeficientes foram obtidos através da minimização do erro médio quadrático de predição [6]. O coeficiente  $a_1$  do preditor de 1ª ordem é 0,98 para os sistemas delta e 0,7901 para os sistemas ADPCM. Os coeficientes  $a_1$  e  $a_2$  do preditor de 2ª ordem são dados, respectivamente, por 1,916 e - 0,946 para os sistemas delta e 1,3066 e - 0,6537 para os sistemas ADPCM. Os valores de  $\sigma_i^2$  e de  $|d_i|^2$  foram limitados a um máximo e mínimo, cobrindo uma faixa de 55 dB.

É importante ressaltar que não foram considerados canais ruidosos com taxas de erro de bit inferiores a  $10^{-4}$ , pois verificou-se que os sistemas mais robustos só apresentam deterioração para taxas mais elevadas. Assim, não há significado em considerar, dentro dessa análise comparativa, taxas de erro entre 0 e  $10^{-4}$ . Isso ficará claro ao longo do texto.

Na Seção 2 deste trabalho são examinados os codificadores delta e na Seção 3 os codificadores ADPCM. A parte final é dedicada às principais conclusões relativas ao comportamento dos diversos métodos de codificação diferencial de voz aqui apresentados.

## 2. Comportamento de Codificadores Delta em Canal Ruidoso

### 2.1. Sistema de Adaptação com Memória de um Bit de Jayant

O método de Jayant [7] é um dos métodos mais simples para codificação digital de sinais de voz. O passo no instante  $i$  é obtido a partir do passo no instante  $i-1$  e dos dois bits de informação mais recentes, através da equação

$$d_i = \begin{cases} P \cdot |d_{i-1}|, & b_i = b_{i-1} \\ Q \cdot |d_{i-1}|, & b_i \neq b_{i-1} \end{cases} \quad (8)$$

Os multiplicadores  $P = 1,5$  e  $Q = 0,66$  são obtidos através de maximização da relação sinal-ruído de quantização, utilizando um procedimento de busca.

Deve-se ressaltar aqui que o emprego de preditores de ordem superior à 1ª torna esse método instável, como já verificado em trabalhos anteriores [8]. Os resultados obtidos com preditor de 1ª ordem serão apresentados posteriormente em conjunto com os resultados obtidos quando se usa um mecanismo de dissipação de erro.

## 2.2. Sistema de Adaptação de Castellino e Scagliola

O sistema de adaptação do CSELT ("Centro Studi e Laboratori Telecomunicazioni", Torino, Itália) [9] é definido por

$$|d_i| = L_i |d_{i-1}| \quad (9)$$

onde  $L_i = A_k + 1$  se os  $(M + 1)$  bits mais recentes correspondem à seqüência  $k$ , sendo  $k$  o número decimal correspondente ao número binário representado pela seqüência de bits  $\{b_{i-M}, \dots, b_{i-1}, b_i\}$ . Note-se que para se obter essa representação é necessário considerar o bit  $-1$  em (7) como 0. Os multiplicadores de adaptação são obtidos a partir da observação de que, idealmente, o multiplicador  $A_k$  associado a uma dada seqüência de bits  $\{b_{i-M}, \dots, b_{i-1}, b_i\}$  deve ser tal que o novo passo é igual ao erro de predição, ou seja,

$$|d_i| = |e_i| \quad (10)$$

Um algoritmo de aprendizagem para determinação dos multiplicadores ótimos é descrito em [8] e [9].

A Tabela 1 mostra o efeito da memória  $M$  sobre o desempenho (RSS) desse sistema em canal não-ruído e em canal com taxa de erro de bit (TEB) igual a  $10^{-4}$ ,  $10^{-3}$ ,  $10^{-2}$  e  $10^{-1}$ . Um preditor de 2ª ordem foi utilizado na obtenção desses resultados. Observa-se dessa tabela que em canais com  $TEB = 10^{-4}$  o desempenho é melhor quando se usa memória de adaptação mais curta.

M	TAXA DE ERRO DE BIT				
	0	10 <sup>-4</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-1</sup>
2	20,6	17,4	1,6	- 17,2	- 40,6
3	20,8	16,7	2,7	- 15,9	- 39,5
4	20,8	16,7	2,6	- 16,5	- 39,5

Tabela 1. RSS em dB para o codificador delta do CSELT usando preditor de 2ª ordem.

Uma explicação para isso é que quanto mais longa a memória mais freqüentemente se terá uma sequência de bits, indicadora do multiplicador, com erro. É fácil ver que para um sistema com memória M, um erro de um bit provoca pelo menos M + 1 multiplicadores incorretos.

A Tabela 2 apresenta o desempenho do sistema do CSELT, com memória de 2 bits, usando preditor de 1ª ordem. Comparando com a Tabela 1, observa-se que os resultados obtidos em canal ruidoso são superiores quando se usa preditor de 1ª ordem. Isso se acentua mais para canais com taxas de erro de bit mais elevadas. Uma possível razão para isso é que o emprego de um preditor de 2ª ordem implica em predições mais incorretas quando o canal é muito ruidoso. Neste caso há acúmulo de dados incorretos nos quais as predições são baseadas.

TEB	0	10 <sup>-4</sup>	10 <sup>-3</sup>	10 <sup>-2</sup>	10 <sup>-1</sup>
RSS	19,8	17,7	11,7	2,9	- 3,7

Tabela 2. RSS em dB para o codificador delta do CSELT com memória de 2 bits e preditor de 1ª ordem.

### 2.3. Sistema de Adaptação Silábica

Em sistemas delta com adaptação silábica [8], [10] – [16] o passo  $|d_i|$  é ajustado lentamente com o tempo, sendo, conseqüentemente, menos sensível a erros no canal. Um algoritmo representativo é o CVSD ("Continuously Variable Slope Delta"), que consiste de uma lógica de adaptação seguida por um filtro passa-baixa simples. O algoritmo que será apresentado aqui é aquele que forneceu melhores resultados em uma análise recente [15]. Esse algoritmo é definido por

$$|d_i| = \beta |d_{i-1}| + K_i d_o \quad (11)$$

$$\text{onde } K_i = \begin{cases} 1 & , \left| \sum_{n=0}^3 b_{i-n} \right| = 4 \\ 0 & , \text{ caso contrário} \end{cases} \quad (12)$$

$$d_o = (1 - \beta) \left| d_i \right|_{\max} \quad (13)$$

$\left| d_i \right|_{\max}$  é a magnitude máxima permitida ao passo e  $\beta = 0,996$ .

A **Tabela 3** mostra o desempenho desse sistema quando são utilizados preditores de 1ª e 2ª ordem na malha de realimentação. Observa-se daí que assim como no sistema delta do CSELT, o uso de preditor de 1ª ordem no sistema CVSD fornece resultados mais atraentes que aqueles obtidos com preditor de 2ª ordem.

ORDEM DO PREDITOR	TAXA DE ERRO DE BIT				
	0	$10^{-4}$	$10^{-3}$	$10^{-2}$	$10^{-1}$
1ª	19,7	19,4	16,5	5,9	- 8,8
2ª	19,9	19,4	14,5	0,8	- 16,7

**Tabela 3.** RSS em dB para o codificador CVSD.

#### 2.4. Codificadores Delta com Mecanismo de Dissipação de Erro

Os codificadores delta com memória de 1 bit de Jayant [7] e com memória de 2 bits do CSELT [9] são bons esquemas em canais de comunicação sem ruído. Codificadores desse tipo podem ser descritos através de (10).

É fácil ver que em sistemas desse tipo com memória de adaptação de M bits, a ocorrência de um único bit com erro causa o descasamento de M+1 multiplicadores (valores de  $L_j$ ) entre codificador e decodificador. Pode-se mostrar [17] ainda que se um único erro ocorreu no instante  $\ell$ , então em qualquer instante  $j > \ell + M - 1$  a versão no decodificador de  $d_j$  será dada por

$$d'_j = \left[ \begin{matrix} M \\ \pi \\ n=0 \end{matrix} \frac{L'_\ell + n}{L_{\ell+n}} \right] \cdot \left| d_j \right| \quad (14)$$

onde  $L'_j$  é a versão no decodificador do multiplicador  $L_j$ .

O significado de (14) é que um único erro de bit causa um descasamento multiplicativo entre codificador e decodificador que pode persistir indefinidamente. Por esse motivo

foi examinada recentemente [17] uma modificação desses codificadores de modo que o efeito desse descasamento sofra um decaimento com o tempo. Essa modificação foi proposta por Goodman e Wilkinson [18] e usada anteriormente com sistemas ADPCM. No caso dos codificadores delta esse mecanismo de dissipação de erro pode ser descrito pela seguinte regra de adaptação do passo:

$$|d_i| = L_i \cdot |d_{i-1}|^\lambda, \quad 0 < \lambda < 1 \quad (15)$$

Nesse caso, se um único erro ocorrer no instante  $\ell$ , então em qualquer instante  $j > \ell + M - 1$  a versão no decodificador de  $|d_j|$  será dada por

$$|d'_j| = \left\{ \begin{array}{c} M \\ \pi \\ n=0 \end{array} \left[ \frac{L' \cdot \ell + n}{L \cdot \ell + n} \right]^\lambda \right\}^{\lambda^{M-n}} \cdot |d_j|^{\lambda^{(i-\ell-M)}} \quad (16)$$

Dessa forma, o descasamento devido a cada erro digital que ocorra na transmissão decairá exponencialmente com o tempo. O valor de  $\lambda$  é um compromisso entre a taxa de dissipação de erro e a faixa dinâmica.

A Tabela 4 mostra o desempenho do sistema de Jayant com mecanismo de dissipação de erro e preditor de 1ª ordem para diversos valores de  $\lambda$ . O valor  $\lambda = 1$  corresponde obviamente ao sistema original. Dessa tabela pode-se ver que quando a taxa de erro de bit aumenta, a taxa de degradação de desempenho é menor para o sistema com  $\lambda = 15/16$  e maior para o sistema original ( $\lambda = 1$ ). Entretanto, como a faixa dinâmica é maior quando  $\lambda$  se encontra mais próximo de 1, e tendo em vista os resultados da Tabela 4, um valor de  $\lambda = 63/64$  parece fornecer os resultados mais satisfatórios.

$\lambda$	TAXA DE ERRO DE BIT				
	0	$10^{-4}$	$10^{-3}$	$10^{-2}$	$10^{-1}$
15/16	15,3	15,1	13,4	6,7	- 1,5
31/32	18,2	17,8	15,3	5,0	- 4,7
63/64	19,5	19,2	15,8	3,9	- 8,3
1	20,3	18,9	11,2	- 2,0	- 14,4

Tabela 4. RSS em dB para o codificador delta de Jayant com mecanismo de dissipação de erro e preditor de 1ª ordem.

O desempenho do codificador delta do CSELT com memória de 2 bits ( $M = 2$ ) e mecanismo de dissipação de erro com  $\lambda = 63/64$  foi então avaliado para canais com taxa de erro de bit de  $0$ ,  $10^{-4}$ ,  $10^{-3}$ ,  $10^{-2}$  e  $10^{-1}$ . Os resultados obtidos são mostrados na **Tabela 5**.

ORDEM DO PREDITOR	TAXA DE ERRO DE BIT				
	0	$10^{-4}$	$10^{-3}$	$10^{-2}$	$10^{-1}$
1ª	19,5	19,2	16,3	6,1	- 2,8
2ª	20,6	20,1	15,5	0,7	- 23,7

**Tabela 5.** RSS em dB para o codificador delta do CSELT com memória de 2 bits, usando mecanismo de dissipação de erro com  $\lambda = 63/64$ .

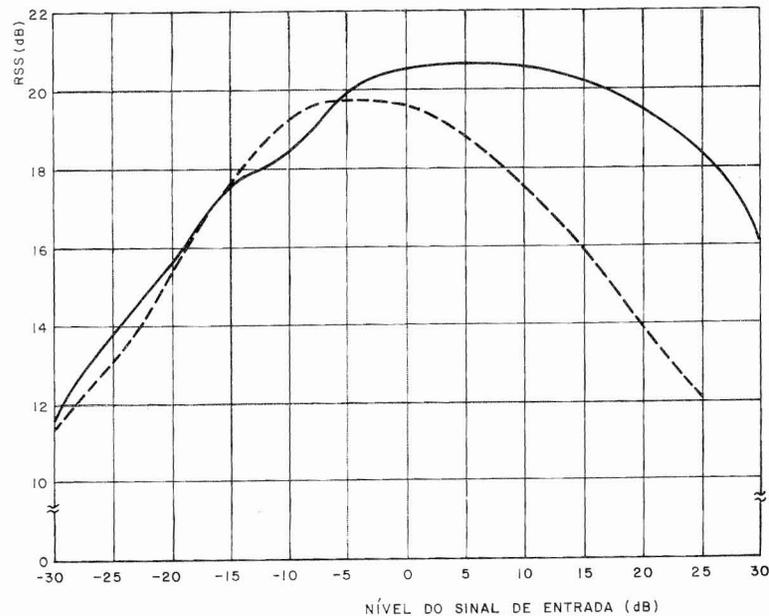
Comparando-se os resultados apresentados na **Tabela 5** com aqueles obtidos para os outros codificadores, observa-se que o sistema delta do CSELT com mecanismo de dissipação de erro e preditor de 2ª ordem é o que melhor se comporta tanto em canal ideal como em canal com  $TEB = 10^{-4}$ . Entretanto, em canais muito ruidosos ( $TEB \geq 10^{-3}$ ) esse sistema é inferior ao codificador CVSD, quando esse último emprega um preditor de 1ª ordem. Porém, quando ambas as técnicas usam um preditor de 1ª ordem os desempenhos são praticamente iguais, mesmo para canais muito ruidosos.

Curvas de RSS versus nível de voz são comparadas na **Fig. 3** para os melhores sistemas delta, ou seja, CSELT com mecanismo de dissipação de erro e preditor de 2ª ordem e CVSD com preditor de 1ª ordem. Observa-se dessa figura que o esquema do CSELT apresenta uma faixa dinâmica maior que 40 dB enquanto o CVSD não alcança 30 dB. Isso representa uma vantagem adicional do codificador delta do CSELT quando um mecanismo de dissipação de erro é incorporado.

### 3. Comportamento de Codificadores ADPCM em Canal Ruidoso

#### 3.1. Comentários sobre o Tipo de Código Utilizado

Em sistemas ADPCM a cada saída  $z_i$  do quantizador é associada uma palavra-código de  $n$  bits. Isso é feito pelo codificador multinível-binário mostrado no diagrama em blocos da **Fig. 1**. No caso da taxa de bits aqui considerada de 32 kbit/s o quantizador possui 16 níveis sendo, portanto,  $n$  igual a 4.



**Figura 3.** Desempenho do sistema delta do CSELT com memória de 2 bits usando mecanismo de dissipação de erro com  $\lambda = 63/64$  e preditor de 2ª ordem (—) e do sistema CVSD usando preditor de 1ª ordem (---).

O desempenho dos codificadores em canal ruidoso depende, obviamente, da forma com que é feita a associação dos  $2N$  níveis de quantização  $v_i$  às palavras-código, ou seja, do código utilizado. Aqui foram considerados, em princípio, dois tipos de código: binário simétrico (CBS) e binário natural (CBN). Esses códigos são ilustrados na Fig. 4 para um quantizador de 3 bits.

Estudos anteriores [21] e simulações realizadas com codificadores ADPCM, cujos resultados não se achou de interesse incluir nesse trabalho, indicam uma forte preferência pelo código binário simétrico. Por essa razão, esse código foi utilizado em todas as simulações realizadas no presente trabalho.

### 3.2. Sistema de Adaptação com Memória de uma Palavra de Jayant

A estimativa  $\sigma_i$  do desvio padrão é obtida no sistema de Jayant [19], [20] a partir da saída anterior  $z_{i-1}$  do quantizador. A regra de adaptação é definida por

$$\sigma_i = [M(l_i)] \sigma_{i-1} \quad (17)$$

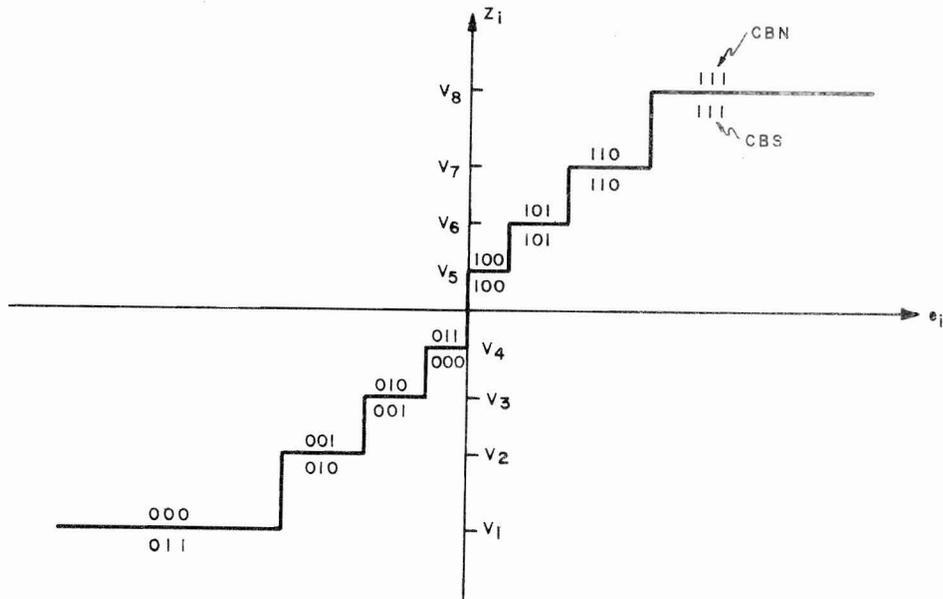


Figura 4. Código CBN e CBS para um quantizador de 3 bits.

onde  $l_i = k$  se  $z_{i-1} = \pm \sigma_{i-1} v_k$  para  $k = 1, \dots, N$ .

Em (17) os multiplicadores  $M(j)$ ,  $j = 1, \dots, N$ , são determinados através de um procedimento de busca, no sentido de maximizar a relação sinal-ruído de quantização. Os valores utilizados nesse trabalho são os mesmos obtidos por Jayant [19].

A Tabela 6 mostra o desempenho desse sistema para preditores de 1ª e 2ª ordem e quantizadores ótimos [2] - [4] projetados para entrada com distribuição de Gauss, Laplace e Gama.

ORDEM DO PREDITOR	QUANTIZADOR	TAXA DE ERRO DE BIT			
		0	$10^{-4}$	$10^{-3}$	$10^{-2}$
1ª	Gauss	22,7	21,6	11,1	- 9,1
	Laplace	22,5	21,3	9,5	- 10,4
	Gama	21,6	21,0	9,0	- 11,5
2ª	Gauss	23,6	23,1	12,2	- 11,1
	Laplace	23,3	22,8	11,6	- 12,0
	Gama	22,4	22,0	11,3	- 12,6

Tabela 6. RSS em dB para o codificador ADPCM de Jayant

Pode ser visto da **Tabela 6** que o quantizador projetado para uma distribuição de Gauss é o que fornece melhores resultados tanto para canal ideal como para canal ruidoso. Com relação ao preditor, observa-se que a utilização de um preditor de 2ª ordem é mais vantajosa que quando é empregado um preditor de 1ª ordem na malha de realimentação. Entretanto, os resultados mostram que esse codificador não pode ser usado para canais com taxas de erro de bit muito elevadas.

### 3.3. Codificador ADPCM de Jayant com Mecanismo de Dissipação de Erro

Uma das principais razões pelas quais o codificador ADPCM de Jayant não funciona em canais muito ruidosos é a propagação de erros. Para se verificar esse efeito considere-se que  $\sigma'$  e  $l'$  são as versões de  $\sigma$  e  $l$  no decodificador. Supondo que no instante  $\ell$  tem-se  $l_\ell = k$ , ou seja,  $M(l_\ell) = M(k)$  e que um erro na transmissão causa  $l'_\ell = j \neq k$ , isto é,  $M(l'_\ell) = M(j)$ , resulta que em qualquer instante  $i > \ell$

$$\sigma'_i = \frac{M(j)}{M(k)} \cdot \sigma_i \quad (18)$$

Isso significa que cada erro que ocorre provoca um descasamento multiplicativo do parâmetro de ajuste do quantizador entre codificador e decodificador. O que é mais importante ainda é que, em princípio, esse descasamento pode persistir indefinidamente.

Com o objetivo de diminuir esse efeito com o tempo, Goodman e Wilkinson [18] propuseram uma modificação da regra de adaptação dada por (17). Essa modificação pode ser escrita como

$$\sigma_i = [M(l_i)] \sigma_{i-1}^\lambda, \quad 0 < \lambda < 1 \quad (19)$$

É fácil mostrar que, nesse caso, o descasamento multiplicativo expresso por (18) passa a ser

$$\sigma'_i = \left[ \frac{M(j)}{M(k)} \right]^\lambda \cdot \sigma_i \quad (20)$$

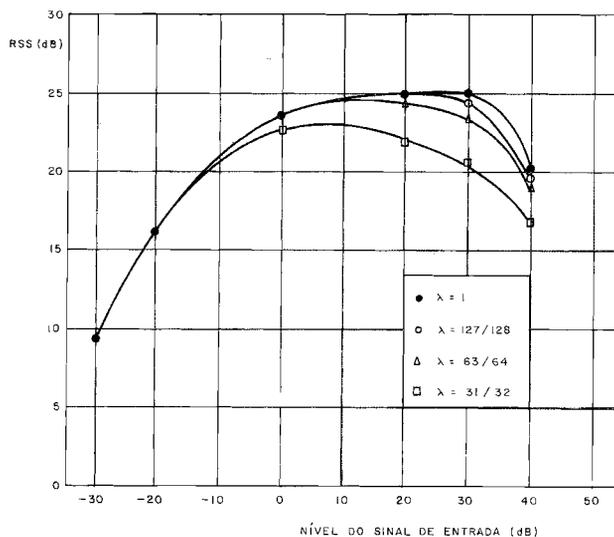
Utilizando-se o algoritmo de adaptação expresso por (19), observa-se que o descasamento multiplicativo devido a cada erro no canal decai exponencialmente com o tempo. A **Tabela 7** mostra o desempenho do sistema ADPCM utilizando esse mecanismo de dissipação de erro, um quantizador ótimo de Gauss e preditor de 2ª ordem. Quando  $\lambda$  é igual a 1 o sistema se reduz obviamente ao original de Jayant.

$\lambda$	TAXA DE ERRO DE BIT			
	0	$10^{-4}$	$10^{-3}$	$10^{-2}$
31/32	22,7	22,5	19,8	8,9
63/64	23,6	23,3	20,0	6,9
127/128	23,7	23,5	19,3	3,8
1	23,6	23,0	11,9	-10,9

**Tabela 7.** RSS em dB para o codificador ADPCM de Jayant com mecanismo de dissipação de erro e preditor de 2ª ordem ( $\lambda = 63/64$ ).

Da **Tabela 7** fica evidente a vantagem da utilização desse método em canal ruidoso. Por exemplo, para canais com taxa de erro de bit de  $10^{-3}$  consegue-se uma melhoria superior a 8 dB quando se usa o sistema com  $\lambda = 63/64$ .

A **Fig. 5** mostra o desempenho desse sistema em função do nível do sinal de voz de entrada. Observa-se que quanto menor  $\lambda$  menor é também a faixa dinâmica. Analisando-se esses resultados juntamente com aqueles da **Tabela 7**, pode-se ver que o sistema usando  $\lambda = 63/64$  é o que fornece resultados mais satisfatórios em termos de um compromisso entre taxa de dissipação de erro e faixa dinâmica.



**Figura 5.** Desempenho do sistema ADPCM de Jayant com mecanismo de dissipação de erro, quantizador ótimo de Gauss e preditor de 2ª ordem em função do nível do sinal de entrada.

### 3.4. Codificador ADPCM com Adaptação através da Energia Segmentar sem Transmissão do Parâmetro de Ajuste do Quantizador

Um método comum [14] de se estimar a variância de um sinal amostrado  $\{e_i\}$  é através de sua energia segmentar  $\sigma_i^2$  definida como a saída de um filtro passa-baixa com resposta impulsional  $\{h_n\}$ , quando a entrada é  $\{e_i^2\}$ , ou seja,

$$\sigma_i^2 = \sum_{n=-\infty}^{\infty} e_n^2 h_{i-n} \quad (21)$$

Se  $\{e_i\}$  for estacionário com média nula e variância  $\sigma_e^2$ , é fácil mostrar que o valor esperado de  $\sigma_i^2$  é proporcional a  $\sigma_e^2$ .

Uma maneira simples de utilizar essa técnica em codificadores ADPCM que não transmitem o parâmetro de ajuste  $\sigma_i$  do quantizador é substituir  $\{e_i\}$  por seu valor quantizado  $\{z_i\}$  em (21) e usar um filtro com resposta impulsional

$$h_n = \begin{cases} C/M, & 1 \leq n \leq M \\ 0, & \text{caso contrário} \end{cases} \quad (22)$$

onde C é uma constante. Nesse caso, as estimativas  $\sigma_i^2$  são dadas por

$$\sigma_i^2 = \frac{C}{M} \sum_{j=1}^M z_{i-j}^2 \quad (23)$$

Esse método foi estudado em canal ideal por Noll [22], para um quantizador de 3 bits, com valores de C e M dados por 1,8 e 3 respectivamente. O comportamento desse sistema para um quantizador de 4 bits é mostrado nas Tabelas 8 e 9 para diferentes valores de C e M. Esses resultados foram obtidos utilizando-se um quantizador ótimo de Gauss e um preditor de 1ª ordem. Os resultados da Tabela 8 se referem ao desempenho em canal ideal e os da Tabela 9 ao desempenho em canal com taxa de erro de bit igual a  $10^{-3}$ .

Da Tabela 8 observa-se que em canal ideal os melhores desempenhos são obtidos para  $M = 8$  (fixando-se C) e para  $C = 1,5$  (fixando-se M). Por outro lado, em canal ruidoso ( $TEB = 10^{-3}$ ) obtêm-se os melhores resultados para  $M = 16$  (fixando-se C) e para  $C = 0,5$  (fixando-se M). Um valor maior de M ser preferível em canal ruidoso é sem dúvida devido às variações lentas de  $\sigma_i^2$  neste caso, o que torna o sistema menos sensível a erros no canal.

Esse sistema mostrou um comportamento significativamente inferior àquele obtido com o codificador de Jayant com mecanismo de dissipação de erro. Por essa razão não se achou de interesse um estudo mais detalhado desse método.

C \ M	3	5	8	16
0,5	16,9	17,7	18,4	18,1
1,0	20,7	21,5	22,1	21,9
1,5	21,7	22,3	22,5	22,4
1,8	21,7	22,2	22,4	22,2
2,0	21,7	22,2	22,2	22,1

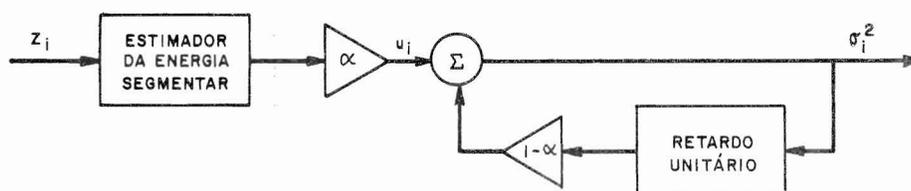
**Tabela 8.** RSS em dB para o codificador ADPCM com adaptação através da energia segmentar em canal ideal (preditor de 1ª ordem).

C \ M	3	5	8	16
0,5	14,5	14,2	15,5	15,9
1,0	13,6	12,1	13,6	15,2
1,5	11,5	10,3	11,1	13,0
1,8	11,4	9,9	10,6	11,8
2,0	10,5	9,9	8,9	10,7

**Tabela 9.** RSS em dB para o codificador ADPCM com adaptação através da energia segmentar em canal com taxa de erro de bit de  $10^{-3}$  (preditor de 1ª ordem).

### 3.5. Codificador ADPCM com Adaptação Silábica

Sistemas de adaptação silábica são, em geral, menos sensíveis a erros no canal. A razão disso é a variação lenta do parâmetro de ajuste  $\sigma_i$  do quantizador. Um método que será analisado aqui consiste de um estimador da energia segmentar a partir das últimas N saídas do quantizador, seguido por um filtro passa-baixa simples como mostrado na Fig. 6.



**Figura 6.** Sistema de adaptação silábica em codificação ADPCM.

A entrada  $u_i$  do filtro passa-baixa mostrado na Fig. 6 é obtida através da relação

$$u_i = \frac{\alpha}{N} \sum_{j=1}^N z_{i-j}^2 \quad (24)$$

e a estimativa  $\sigma_i^2$  é dada por

$$\sigma_i^2 = (1 - \alpha) \sigma_{i-1}^2 + u_i \quad (25)$$

Note-se que  $\sigma_i^2$  é a saída do filtro passa-baixa com resposta impulsional

$$h_n = \begin{cases} (1 - \alpha)^{n-1} & , n \geq 1 \\ 0 & , n < 1 \end{cases} \quad (26)$$

onde  $0 < \alpha < 1$  para garantir estabilidade.

Note-se que (25) também pode ser escrita como

$$\sigma_i^2 = \sigma_{i-1}^2 + \alpha \left[ \left( \frac{1}{N} \sum_{j=1}^N z_{i-j}^2 \right) - \sigma_{i-1}^2 \right] \quad (27)$$

de onde pode-se ver que  $\sigma_i^2$  é obtida através de um incremento a  $\sigma_{i-1}^2$  positivo ou negativo, proporcional à diferença entre uma estimativa de bloco (a partir das últimas N saídas) e  $\sigma_{i-1}^2$ .

Um método semelhante havia sido considerado anteriormente [24] para o caso em que  $N = 1$ , ou seja  $u_i = \alpha z_{i-1}^2$ . Aqui será analisado conjuntamente o efeito de  $\alpha$  e N sobre o desempenho do codificador ADPCM silábico em canal ruidoso. Em todas as simulações foi utilizado um quantizador ótimo de Gauss.

O comportamento do sistema ADPCM silábico com preditor de 1ª ordem em função de N e  $\alpha$  é mostrado na Tabela 10 para canal ideal e na Tabela 11 para canal com taxa de erro de bit de  $10^{-3}$ . Em canal ideal, fixando-se  $\alpha$ , o desempenho é deteriorado à medida que se aumenta o valor de N. Isso é devido a uma variação mais lenta de  $\sigma_i^2$  e, em consequência, a uma maior dificuldade de acompanhar variações rápidas da variância do erro, resultando em um excesso de sobrecarga. O mesmo raciocínio se aplica para valores de  $\alpha$  muito pequenos.

Por outro lado, em canais ruidosos, como pode ser visto da Tabela 11, esses comentários não são mais válidos, pois uma variação mais lenta de  $\sigma_i^2$  torna o sistema mais resistente aos erros de transmissão.

A partir dos resultados obtidos observa-se que  $\alpha = 1/32$  é um bom compromisso entre desempenho em canal ideal e sensibilidade a erros de transmissão. Utilizando-se esse valor de  $\alpha$  determinou-se o comportamento desse sistema para  $N = 1, 2, 4, 8$  e  $16$  em

canal ideal e canal com taxas de erro de bit de  $10^{-4}$ ,  $10^{-3}$  e  $10^{-2}$ . Os resultados são mostrados na **Tabela 12**, de onde se observa que o desempenho varia pouco com N. Isso é verdade também em termos da faixa dinâmica. A curva da **Fig. 7** representa o desempenho em função do nível de entrada para o sistema com  $\alpha = 1/32$  e  $N = 2$ . Para outros valores de N as curvas obtidas são praticamente iguais à da **Fig. 7**. Para taxa de erro de bit igual a  $10^{-1}$ , não considerada na **Tabela 12**, obtém-se RSS negativa em todos os casos.

$\alpha \backslash N$	2	4	8	16	32
1/2	22,1	22,0	21,9	21,6	21,4
1/4	22,5	22,2	21,7	21,3	21,2
1/8	22,5	22,1	22,0	21,4	21,2
1/16	22,5	22,2	22,0	21,6	21,4
1/32	22,2	22,1	21,9	21,7	21,4
1/64	21,5	21,4	21,3	21,1	20,8
1/128	20,2	20,1	20,0	19,8	19,6
1/256	18,3	18,3	18,2	18,1	17,9

**Tabela 10.** RSS em dB para o codificador ADPCM silábico com preditor de 1ª ordem em canal ideal.

A relação sinal-ruído segmentada média que se obtém com esse sistema com  $\alpha = 1/32$  e  $N = 2$  quando se utiliza preditor de 2ª ordem é superior àquela obtida com preditor de 1ª ordem para  $TEB \leq 10^{-4}$ . Essa melhoria é de 0,6 dB em canal ideal e 0,4 dB em canal com  $TEB = 10^{-4}$ . Para canais muito ruidosos o desempenho com preditor de 2ª ordem é inferior. Isso é possivelmente devido ao acúmulo de dados incorretos nos quais as predições são baseadas.

$\alpha \backslash N$	2	4	8	16	32
1/2	12,2	12,3	14,4	15,5	16,8
1/4	12,6	14,0	15,7	15,8	16,4
1/8	15,5	15,7	16,2	16,8	17,0
1/16	17,5	17,8	17,4	17,3	17,6
1/32	18,8	18,5	18,9	18,5	18,3
1/64	18,5	18,4	18,4	18,4	18,4
1/128	18,0	18,0	17,9	17,7	17,6
1/256	16,1	16,1	16,1	16,0	15,9

**Tabela 11.** RSS em dB para o codificador ADPCM silábico com preditor de 1ª ordem em canal com taxa de erro de bit de  $10^{-3}$ .

N	TAXA DE ERRO DE BIT			
	0	$10^{-4}$	$10^{-3}$	$10^{-2}$
1	22,3	22,1	18,6	4,6
2	22,2	22,0	18,8	4,8
4	22,1	21,9	18,5	5,1
8	21,9	21,7	18,9	5,5
16	21,7	21,5	18,5	5,7

Tabela 12. RSS em dB para o codificador ADPCM silábico com preditor de 1ª ordem e  $\alpha = 1/32$ .

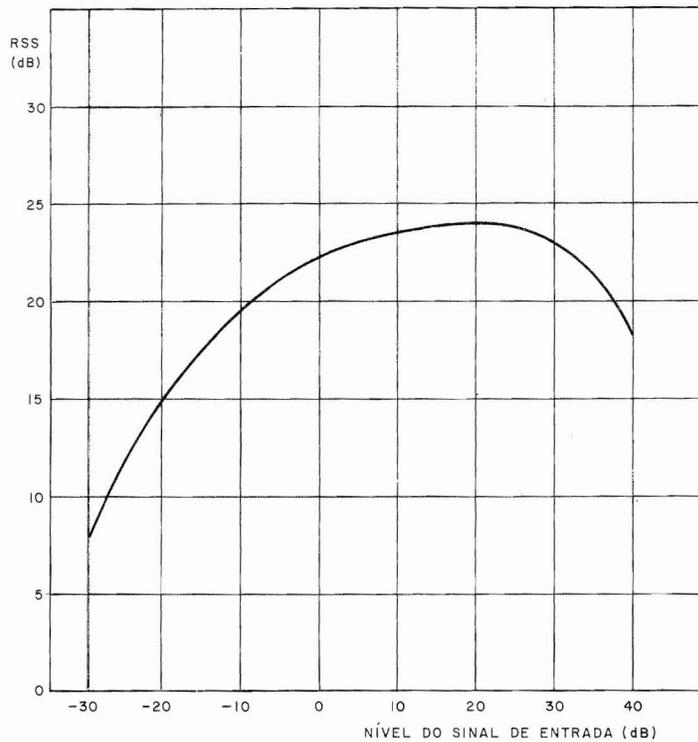


Figura 7. Desempenho do sistema ADPCM silábico com quantizador ótimo de Gauss e preditor de 1ª ordem em função do nível do sinal de entrada.

Os resultados obtidos com o sistema silábico mostram sua relativa imunidade a erros no canal. Entretanto, seu desempenho em canais ruidosos é um pouco inferior ao sis-

tema de Jayant com mecanismo de dissipação de erro e sua faixa dinâmica é também um pouco inferior.

#### 4. Conclusões

Codificadores diferenciais de sinais de voz, operando a uma taxa de transmissão de 32 kbit/s, foram avaliados em presença de erros no canal de comunicações. Essa avaliação foi feita através de uma medida fortemente correlacionada com medidas subjetivas de desempenho: a relação sinal-ruído segmentada média.

Os sistemas delta são os mais simples codificadores digitais de voz. Sistemas desse tipo que usam adaptação lenta do quantizador (delta silábico) fornecem resultados razoavelmente bons em canais ruidosos. Por outro lado, codificadores delta com adaptação instantânea são muito sensíveis a erros de transmissão, embora sejam melhores que os codificadores silábicos em canal ideal, além de apresentar uma maior faixa dinâmica.

O desempenho de codificadores delta com adaptação instantânea em canal ruidoso pode ser muito melhorado se um mecanismo de dissipação de erro do tipo proposto por Goodman e Wilkinson [18] for incorporado à técnica de adaptação do quantizador. Com essa melhoria os codificadores chegam inclusive a superar o desempenho dos sistemas delta silábicos para canais com taxa de erro de bit moderadas (TEB da ordem de  $10^{-4}$ ). Nesses casos o uso de um preditor de 2ª ordem é preferível ao de 1ª ordem.

Codificadores ADPCM com adaptação instantânea de Jayant [19], [20] fornecem resultados melhores quando o quantizador é projetado para uma distribuição de entrada Gaussiana e o preditor é de 2ª ordem. Entretanto, seu desempenho é muito fraco em canais de comunicação com taxas de erro de bit elevadas. Aqui também, como no caso dos sistemas delta, se um mecanismo de dissipação de erro for utilizado, essa sensibilidade a erros no canal é consideravelmente diminuída e o desempenho passa a ser bastante satisfatório.

A adaptação do quantizador de um sistema ADPCM através da energia segmentar, embora possa fornecer resultados de certa forma comparáveis aos de um sistema delta, apresenta um desempenho que é significativamente inferior ao ADPCM de Jayant com mecanismo de dissipação de erro. Por outro lado, o sistema ADPCM silábico fornece resultados inferiores embora próximos a esse último.

Uma comparação geral dos diversos sistemas permite finalmente concluir que o codificador ADPCM com memória de uma palavra de Jayant com mecanismo de dissipação de erros é o que apresenta melhor comportamento em canais de comunicação ruidosos. Além disso, verificou-se que a melhoria obtida quando se usa um preditor de 2ª ordem em relação ao desempenho alcançado com um preditor de 1ª ordem não é significativa

## Referências

- [1] A. Alcaim e J.R. Boisson de Marca, "Técnicas de Adaptação em Codificadores Diferenciais: Aplicações a Sinais de Voz", Revista TELEBRÁS, Ano VIII, nº 1, Março 1984, pp. 27-35.
  - [2] J.Max, "Quantizing for Minimum Distortion", IRE Transactions on Information Theory, vol. IT-6, nº 1, Março 1960, pp. 7-12.
  - [3] M.D. Paez e T.H. Glisson, "Minimum Mean Squared-Error Quantization in Speech PCM and DPCM Systems", IEEE Transactions on Communications, vol. COM-20, nº 2, Abril 1972, pp. 225-230.
  - [4] P. Kabal, "Minimum Mean Square Error Quantizers", Institut National des Recherches Scientifiques (INRS), Canadá, Telecommunications Technical Report nº 80-09, Maio 1980.
  - [5] B.J. McDermott, C. Scagliola e D.J. Goodman, "Perceptual and Objective Evaluation of Speech Processed by Adaptive Differential PCM", Bell System Technical Journal, vol. 57, nº 5, Maio-Junho 1978, pp. 1597-1618.
  - [6] R.A. McDonald, "Signal-to-Noise and Idle Channel Performance of Differential Pulse Code Modulation Systems-Particular Applications to Voice Signals", Bell System Technical Journal, vol. 45, nº 7, Setembro 1966, pp. 1123-1151.
  - [7] N.S. Jayant, "Adaptive Delta Modulation with a One-Bit Memory", Bell System Technical Journal, vol. 49, nº 3, Março 1970, pp. 321-342.
  - [8] A. Alcaim e J.R. Boisson de Marca, "A Comparative Study of 32 kbit/s Adaptive Delta Modulation Systems Using Higher Order Predictors", Proceedings of the IEEE Global Telecommunications Conference, Miami, Florida, Estados Unidos, Novembro 1982.
  - [9] P. Castellino e C. Scagliola, "Design of Instantaneously Companded Delta Modulators with M-Bit Memory", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Hartford, Connecticut, Estados Unidos, Maio 1977.
  - [10] J.A. Greefkes e F. de Jager, "Continuous Delta Modulation", Philips Research Report, vol. 23/2, 1968, pp. 233-246.
  - [11] J.A. Greefkes e K. Riemens, "Code Modulation with Digitally Controlled Companding for Speech Transmission", Philips Technical Review, vol. 31, nº 11/12, 1970, pp. 335-353.
-

- [12] D.L. Schilling, J. Garodnick e H.A. Vang, "Voice Encoding for the Space Shuttle Using Adaptive Delta Modulation", IEEE Transactions on Communications, vol. COM-26, nº 11, Novembro 1978, pp. 1652-1659.
- [13] J.L. Flanagan, M.R. Schröder, B.S. Atal, R.E. Crochiere, N.S. Jayant e J.M. Tribolet, "Speech Coding", IEEE Transactions on Communications, vol. COM-17, nº 4, Abril 1979, pp. 710-736.
- [14] L.R. Rabiner e R.W. Schafer, "Digital Processing of Speech Signals", Prentice-Hall, Inc. 1978.
- [15] A. Alcaim, "Codificadores Delta em Canais de Comunicação Ruidosos", Publicação CETUC-D-ST-06/83, Dezembro 1983.
- [16] N.S. Jayant, R.W. Schafer e M.R. Karim, "Step-Size-Transmitting Differential Coders for Mobile Telephony", Proceedings of the IEEE International Conference on Communications, San Francisco, California, Estados Unidos, Junho 1975.
- [17] A. Alcaim, "Instantaneously Companded Delta Coders with Error Dissipating Mechanism", Proceedings of the International Conference on Digital Signal Processing, Florença, Itália, Setembro 1984.
- [18] D.J. Goodman e R.M. Wilkinson, "A Robust Adaptive Quantizer", IEEE Transactions on Communications, vol. COM-23, nº 11, Novembro 1975, pp. 1362-1365.
- [19] N.S. Jayant, "Adaptive Quantization with a One-Word Memory", Bell System Technical Journal, vol. 52, nº 7, Setembro 1973, pp. 1119-1144.
- [20] P. Cummiskey, N.S. Jayant e J.L. Flanagan, "Adaptive Quantization in Differential PCM Coding of Speech", Bell System Technical Journal, vol. 52, nº 7, Setembro 1973, pp. 1105-1118.
- [21] P. Noll, "Effects of Channel Errors on the Signal-to-Noise Performance of Speech-Encoding Systems", Bell System Technical Journal, vol. 54, nº 9, Novembro 1975, pp. 1615-1636.
- [22] P. Noll, "Adaptive Quantizing in Speech Coding Systems", Proceedings of the IEEE Zurich Seminar on Digital Communications, Suíça, Março 1974.
- [23] P. Noll, "A Comparative Study of Various Quantization Schemes for Speech Encoding", Bell System Technical Journal, vol. 54, nº 9, Novembro 1975, pp. 1597-1614.

- [24] P. Castellino, G. Modena, L. Nebbia e C. Scagliola, "Bit Rate Reduction by Automatic Adaptation of Quantizer Step-Size in DPCM Systems", Proceedings of the IEEE Zurich Seminar on Digital Communications, Suíça, Março 1974.



**ABRAHAM ALCAIM** formou-se em Engenharia Elétrica pela PUC/RJ, com mestrado pela mesma universidade e doutorado pelo Imperial College – Universidade de Londres. Tem mais de 10 anos de experiência nas áreas de codificação digital e transmissão de formas de onda e processamento digital de sinais de voz. Nos últimos anos tem atuado em projetos de codificadores digitais de voz na faixa de 16 a 40 kbit/s, orientado teses de mestrado também nessa área e ministrado cursos de graduação e pós-graduação na PUC do Rio de Janeiro. Atualmente é professor do Centro de Estudos em Telecomunicações da Universidade Católica, onde trabalha em ensino e pesquisa no grupo de Sistemas de Telecomunicações.