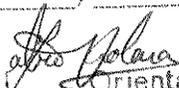


UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA
DEPARTAMENTO DE COMUNICAÇÕES

Este exemplar corresponde à redação final da tese
defendida por Flávia Martinho Ferreira
e aprovada pela Comissão
Ju'gadora em _____


Orientador

CODEC ADPCM A 16 KBIT/S COM QUANTIZAÇÃO DE ÁRVORE

FLÁVIA MARTINHO FERREIRA
Orientador: *Prof. Doutor FÁBIO VIOLARO*

Banca Examinadora: *FÁBIO VIOLARO (UNICAMP)*
OSVALDO CATSUMI IMAMURA (CTA-ITA)
JOÃO MARCOS T. ROMANO (UNICAMP)
CELSO DE ALMEIDA (UNICAMP)

Tese apresentada à Faculdade de Engenharia Elétrica
da Universidade Estadual de Campinas - UNICAMP,
como parte dos requisitos exigidos para a obtenção do
título de MESTRE EM ENGENHARIA ELÉTRICA.

OUTUBRO - 1990

Bo/9102553

Resumo

Neste trabalho é descrito um codec ADPCM à taxa de 16 kbit/s que emprega *codificação/quantização de árvore* e técnicas que permitem a redução da perceptibilidade do ruído de reconstrução, especificamente *pós-filtragem* e *filtragem de ponderação de ruído*. A motivação do trabalho reside na obtenção de um codec a 16 kbit/s com atraso e qualidade subjetiva compatíveis para aplicação telefônica, mediante a utilização da estrutura do codec ADPCM a 32 kbit/s padronizado pelo CCITT (Recomendação G.721). O bom desempenho subjetivo obtido para tal codec conduziu à redução da taxa de bits de 16 para 9.6 kbit/s. O atraso do algoritmo é 2 ms.

*AOS MEUS PAIS, JÚLIO E NAIR.
AO ZECA.*

Agradecimentos

Ao *Professor Doutor Fábio Violaro*, meu grande agradecimento por viabilizar este trabalho através de sua dedicação, atenção e das inúmeras discussões motivadas ao longo do mesmo. Agradeço profundamente ao *Engenheiro José Sindi Yamamoto*, chefe da Área de Processamento Digital de Voz do CPqD/TELEBRÁS, pela motivação do estudo da técnica de quantização de árvore e pelas discussões, que, ao longo do trabalho, muito contribuíram para seu êxito. Meus agradecimentos também aos *Engenheiros José Antônio Martins, Margarete Mitiko Iramina e Mônica Lacerda Rocha* pelos desenhos, figuras, incentivo e apoio; ao *Analista de Sistemas Claudir Antônio Moretti*, pela otimização computacional; ao *Fernando Oscar Runstein* e aos *companheiros de trabalho*, pela convivência e ajuda na realização deste trabalho.

Índice

1. INTRODUÇÃO	1
2. CODEC ADPCM A 32 KBIT/S	4
2.1 Introdução	4
2.2 Codificador ADPCM	4
2.2.1 Quantizador Adaptativo	6
2.2.1.1 Fator de escala do quantizador	7
2.2.1.2 Controle da velocidade de adaptação	11
2.2.2 Preditor Adaptativo	12
2.3 Decodificador ADPCM	15
2.3.1 Conversão para PCM	15
2.3.2 Ajuste de Codificação Síncrona	15
2.4 Relação S/R e ganho de predição para sistemas ADPCM	16
2.5 Desempenho do codec	17
3. CODEC ADPCM A 16 KBIT/S	20
3.1 Introdução	20
3.2 Princípio das modificações	20
3.3 Descrição das modificações	21
3.3.1 Quantizador	21
3.3.2 Adaptação do Quantizador	22
3.4 Desempenho do codec	22
4. CODEC ADPCM A 16 KBIT/S COM QUANTIZAÇÃO DE ÁRVORE	23
4.1 Introdução	23
4.2 Codificação de árvore - Idéia Básica	24
4.3 Codificador ADPCM com quantização de árvore	26
4.3.1 Gerador de Código	28
4.3.2 Algoritmo de Busca	29
4.3.3 Cálculo da Distorção	31
4.4 Decodificador ADPCM com quantização de árvore	32
4.5 Atraso do Algoritmo	33
4.6 Desempenho do codec ADPCM com quantização de árvore	33

5. CODEC ADPCM COM QUANTIZAÇÃO DE ÁRVORE E PÓS-FILTRAGEM	46
5.1 Introdução	46
5.2 Pós-filtragem - Idéia Básica	46
5.3 Pós-filtro - Função de Transferência	48
5.4 Resultados Experimentais	49
6. CODEC ADPCM COM QUANTIZAÇÃO DE ÁRVORE E PONDERAÇÃO DE RUÍDO	55
6.1 Introdução	55
6.2 Filtro de Ponderação de Ruído - Idéia Básica	55
6.3 Filtro de Ponderação de Ruído - Função de Transferência	57
6.4 Resultados Experimentais	58
7. CODEC ADPCM COM QUANTIZAÇÃO DE ÁRVORE A 9.6 KBIT/S	62
7.1 Introdução	62
7.2 Codificação Multi-árvore - Idéia Básica	62
7.3 Estrutura do Codec a 9.6 kbit/s	64
7.4 Desempenho do codec	65
8. TESTES E RESULTADOS	69
8.1 Introdução	69
8.2 Ambiente de trabalho	69
8.2.1 Sistema de aquisição de dados da UNICAMP	70
8.2.2 Sistema de aquisição de dados do CPqD/TELEBRÁS	70
8.3 Descrição dos esquemas ADPCM simulados	70
8.3.1 Codec ADPCM a 16 kbit/s	71
8.3.2 Codec ADPCM a 9.6 kbit/s	72
8.4 Procedimento de otimização dos parâmetros dos codecs simulados .	72
8.5 Testes objetivos e subjetivos informais	74
9. CONCLUSÕES	78
Apêndice I - QUANTIZAÇÃO PARA S/R ÓTIMA	80
Apêndice II - ADAPTAÇÃO BIMODAL	84
Apêndice III - CONSTANTE DE TEMPO DOS FILTROS DIGITAIS	92
Apêndice IV - ESTRUTURA DO PREDITOR ADAPTATIVO	94

Apêndice V - COMPLEXIDADE EM CODIFICAÇÃO COM BUSCA MULTI-CAMINHO	96
Apêndice VI - EFEITO DO PARÂMETRO α NA CONFORMAÇÃO DO ESPECTRO DO SINAL FILTRADO	100
Apêndice VII - MEDIDAS DE DESEMPENHO OBJETIVO	103
REFERÊNCIAS BIBLIOGRÁFICAS	105

Capítulo 1

INTRODUÇÃO

Através dos avanços obtidos na área de componentes, a codificação de voz rapidamente tornou-se uma tecnologia bastante viável e atrativa para sistemas de comunicação, à medida que permite que um sinal de voz seja eficientemente armazenado e transmitido.

A transmissão de um sinal de voz pode ser em tempo real, como é o caso de uma conversação telefônica, ou em tempo não real, como em armazenamento de voz para envio eletrônico de mensagens e para máquinas anunciadoras. Nestes sistemas, sabe-se que a largura de faixa de um canal de transmissão limita o número de sinais que podem ser transmitidos simultaneamente e que, no caso de armazenamento de voz, baixas taxas reduzem a quantidade de memória necessária para tal armazenamento. Desta forma, com o objetivo de reduzir a taxa de transmissão, técnicas de codificação de voz são avaliadas considerando-se custos de transmissão e armazenamento, complexidade, atraso e qualidade.

Baseando-se no compromisso entre estes requisitos, a redução da informação para a representação do sinal é atingida à medida que os codificadores exploram características inerentes ao sinal de voz, como, por exemplo, distribuição de amplitudes, correlação entre amostras e características espectrais (estrutura de formantes, “pitch” e ruído). Neste contexto, destacam-se os seguintes métodos de codificação de voz: codificação de forma de onda e codificação paramétrica. Na

codificação de forma de onda, preserva-se o “formato da onda” do sinal de voz analógico ao longo dos processos de amostragem e quantização, enquanto que na codificação paramétrica, o sinal de voz é representado por parâmetros de saída de um modelo para produção de voz, onde o sinal de voz é amostrado e processado para a obtenção destes parâmetros.

Tratando-se de aplicação telefônica, as taxas padronizadas mundialmente e atualmente em uso são 64 kbit/s (PCM - Pulse Code Modulation - Lei A e Lei μ) e 32 kbit/s (ADPCM - Adaptive Differential Pulse Code Modulation), ambos codificadores de forma de onda. Além disso, quanto a 16 kbit/s, encontra-se ainda em fase de padronização o codificador paramétrico LD-CELP (Low Delay - Code-Excited Linear Prediction).

Quando do início do processo de padronização do codec à taxa de 16 kbit/s, o CCITT, considerando as diversas aplicações deste codec em sistemas de comunicação, adotou, como requisito, que o atraso do codec não deveria exceder 5 ms, e, como objetivo, um atraso menor ou igual a 2 ms. Assim, em março de 1989, surgem como candidatos à padronização apenas dois codificadores (ambos paramétricos): por parte da AT&T, o LD-CELP anteriormente mencionado, e, por parte de um consórcio entre a Universidade Simon Fraser, a Universidade da Califórnia e a Voicecraft Inc., o LD-VXC (Low Delay - Vector Excitation Coder).

Neste cenário, tendo como motivação o fato de que à taxa de 16 kbit/s o número de candidatos a padronização era reduzido, o que refletia o grande desafio de se obter um codificador de pequeno atraso com bom desempenho, surgiu este trabalho, onde, baseando-se no codec ADPCM a 32 kbit/s padronizado pelo CCITT, procurou-se obter um codificador a 16 kbit/s que atendesse tais requisitos¹. Assim, no Capítulo 2 é apresentado o codec ADPCM a 32 kbit/s segundo a Recomendação G.721 do CCITT. No Capítulo 3 encontram-se as modificações introduzidas neste algoritmo de modo a reduzir a taxa para 16 kbit/s. Ressalta-se que a estrutura do preditor adaptativo foi mantida, dado seu desempenho satisfatório.

Uma vez obtido o codec ADPCM a 16 kbit/s, com o objetivo de se obter

¹Vislumbrou-se inicialmente a aplicação de tal codificador em Sistema de Comunicação Celular Digital, Sistema de Rádio-Acesso Celular Digital (telefonia rural), Equipamentos Multiplicadores de Circuitos Digitais (DCME - Digital Circuits Multiplication Equipaments), bem como em sistemas de armazenamento e transmissão e máquinas anunciadoras.

uma melhoria em termos de qualidade objetiva e subjetiva e, portanto, torná-lo compatível para aplicação telefônica, no Capítulo 4 é introduzido o conceito de quantização não instantânea, especificamente a quantização/codificação de árvore. Através desta técnica, um atraso na codificação dá ao sistema capacidade de busca multi-caminho, a qual permite identificar, dentre um conjunto de alternativas, a melhor amostra do sinal diferença que será codificada e transmitida. Ao contrário, no sistema ADPCM convencional, tem-se a busca através de um caminho único, onde a decisão da melhor amostra do sinal diferença é instantânea e, portanto, irrevogável.

Apesar da melhoria obtida através da quantização/codificação de árvore, o ruído de fundo ainda era bastante significativo, de modo que nos Capítulos 5 e 6 são apresentadas duas alternativas para redução da perceptibilidade do erro de reconstrução : pós-filtragem e filtragem de ponderação de ruído, respectivamente. A pós-filtragem é realizada após o sinal ser processado pelo codificador/decodificador, sendo que a redução da perceptibilidade do ruído é obtida às custas da distorção do sinal reconstruído. Já o filtro de ponderação de ruído encontra-se na seção de transmissão embutido no processo de seleção da melhor amostra do sinal diferença a ser codificada e transmitida.

A combinação de todas as técnicas apresentadas, ou seja, quantização de árvore, pós-filtragem e filtragem de ponderação de ruído, permitiu a obtenção de um codec a 16 kbit/s com um desempenho satisfatório para aplicação telefônica. Tal fato motivou a redução da taxa de 16 para 9.6 kbit/s, sendo que a investigação do codec resultante encontra-se no Capítulo 7.

No Capítulo 8 são descritas as simulações dos codificadores apresentados, bem com os resultados obtidos em termos de medidas objetivas e subjetivas. Finalmente, no Capítulo 9 conclui-se o trabalho, realizando-se uma análise dos resultados obtidos bem como vislumbrando-se sugestões para novos trabalhos.

Capítulo 2

CODEC ADPCM A 32 KBIT/S

2.1 Introdução

O algoritmo ADPCM (Adaptive Differential Pulse Code Modulation) à taxa de 32 kbit/s foi padronizado internacionalmente pelo CCITT através da Recomendação G.721 [1]. Trata-se de uma técnica de transcodificação digital que, a partir do PCM à taxa de 64 kbit/s (Lei A ou Lei μ), gera um sinal ADPCM a 32 kbit/s.

Através de quantização e predição adaptativas, o algoritmo oferece uma redução na taxa de bits na razão 2:1 sobre o PCM à taxa de 64 kbit/s, mantendo o mesmo desempenho de transmissão do PCM.

2.2 Codificador ADPCM

Na figura 2.1 encontra-se o diagrama em blocos do codificador. Este codificador é representado de maneira mais detalhada na figura 2.2¹ onde, para cada variável a ser descrita, k é o índice da amostra. As amostras são atualizadas a cada intervalo de 125 μ s.

¹Ainda no período de validação deste codec, verificou-se uma degradação no desempenho do codec quando da codificação de sinais de dados na faixa de voz, especificamente do tipo FSK operando à taxa de 1200 bit/s. Foi proposta uma modificação no algoritmo [2] sem impacto significativo no desempenho do codec para outros sinais, sendo, portanto, incorporada à Recomendação G.721. Entretanto, neste trabalho tal modificação não foi implementada, uma vez que o interesse primário reside na codificação de sinais de voz.

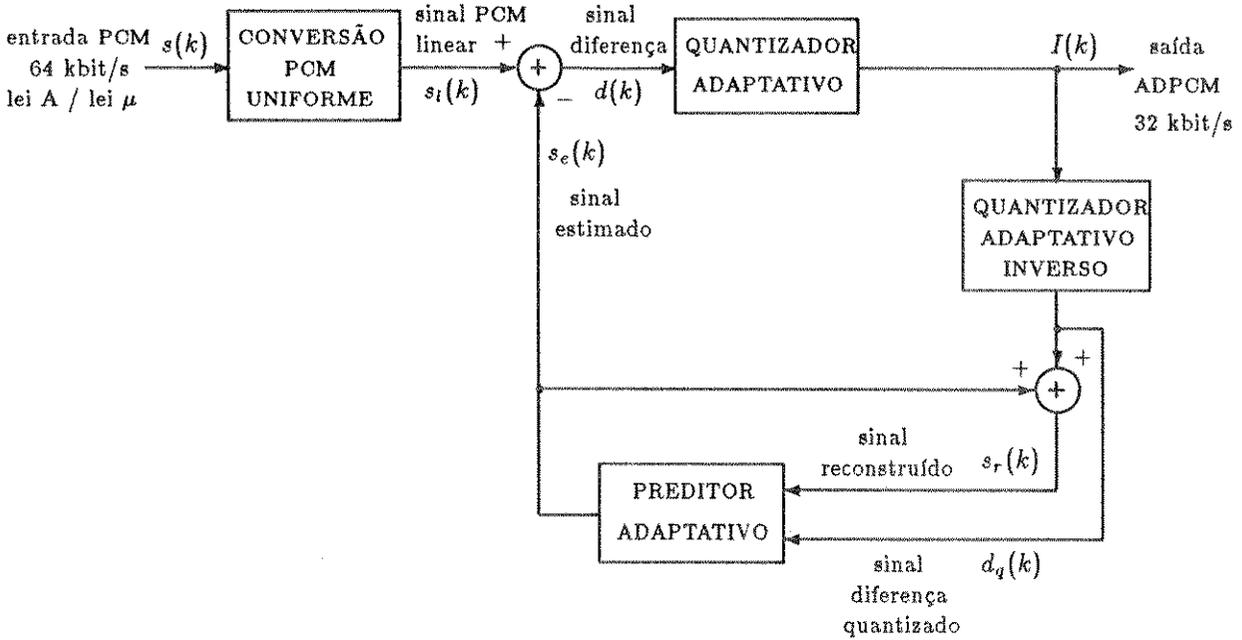


Figura 2.1: Diagrama em blocos simplificado do codificador ADPCM.

O sinal PCM Lei A (Lei μ) $s(k)$ é inicialmente convertido para o sinal PCM uniforme $s_l(k)$. Esta conversão é implementada segundo a Recomendação G.711 do CCITT [3], onde um sinal PCM 8 bits Lei A (Lei μ) é convertido para o formato PCM uniforme com resolução de 13 bits. Em seguida, o sinal diferença $d(k)$ é obtido calculando-se a diferença entre o sinal PCM uniforme $s_l(k)$ e o sinal estimado pelo preditor $s_e(k)$:

$$d(k) = s_l(k) - s_e(k). \quad (2.1)$$

Esta diferença é então submetida a um quantizador adaptativo de 16 níveis que fornece à sua saída o sinal ADPCM $I(k)$ constituído de 4 bits/amostra.

A palavra binária $I(k)$, após ser processada pelo quantizador adaptativo inverso, gera o sinal diferença quantizado. O sinal estimado $s_e(k)$ é obtido submetendo-se $d_q(k)$ a um preditor adaptativo. O sinal reconstruído $s_r(k)$ é obtido somando-se o sinal estimado ao sinal diferença quantizado:

$$s_r(k) = s_e(k) + d_q(k). \quad (2.2)$$

Para minimizar o erro de quantização ($d_q(k) - d(k)$) e conseqüentemente fazer

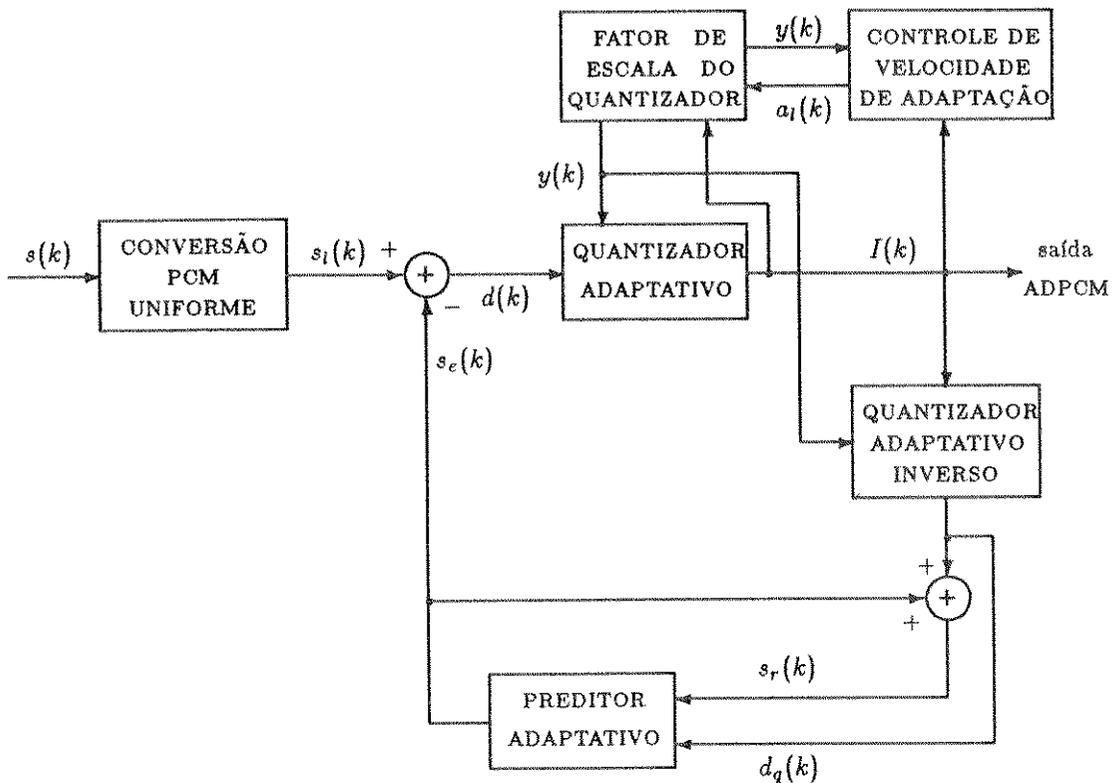


Figura 2.2: Diagrama em blocos detalhado do codificador ADPCM.

com que $s_r(k)$ tenda a uma réplica de $s_1(k)$, tanto o quantizador quanto o preditor operam de maneira adaptativa.

2.2.1 Quantizador Adaptativo

Este bloco é responsável pela quantização/codificação do sinal diferença $d(k)$. Segundo a Recomendação G.721 do CCITT, com o objetivo de manter-se uma grande faixa dinâmica e minimizar a complexidade (substituição de divisão por subtração), a quantização e a adaptação são feitas no domínio logarítmico \log_2 , como mostra a figura 2.3.

Os bits $I_i(k)$, $i = 1, 2, 3$ e 4 , na saída do quantizador constituem o sinal ADPCM à taxa de 32 kbit/s. Como a curva de quantização apresenta uma simetria ímpar em relação à origem (Apêndice I, figura I.1), a palavra binária $I(k)$ é definida de modo que o bit $I_1(k)$ corresponda à polaridade e os demais bits correspondam à magnitude do sinal diferença quantizado $d_q(k)$. Assim, $I_1(k) = 0$

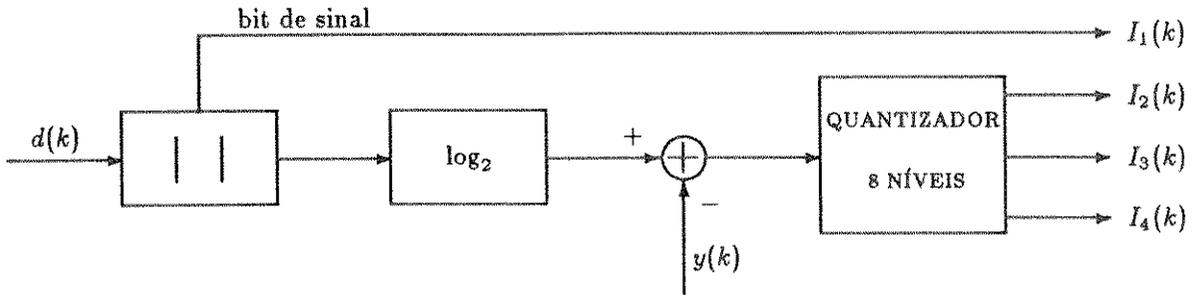


Figura 2.3: Detalhes do quantizador adaptativo.

para $d_q(k) \geq 0$ e $I_1(k) = 1$ para $d_q(k) < 0$. Ao longo do trabalho será adotada a notação $|I(k)|$ para indicar o equivalente decimal dos bits $I_2(k)$ a $I_4(k)$. Assim, $|I(k)|$ assume valores no intervalo $[0,7]$, sendo que os valores 0 e 7 correspondem, respectivamente, às amplitudes mínima e máxima de $d_q(k)$.

O fator de escala $y(k)$ é responsável pela operação de adaptação, sendo computado no bloco Fator de Escala do Quantizador. Segundo a Recomendação G.721, o quantizador ótimo no domínio linear é caracterizado segundo a tabela 2.1. No domínio logarítmico, este quantizador encontra-se na tabela 2.2, onde definiu-se o fator de escala da seguinte forma:

$$y(k) = \log_2 Y(k), \quad (2.3)$$

onde $Y(k)$ é o fator de escala no domínio linear.

Deste modo, o quantizador opera com o sinal diferença normalizado $|d(k)|/Y(k)$, ou, no domínio logarítmico, $\log_2 |d(k)| - y(k)$.

Comparando-se as tabelas 2.1 e I.1 (Apêndice I), verifica-se que o quantizador da Recomendação G.721 é o quantizador ótimo de Max [4] para a variância $\sigma \simeq 3.7$ do sinal diferença normalizado.

2.2.1.1 Fator de escala do quantizador

O fator de escala $y(k)$ é atualizado a cada amostra k através da técnica de adaptação bimodal. Para sinais que produzem sinais diferença $d(k)$ com grandes flutuações de amplitude (por exemplo, voz), a adaptação é rápida, isto é, $y(k)$ deve variar rapidamente de forma a acompanhar a evolução da envoltória do sinal diferença. No caso de sinais que produzem sinais diferença $d(k)$ com pequenas flutuações de amplitude (por exemplo, tons e dados na faixa de voz), a adaptação é

Faixa de Entrada Normalizada do Quantizador $ d(k) /Y(k)$	$ I(k) $	Saída Normalizada do Quantizador Inverso $ d_q(k) /Y(k)$
$[8.94; +\infty)$	7	10.13
$[6.87; 8.94)$	6	7.73
$[5.35; 6.87)$	5	6.02
$[4.11; 5.35)$	4	4.69
$[2.99; 4.11)$	3	3.51
$[1.95; 2.99)$	2	2.46
$[0.97; 1.95)$	1	1.44
$[0; 0.97)$	0	0.48

Tabela 2.1: Característica Entrada/Saída Normalizada do Quantizador no Domínio Linear.

Faixa de Entrada Normalizada do Quantizador $\log_2 d(k) - y(k)$	$ I(k) $	Saída Normalizada do Quantizador Inverso $\log_2 d_q(k) - y(k)$
$[3.16; +\infty)$	7	3.34
$[2.78; 3.16)$	6	2.95
$[2.42; 2.78)$	5	2.59
$[2.04; 2.42)$	4	2.23
$[1.58; 2.04)$	3	1.81
$[0.96; 1.58)$	2	1.29
$[-0.05; 0.96)$	1	0.53
$(-\infty; -0.05)$	0	-1.05

Tabela 2.2: Característica Entrada/Saída Normalizada do Quantizador no Domínio Logarítmico.

$ I $	$W[I]$
7	69.25
6	21.25
5	11.50
4	6.12
3	3.12
2	1.70
1	0.25
0	-0.75

Tabela 2.3: Função $W[I]$.

lenta, correspondendo basicamente a um quantizador fixo com um escalonamento apropriado. No Apêndice II encontra-se um estudo sobre a técnica de adaptação bimodal, onde o fator de escala é resultante da combinação linear de dois termos, o fator de escala “unlocked”, $y_u(k)$, e o fator de escala “locked”, $y_l(k)$.

• Fator de escala “unlocked” y_u

O fator de escala Y_u , segundo a expressão (II.9) do Apêndice II, é dado por

$$Y_u(k) = M(I(k))Y^\beta(k). \quad (2.4)$$

Segundo a expressão (2.3), $y(k) = \log_2 Y(k)$. Aplicando-se \log_2 em (2.4), obtêm-se

$$\log_2 Y_u(k) = \log_2 [M(I(k))Y^\beta(k)] \quad (2.5)$$

e

$$y_u(k) = \beta y(k) + \log_2 M(I(k)). \quad (2.6)$$

Segundo a Recomendação G.721,

$$y_u(k) = (1 - 2^{-5})y(k) + 2^{-5}W[I(k)], \quad 1.06 \leq y_u(k) \leq 10, \quad (2.7)$$

onde a função discreta $W[I(k)]$ é definida de acordo com a tabela 2.3.

Os valores de $W[I]$ foram obtidos empiricamente. A função $W[I(k)]$ faz com que $y_u(k)$ se adapte através de degraus maiores para valores crescentes de $I(k)$.

Comparando-se as expressões (2.6) e (2.7), verifica-se que $\beta = 31/32$ e $\log_2 M(I(k)) = 2^{-5}W[I(k)]$, isto é,

$$M(I(k)) = 2^{2^{-5}W[I(k)]}. \quad (2.8)$$

Na tabela 2.4 têm-se os valores dos multiplicadores $M[I]$ no domínio linear. Tais resultados podem parecer contraditórios, uma vez que apenas para $|I| = 0$

I	$M[I]$
7	4.482
6	1.585
5	1.283
4	1.142
3	1.070
2	1.037
1	1.005
0	0.984

Tabela 2.4: Função $M[I]$.

tem-se $M[I] < 1$. Entretanto, tal contradição é apenas aparente, pois, segundo a expressão (2.4), $Y(k)$ é elevado ao expoente β antes de sofrer a ação dos multiplicadores $M[I]$.

Referenciando-se ainda à expressão (2.7), verifica-se que a função $W[I(k)]$ está sendo filtrada. Quando ocorrer um erro em $I(k)$, o efeito deste erro decresce exponencialmente com o tempo. O cálculo da constante de tempo é apresentado no Apêndice III, e, para a frequência de amostragem de 8 kHz ($T = 125 \mu s$), ela é aproximadamente 4 ms.

- Fator de escala "locked" y_l

Como mencionado no Apêndice II, segundo a expressão (II.11), $y_l(k)$ é obtido através de filtragem passa-baixas de $y_u(k)$:

$$y_l(k) = \alpha y_l(k-1) + (1 - \alpha) y_u(k). \quad (2.9)$$

Segundo a Recomendação G.721,

$$y_l(k) = (1 - 2^{-6}) y_l(k-1) + 2^{-6} y_u(k). \quad (2.10)$$

Neste caso, $\alpha = 63/64$ e, de acordo com o Apêndice III, a constante de tempo do filtro é da ordem de 8 ms, de modo que $y_l(k)$ se adapta mais lentamente, mas o suficiente para conseguir acompanhar sinais com pequenas flutuações de amplitude.

- Fator de escala y

Como também já foi mencionado no Apêndice II, o fator de escala y é resultante da combinação linear de $y_l(k)$ e $y_u(k)$. Segundo a Recomendação G.721,

$$y(k) = a_l(k) y_u(k-1) + (1 - a_l(k)) y_l(k-1), \quad (2.11)$$

onde a_l é o parâmetro de controle de velocidade de adaptação.

I	$F[I]$
7	7
6	3
5	1
4	1
3	1
2	0
1	0
0	0

Tabela 2.5: Função $F[I]$.

2.2.1.2 Controle da velocidade de adaptação

O parâmetro $a_l(k)$ da expressão (2.11) pode assumir valores no intervalo $[0,1]$. Ele tende a zero para sinais de dados na faixa de voz e para tons, onde o sinal diferença apresenta pequenas flutuações de amplitude (adaptação lenta: $y_l(k)$); e tende à unidade para sinais de voz, onde o sinal diferença apresenta grandes flutuações de amplitude (adaptação rápida: $y_u(k)$). Descreve-se a seguir o procedimento para a obtenção do parâmetro $a_l(k)$.

Primeiramente, são computadas as médias a curto ($d_{ms}(k)$) e longo prazo ($d_{ml}(k)$) do sinal ADPCM transmitido. Segundo a Recomendação G.721,

$$d_{ms}(k) = (1 - 2^{-5})d_{ms}(k-1) + 2^{-5}F[I(k)] \quad (2.12)$$

e

$$d_{ml}(k) = (1 - 2^{-7})d_{ml}(k-1) + 2^{-7}F[I(k)], \quad (2.13)$$

onde $F[I(k)]$ é definida de acordo com a tabela 2.5.

$F[I(k)]$ é uma função definida empiricamente. Ela amplifica as diferenças na sequência $I(k)$ e reduz o tempo de transição entre os modos lento e rápido de adaptação.

Novamente pode-se determinar as constantes de tempo destes filtros. No caso de filtragem a curto prazo, segundo o Apêndice III, a constante de tempo é aproximadamente 4 ms. Para a filtragem a longo prazo, a constante de tempo é aproximadamente 16 ms.

Os valores de $d_{ms}(k)$ e $d_{ml}(k)$ são comparados. Se a diferença entre eles é grande, isto significa que um sinal de voz está presente. Caso contrário, é provável que um outro tipo de sinal esteja presente (por exemplo, dados na faixa de voz, sinalização, etc). Então, um parâmetro intermediário $a_p(k)$ é definido segundo a

Recomendação G.721:

$$a_p(k) = \begin{cases} (1 - 2^{-4})a_p(k-1) + 2^{-3}, & \text{se } |d_{ms}(k) - d_{ml}(k)| > 2^{-3}d_{ml}(k) \\ (1 - 2^{-4})a_p(k-1) + 2^{-3}, & \text{se } y(k) < 3 \\ (1 - 2^{-4})a_p(k-1), & \text{c.c..} \end{cases} \quad (2.14)$$

A condição $y(k) < 3$ indica que o canal está vazio e portanto necessitando uma adaptação rápida se houver presença de sinal e se este for de voz.

Finalmente, o parâmetro de controle de velocidade será

$$a_l(k) = \begin{cases} 1, & a_p(k-1) > 1 \\ a_p(k-1), & a_p(k-1) \leq 1. \end{cases} \quad (2.15)$$

Isto permite que seja introduzido um atraso no início de transição do modo lento para o modo rápido, até que o valor absoluto de $I(k)$ permaneça constante por um determinado tempo, eliminando assim transições prematuras para sinais de entrada pulsados (por exemplo, dados na faixa de voz com portadora chaveada).

2.2.2 Preditor Adaptativo

No esquema ADPCM ilustrado pelas figuras 2.1 e 2.5, verifica-se que a estrutura do decodificador é constituída da malha de realimentação do codificador, ou seja, na seção de transmissão dispõe-se do sinal reconstruído $s_r(k)$, que é o mesmo obtido na seção de recepção, quando da ausência de erros de transmissão. Assim, para garantir o alinhamento entre o codificador (transmissão) e o decodificador (recepção), a estimativa do sinal gerada pelo preditor é baseada em amostras passadas do sinal reconstruído $s_r(k)$ e não do sinal PCM linear $s_l(k)$, uma vez que este não encontra-se disponível na seção de recepção. Neste sentido, na figura 2.4 encontra-se o preditor adaptativo presente na Recomendação G.721 do CCITT, onde o sinal estimado $s_e(k)$ é resultante da combinação de duas estruturas adaptativas, $A(z)$ (segunda ordem) e $B(z)$ (sexta ordem), que modelam os pólos e os zeros, respectivamente.

A escolha final deste preditor é resultante de investigações de vários fatores: estabilidade, alinhamento entre o codificador e o decodificador na presença de erros de transmissão, e eficiência computacional. Tais estudos encontram-se no Apêndice IV.

O sinal reconstruído é obtido a partir da expressão

$$s_r(k) = s_e(k) + d_q(k), \quad (2.16)$$

ou seja,

$$d_q(k) = s_r(k) - s_e(k). \quad (2.17)$$

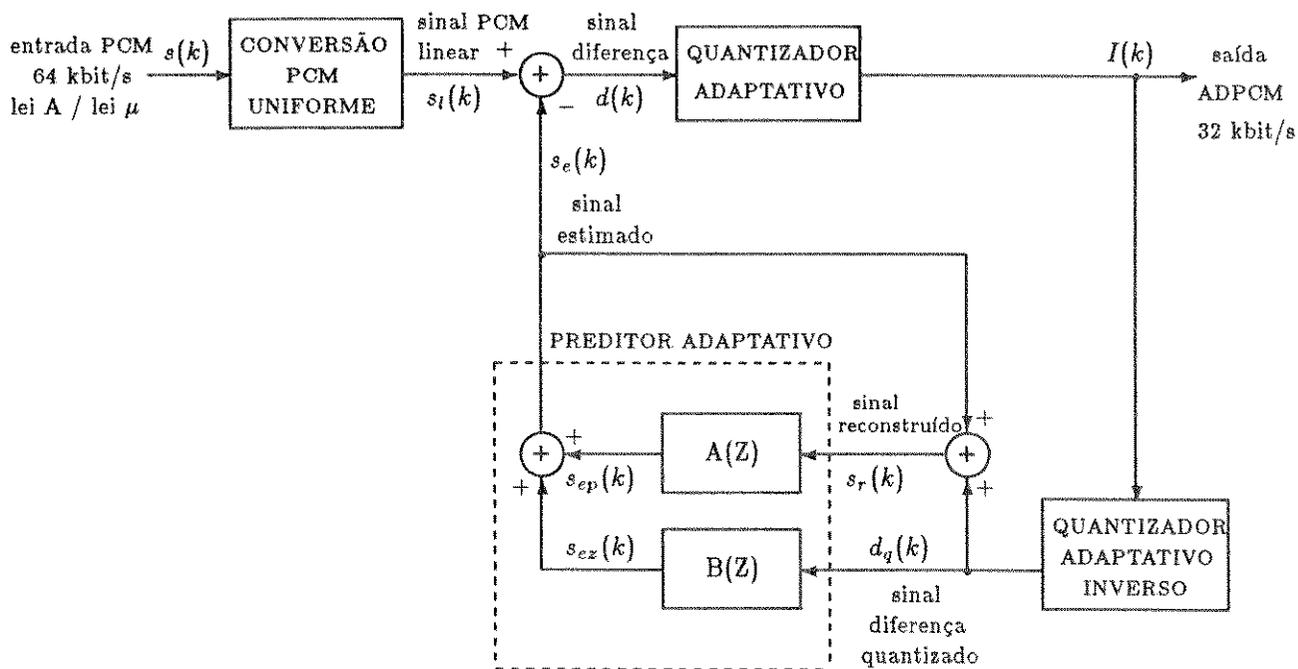


Figura 2.4: Diagrama em blocos do Codificador ADPCM - Preditor Adaptativo.

Por outro lado, tem-se o erro de predição dado por

$$d(k) = s_l(k) - s_e(k). \quad (2.18)$$

Baseando-se nas expressões (2.17) e (2.18), verifica-se que minimizando a diferença entre o sinal PCM linear $s_l(k)$ e sua estimativa $s_e(k)$, também está-se minimizando a diferença entre $s_l(k)$ e o sinal reconstruído $s_r(k)$, o que é desejável.

As equações básicas que descrevem a operação deste preditor são as seguintes:

$$s_e(k) = s_{ep}(k) + s_{ez}(k), \quad (2.19)$$

onde

$$s_{ep}(k) = \sum_{i=1}^2 a_i(k-1)s_r(k-i) \quad (2.20)$$

e

$$s_{ez}(k) = \sum_{i=1}^6 b_i(k-1)d_q(k-i). \quad (2.21)$$

Os coeficientes a_i e b_i são adaptados de acordo com o Algoritmo do Sinal [5], que é uma versão simplificada do Algoritmo do Gradiente.

As equações de adaptação dos coeficientes do preditor que serão apresentadas a seguir seguem a notação da Recomendação G.721 do CCITT, onde, no instante k , inicialmente são calculados a predição $s_e(k)$ e o sinal diferença $d(k)$ e posteriormente são realizadas a quantização e a quantização inversa do sinal diferença, resultando no sinal diferença quantizado $d_q(k)$. A atualização dos coeficientes do preditor é feita a partir destes sinais, obtendo-se então $a_i(k)$ e $b_i(k)$. Estes novos coeficientes serão utilizados no instante $k+1$ para o cálculo da nova predição $s_e(k+1)$.

• Adaptação dos coeficientes do preditor de pólos

Fazendo-se $\delta_{a_1} = 2^{-8}$ e $\alpha_{a_1} = 3 \times 2^{-8}$ na expressão (IV.5) do Apêndice IV, obtem-se a regra de adaptação da Recomendação G.721 para o coeficiente a_1 :

$$a_1(k) = (1 - 2^{-8})a_1(k-1) + 3 \times 2^{-8} \text{sgn}[p(k)]\text{sgn}[p(k-1)]. \quad (2.22)$$

Analogamente, considerando-se $\delta_{a_2} = 2^{-7}$ e $\alpha_{a_2} = 2^{-7}$ na expressão (IV.6), obtem-se para o coeficiente a_2 :

$$a_2(k) = (1 - 2^{-7})a_2(k-1) + 2^{-7} \{ \text{sgn}[p(k)]\text{sgn}[p(k-2)] - 2a_1(k-1)\text{sgn}[p(k)]\text{sgn}[p(k-1)] \}, \quad (2.23)$$

onde

$$p(k) = d_q(k) + s_{ez}(k), \quad (2.24)$$

com $|a_2(k)| \leq 0.75$ e $|a_1(k)| < 1 - 2^{-4} - a_2(k)$ como condições de estabilidade.

Entretanto, segundo a Recomendação G.721, a regra de adaptação do coeficiente a_2 é dada por

$$a_2(k) = (1 - 2^{-7})a_2(k-1) + 2^{-7} \{ \text{sgn}[p(k)]\text{sgn}[p(k-2)] - f[a_1(k-1)]\text{sgn}[p(k)]\text{sgn}[p(k-1)] \}, \quad (2.25)$$

onde

$$f(a_1) = \begin{cases} 4a_1, & |a_1| \leq 1/2 \\ 2\text{sgn}(a_1), & |a_1| > 1/2 \end{cases} \quad (2.26)$$

A diferença entre as expressões (2.23) e (2.25) reside na substituição de $a_1(k-1)$ pela função $f[a_1(k-1)]$, obtida empiricamente quando da tentativa de alcançar um melhor desempenho do sistema.

• Adaptação dos coeficientes do preditor de zeros

Fazendo-se $\delta_b = 2^{-8}$ e $\alpha_b = 2^{-7}$ na expressão (IV.4) do Apêndice IV, obtém-se a regra de adaptação da Recomendação G.721 para os coeficientes b_i :

$$b_i(k) = (1 - 2^{-8})b_i(k-1) + 2^{-7} \text{sgn}[d_q(k)]\text{sgn}[d_q(k-i)], i = 1, \dots, 6. \quad (2.27)$$

2.3 Decodificador ADPCM

Na figura 2.5 encontra-se o diagrama em blocos do decodificador. Uma representação mais detalhada deste decodificador encontra-se na Figura 2.6. Verifica-se que a estrutura do decodificador é similar à da malha de realimentação do codificador, sendo que a diferença entre elas reside na presença dos blocos Conversão para PCM e Ajuste de Codificação Síncrona.

2.3.1 Conversão para PCM

Este bloco converte o sinal reconstruído PCM uniforme $s_r(k)$ para o sinal PCM Lei A (Lei μ) $s_p(k)$, de acordo com a Recomendação G.711 do CCITT [3].

2.3.2 Ajuste de Codificação Síncrona

Na figura 2.7 tem-se a estrutura mais detalhada deste bloco, cuja função é evitar possível distorção acumulativa em ligações TANDEM (ADPCM-PCM-ADPCM-PCM-...).

Verifica-se que no próprio decodificador, o próximo codificador da ligação TANDEM é simulado antes da transmissão da amostra PCM 64 kbit/s. Caso esta simulação resulte em um valor $I(k)$ diferente do recebido pelo decodificador, a saída PCM $s_p(k)$ é mudada de modo a eliminar esta diferença, garantindo assim que os valores de todas as variáveis de estado nos sucessivos estágios ADPCM sejam idênticos, e, portanto, eliminando distorção acumulativa.

Primeiramente, o sinal $s_p(k)$ é convertido para um sinal PCM uniforme $s_{lx}(k)$ e a diferença entre este e o sinal estimado $s_e(k)$ anteriormente descrito é calculada:

$$d_x(k) = s_{lx}(k) - s_e(k). \quad (2.28)$$

Esta diferença é então comparada com o intervalo de decisão do quantizador ADPCM determinado por $I(k)$ e $y(k)$. O valor da amostra PCM $s_d(k)$ é então definido da seguinte forma:

$$s_d(k) = \begin{cases} s_p^+(k), & d_x(k) < \text{limitante inferior do intervalo} \\ s_p^-(k), & d_x(k) \geq \text{limitante superior do intervalo} \\ s_p(k), & \text{c.c.} \end{cases} \quad (2.29)$$

onde $s_d(k)$ é a saída PCM do decodificador; $s_p^+(k)$, o nível PCM imediatamente superior a $s_p(k)$; e $s_p^-(k)$, o nível PCM imediatamente inferior a $s_p(k)$.

2.4 Relação S/R e ganho de predição para sistemas AD-PCM

Considerando-se a figura 2.1 e definindo-se $e(k)$ como ruído de quantização, no esquema ADPCM têm-se

$$d(k) = s_l(k) - s_e(k), \quad (2.30)$$

$$d_q(k) = d(k) + e(k), \quad (2.31)$$

e

$$sr(k) = se(k) + dq(k), \quad (2.32)$$

que correspondem, respectivamente, às expressões do sinal diferença, sinal diferença quantizado e sinal reconstruído.

Combinando-se as equações acima pode-se expressar o sinal reconstruído da seguinte forma:

$$s_r(k) = s_l(k) + e(k). \quad (2.33)$$

Define-se a relação sinal/ruído como

$$S/R \triangleq \frac{E[s_1^2(k)]}{E[e^2(k)]} = \frac{\sigma_{sl}^2}{\sigma_e^2}. \quad (2.34)$$

Definindo-se $E[d^2(k)] = \sigma_d^2$, a expressão anterior pode ser reescrita como

$$S/R = \frac{\sigma_{sl}^2}{\sigma_d^2} \cdot \frac{\sigma_d^2}{\sigma_e^2} = G_p \cdot (S/R)_Q, \quad (2.35)$$

onde $G_p = \frac{\sigma_{sl}^2}{\sigma_d^2}$ é o ganho de predição e $(S/R)_Q = \frac{\sigma_d^2}{\sigma_e^2}$ é a relação sinal/ruído do quantizador.

2.5 Desempenho do codec

No Capítulo 8 encontram-se os procedimentos e resultados de medidas de desempenho de todos os codecs analisados neste trabalho. Entretanto, medidas preliminares de desempenho foram realizadas considerando-se um arquivo de voz de duração de aproximadamente 6 segundos, constituído de sentenças masculinas e femininas, a saber:

- frases masculinas: *He suspected his brother. She danced next with an officer.*
- frases femininas: *He spoke with a weary voice. I suppose you mean my mother.*

Para o sistema ADPCM a 32 kbit/s descrito, foram obtidas as seguintes medidas objetivas:

⇒ relação sinal/ruído segmental² do codec: $[S/R]_{seg} = 27.09$ dB

⇒ ganho de predição segmental: $G_p = 12$ dB

Por outro lado, a expressão (2.35) no domínio logarítmico é

$$[S/R]_{\log} = [G_p]_{\log} + [(S/R)_Q]_{\log}, \quad (2.36)$$

de modo que a relação sinal/ruído segmental do quantizador é $[(S/R)_Q]_{seg} \simeq 15$ dB.

²A definição e o procedimento para a abtenção da relação sinal/ruído segmental encontram-se no Apêndice VII

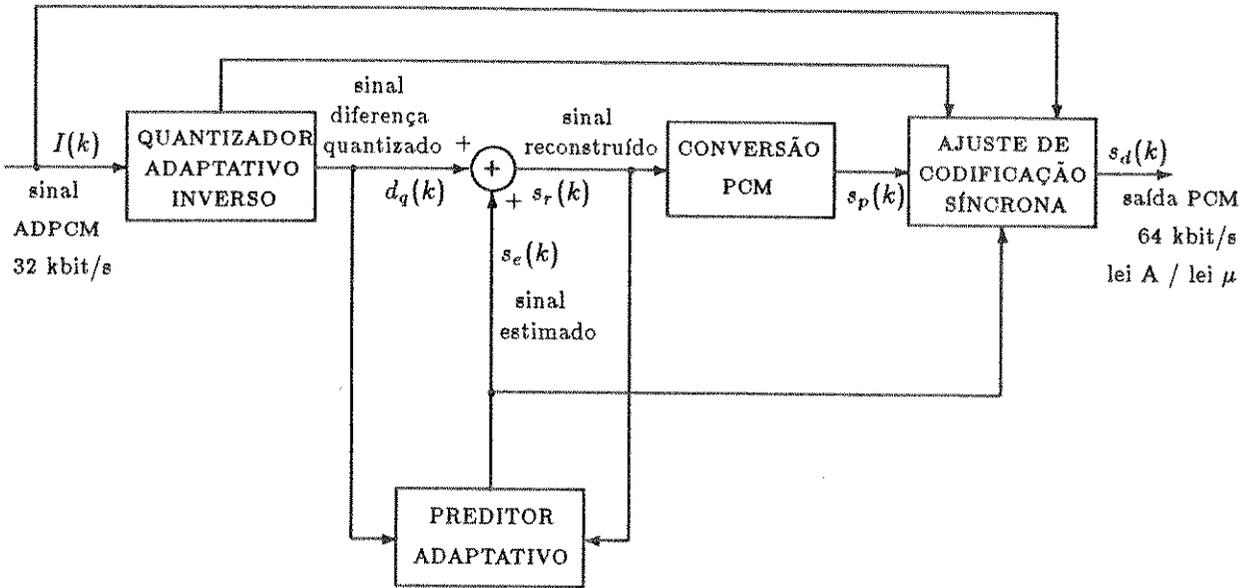


Figura 2.5: Diagrama em blocos do decodificador ADPCM.

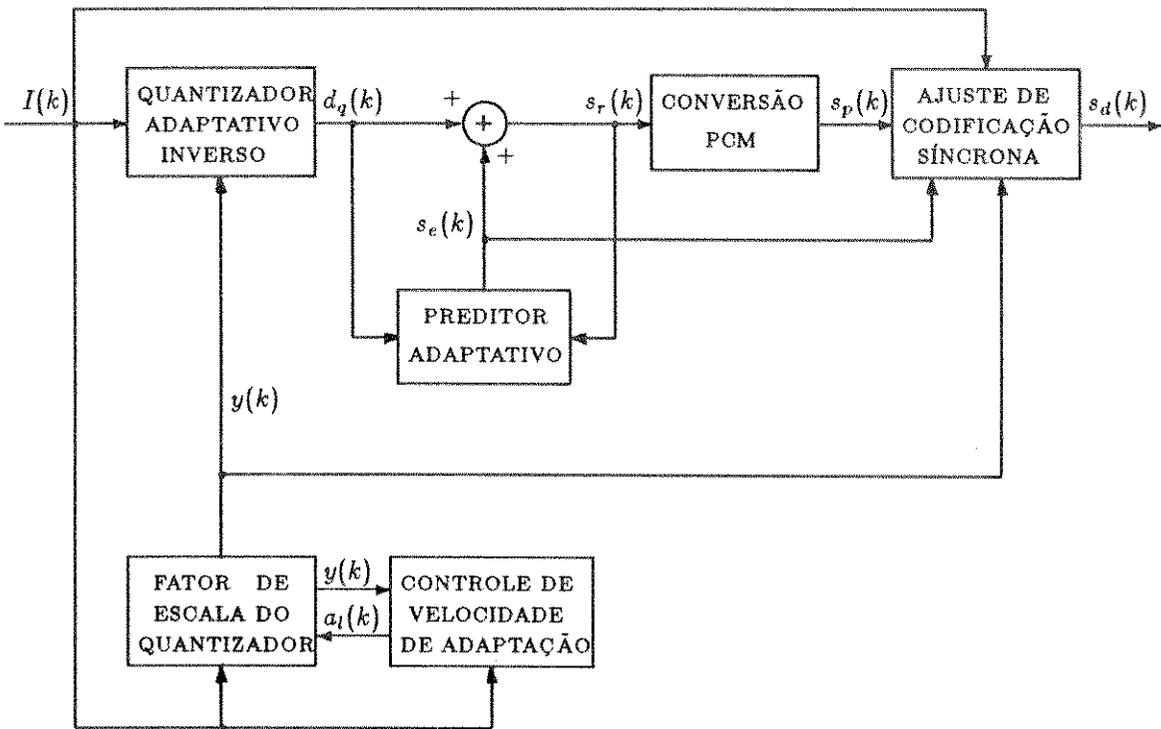


Figura 2.6: Diagrama em blocos detalhado do decodificador ADPCM.

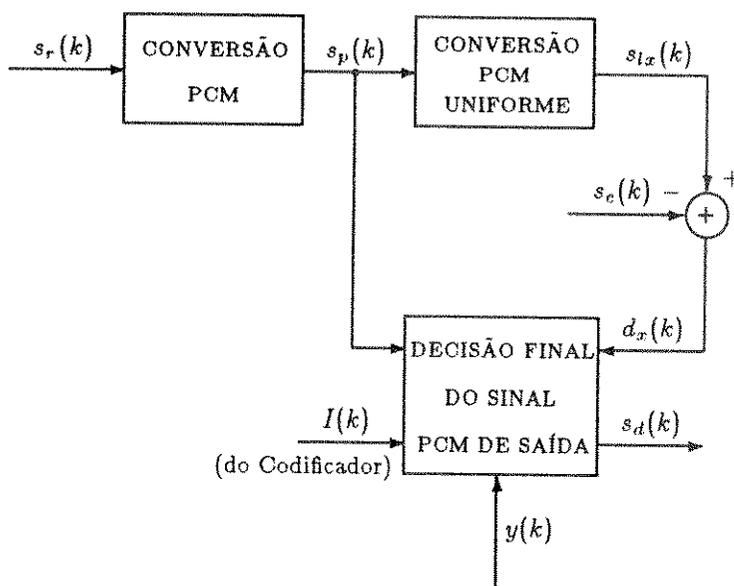


Figura 2.7: Diagrama em blocos do Ajuste de Codificação Síncrona.

Capítulo 3

CODEC ADPCM A 16 KBIT/S

3.1 Introdução

Neste capítulo são apresentadas as modificações introduzidas no algoritmo ADPCM 32 kbit/s padronizado pelo CCITT (Recomendação G.721) com a finalidade de reduzir a taxa de bits para 16 kbit/s.

3.2 Princípio das modificações

A mudança básica em termos de taxa de bits reside na relação *número de bits/amostra ADPCM*, ou seja, à taxa de 32 kbit/s têm-se 4 bits/amostra, enquanto que a 16 kbit/s, a relação é 2 bits/amostra. Considerando-se que a estrutura do preditor do codec a 32 kbit/s será mantida, tal mudança residirá no bloco responsável pela quantização/codificação do sinal diferença $d(k)$.

Assim, considerando-se as figuras 2.2 e 2.6 presentes no Capítulo 2, representando, respectivamente, o codificador e o decodificador ADPCM a 32 kbit/s, verifica-se que para a taxa de 16 kbit/s apenas as especificações do quantizador e dos blocos envolvidos em seu processo de adaptação (Fator de Escala do Quantizador e Controle de Velocidade de Adaptação) devem ser reformuladas.

Uma vez que o algoritmo a 16 kbit/s destina-se unicamente ao processamento de sinais de voz, eliminou-se do fator de escala do quantizador $y(k)$ a componente “locked” $y_l(k)$, ou seja, $y(k)$ é igual ao fator de escala “unlocked” $y_u(k)$, sendo este

responsável pela adaptação rápida do quantizador. Assim, em qualquer instante de amostragem, o parâmetro de controle de velocidade, $a_l(k)$, é sempre igual a 1.

Foram suprimidos, por simplicidade, os blocos responsáveis pelas conversões PCM (Lei A ou Lei μ) para PCM uniforme e vice-versa, de modo que o sinal de entrada do codec é um sinal PCM linear de 13 bits/amostra.

3.3 Descrição das modificações

Conforme mencionado na seção anterior, para a implementação do codec a 16 kbit/s baseado na Recomendação G.721 do CCITT, dado que a estrutura do preditor será mantida (embora outras estruturas pudessem contribuir para um melhor desempenho do codec), faz-se necessária a definição de um quantizador de 2 bits e de sua regra de adaptação.

3.3.1 Quantizador

O quantizador utilizado é o quantizador ótimo de Max [4] de 4 níveis para sinais com densidade de probabilidade gaussiana e casado com a variância $\sigma \simeq 3.7$ do sinal diferença normalizado [12]. Sua característica no domínio linear e logarítmico encontra-se nas tabelas 3.1 e 3.2, respectivamente.

Faixa de Entrada Normalizada do Quantizador $ d(k) /Y(k)$	$ I(K) $	Saída Normalizada do Quantizador Inverso $ dq(k) /Y(k)$
$[3.65; +\infty)$	1	5.62
$[0; 3.65)$	0	1.68

Tabela 3.1: Característica Entrada/Saída Normalizada do Quantizador no Domínio Linear.

Faixa de Entrada Normalizada do Quantizador $\log_2 d(k) - y(k)$	$ I(K) $	Saída Normalizada do Quantizador Inverso $\log_2 dq(k) - y(k)$
$[1.868; +\infty)$	1	2.49
$(-\infty; 1.868)$	0	0.75

Tabela 3.2: Característica Entrada/Saída Normalizada do Quantizador no Domínio Logarítmico.

3.3.2 Adaptação do Quantizador

Considerando-se que a 16 kbit/s o fator de escala $y(k)$ é igual a $y_u(k)$ e que este é definido segundo a expressão (2.7), a definição da regra de adaptação do quantizador reside na escolha da função $W(I(k))$. Uma vez definido o quantizador ótimo, a função $W(I(k))$ foi obtida empiricamente através de inúmeras simulações, onde as relações sinal/ruído e sinal/ruído segmental foram monitoradas [12].

A função $W(I(k))$ que resultou na melhor relação sinal/ruído segmental foi escolhida como sendo aquela que fornece o parâmetro ótimo de adaptação do quantizador de 2 bits. Tal função encontra-se definida na tabela 3.3.

I	$W(I)$
1	22.0
0	-0.25

Tabela 3.3: Função $W(I)$.

3.4 Desempenho do Codec

A partir das alterações anteriormente descritas, medidas preliminares de desempenho foram realizadas considerando-se o arquivo de voz usado no Capítulo 2 para avaliação do desempenho do codec a 32 kbit/s.

A relação sinal/ruído segmental obtida para o codec a 16 kbit/s foi aproximadamente 14.9 dB. Comparando-se os resultados preliminares referentes ao desempenho do codec para as taxas de 32 e 16 kbit/s, verifica-se que diminuindo-se o número de bits usados na quantização/codificação do sinal diferença $d(k)$ de 4 para 2, a degradação da relação sinal/ruído é próxima de 12 dB. Testes subjetivos informais mostram que, em relação à taxa de 32 kbit/s, tal algoritmo apresenta uma degradação significativa na qualidade de voz. Uma possível solução para este problema encontra-se no capítulo seguinte, onde é apresentada uma nova estrutura para o quantizador.

Capítulo 4

CODEC ADPCM A 16 KBIT/S COM QUANTIZAÇÃO DE ÁRVORE

4.1 Introdução

No capítulo anterior foram apresentadas modificações introduzidas no algoritmo do codec ADPCM a 32 kbit/s, o que resultou num codec operando à taxa de 16 kbit/s.

Entretanto, nesta taxa tal algoritmo apresenta uma degradação significativa em termos de qualidade objetiva e subjetiva, tornando-se incompatível com a aplicação telefônica.

Comparando-se os algoritmos a 32 kbit/s e a 16 kbit/s em termos de desempenho objetivo, verifica-se que, para um mesmo arquivo de voz, o primeiro apresenta uma relação sinal/ruído aproximadamente 12 dB superior à do segundo. Tal degradação pode ser explicada da seguinte forma: uma vez que o quantizador e o preditor se interagem através de uma malha de realimentação, o ruído de quantização aumentando (têm-se apenas 2 bits para a quantização de cada amostra do sinal diferença), a predição do sinal desejado também é prejudicada. Portanto, sendo o sistema realimentado, são crescentes os efeitos do ruído de quantização na operação do codec.

Considerando-se, conforme mencionado no capítulo anterior, que neste trabalho será mantida a mesma estrutura do preditor da Recomendação G.721, para melhorar o desempenho do algoritmo a 16 kbit/s pode-se recorrer à quantização não instantânea ou multi-dimensional.

Neste sentido, esquemas ADPCM que empregam quantização de árvore (*codificação de árvore*) [13,14], quantização de treliça bem como quantização vetorial, apresentam-se como técnicas de codificação de voz a taxas menores ou iguais a 16 kbit/s altamente promissoras em termos de desempenho e pequeno atraso (atraso do algoritmo).

De fato, um atraso na codificação dá ao sistema capacidade de *busca multicaminho*, que permite identificar, dentre um conjunto de alternativas, a melhor amostra do sinal diferença que será codificada e transmitida. Ao contrário, no sistema ADPCM convencional, tem-se a *busca através de um caminho único*, onde a decisão da melhor amostra do sinal diferença é instantânea e imutável.

Neste capítulo é apresentado um codec ADPCM à taxa de 16 kbit/s empregando quantização de árvore.

4.2 Codificação de árvore - Idéia Básica

Os esquemas ADPCM apresentados nos capítulos anteriores operam sem atraso no sentido de que, para uma amostra do sinal de entrada no instante k , são necessárias informações somente dos instantes $j \leq k$ para viabilizar o processo de codificação.

Na quantização de árvore é utilizado um atraso de L amostras para a codificação de cada amostra do sinal diferença. Desta forma, a codificação de uma amostra deste sinal no instante k é realizada no instante $k + L$, após serem examinadas várias possíveis seqüências do sinal reconstruído neste intervalo de $L + 1$ amostras e se efetuar uma medida de distorção entre estas seqüências e a seqüência correspondente ao sinal fonte. A menor distorção define a k -ésima amostra do sinal diferença a ser transmitida. Em seguida, o procedimento é repetido para a codificação da amostra $k + 1$, e assim sucessivamente.

O diagrama básico do codificador de árvore encontra-se na figura 4.1. Na

quantização de árvore, a estrutura do *gerador de código* é em forma de árvore. Um exemplo de tal estrutura encontra-se na figura 4.2, que ilustra uma árvore associada à palavra código $I(k)$. O sinal ADPCM é constituído por uma sucessão de palavras código $I(k)$ dispostas sequencialmente num CAMINHO específico ao longo da árvore. Por exemplo, nesta figura, o sinal ADPCM correspondente ao caminho assinalado é 1011. Para o codificador ADPCM com quantização de árvore em estudo, tem-se uma árvore associada a cada um dos parâmetros responsáveis pela operação do codec, ou seja, a cada um dos nós de tais árvores é associada uma amostra da medida de distorção, uma amostra do sinal ADPCM codificado $I(k)$, uma amostra do sinal diferença quantizado $d_q(k)$, uma amostra do sinal estimado $s_e(k)$, uma amostra do sinal reconstruído $s_r(k)$, uma amostra do sinal fator de escala de adaptação $y(k)$, etc.

Uma vez que na quantização de árvore emprega-se um atraso de L amostras, define-se que a árvore tem PROFUNDIDADE L . Define-se também *taxa de codificação de árvore* como [15]

$$R = \frac{1}{\beta_t} \log_2 \alpha_t, \quad (4.1)$$

onde β_t é o número de amostras associado a cada nó da árvore e α_t é o número de RAMOS oriundos de cada nó da árvore. Verifica-se que R é independente da profundidade da árvore L . Especificamente para o sinal ADPCM a 16 kbit/s com amostragem a 8 kHz, $\alpha_t = 4$ e $\beta_t = 1$, de modo que a taxa de codificação de árvore é $R = 2$ bit/amostra.

Considerando-se ainda a figura 4.1, a função do *algoritmo de busca* é seleccionar, dentre várias possibilidades, a melhor seqüência do sinal reconstruído no intervalo de $L + 1$ amostras. Ou seja, neste intervalo efetuam-se *medidas de distorção* entre a seqüência fonte e as possíveis seqüências do sinal reconstruído e selecciona-se a seqüência correspondente à menor distorção. Após esta seleção, fica decidido que no instante $k + L$ a palavra código com a informação do instante k a ser transmitida é a $I(k)$ correspondente ao nó inicial da seqüência de menor distorção ao longo da árvore.

Através do que foi apresentado, verifica-se que a idéia básica de codificação de árvore reside no fato de que um atraso de codificação dá ao sistema capacidade de *busca multi-caminho* ("*multipath search*"), tornando-o mais eficiente na codificação do sinal diferença. Tal técnica é conhecida na literatura como *Codificação com*

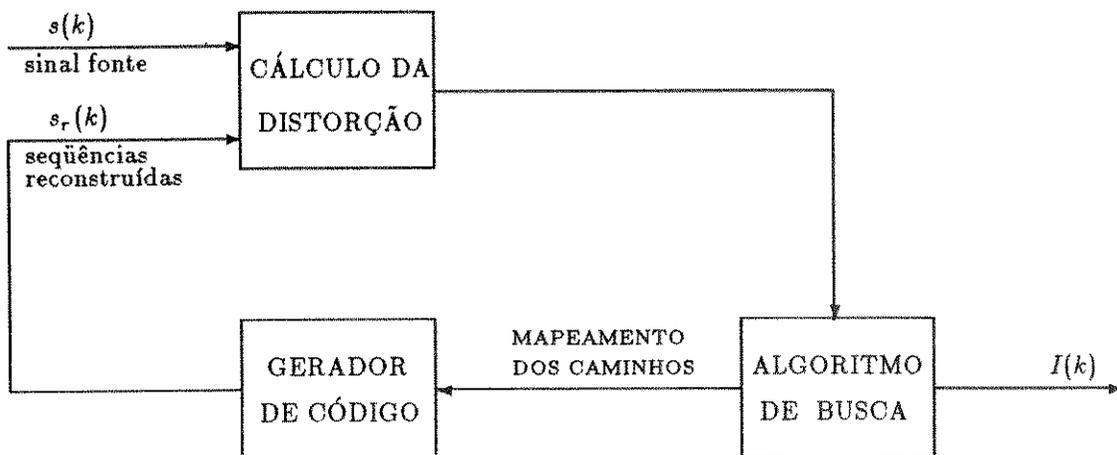


Figura 4.1: Diagrama funcional básico do codificador de árvore.

Decisão Atrasada (“*Delayed Decision Coding*”) ou *Codificação com Busca Multi-caminho* (“*Multipath Search Coding*”). Assim, uma vez que no codec ADPCM a 16 kbit/s o sinal diferença é codificado com 2 bits por amostra ($R = 2$), têm-se 4^{L+1} possíveis seqüências de comprimento $L + 1$, sendo testado um número restrito delas para se definir a melhor amostra do sinal diferença no instante k . Para $L = 0$, recai-se no sistema ADPCM convencional apresentado no capítulo anterior, onde a quantização/codificação do sinal diferença $d(k)$ é instantânea, isto é, realizada no instante k . Especificamente neste caso, tem-se uma *Codificação com busca através de um único caminho* (“*Single Path Search Coding*”), isto é, escolhas instantâneas e irrevogáveis são executadas. Concluindo, o atraso de codificação L permite que na definição de uma amostra do sinal diferença no instante k sejam consideradas amostras futuras $(k, k + 1, \dots, k + L)$, justificando-se, então, a melhoria de desempenho em relação ao esquema ADPCM convencional ($L = 0$).

4.3 Codificador ADPCM com quantização de árvore

Segundo a figura 4.1, que representa a estrutura básica de um codificador de árvore, verifica-se que para o projeto deste codificador faz-se necessária a escolha de um gerador de código, de um método para medida de distorção e de um algoritmo de busca através da árvore.

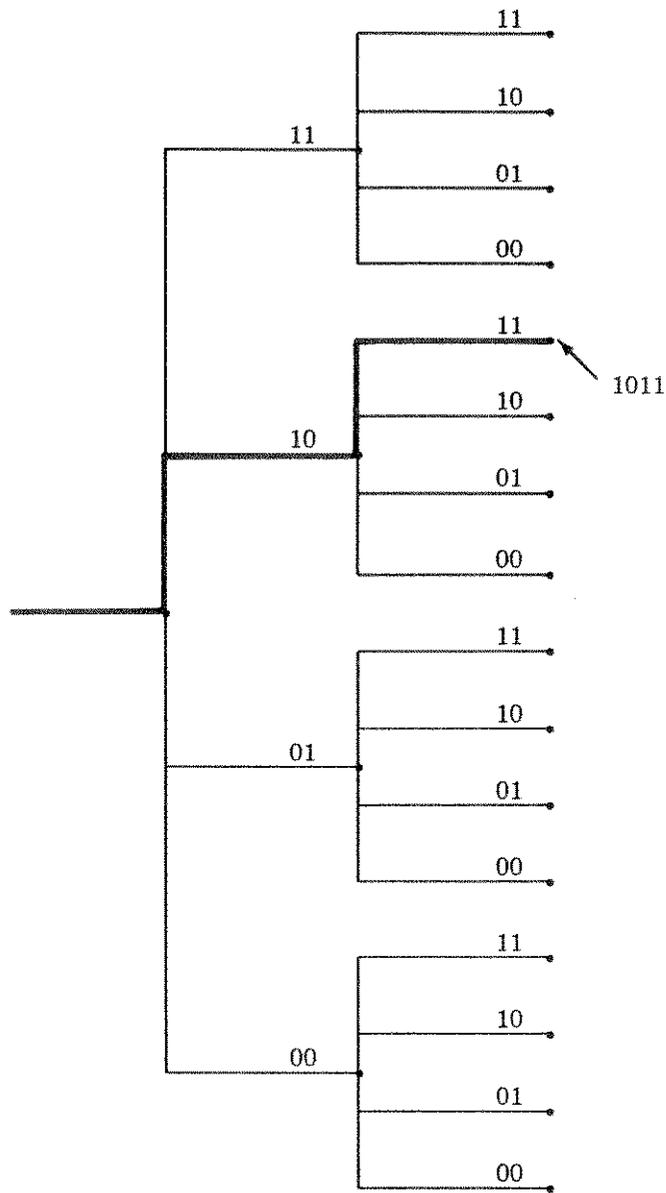


Figura 4.2: Estrutura de árvore binária.

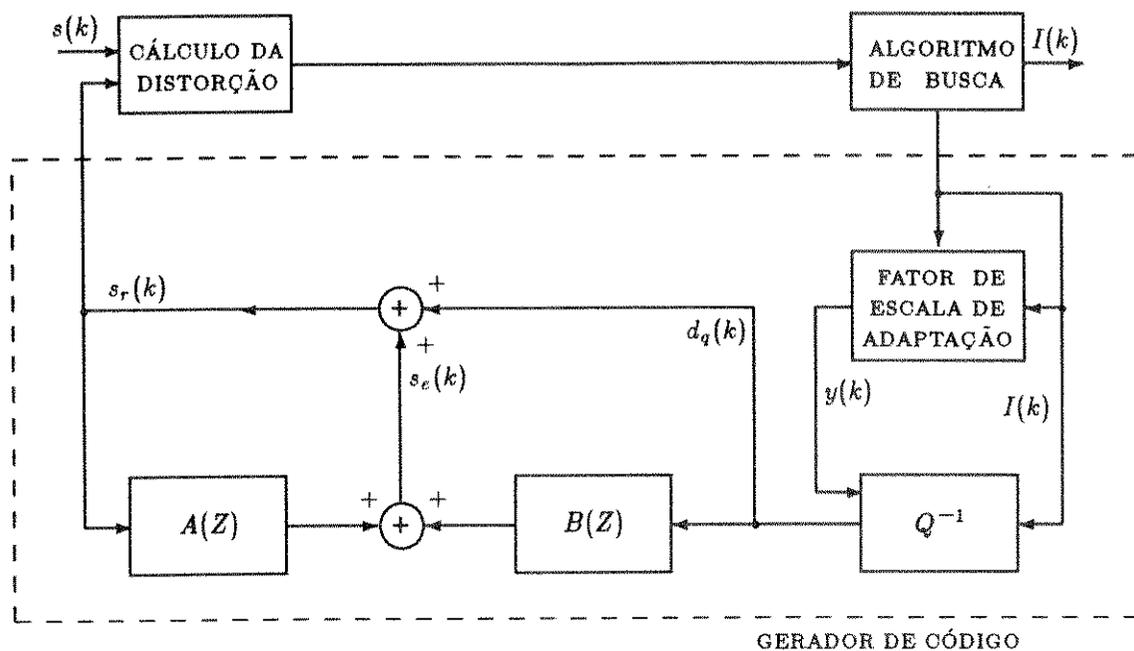


Figura 4.3: Diagrama de blocos do codificador ADPCM com quantização de árvore.

Na figura 4.3 é mostrado o diagrama de blocos do codificador ADPCM com quantização de árvore, objeto de estudo desta seção.

4.3.1 Gerador de Código

Nos capítulos anteriores verificou-se que no esquema ADPCM a estrutura do decodificador é idêntica à malha de realimentação do codificador. Na codificação ADPCM com quantização de árvore isto também ocorre, de modo que a malha de realimentação do codificador pode funcionar como um gerador de código, uma vez que está presente na seção de transmissão e também na seção de recepção. Desta forma, de acordo com a figura 4.3, o gerador de código é constituído de um preditor e de um quantizador, ambos adaptativos.

Por se tratar de esquema ADPCM, o gerador de código é classificado como *determinístico*, pois é responsável pela geração de um conjunto extremamente restrito de seqüências do sinal reconstruído $s_r(k)$. De fato, cada amostra do sinal diferença quantizado $d_q(k)$ pode assumir um dos 2^R valores, onde R , como foi

mencionado, é o número de bits por amostra usado na quantização/codificação do sinal diferença $d(k)$.

As equações básicas que descrevem o preditor são as mesmas presentes na seção 2.2.2 do Capítulo 2, e o quantizador adaptativo é o mesmo descrito no Capítulo 3.

4.3.2 Algoritmo de Busca

Uma vez selecionado o gerador de código, o próximo passo é selecionar o algoritmo responsável pela busca da melhor seqüência de excitação na árvore.

Primeiramente pensou-se num procedimento de *busca exaustiva*, onde *todas* as possíveis seqüências de saída do gerador de código $\{s_r(k) \dots s_r(k+L)\}$ seriam examinadas. Entretanto, se o número destas seqüências é elevado, a complexidade do algoritmo também aumenta (no Apêndice V encontra-se um estudo sobre complexidade em codificação com busca multi-caminho).

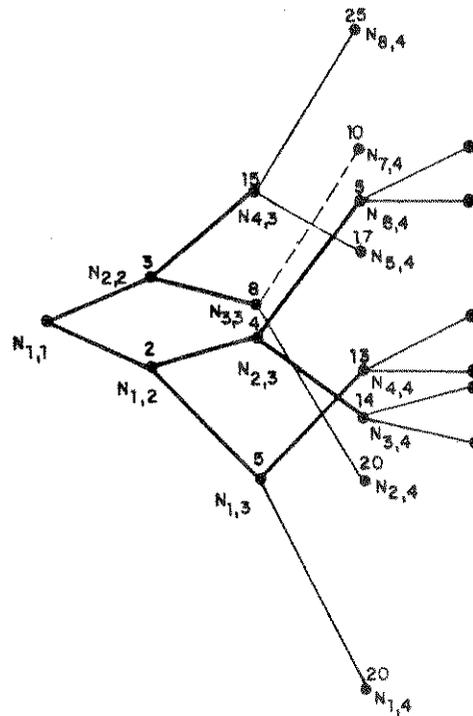
Optou-se, portanto, por um procedimento de *busca não exaustiva*, onde são examinadas apenas as seqüências $\{s_r(k) \dots s_r(k+L)\}$ mais promissoras, isto é, aquelas que, numa janela de observação de $L+1$ amostras, apresentam as menores distorções em relação à seqüência fonte $\{s(k) \dots s(k+L)\}$. Assim, tal procedimento resulta numa seleção sub-ótima, que, por outro lado, pode não representar uma perda significativa no cômputo geral do desempenho do codec.

O procedimento de busca não exaustiva adotado para o codificador em questão foi o *Algoritmo (M, L)* [16,17,18,19], que é descrito sucintamente a seguir.

1. No início do processo, as 4^{L+1} seqüências $\{I(k) \dots I(k+L)\}$ são submetidas ao gerador de código, originando 4^{L+1} possíveis seqüências reconstruídas $\{s_r(k) \dots s_r(k+L)\}$. Cada uma destas seqüências corresponde a um caminho de comprimento $L+1$ na árvore (L é definido como profundidade da árvore).

2. Efetua-se uma medida de distorção entre a seqüência fonte $\{s(k) \dots s(k+L)\}$ e as possíveis seqüências reconstruídas $\{s_r(k) \dots s_r(k+L)\}$. O valor de distorção obtido é associado ao nó $k+L$ de cada seqüência reconstruída.

3. A seqüência reconstruída de menor distorção define um caminho de comprimento $L+1$ na árvore e o valor de $I(k)$ associado ao nó inicial deste caminho



INSTANTE k	:	1	2	3	4	5
NÓ TRANSMITIDO	:			$N_{1,1}$	$N_{1,2}$	

Figura 4.4: Operação do Algoritmo (M, L) para $M = 4, L = 2$ e codificação com 1 bit por amostra.

é enviado ao receptor.

4. Antes da geração da amostra seguinte $I(k + 1)$, são retidas as M seqüências $\{I(k) \dots I(k + L)\}$ (ou caminhos) de menor distorção. Destas M seqüências, são selecionadas apenas as que se originam do nó definido no passo 3.

5. Retorna-se ao passo 1, mas agora considerando-se as possíveis seqüências $\{I(k + 1) \dots I(k + 1 + L)\}$ cujos elementos $\{I(k + 1) \dots I(k + L)\}$ correspondem às seqüências definidas no passo 4.

O procedimento se repete, sendo que para a geração de cada amostra $I(k)$ faz-se necessário efetuar uma medida de distorção entre a seqüência fonte e no máximo $4M$ seqüências reconstruídas.

A figura 4.4 ilustra a operação do Algoritmo (M, L) , para $M = 4, L = 2$ e considerando-se um número de bits por amostra igual a 1 (que corresponde à taxa de 8kbit/s) para facilitar a explicação. Nesta figura, é indicado sobre cada nó, o valor da distorção a ele associado.

No instante $k = 4$, os $M = 4$ nós que apresentam as menores distorções e que deveriam ser selecionados são $N_{6,4}, N_{7,4}, N_{4,4}$ e $N_{3,4}$, sendo que o nó $N_{6,4}$

corresponde à mínima distorção. De acordo com o Algoritmo (M, L) , verifica-se que à profundidade $L = 2$, o nó $N_{1,2}$ foi quem originou o nó $N_{6,4}$. Desta forma, no instante $k = 4$, o valor de $I(k)$ no nó $N_{1,2}$ (referente ao instante $k = 2$) é transmitido.

Verifica-se, então, se os nós selecionados $N_{7,4}$, $N_{4,4}$ e $N_{3,4}$ são também oriundos do nó $N_{1,2}$. Uma vez que o nó $N_{7,4}$ não o é, ele e o ramo a ele ligado são eliminados, embora o caminho correspondente apresente a segunda menor distorção. Para o próximo instante de amostragem $k = 5$, dado que cada nó selecionado origina 2 ramos, têm-se 6 ramos e, portanto, 6 seqüências reconstruídas de comprimento $L + 1 = 3$ a serem consideradas.

4.3.3 Cálculo da Distorção

O critério adotado para o cálculo da distorção entre a seqüência fonte $\{s(k)...s(k + L)\}$ e as possíveis seqüências reconstruídas $\{s_r(k)...s_r(k + L)\}$ foi o erro quadrático definido como

$$dist(k + L) = \sum_{i=k}^{k+L} [s(i) - s_r(i)]^2, \quad (4.2)$$

onde L , conforme descrito anteriormente, é a profundidade da árvore. Verifica-se ainda que o valor obtido é associado ao instante $k + L$ da árvore de cada seqüência reconstruída.

A seguir é apresentado o método adotado para reduzir a complexidade do cálculo da distorção [16].

Seja $e(k) \triangleq s(k) - s_r(k)$, de modo que pode-se reescrever a expressão (4.2) da seguinte forma:

$$dist(k + L) = \sum_{i=k}^{k+L} e^2(i). \quad (4.3)$$

Expandindo-se a expressão (4.3), tem-se

$$dist(k + L) = e^2(k) + e^2(k + 1) + e^2(k + 2) + \dots + e^2(k + L). \quad (4.4)$$

A distorção associada ao instante $k + L + 1$ (próximo instante de amostragem em relação a $k + L$) é

$$dist(k + L + 1) = \sum_{i=k+1}^{k+L+1} e^2(i), \quad (4.5)$$

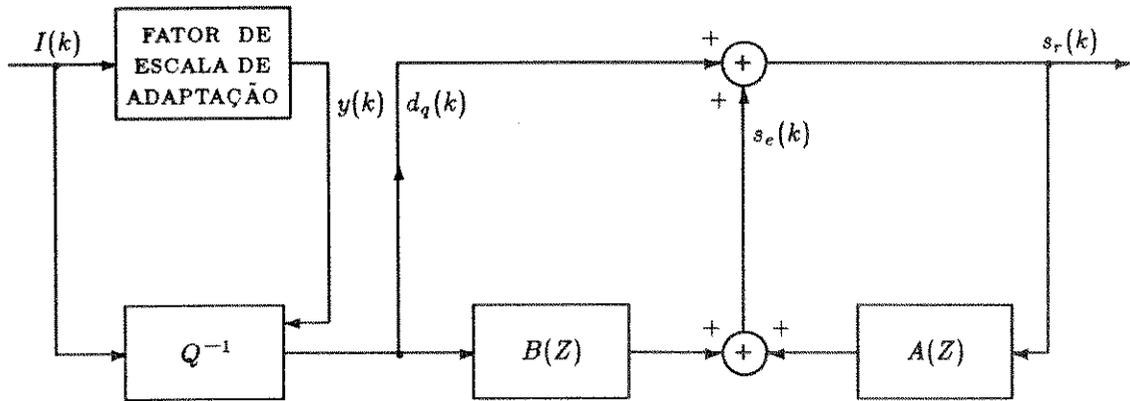


Figura 4.5: Diagrama de blocos do decodificador ADPCM com quantização de árvore.

cuja expressão na forma expandida é dada por

$$dist(k + L + 1) = e^2(k + 1) + e^2(k + 2) + \dots + e^2(k + L) + e^2(k + L + 1). \quad (4.6)$$

Combinando-se as expressões (4.4) e (4.6), obtém-se

$$dist(k + L + 1) = dist(k + L) + e^2(k + L + 1) - e^2(k), \quad (4.7)$$

que é a expressão empregada na implementação do algoritmo em questão. Portanto, verifica-se que através do procedimento iterativo descrito pela expressão (4.7) obtém-se uma redução significativa no número de multiplicações e de adições.

4.4 Decodificador ADPCM com quantização de árvore

Na figura 4.5 encontra-se o diagrama de blocos detalhado do decodificador de árvore. Verifica-se, conforme mencionado anteriormente, que a estrutura do decodificador é idêntica à malha de realimentação do codificador de árvore.

O sinal ADPCM $I(k)$ correspondente ao melhor caminho é recebido, e, através de processamentos análogos aos realizados no bloco codificador, o sinal reconstruído $s_r(k)$ é obtido.

4.5 Atraso do algoritmo

A princípio, vislumbrou-se que o codec em estudo teria aplicação primária no Sistema Rádio Acesso Celular Digital (telefonia rural), bem como no Sistema de Comunicação Celular Digital (rádio móvel digital). Pensou-se também em Equipamentos Multiplicadores de Circuitos Digitais (“Digital Circuits Multiplication Equipaments - DCME”), sistemas de armazenamento e transmissão (“store and forward systems”), serviços de videofone e mensagens de voz para anúncios gravados.

Para as três primeiras aplicações acima deve-se considerar a possibilidade do codec ADPCM em questão fazer parte de uma conexão onde uma série de outros equipamentos de transmissão estão dispostos em cascata. Neste caso, um requisito importante é o atraso total de transmissão ponta a ponta, que é justificado considerando-se os seguintes aspectos: este atraso de transmissão não deve exceder o limite de 400 ms especificado pelo CCITT através da Recomendação G.114 [20]; se este atraso de propagação ultrapassar 25 ms e conseqüentemente o eco que porventura possa existir na conexão também ultrapassar 50 ms, a inteligibilidade de uma conversação poderá ficar comprometida¹. Sob este último aspecto, dependendo da aplicação, quanto menor for o atraso do codec, mais equipamentos podem ser conectados em cascata numa conexão, sem a necessidade de se recorrer a dispositivos de controle de eco. Neste sentido, quando da definição dos requisitos de aplicação e desempenho para o codec a 16 kbit/s a ser padronizado pelo CCITT [22], enfatizou-se que em aplicações como RDSI e telefone sem fio o atraso total do codec deveria ser inferior a 5 ms.

Em função destas considerações e levando-se em conta a complexidade computacional, para o codec em estudo adotou-se um atraso máximo de 2 ms. Considerando-se que a freqüência de amostragem é 8000 Hz, este atraso corresponde a $L = 16$ amostras.

4.6 Desempenho do codec ADPCM com quantização de árvore

O arquivo de voz usado neste experimento é o mesmo que foi usado nos capítulos anteriores para a avaliação preliminar dos codecs apresentados.

¹Nesta situação, segundo a Recomendação G.161 do CCITT [21], recomenda-se a utilização de supressores ou canceladores de eco.

Inicialmente analisou-se o desempenho do codec em função de M e L através da figura 4.6 [23], onde encontram-se as curvas de relação sinal/ruído segmental para diferentes valores de M e L . Observa-se que a relação sinal/ruído segmental é uma função crescente de L e tende a uma saturação em função do parâmetro M . Para $M = 16$ e $L = 16$, a relação sinal/ruído é aproximadamente 18.5 dB e portanto 3.5 dB superior ao do codec ADPCM convencional a 16 kbit/s ($M = 1, L = 0$).

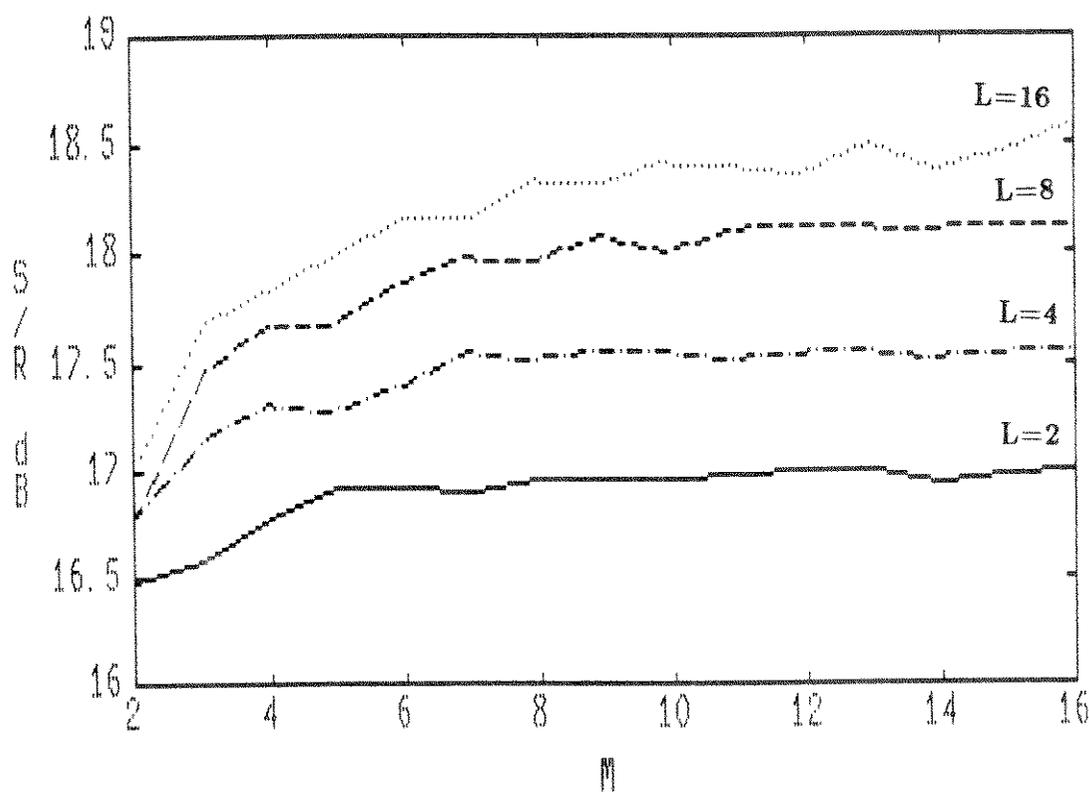


Figura 4.6: Desempenho objetivo do codec em função de M e L .

Como ilustração, nas figuras 4.7 e 4.8 têm-se, respectivamente, as formas de onda e espectros LPC² de um segmento sonoro do sinal de voz original e do sinal reconstruído para os codecs ADPCM citados, convencional ($M = 1, L = 0$) e com quantização de árvore ($M = 16, L = 16$), ambos operando a 16 kbit/s. Tal segmento é constituído de 4 blocos, cada um com 64 amostras, tendo portanto

²O "software" utilizado para a geração das figuras é o ILS (Iterative Laboratory System) versão 6.0, desenvolvido pela Signal Technology Incorporation. Tal ferramenta é capacitada para atuar em análise de voz bem como processamento digital de sinais.

uma duração de 32 ms.

Testes subjetivos informais comprovam o resultado dos testes objetivos, ou seja, o desempenho do codec tende a uma saturação em função de M , notadamente para $M \geq 8$ e a qualidade do codec é melhor para valores crescentes de L , sendo bastante evidente a superioridade deste codec em relação ao codec ADPCM convencional a 16 kbit/s. Entretanto, apesar da alta relação sinal/ruído segmental e da melhoria subjetiva obtida, verifica-se a presença de um ruído de fundo ainda significativo.

A seguir, o comportamento do ruído de quantização nos sistemas ADPCM estudados é analisado em função do preditor e dos quantizadores empregados. Considerando-se o fato de que se a ordem do preditor é suficientemente elevada, o erro de predição consiste de amostras incorrelatas (ruído branco), no caso dos sistemas ADPCM apresentados, sendo o preditor de curto prazo, ele não consegue tirar a correlação de longo prazo que é característica de sons sonoros e que está associada ao período de “pitch”. Tal fato é ilustrado através das figuras 4.9 e 4.10, onde encontram-se, respectivamente, as formas de onda e os espectros FFT³ de um segmento sonoro do sinal de voz original, sinal reconstruído e erro de quantização, para o codec ADPCM convencional a 16 kbit/s ($M = 1, L = 0$). Neste caso, o segmento de análise é constituído de 2 blocos de 64 amostras, correspondendo à duração de 16 ms. Ao se observar tanto a figura 4.9 como a figura 4.10, verifica-se que o erro de quantização apresenta uma componente correlacionada com o sinal. Já para sons não sonoros, que não apresentam um comportamento periódico, o ruído de quantização apresenta uma característica espectral mais plana. Tal fato é ilustrado pelas figuras 4.11 e 4.12, onde encontram-se, respectivamente, as formas de onda e os espectros FFT de um segmento não sonoro do sinal de voz original, sinal reconstruído e erro de quantização, para o codec ADPCM convencional a 16 kbit/s ($L = 0, M = 1$). A duração de tal segmento é a mesma das figuras 4.9 e 4.10.

Com a introdução da quantização de árvore no codec ADPCM convencional, o sistema incorpora a capacidade de busca multi-caminho, o que contribui para uma redução do ruído de quantização. Nas figuras 4.13 e 4.14 encontram-se, para o codec ADPCM com quantização de árvore ($M = 16, L = 16$), respectivamente, as formas de onda e os espectros FFT do sinal reconstruído e do erro de quantização correspondentes ao segmento de voz sonoro das figuras 4.9 e 4.10. Comparando-se

³Idem 2.

as figuras 4.13 e 4.14 com 4.9 e 4.10, fica evidente a redução do ruído de quantização, embora ele ainda continue correlacionado com o sinal (para sons sonoros).

É sabido que o sinal de voz reproduzido por um sistema ADPCM convencional pode apresentar uma boa qualidade subjetiva, embora com um elevado nível de ruído de quantização, devido ao fato de que em sons sonoros este ruído é correlacionado com o sinal e, portanto, por ele mascarado. Nas primeiras investigações relatadas sobre o emprego de quantização de árvore em sistemas DPCM [13,14], dizia-se que, apesar da melhoria na relação sinal/ruído, a qualidade subjetiva da voz era inferior. O motivo alegado era que o ruído de quantização produzido pela quantização de árvore era branco. Os resultados por nós obtidos permitem refutar estas informações e instituir que o mau resultado dessas investigações deveu-se basicamente ao uso de preditores fixos.

Na figura 4.15 são ilustradas as formas de onda do erro de quantização para os codecs ADPCM convencional e com quantização de árvore correspondentes ao segmento sonoro da figura 4.7 (duração 32 ms) com o objetivo de enfatizar a redução do ruído de quantização e a existência de uma correlação residual na quantização de árvore.

Apesar da melhoria objetiva e subjetiva atingida com a quantização de árvore, o ruído do codec ainda é perceptível. Para diminuir a percepção deste ruído pode-se recorrer a uma pós-filtragem do sinal reconstruído no receptor ou uma ponderação do ruído no transmissor. Tais técnicas serão abordadas nos próximos capítulos.

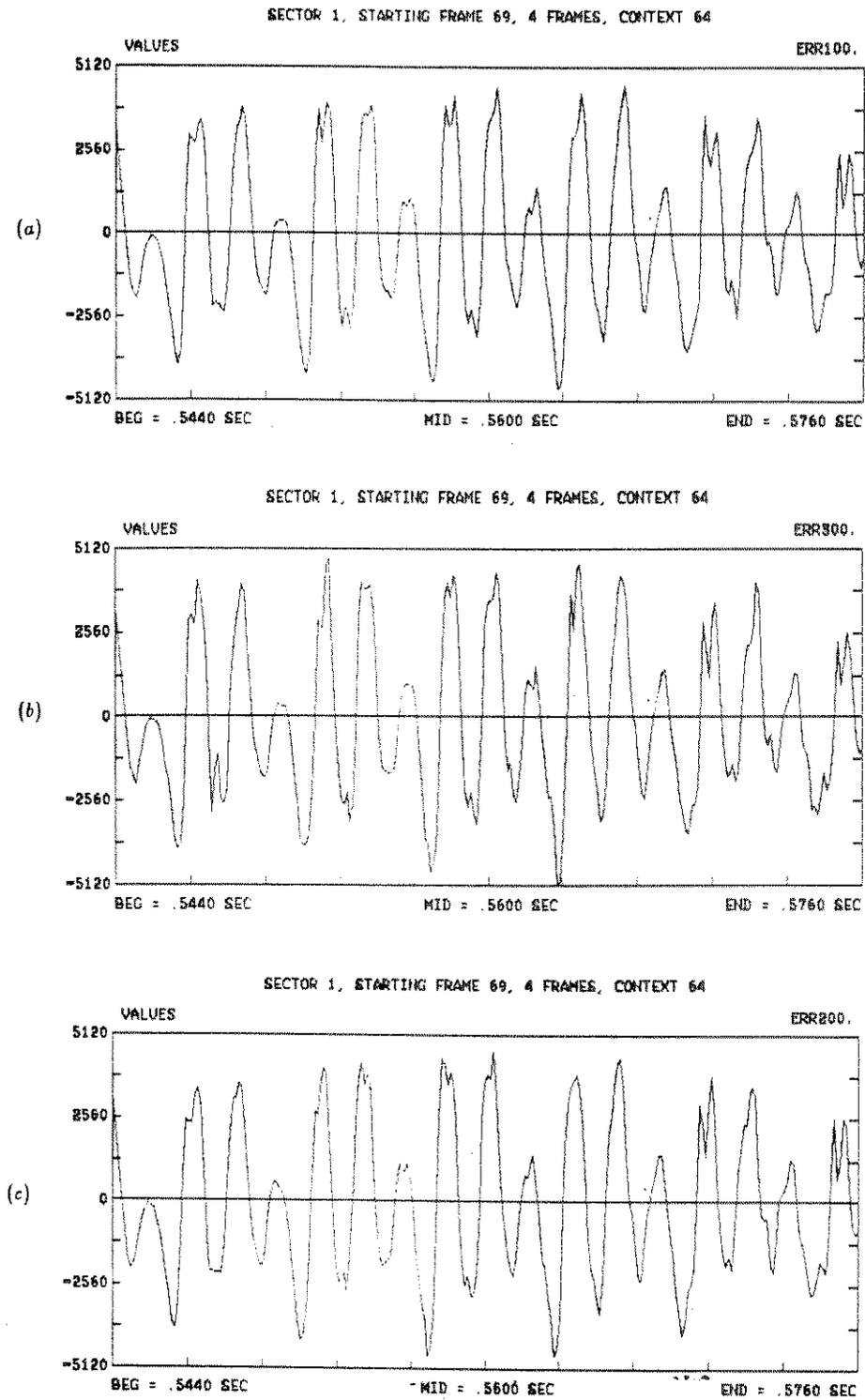


Figura 4.7: Formas de onda de um segmento sonoro de voz: (a) sinal de voz original; (b) sinal reconstruído do codec ADPCM convencional a 16 kbit/s ($M = 1, L = 0$); (c) sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$).

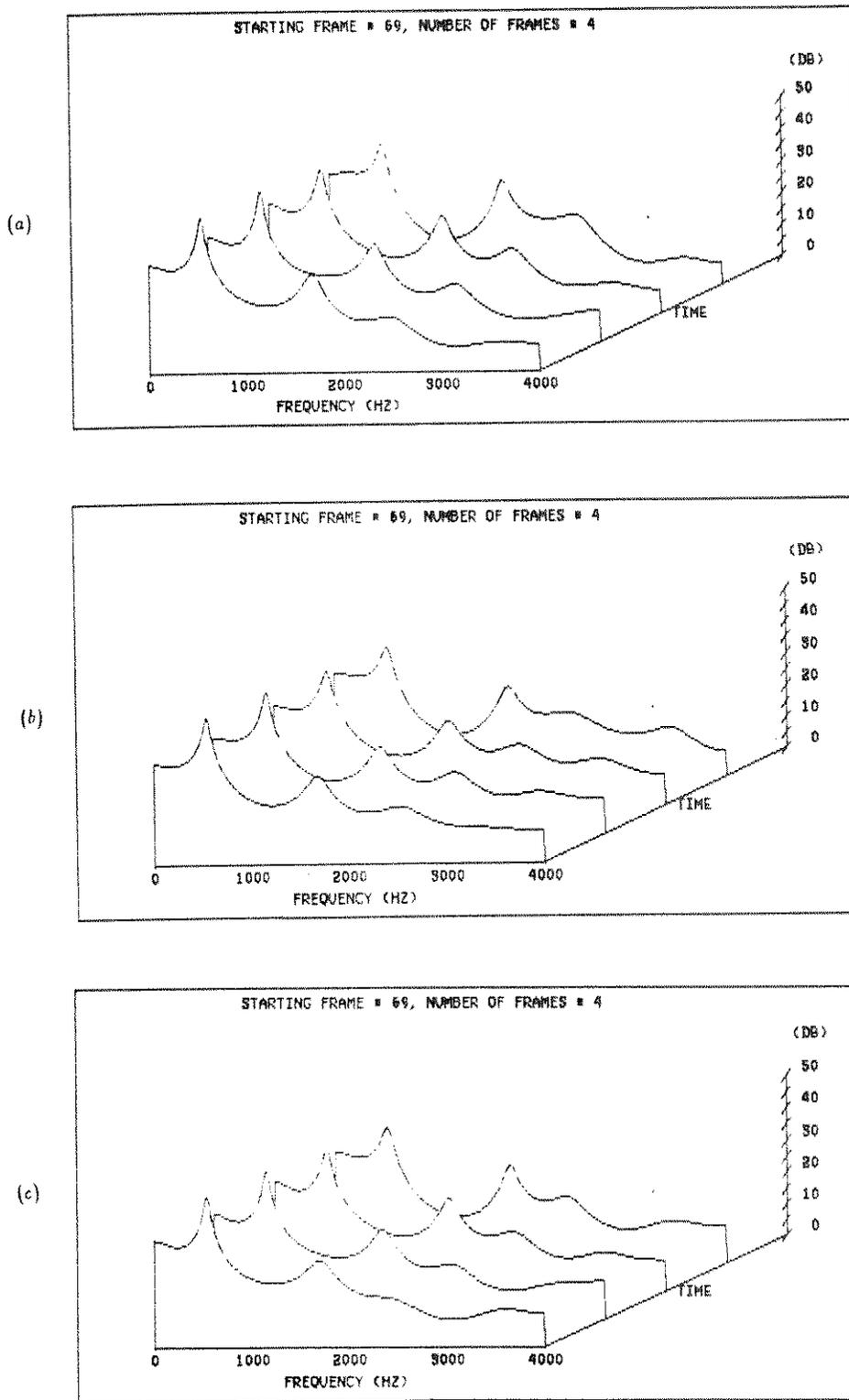


Figura 4.8: Espectros LPC de um segmento sonoro de voz: (a) sinal de voz original; (b) sinal reconstruído do codec ADPCM convencional a 16 kbit/s ($M = 1, L = 0$); (c) sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$).

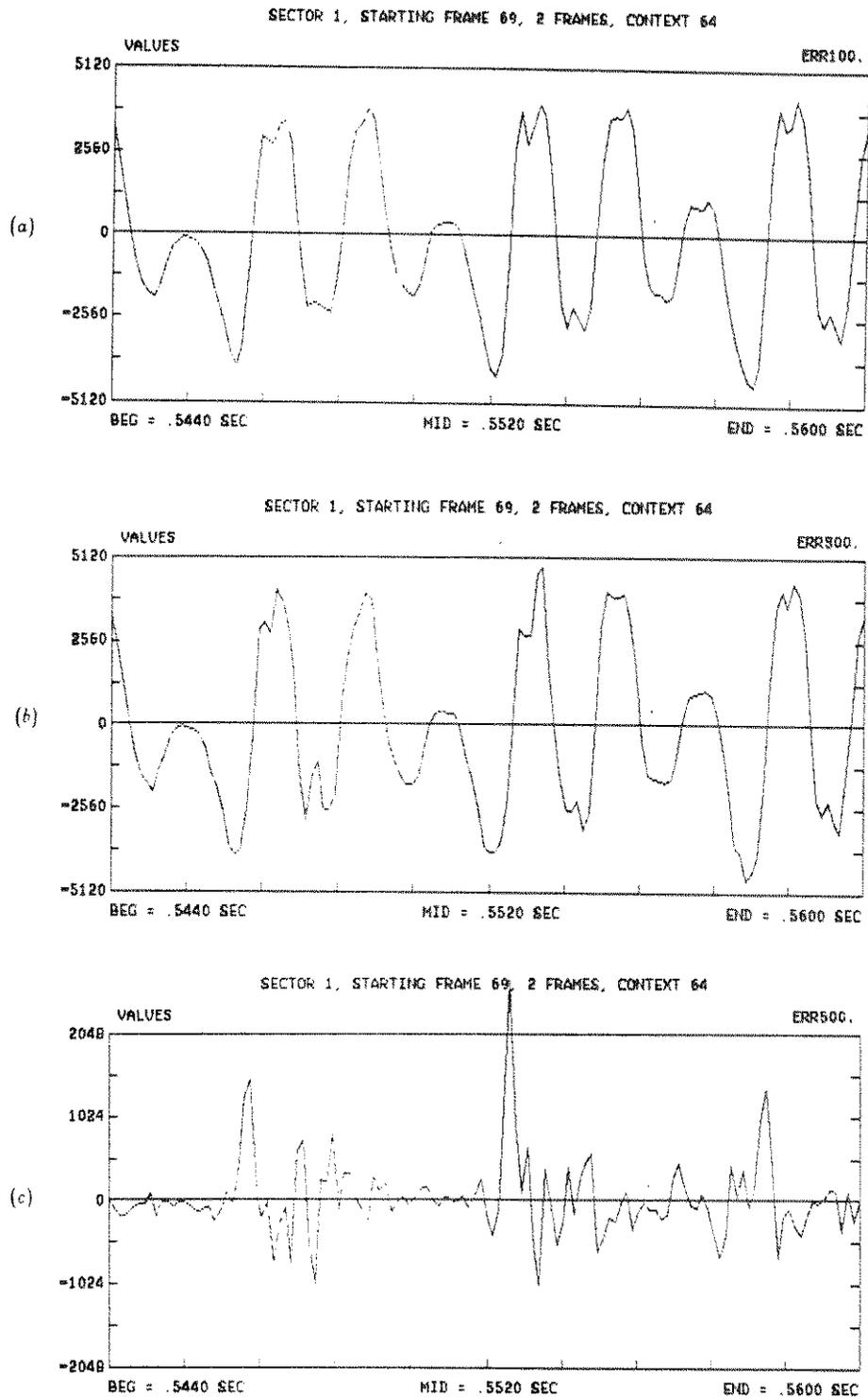


Figura 4.9: Formas de onda de um segmento sonoro de voz: (a) sinal de voz original; (b) sinal reconstruído e (c) erro de quantização referentes ao codec ADPCM convencional a 16 kbit/s ($M = 1, L = 0$).

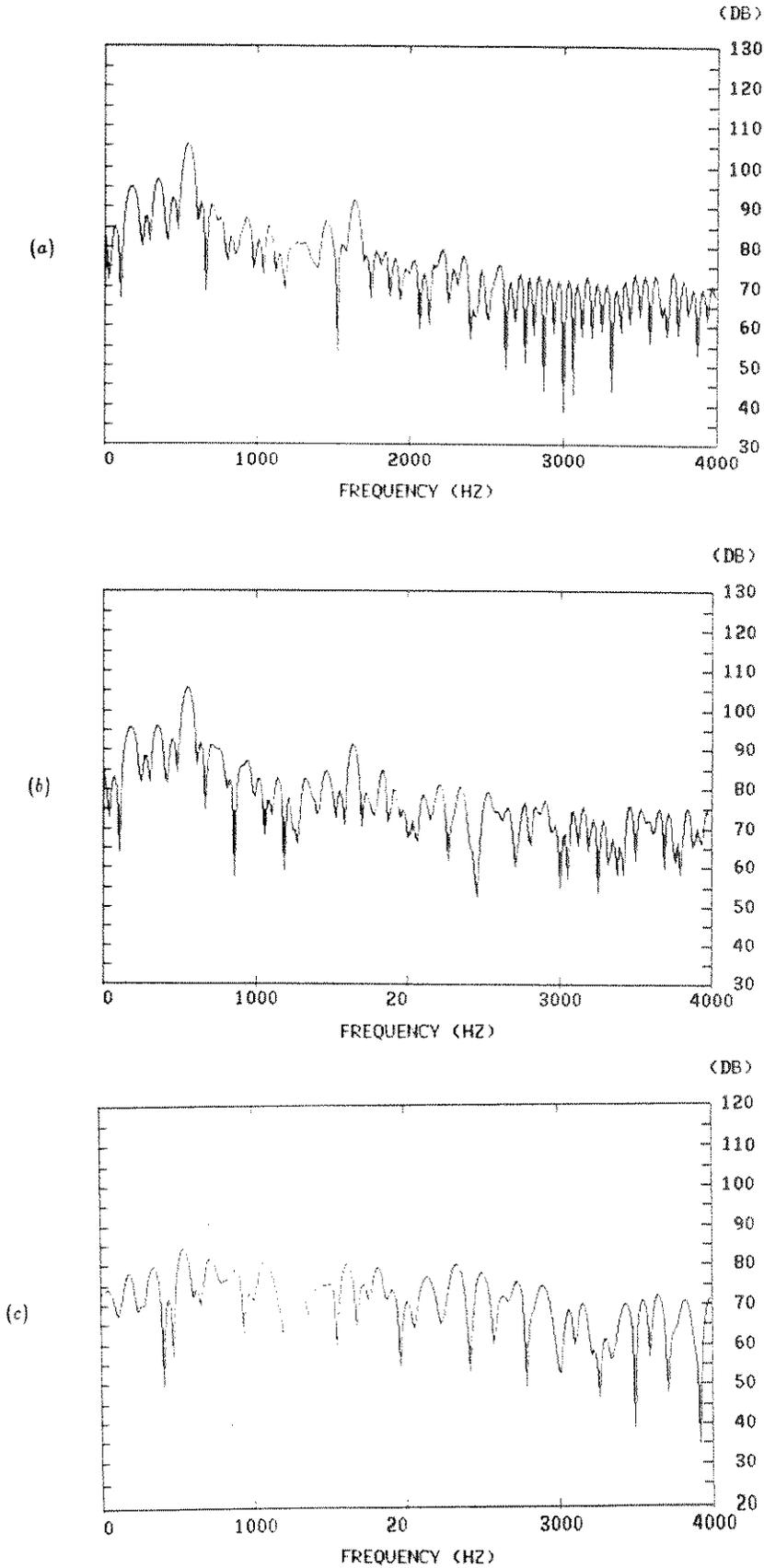


Figura 4.10: Espectros FFT de um segmento sonoro de voz: (a) sinal de voz original; (b) sinal reconstruído e (c) erro de quantização referentes ao codec ADPCM convencional a 16 kbit/s ($M = 1, L = 0$).

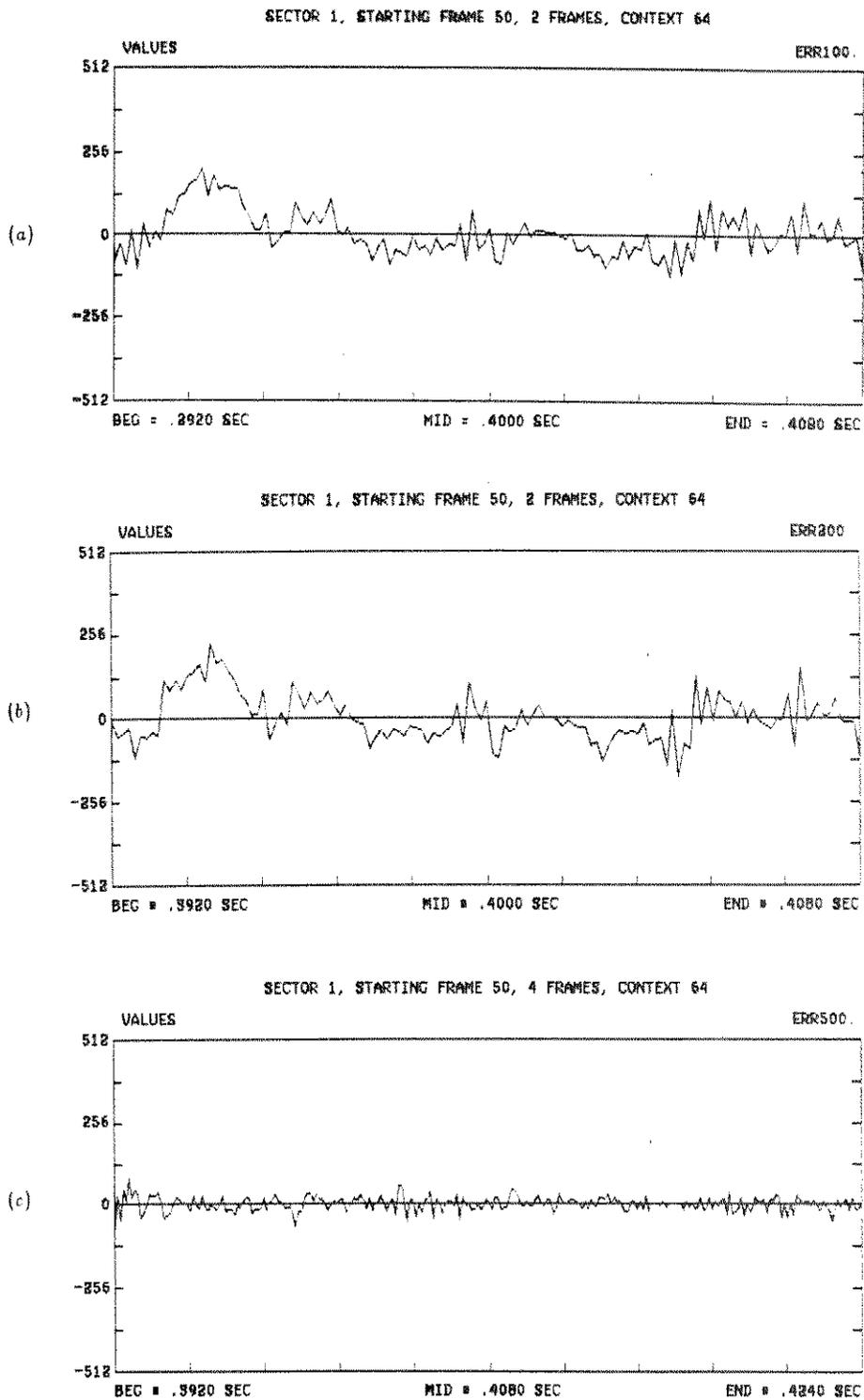


Figura 4.11: Formas de onda de um segmento não sonoro de voz: (a) sinal de voz original; (b) sinal reconstruído e (c) erro de quantização referentes ao codec ADPCM convencional a 16 kbit/s ($M = 1, L = 0$).

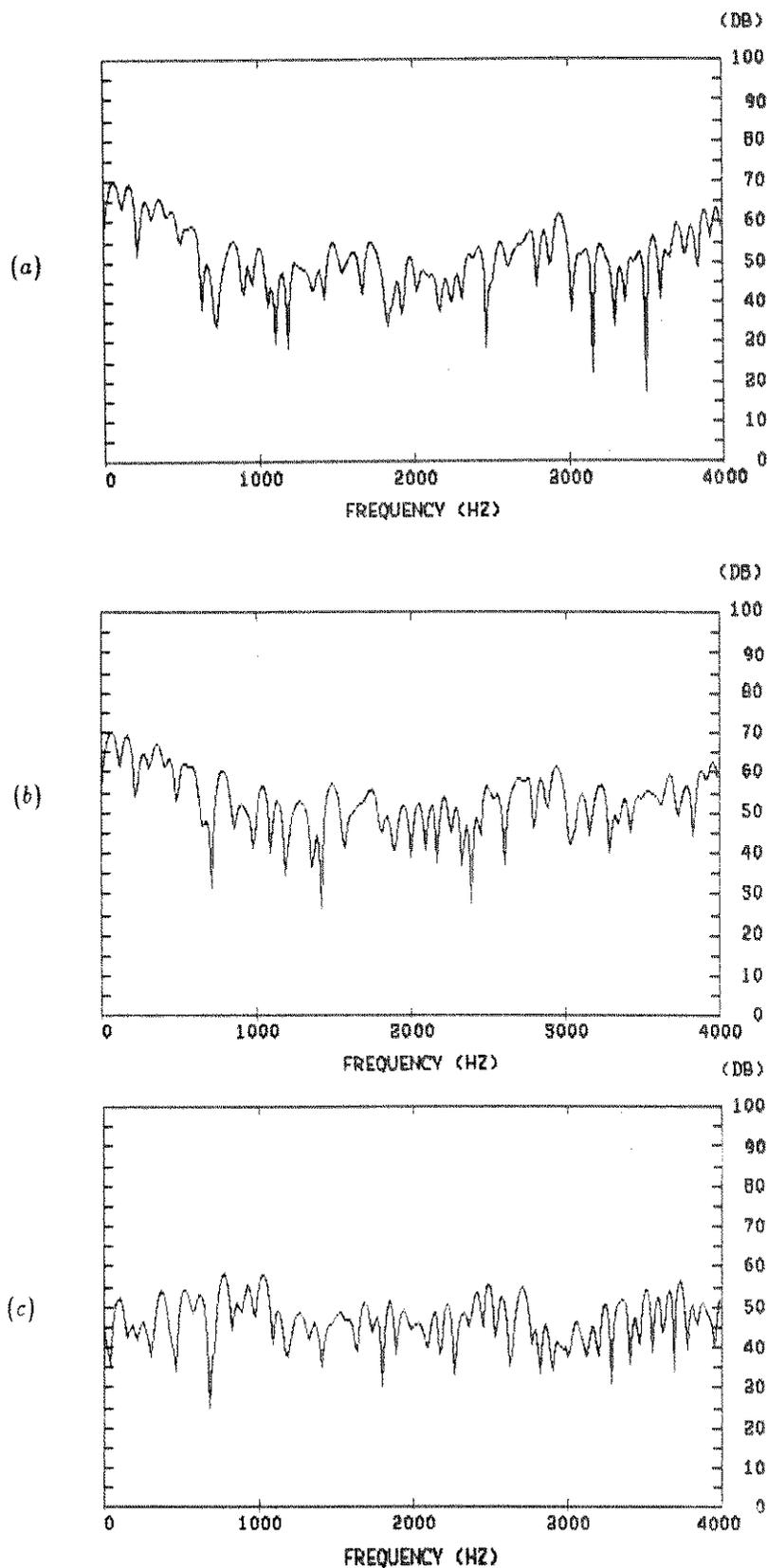


Figura 4.12: Espectros FFT de um segmento não sonoro de voz: (a) sinal de voz original; (b) sinal reconstruído e (c) erro de quantização referentes ao codec ADPCM convencional a 16 kbit/s ($M = 1, L = 0$).

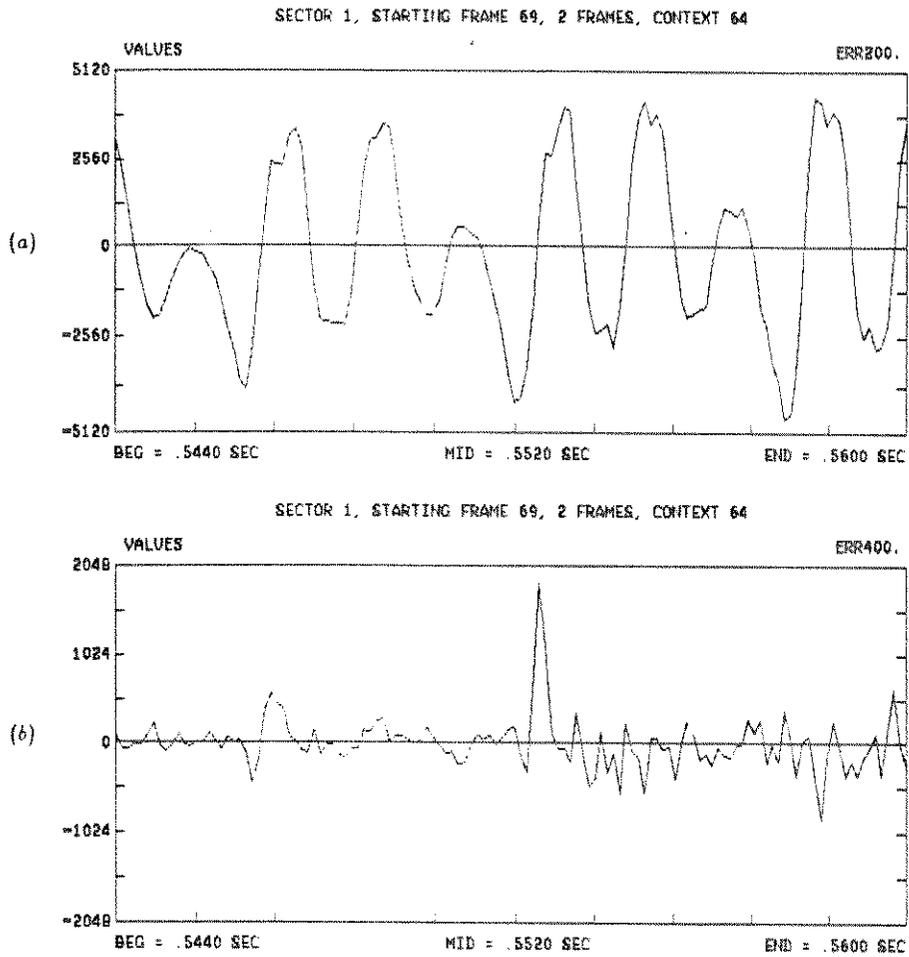


Figura 4.13: Formas de onda de um segmento sonoro de voz: (a) sinal reconstruído e (b) erro de quantização referentes ao codec ADPCM com quantização de árvore ($M = L = 16$).

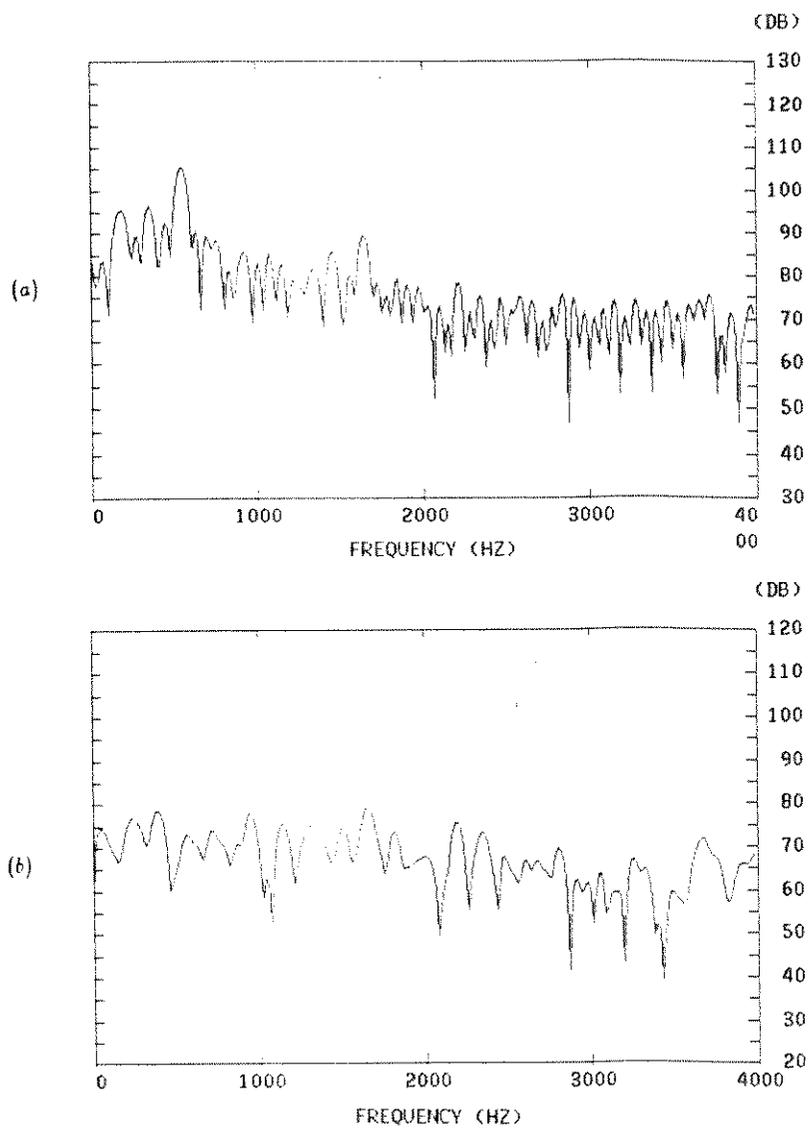


Figura 4.14: Espectros FFT de um segmento sonoro de voz: (a) sinal reconstruído e (b) erro de quantização referentes ao codec ADPCM com quantização de árvore ($M = L = 16$).

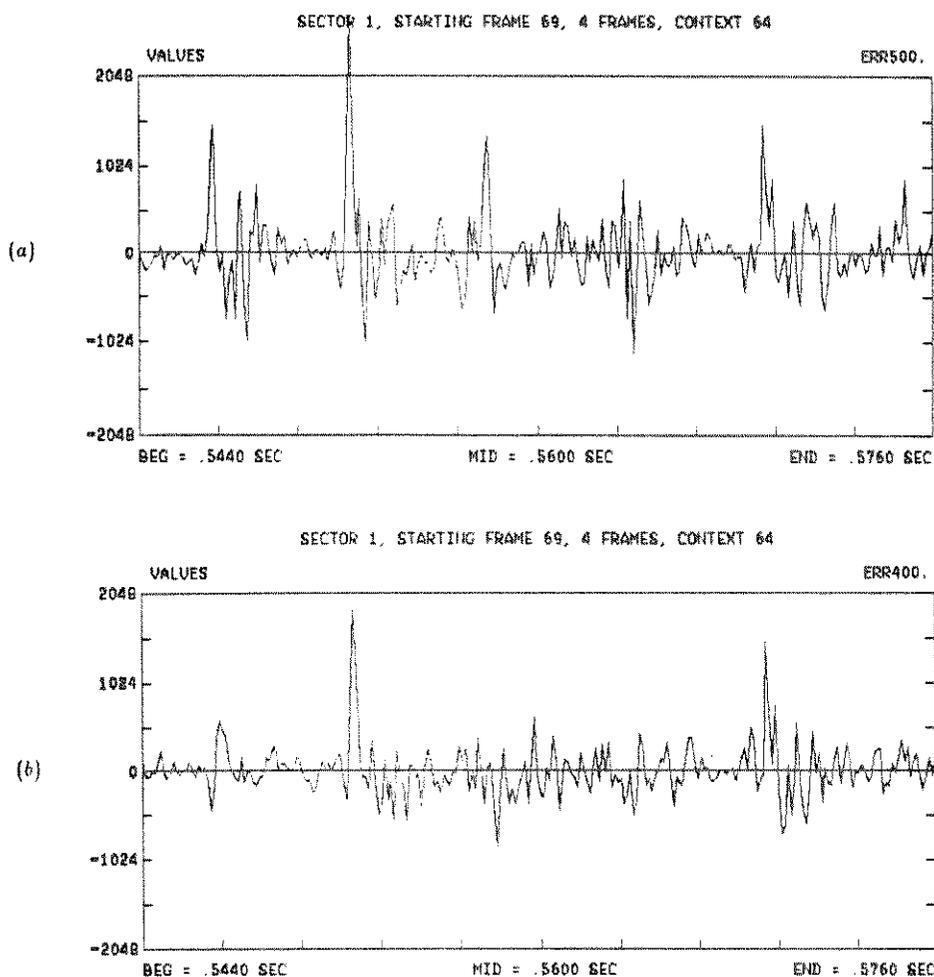


Figura 4.15: Formas de onda de um segmento sonoro de voz: (a) erro de quantização do codec ADPCM convencional a 16 kbit/s ($M = 1, L = 0$) e (b) erro de quantização do codec ADPCM com quantização de árvore ($M = L = 16$).

Capítulo 5

CODEC ADPCM COM QUANTIZAÇÃO DE ÁRVORE E PÓS-FILTRAGEM

5.1 Introdução

No capítulo anterior, apesar da quantização de árvore melhorar o desempenho do codec ADPCM a 16 kbit/s, verificou-se que o ruído de quantização ainda era bastante perceptível, tornando o codec incompatível para aplicação telefônica. Para tornar o efeito deste ruído menos danoso à inteligibilidade do sinal, duas alternativas são citadas: *pós-filtragem* [24,25,26] e *filtro de ponderação de ruído* [26,27,28]. Este capítulo é dedicado ao estudo da técnica de pós-filtragem.

5.2 Pós-filtragem - Idéia Básica

Quando da avaliação preliminar do codec com quantização de árvore, apesar da diminuição do ruído, ele ainda era bastante perceptível. Considerando-se tal fato, uma filtragem do sinal reconstruído é sugerida de modo a atenuar a perceptibilidade deste ruído. Tal técnica é conhecida na literatura como *pós-filtragem*, sendo que o termo *pós* deve-se ao fato de tal filtragem ser realizada após o sinal ser processado pelo codificador/decodificador, conforme mostra a figura 5.1.

A idéia básica de tal técnica pode ser descrita com a auxílio da figura 5.2 [25], onde encontra-se uma representação idealizada dos espectros de um sinal

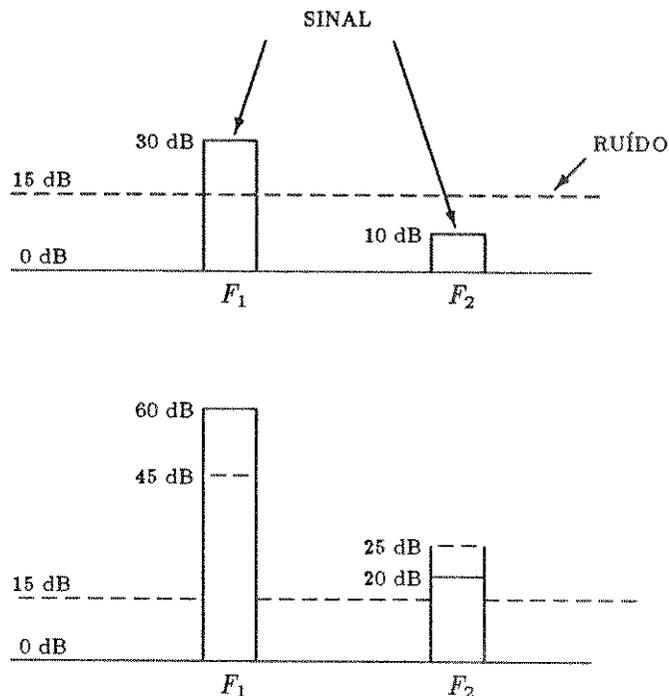


Figura 5.2: Efeito da pós-filtragem. Espectro idealizado de um sinal reconstruído $s_r(k)$: (a) sem pós-filtragem; (b) após filtragem.

espectro do sinal reconstruído².

5.3 Pós-filtro - Função de Transferência

Através da figura 5.2, verifica-se que o processamento do sinal reconstruído pelo pós-filtro tem por objetivo fazer com que a componente ruído fique, de certa forma, mais correlacionada com a componente voz, sendo que nas regiões de formantes, um aumento na potência de ruído não é perceptível, uma vez que a potência da componente voz aumenta proporcionalmente. Por outro lado, tal aumento na potência do sinal produz uma melhoria na relação sinal/(ruído na região entre formantes). Para se obter na saída do pós-filtro a componente ruído correlacionada

²Na figura 5.2(a), a diferença entre as potências do sinal nas regiões F_1 e F_2 é 20 dB, enquanto que na figura 5.2(b), tal diferença é 40 dB.

com a componente voz, faz-se necessário um pós-filtro com a seguinte função de transferência (filtro de síntese):

$$F(z) = \frac{1 + B(z/\beta)}{1 - A(z/\alpha)}, \quad (5.1)$$

onde

$$A(z/\alpha) = \sum_{i=1}^2 a_i \alpha^i z^{-i}, \quad 0 \leq \alpha \leq 1 \quad (5.2)$$

e

$$B(z/\beta) = \sum_{i=1}^6 b_i \beta^i z^{-i}, \quad 0 \leq \beta \leq 1. \quad (5.3)$$

Salienta-se que $A(z)$ e $B(z)$ são obrigatoriamente as estruturas adaptativas de segunda e sexta ordens que constituem o preditor dos codecs analisados até então.

As formantes do sinal de voz são modeladas pelos pólos da função de transferência $F(z)$, enquanto que os zeros de $F(z)$ modelam as regiões do espectro onde o conteúdo espectral é baixo (região entre formantes).

Os parâmetros α e β são responsáveis pelo escalonamento dos coeficientes do filtro, de modo que quando $\alpha = \beta = 1$, tem-se uma filtragem “máxima”, enquanto que quando $\alpha = \beta = 0$, tem-se o sinal reconstruído sem pós-filtragem. Um estudo sobre o efeito do parâmetro α na conformação do espectro do sinal filtrado encontra-se no Apêndice VI, onde verifica-se que conforme α se aproxima da unidade, o espectro na região das formantes torna-se mais estreito.

5.4 Resultados Experimentais

Para o codec ADPCM com quantização de árvore em estudo, considerando-se o compromisso entre redução da perceptibilidade do ruído e distorção do sinal, o melhor desempenho foi alcançado através de $\alpha = 0.0$ e $\beta = 1.0$, o que significa que a pós-filtragem basicamente está atenuando a região entre formantes. Nas figuras 5.3 e 5.4 encontram-se, respectivamente, as formas de onda e os espectros LPC de um segmento do sinal de voz original e do sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$) após esta pós-filtragem ($\alpha = 0.0$ e $\beta = 1.0$). O segmento da figura 5.3 é constituído de 2 blocos de 64 amostras cada um, correspondendo à duração de 16 ms, enquanto que na figura 5.4, tem-se um

segmento de 4 blocos e duração 32 ms. Devido à distorção introduzida pelo pós-filtro, as medidas objetivas do desempenho do codec só foram possíveis mediante a passagem do sinal de voz original pelo mesmo pós-filtro, como mostra a figura 5.5, ou seja, as relações sinal/ruído e sinal/ruído segmental após a pós-filtragem são medidas entre o sinal reconstruído filtrado $s_{rp}(k)$ e o sinal de voz original idênticamente filtrado $s_p(k)$. Na figura 5.6 têm-se, para o mesmo segmento da figura 5.3, as formas de onda de um segmento do sinal de voz original e do sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$), ambos após pós-filtragem ($\alpha = 0.0$ e $\beta = 1.0$).

Na tabela 5.1 encontram-se os resultados³ objetivos (relação sinal/ruído segmental) obtidos para os codecs ADPCM convencional ($M = 1, L = 0$) e com quantização de árvore ($M = L = 16$) em função dos parâmetros α e β . Foram incluídos na análise os resultados obtidos para tais codecs sem pós-filtragem ($\alpha = \beta = 0$).

O maior desempenho objetivo é obtido para $\alpha = 1.0$ e $\beta = 1.0$, e, devido à distorção introduzida no sinal, verifica-se que há uma melhoria de aproximadamente 4.7 dB e 5 dB, respectivamente em relação aos codecs convencional e com quantização de árvore sem pós-filtragem. As figuras 5.2(a) e 5.2(b), onde encontram-se respectivamente, os espectros de um sinal de voz reconstruído sem e com pós-filtragem, comprovam que a pós-filtragem pode resultar numa considerável melhoria na relação sinal/ruído. De fato, na figura 5.2(b) tem-se o espectro idealizado do sinal reconstruído filtrado que mais se aproxima da situação de máxima filtragem, isto é, quando $\alpha = 1.0$ e $\beta = 1.0$. Entretanto, para esta filtragem medidas subjetivas informais mostram que a distorção introduzida no sinal é significativa, o que resulta num efeito passa-baixas bastante perceptível quando da comparação do sinal de voz reconstruído filtrado com o sinal de voz original.

Para $\alpha = 0.0$ e $\beta = 1.0$, apesar do desempenho objetivo inferior em relação aos demais codecs com pós-filtragem ($\alpha = 0.0$ e $\beta = 1.0$ significa que a pós-filtragem está basicamente atenuando a região entre formantes), testes subjetivos informais mostram a superioridade desta pós-filtragem, dado que a redução da perceptibilidade do ruído bem como a diminuição do efeito passa-baixas anteriormente mencionado são bastante marcantes em relação às demais filtrações.

³O arquivo de voz usado neste experimento é o mesmo que foi usado nos capítulos anteriores para a avaliação preliminar dos codecs apresentados.

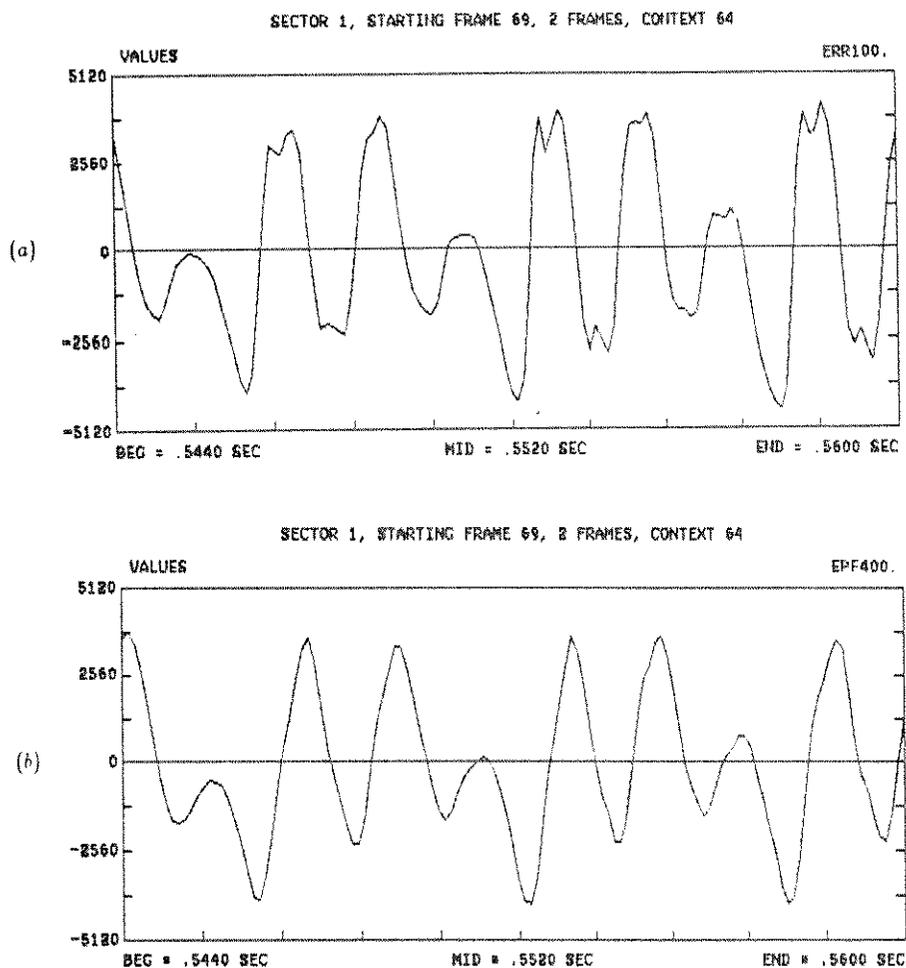


Figura 5.3: Formas de onda: (a) sinal de voz original; (b) sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$) após pós-filtragem ($\alpha = 0.0$ e $\beta = 1.0$).

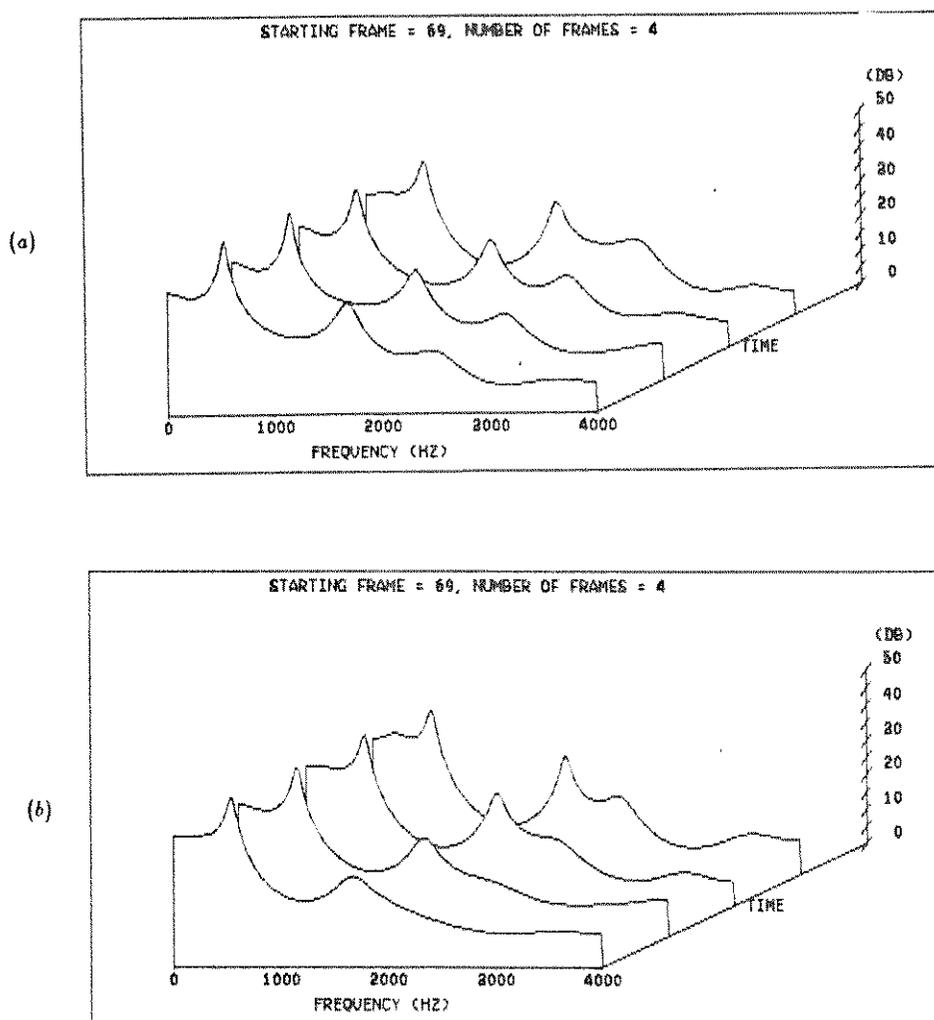


Figura 5.4: Espectros LPC: (a) sinal de voz original; (b) sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$) após pós-filtragem ($\alpha = 0.0$ e $\beta = 1.0$).

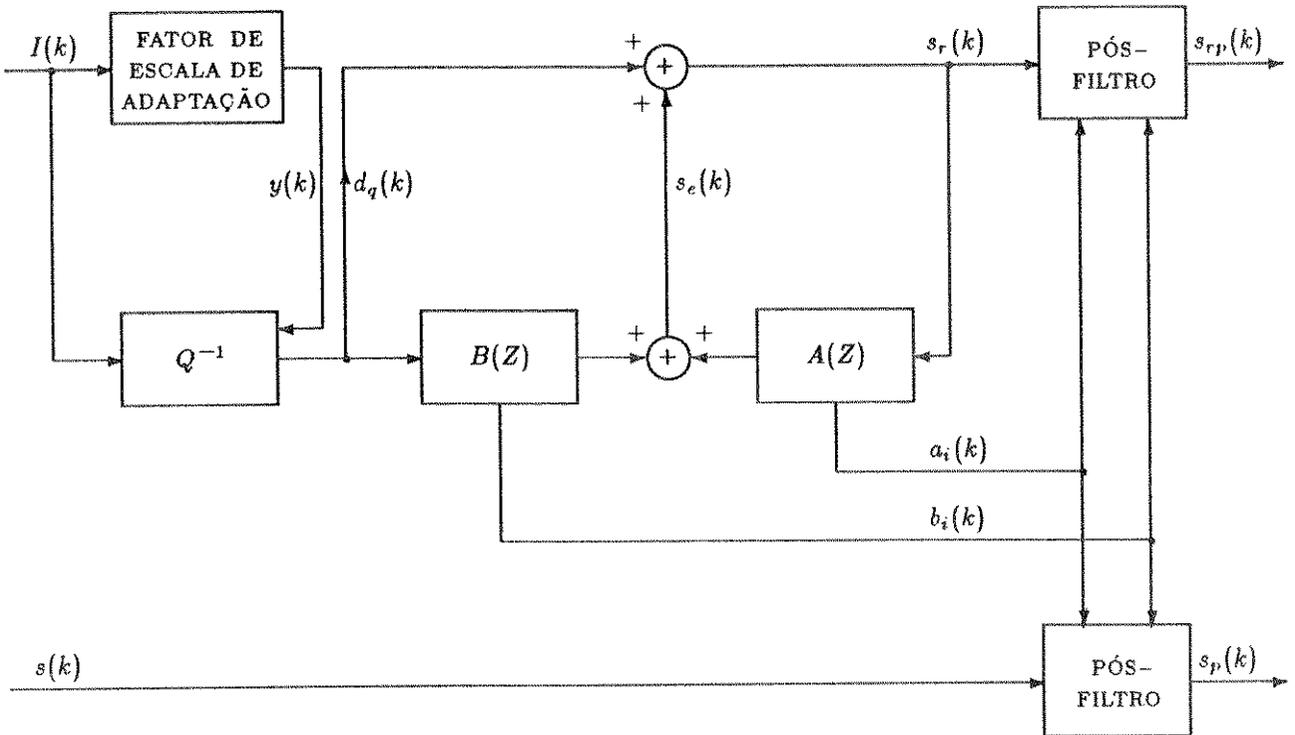


Figura 5.5: Diagrama de blocos para medida de desempenho objetivo do codec ADPCM com pós-filtragem.

Baseando-se nestes resultados, quando da escolha dos parâmetros α e β associados ao pós-filtro, o compromisso entre redução da perceptibilidade do ruído e distorção do sinal deve ser considerado. Ressalta-se também que para aplicações do codec ADPCM com pós-filtragem em sistemas com ligações TANDEM (ligações com codificação/decodificação sucessivas), a distorção do sinal torna-se acumulativa e o desempenho de tais sistemas é bastante degradado. De fato, testes subjetivos informais mostram que, para um número de codecs colocados em cascata maior ou igual a três, a degradação é muito significativa. Desta forma, no próximo capítulo é proposta outra alternativa para a redução da perceptibilidade do ruído, sendo que a distorção do sinal reconstruído na saída do decodificador é significativamente menor quando comparada à técnica de pós-filtragem. Trata-se do filtro de ponderação de ruído.

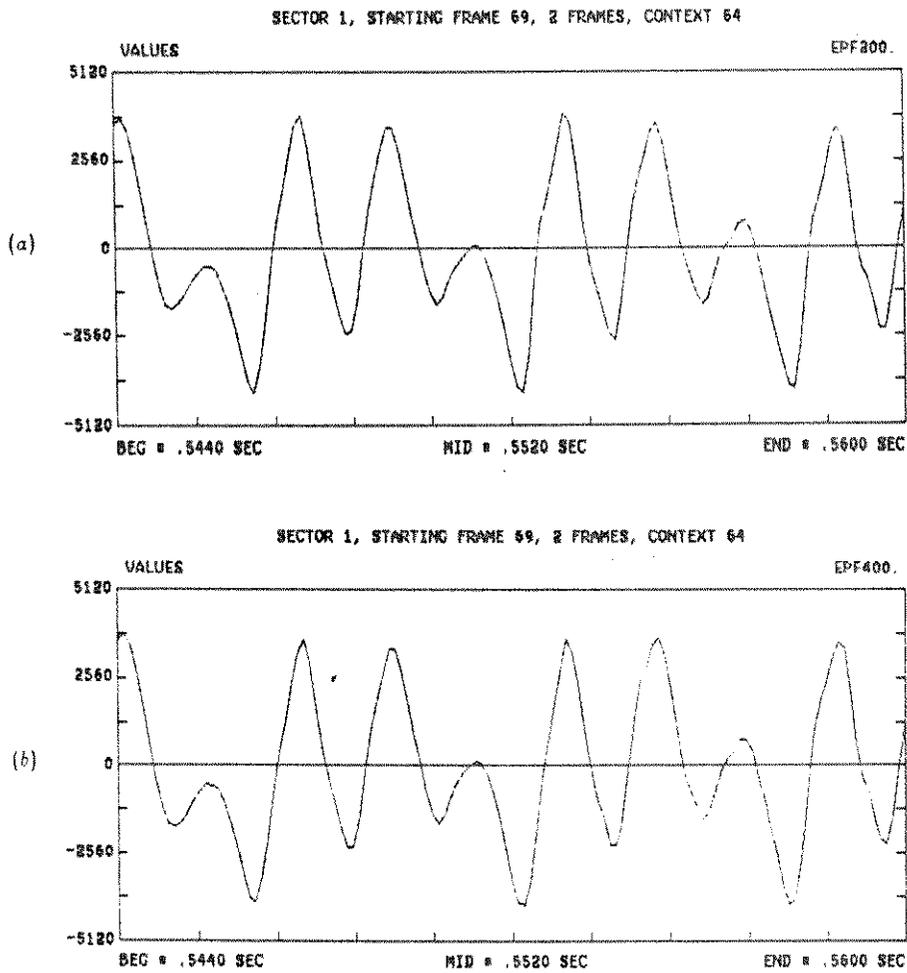


Figura 5.6: Formas de onda: (a) sinal de voz original e (b) sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$), ambos após pós-filtragem ($\alpha = 0.0$ e $\beta = 1.0$).

CODEC ADPCM	S/R_{seg} sem pós-filtro ($\alpha = \beta = 0$)	S/R_{seg} com pós-filtro		
		$\alpha = 0.0$ $\beta = 1.0$	$\alpha = 0.5$ $\beta = 1.0$	$\alpha = 1.0$ $\beta = 1.0$
Convencional ($M = 1, L = 0$)	14.84	18.45	19.20	19.50
Árvore ($M = L = 16$)	18.27	22.20	23.05	23.33

Tabela 5.1: Desempenho objetivo dos codecs ADPCM a 16 kbit/s. Efeito da pós-filtragem.

Capítulo 6

CODEC ADPCM COM QUANTIZAÇÃO DE ÁRVORE E PONDERAÇÃO DE RUÍDO

6.1 Introdução

No capítulo anterior, verificou-se que, através da técnica de pós-filtragem, a perceptibilidade do ruído de quantização foi bastante reduzida. Verificou-se também que quando da aplicação destes codecs em sistemas com ligações TANDEM, devido à distorção introduzida no sinal, o desempenho de tais sistemas era degradado. Assim, é proposto o emprego de *filtro de ponderação de ruído* nos esquemas ADPCM apresentados como outra possibilidade de tornar o efeito do ruído de quantização menos danoso à inteligibilidade do sinal, sendo que a distorção introduzida no sinal após seu processamento pelo codificador/decodificador é bastante inferior em relação ao emprego da pós-filtragem.

6.2 Filtro de Ponderação de Ruído - Idéia Básica

O filtro de ponderação de ruído visa modificar o espectro do ruído de quantização de modo a torná-lo menos perceptível. Segundo a figura 6.1, onde encontra-se o digrama de blocos do codificador ADPCM com quantização de árvore e ponderação de ruído, tal filtro encontra-se na seção de transmissão e está embutido no processo de seleção da melhor amostra do sinal diferença que será codificada e transmitida. Considerando-se ainda que a estrutura do decodificador é a mesma

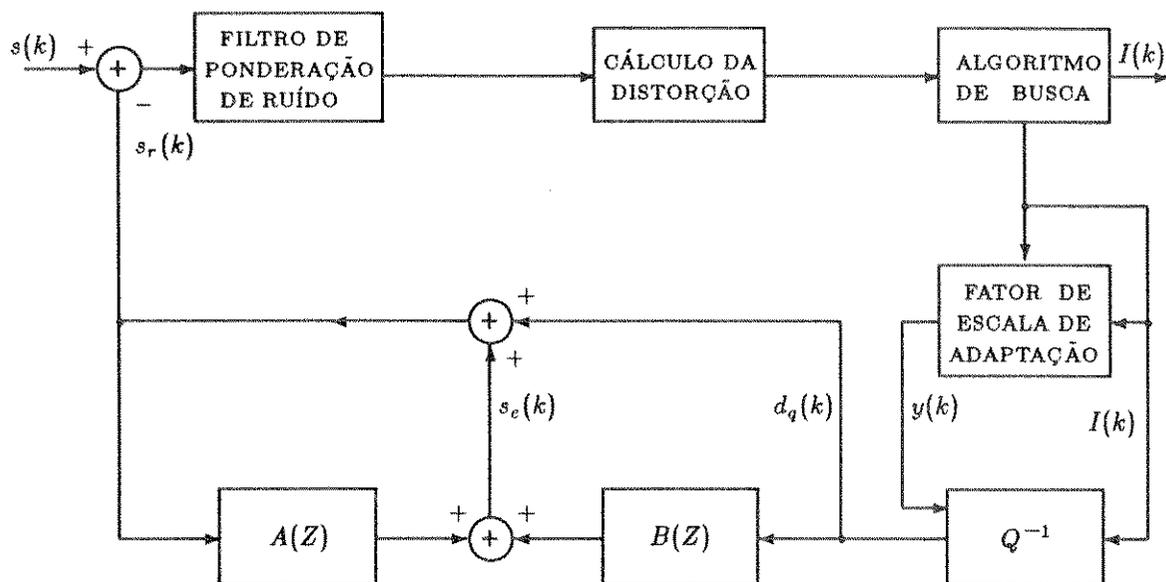


Figura 6.1: Diagrama de blocos do codificador ADPCM com quantização de árvore e ponderação de ruído.

do esquema ADPCM com quantização de árvore (figura 4.5), conclui-se que em sistemas com ligações TANDEM a distorção acumulativa associada ao fato do espectro do ruído de quantização ter sido modificado é significativamente menor quando comparada com a técnica de pós-filtragem.

A idéia básica da técnica de filtragem de ponderação de ruído reside no fato do erro de reconstrução $s(k) - s_r(k)$, neste caso igual ao ruído de quantização, não ser uma medida verdadeira da diferença perceptível entre o sinal de voz original $s(k)$ e o sinal reconstruído $s_r(k)$. Ou seja, em termos de análise subjetiva, pode-se não realizar a seleção da melhor amostra do sinal diferença baseando-se exclusivamente no erro quadrático¹.

Tais afirmações são justificadas considerando-se que na região de formantes o ruído de quantização é mascarado pelo sinal de voz e que grande parte do ruído perceptível do codec é proveniente de regiões do espectro entre formantes. Assim, o objetivo da ponderação de ruído é fazer com que o ruído de quantização seja de-

¹Tal fato justifica não só o emprego da técnica de filtragem de ponderação de ruído, como também da pós-filtragem.

senfatizado na região de formantes, uma vez que, devido à concentração de energia do sinal de voz nesta região, pode-se tolerar erros maiores em comparação com as regiões entre formantes, onde, ao contrário, através desta técnica, o ruído é enfatizado. Tal fato é ilustrado através da figura 6.2 [27], onde encontram-se os espectros de voz original e do filtro de ponderação de ruído.

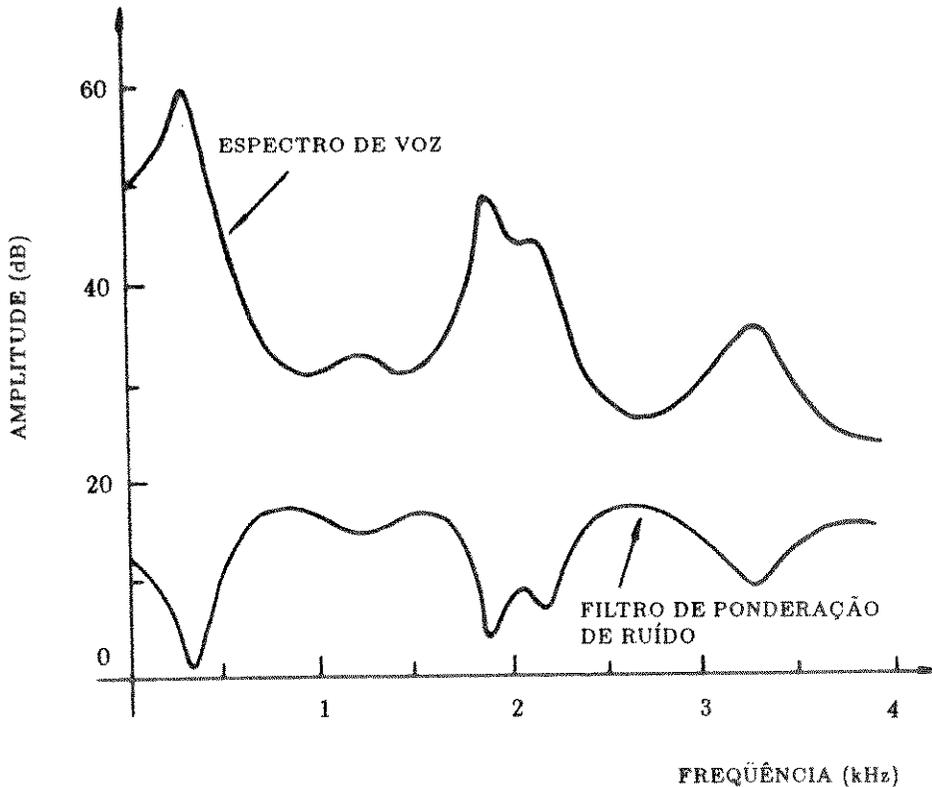


Figura 6.2: Efeito da ponderação de ruído. Espectros do sinal voz original e do filtro de ponderação de ruído.

Através do que foi exposto e referenciando-se à figura 6.1, diz-se que a busca da melhor amostra do sinal diferença é baseada na minimização do erro ponderado, permitindo, portanto, que tal erro reflita uma melhor medida da diferença de perceptibilidade entre $s(k)$ e $s_r(k)$.

6.3 Filtro de Ponderação de Ruído - Função de Transferência

O filtro de síntese dos esquemas ADPCM apresentados tem como função de transferência

$$F(z) = \frac{1 + B(z)}{1 - A(z)}, \quad (6.1)$$

onde

$$A(z) = \sum_{i=1}^2 a_i z^{-i} \quad (6.2)$$

e

$$B(z) = \sum_{i=1}^6 b_i z^{-i}. \quad (6.3)$$

Mencionou-se anteriormente que a função do filtro de ponderação de ruído é modificar o espectro do erro de reconstrução de modo a desenfazá-lo na região de formantes e enfatizá-lo nas regiões entre formantes. Desta forma, sugere-se um filtro com a seguinte função de transferência [29]:

$$F_w(z) = \frac{1 - A(z/\gamma_1)}{1 + B(z/\gamma_1)} \frac{1 + B(z/\gamma_2)}{1 - A(z/\gamma_2)}, \quad 0 < \gamma_2 < \gamma_1 < 1, \quad (6.4)$$

onde

$$A(z/\gamma_1) = \sum_{i=1}^2 a_i \gamma_1^i z^{-i}, \quad (6.5)$$

$$A(z/\gamma_2) = \sum_{i=1}^2 a_i \gamma_2^i z^{-i}, \quad (6.6)$$

$$B(z/\gamma_1) = \sum_{i=1}^6 b_i \gamma_1^i z^{-i}, \quad (6.7)$$

e

$$B(z/\gamma_2) = \sum_{i=1}^6 b_i \gamma_2^i z^{-i}. \quad (6.8)$$

Os parâmetros γ_1 e γ_2 são responsáveis pelo escalonamento dos coeficientes a_i e b_i do filtro², de modo que controlam o grau de ponderação de ruído. Para $\gamma_1 = \gamma_2$, $F_w(z) = 1$ e não há ponderação. Para $\gamma_1 = 1$ e $\gamma_2 = 0$, a ponderação do ruído é feita através de um filtro inverso $F_w(z) = 1/F(z)$. Para $\gamma_2 < \gamma_1 < 1$, a ponderação ocorre num grau intermediário, sendo que os valores ótimos de γ_1 e γ_2 são obtidos empiricamente através de uma série de testes subjetivos de avaliação de desempenho.

6.4 Resultados Experimentais

Para o codec ADPCM com quantização de árvore em estudo, o melhor desempenho foi obtido através de $\gamma_1 = 0.94$ e $\gamma_2 = 0.70$. Nas figuras 6.3 e 6.4 encontram-se, respectivamente, as formas de onda e os espectros LPC de um segmento de voz

²Ressalta-se que os coeficientes a_i e b_i , como no caso da pós-filtragem, são os mesmos do filtro predictor (filtro de síntese).

original e do sinal reconstruído para o codec ADPCM com quantização de árvore ($M = L = 16$) e tal filtro de ponderação de ruído. O segmento da figura 6.3 é constituído de 2 blocos de 64 amostras cada um, correspondendo à duração de 16 ms, enquanto que na figura 6.4, tem-se um segmento de 4 blocos e duração de 32 ms.

Na tabela 6.1 encontram-se os resultados³ objetivos obtidos para os codecs ADPCM convencional ($M = 1, L = 0$) e com quantização de árvore ($M = L = 16$), ambos operando a 16 kbit/s. Foram incluídos na análise os resultados obtidos para os mesmos codecs sem ponderação de ruído ($\gamma_1 = \gamma_2 = 0.0$).

CODEC ADPCM	S/R_{seg} sem ponderação de ruído ($\gamma_1 = 0, \gamma_2 = 0$)	S/R_{seg} com ponderação de ruído ($\gamma_1 = 0.94, \gamma_2 = 0.70$)
Convencional ($M = 1, L = 0$)	14.84	14.82
Árvore ($M = L = 16$)	18.27	17.26

Tabela 6.1: Desempenho objetivo dos codecs ADPCM a 16 kbit/s.

Comparando-se o desempenho objetivo dos codecs com e sem ponderação de ruído, verifica-se que através do emprego do filtro de ponderação a relação sinal/ruído segmental para ambos os codecs praticamente não se alterou⁴. Entretanto, através de medidas subjetivas informais, verifica-se a superioridade do codec que emprega o filtro de ponderação de ruído em relação ao que não emprega tal técnica.

Considerando-se o efeito das técnicas de pós-filtragem e ponderação de ruído na conformação do espectro do sinal reconstruído, através da comparação entre as figuras 5.3 e 5.4 e as figuras 6.3 e 6.4, verifica-se que a distorção do sinal após a pós-filtragem é maior que no caso da filtragem de ponderação de ruído. Entretanto, testes subjetivos informais mostram que o emprego da técnica de pós-filtragem é mais eficiente na redução da perceptibilidade do ruído. Neste contexto, uma combinação de ambas as técnicas é sugerida, sendo que testes subjetivos informais

³O arquivo de voz usado neste experimento é o mesmo que foi usado nos capítulos anteriores para a avaliação preliminar dos codecs apresentados.

⁴Quando do processamento de vários sinais de voz pelos codecs ADPCM convencional e com quantização de árvore, ambos operando à taxa de 16 kbit/s e com ponderação do ruído de quantização, verificou-se que a máxima degradação na relação sinal/ruído segmental correspondia a aproximadamente 1 dB.

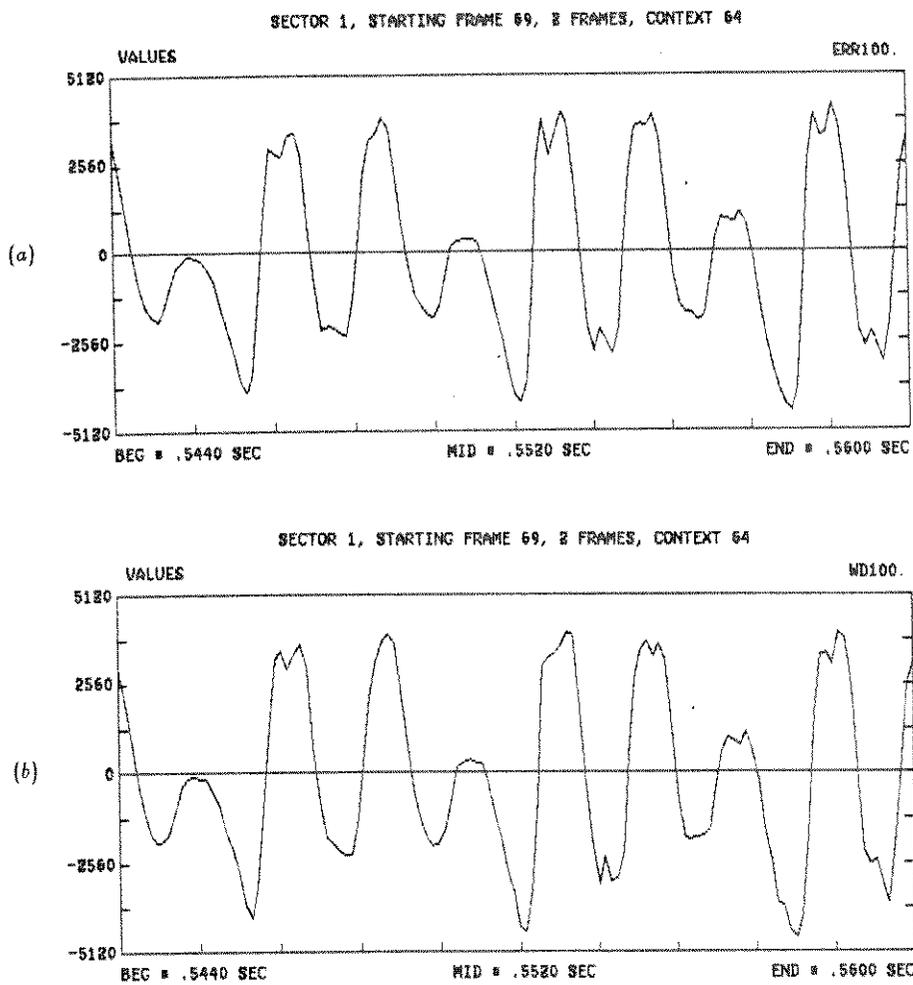


Figura 6.3: Formas de onda: (a) sinal de voz original; (b) sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$) e ponderação de ruído ($\gamma_1 = 0.94$ e $\gamma_2 = 0.70$).

preliminares mostram que o desempenho de tal codec é superior ao dos codecs que empregam isoladamente tais técnicas.

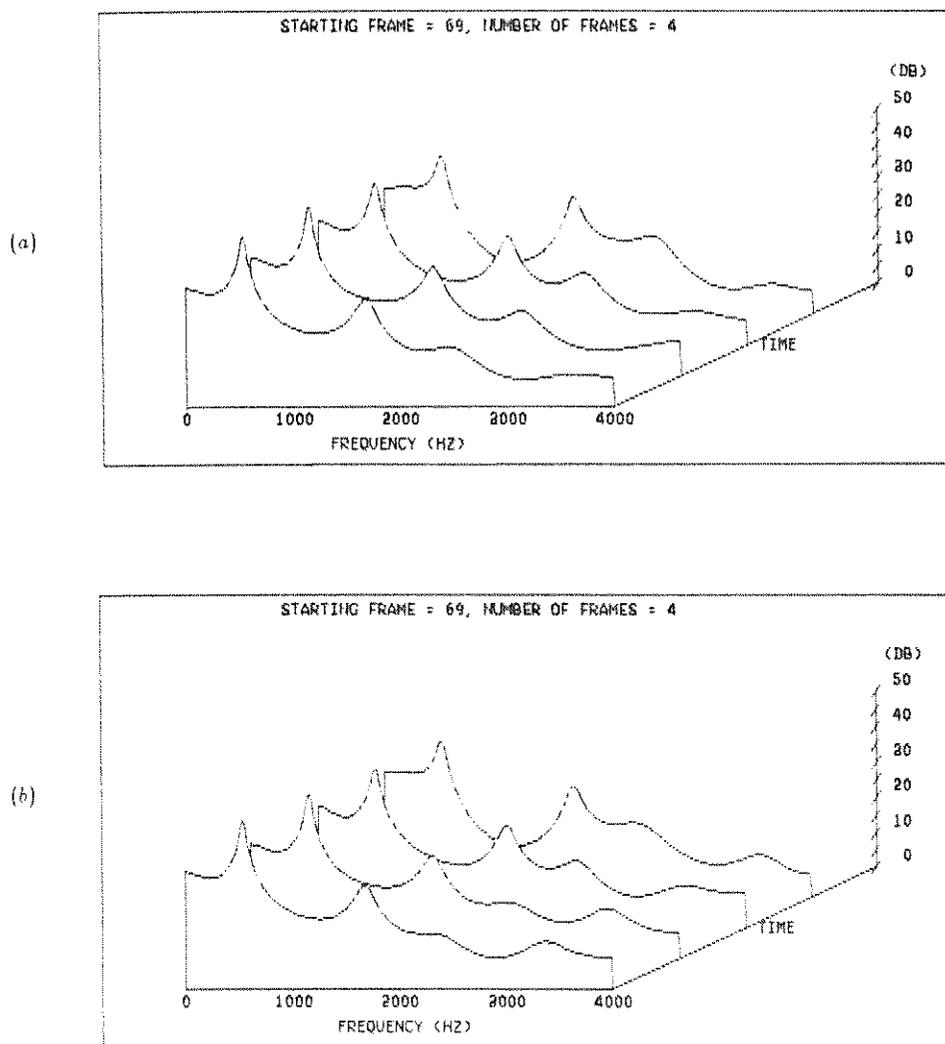


Figura 6.4: Espectros LPC: (a) sinal de voz original; (b) sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$) e ponderação de ruído ($\gamma_1 = 0.94$ e $\gamma_2 = 0.70$).

Capítulo 7

CODEC ADPCM COM QUANTIZAÇÃO DE ÁRVORE A 9.6 KBIT/S

7.1 Introdução

Nos capítulos anteriores, verificou-se que através do emprego de quantização de árvore bem como técnicas de pós-filtragem e ponderação de ruído, obteve-se um codec ADPCM à taxa de 16 kbit/s que apresentava um bom desempenho em termos de qualidade subjetiva. Tal fato conduziu a uma investigação à taxa de 9.6 kbit/s¹ [30]. Tal taxa foi obtida mediante o emprego do conceito de *codificação/quantização multi-árvore* (“*Multi-tree source code*” [31]), onde tem-se uma estrutura constituída de árvores com diferentes taxas de codificação e alternadas entre si.

7.2 Codificação Multi-árvore - Idéia Básica

Na figura 7.1 encontra-se a estrutura multi-árvore para a taxa de 9.6 kbit/s. Definindo-se *nós principais* como sendo os nós da árvore onde a estrutura multi-árvore se repete, a taxa de codificação de tal estrutura, em bit/amostra, é dada por [31]

$$R = \frac{1}{\beta_{mt}} \log_2 \alpha_{mt}, \quad (7.1)$$

¹Inicialmente investigou-se uma estrutura de menor complexidade correspondendo à taxa de 8 kbit/s, e, devido ao fato do codec resultante apresentar um desempenho acentuadamente inferior em relação a 16 kbit/s, partiu-se para o estudo do codec à taxa de 9.6 kbit/s.

Faixa de Entrada Normalizada do Quantizador $\log_2 d(k) - y(k)$	$I(K)$	Saída Normalizada do Quantizador Inverso $\log_2 dq(k) - y(k)$
$[0; +\infty)$	1	1.562
$(-\infty; 0)$	0	-1.562

Tabela 7.1: Característica Entrada/Saída Normalizada do Quantizador no Domínio Logarítmico.

$I(k)$	$W(I(k))$
1	19.0
0	-7.0

Tabela 7.2: Função $W(I(k))$.

7.3 Estrutura do codec a 9.6 kbit/s

Considerando-se a estrutura do codec, manteve-se a mesma do codec a 16 kbit/s, exceto pela inclusão de um quantizador inverso adaptativo correspondente aos ramos onde tem-se $R = 1$ bit/amostra. Neste caso, o quantizador utilizado foi o quantizador ótimo de Max [4] de 2 níveis para sinais com densidade de probabilidade gaussiana e casado com a variância $\sigma \simeq 3.7$ do sinal diferença normalizado. Sua característica no domínio logarítmico encontra-se na tabela 7.1.

O fator de escala de adaptação é dado por

$$y(k) = \begin{cases} (1 - 2^{-7})y(k-1) + 2^{-7}W(I(k) = 1), \\ \quad \text{se } \text{sgn}(dq(k-1)) \cdot \text{sgn}(dq(k)) = 1 \\ (1 - 2^{-7})y(k-1) + 2^{-7}W(I(k) = 0), \\ \quad \text{se } \text{sgn}(dq(k-1)) \cdot \text{sgn}(dq(k)) = -1 \end{cases} \quad (7.2)$$

onde $1.06 \leq y(k) \leq 12.0$ e $W(I(k))$ é uma função discreta definida na tabela 7.2. Tanto o fator de escala de adaptação $y(k)$ como a função $W(I(k))$ foram obtidos empiricamente através de inúmeras simulações, onde as relações sinal/ruído e sinal/ruído segmental bem como medidas subjetivas informais foram monitoradas.

Ressalta-se que para os ramos onde $R = 2$ bit/amostra, a característica do quantizador inverso adaptativo é a mesma do codec a 16 kbit/s. Além disso, a 9.6 kbit/s, as estruturas do pós-filtro e do filtro de ponderação de ruído são as mesmas do codec a 16 kbit/s, o que significa que **NÃO** houve otimização dos mesmos para

esta nova taxa de codificação.

7.4 Desempenho do Codec

O arquivo de voz utilizado neste experimento é o mesmo que foi usado nos capítulos anteriores para a avaliação preliminar dos codes apresentados. Na tabela 7.3 encontram-se os resultados objetivos obtidos para o codec ADPCM convencional ($M = 1, L = 0$) e com quantização de árvore ($M = L = 16$), ambos operando a 9.6 kbit/s. Foi incluído na análise o efeito do filtro de ponderação de ruído.

CODEC ADPCM	S/R_{seg} sem ponderação de ruído ($\gamma_1 = 0, \gamma_2 = 0$)	S/R_{seg} com ponderação de ruído ($\gamma_1 = 0.94, \gamma_2 = 0.70$)
Convencional ($M = 1, L = 0$)	7.74	7.88
Árvore ($M = L = 16$)	10.09	9.93

Tabela 7.3: Desempenho objetivo dos codecs ADPCM a 9.6 kbit/s.

Comparando-se os resultados das tabelas 7.3 e 6.1, onde encontram-se os resultados correspondentes à taxa de 16 kbit/s, verifica-se que o desempenho do codec a 9.6 kbit/s é bastante degradado. De fato, subjetivamente tal degradação é bastante perceptível. Entretanto, uma boa inteligibilidade ainda é mantida.

Nas figuras 7.2 e 7.3 encontram-se, respectivamente, as formas de onda e os espectros LPC de um segmento de voz original e do sinal reconstruído ADPCM com quantização de árvore ($M = L = 16$) sem ponderação de ruído. Tal segmento é constituído de 4 blocos, cada um com 64 amostras, tendo, portanto uma duração de 32 ms. Já na figura 7.4 são ilustradas as formas de onda de um segmento de voz original, do sinal reconstruído e do erro de quantização, ambos sem ponderação de ruído, sendo que neste caso tem-se um segmento de duração de 16 ms.

Comparando-se as figuras 7.4(c) e 4.13(b), verifica-se que ao se reduzir a taxa de bits de 16 para 9.6 kbit/s, o erro de quantização aumentou consideravelmente, sendo que tal fato é compatível com os resultados objetivos e subjetivos obtidos.

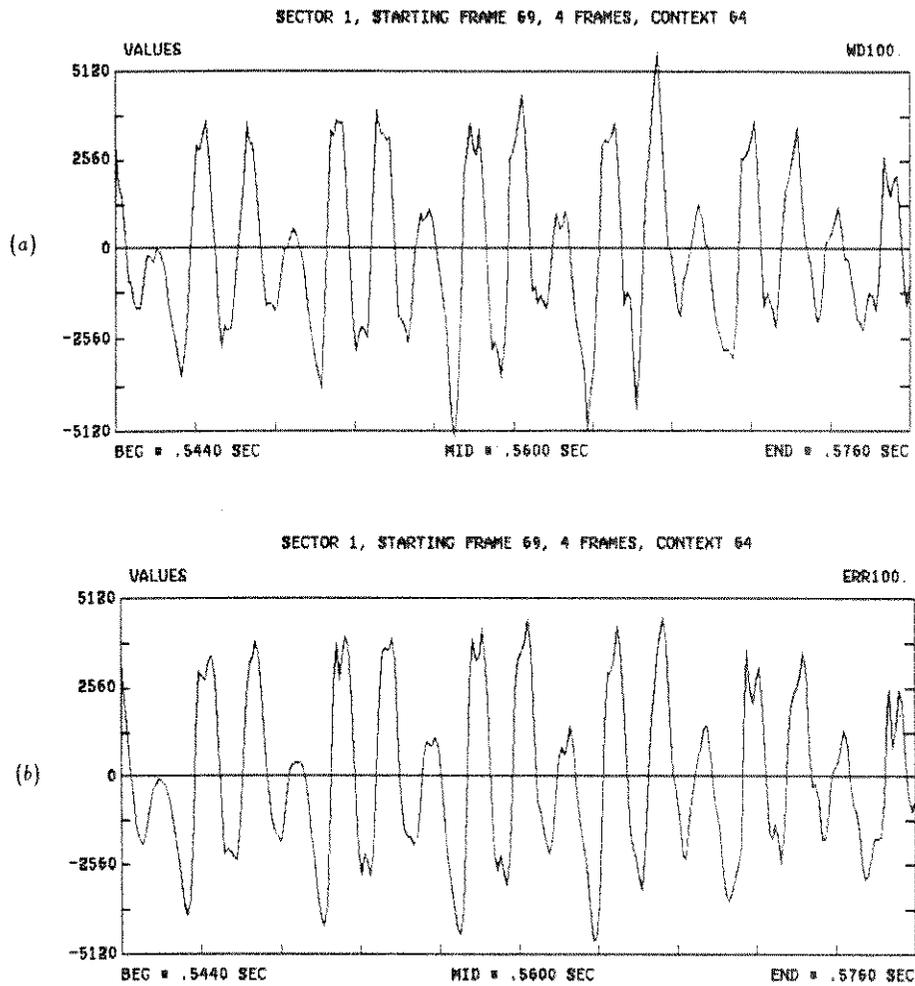


Figura 7.2: Formas de onda: (a) sinal de voz original; (b) sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$) a 9.6 kbit/s.

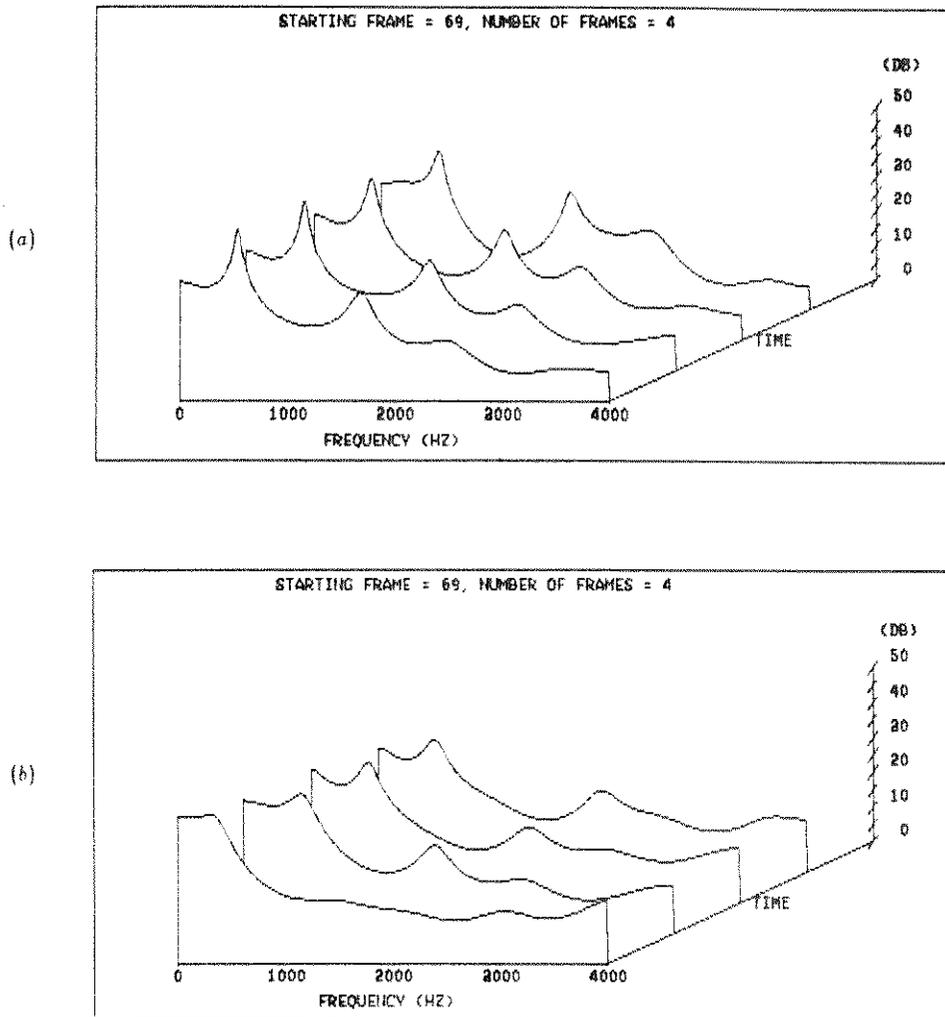


Figura 7.3: Espectros LPC: (a) sinal de voz original; (b) sinal reconstruído do codec ADPCM com quantização de árvore ($M = L = 16$) a 9.6 kbit/s.

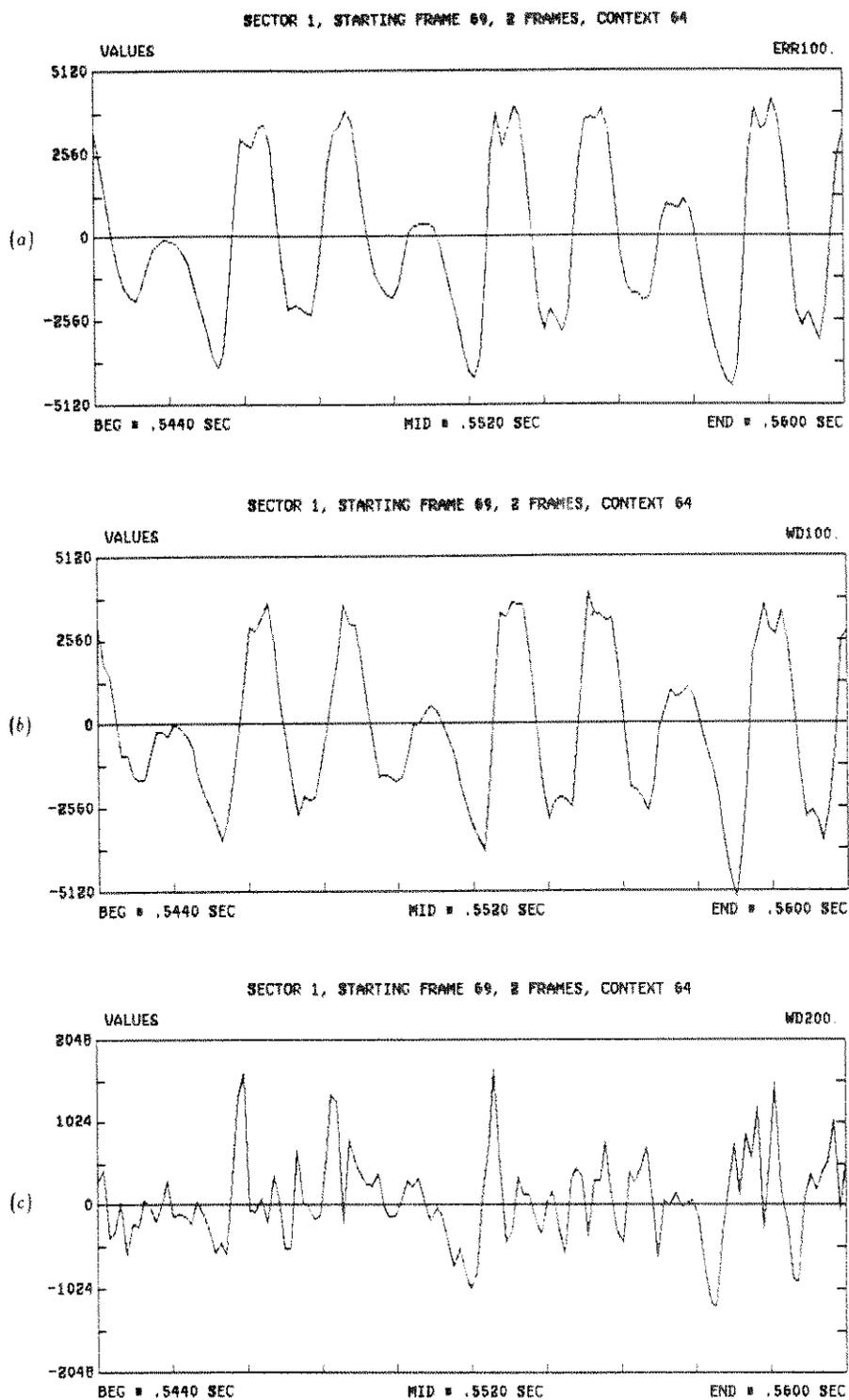


Figura 7.4: Formas de onda: (a) sinal de voz original; (b) sinal reconstruído e (c) erro de quantização do codec ADPCM com quantização de árvore ($M = L = 16$) a 9.6 kbit/s.

Capítulo 8

TESTES E RESULTADOS

8.1 Introdução

Neste capítulo são descritas as simulações realizadas ao longo deste trabalho e os respectivos resultados obtidos. São apresentadas as metodologias empregadas para determinação destes resultados, os quais encontram-se sob a forma de medidas de desempenho objetivo e subjetivo.

Inicialmente é descrito o ambiente de trabalho, onde são caracterizados o sistema de aquisição de dados bem como o sistema onde as simulações foram realizadas.

Posteriormente são apresentados os esquemas ADPCM simulados e os vários experimentos realizados, sendo que numa primeira etapa é descrito o procedimento de otimização de cada um destes esquemas, e, numa segunda etapa, é descrito o procedimento para avaliação dos mesmos para posterior seleção do melhor deles.

8.2 Ambiente de Trabalho

As simulações foram realizadas em tempo não real no computador VAX 8700 da Digital, tendo como ambiente o sistema VMS versão 5.2.

Ao longo deste trabalho foram utilizados dois sistemas de aquisição de dados, um alocado na UNICAMP e outro no CPqD/Telebrás, ambos descritos a seguir.

8.2.1 Sistema de aquisição de dados da UNICAMP

Denominado SAPDV-A (Sistema de Análise e Processamento Digital de Voz - Versão A) [32], tal sistema pertence ao Laboratório de Comunicações da Faculdade de Engenharia Elétrica da UNICAMP. O SAPDV-A permite a geração e gravação direta de arquivos de voz em disco Winchester, bem como sua reprodução através de alto-falante ou telefone. A duração dos arquivos de voz é dependente da capacidade de armazenamento do disco.

Os arquivos de voz, sob forma analógica, são limitados na faixa de frequências de (300 a 3400) Hz, amostrados a 8000 Hz e quantizados com 12 bits por amostra, o que corresponde à excursão em amplitude no intervalo de -2048 a 2047 .

8.2.2 Sistema de aquisição de dados do CPqD/TELEBRÁS

Tal sistema faz parte da placa denominada DSP-16 Data Acquisition Processor, fabricada pela Ariel, que está acoplada a um microcomputador PC-AT.

Considerando-se as aplicações atuais em processamento digital de voz, têm-se as seguintes configurações para aquisição de dados:

(a) os arquivos de voz, sob forma analógica, são limitados na faixa de frequências de (100 a 7000) Hz, amostrados a 16000 Hz e quantizados com 16 bits por amostra;

(b) os arquivos de voz, sob forma analógica, são limitados faixa de frequências de (300 a 3400) Hz, amostrados a 8000 Hz e também quantizados com 16 bits por amostra.

Para ambas as configurações, os conversores A/D-D/A de 16 bits permitem a excursão em amplitude dos sinais no intervalo de -32768 a 32767 .

Assim, uma vez gerados os arquivos de voz, eles são processados em tempo não real no computador e reproduzidos em alto-falante ou telefone, com o objetivo de permitir uma avaliação subjetiva do desempenho do sinal processado.

8.3 Descrição dos esquemas ADPCM simulados

As simulações em computador dos diferentes esquemas ADPCM foram realizadas mediante um "software" específico que permite a variação de diferentes parâmetros dos codecs diretamente associados ao desempenho dos mesmos. O "software" foi desenvolvido em linguagem de alto nível FORTRAN para ser processado, conforme mencionado, em ambiente VMS, versão 5.2 do computador VAX 8700 da Digital.

O programa desenvolvido é constituído basicamente de dois sub-programas correspondentes à seção de transmissão (codificador) e à seção de recepção (decodificador).

8.3.1 Codec ADPCM a 16 kbit/s

- Codificador

Sabe-se que o codificador é constituído basicamente de um gerador de código (quantizador inverso e preditor adaptativos), uma medida de distorção, um algoritmo de busca conhecido como Algoritmo (M, L) e um filtro de ponderação de ruído, que é ativado somente quando desejado. Assim, este sub-programa responsável pela simulação do codificador apresenta as seguintes características:

1. Entrada/saída de dados: a entrada do codificador é um sinal de voz que se encontra sob a forma de arquivo não formatado do tipo inteiro, constituído de blocos de 256 amostras, ou seja, blocos de 512 bytes. Conforme mencionado no Capítulo 3, sabe-se que o sinal de entrada do codificador é um sinal PCM linear de 13 bits por amostra, cuja excursão em amplitude corresponde ao intervalo de -4096 a 4095. Na saída do codificador tem-se o sinal ADPCM com 4 bits por amostra, encontrando-se sob a mesma forma de arquivo não formatado do sinal de entrada;
2. Possibilidade de seleção de M e L , $1 \leq M \leq 16$ e $0 \leq L \leq 16$, onde M é o número máximo de caminhos selecionados na janela de observação de $L + 1$ amostras (L é a profundidade da árvore);
3. Possibilidade de ativação/não ativação do filtro de ponderação de ruído. Caso seja ativado, o programa permite que sejam escolhidos os parâmetros γ_1 e γ_2 do filtro, onde $0 < \gamma_2 < \gamma_1 < 1$;
4. Possibilidade de seleção do número de amostras a serem processadas.

- Decodificador

O decodificador é constituído do gerador de código anteriormente mencionado e do pós-filtro, que é ativado quando desejado. Suas características são as seguintes:

1. Entrada/saída de dados: como entrada e saída, têm-se, respectivamente, o sinal ADPCM e o sinal de voz reconstruído, ambos inteiros e com as mesmas características dos respectivos sinais descritos no sub-programa codificador;

2. Possibilidade de ativação/não ativação do pós-filtro. Caso seja ativado, o programa permite a seleção dos parâmetros α e β , onde $0 \leq \alpha, \beta \leq 1$.

8.3.2 Codec ADPCM a 9.6 kbit/s

Uma vez que este codec apresenta a mesma estrutura do codec ADPCM a 16 kbit/s exceto pela adição de um quantizador inverso adaptativo (1 bit/amostra), as características dos sub-programas desenvolvidos para sua simulação são as mesmas apresentadas para o codec a 16 kbit/s.

8.4 Procedimento de otimização dos parâmetros dos codecs simulados

O objetivo deste trabalho é apresentar um esquema de codificação a 16 kbit/s com atraso e qualidade compatíveis com a aplicação telefônica. Através do que foi apresentado até o presente capítulo, conclui-se que o esquema ADPCM resultante da combinação de técnicas de quantização de árvore, pós-filtragem e filtragem de ponderação de ruído apresenta melhor desempenho em termos de qualidade subjetiva. Tal resultado foi obtido às custas de otimização de tais técnicas isoladamente.

No processo de otimização foram utilizados como sinais de entrada do codificador dois arquivos de voz, um em Português e outro em Inglês, ambos constituídos de sentenças masculinas e femininas. Procurou-se utilizar sentenças foneticamente equilibradas, isto é, constituídas de sons fricativos (f, s, sh, z, ...), oclusivos (b, d, p, ...) e vogais (a, e, ...).

- Arquivo de voz em Português

Este arquivo foi gerado pelo sistema SAPDV-A e tem duração de aproximadamente 16 s. É importante enfatizar que o sinal de voz gerado pelo SAPDV-A excursiona no intervalo de amplitudes de -2048 a 2047, e, portanto, apresenta resolução de 12 bits por amostra, o que não é compatível com o nível de sinal de entrada do codificador descrito anteriormente. Ele é constituído das seguintes sentenças:

- Sentenças masculinas: *Festas surpresa são difíceis de realizar. Muitas coisas podem sair erradas. Algumas vezes os erros são realmente engraçados.*
- Sentenças femininas : *A profissão de aeromoça parece ser excitante, mas é também muito cansativa. Lidar com pessoas faz parte do trabalho.*

- Arquivo de voz em Inglês

Este arquivo tem duração de aproximadamente 7 s e foi cedido à Área de Processamento Digital de Voz do CPqD/TELEBRÁS pelo CSELT (Centro Studi e Laboratori Telecomunicazioni SpA), Itália. Ressalta-se que este arquivo foi quantizado com 13 bits por amostra, o que corresponde à excursão em amplitude no intervalo de -4096 a 4095, e, portanto, está de acordo com o nível esperado na entrada esperado para o codificador. É constituído das seguintes sentenças:

- Sentenças masculinas: *He suspected his brother. He danced next to an officer.*
- Sentenças femininas : *He spoke with a weary voice. I suppose you mean my mother.*

De modo geral, dados os diferentes parâmetros que interferem no desempenho dos codecs, o procedimento adotado para a sua otimização é descrito a seguir. É importante ressaltar que as medidas de desempenho são resultantes de testes objetivos e subjetivos informais. As medidas subjetivas são enfatizadas em relação às objetivas, uma vez que a baixas taxas as últimas muitas vezes não refletem a qualidade e inteligibilidade reais dos esquemas de codificação.

Na otimização de cada parâmetro, os testes subjetivos informais foram caracterizados pela audição, por parte de aproximadamente 5 pessoas, dos sinais de voz processados pelos codecs simulados para diferentes valores deste parâmetro, e pela posterior seleção do melhor entre eles, tendo sempre como referência o sinal de voz original. Os arquivos de voz foram reproduzidos várias vezes, através de telefone e alto-falante. Já o desempenho objetivo foi medido através da relação sinal/ruído (S/R) e relação sinal/ruído segmental (S/R_{seg}), cujas definições encontram-se no Apêndice VII.

Para os codecs simulados, inicialmente foram otimizados os parâmetros (M, L) do algoritmo de busca; em seguida, os parâmetros α e β que controlam o grau de pós-filtragem; e, finalmente, os parâmetros γ_1 e γ_2 que controlam o grau de ponderação de ruído.

Os valores ótimos obtidos referem-se à taxa de 16 kbit/s. Entretanto, a 9.6 kbit/s tais valores foram mantidos, dado o caráter apenas investigatório associado a esta taxa. Fixados estes valores, testes objetivos e subjetivos informais foram

realizados, sob diversas condições de contorno, com o objetivo de se selecionar o melhor codec simulado.

8.5 Testes objetivos e subjetivos informais

Inicialmente, com o objetivo de se estabelecer o desempenho dos codecs ADPCM com quantização de árvore descritos nos capítulos anteriores, medidas objetivas (S/R_{seg}) e medidas subjetivas informais foram conduzidas considerando-se três diferentes níveis de entrada dos codecs: -12 dB, -22 dB e -32 dB abaixo do nível de saturação de conversores A/D-D/A de 16 bits. Nesta etapa foram utilizados dois arquivos de voz gerados pelo sistema de aquisição de dados do CPqD/TELEBRÁS, sendo um constituído de sentenças masculinas e outro de sentenças femininas, cada um com duração de aproximadamente 7 s. São eles:

- Arquivo de sentenças masculinas:
Navegar é preciso, viver não. A natureza é a mãe da vida: respeite-a.
- Arquivo de sentenças femininas :
Primeiro Deus criou a terra e o céu. Antes eu queria me desculpar com a senhora.

Na tabela 8.1 encontram-se as medidas objetivas para o codec ADPCM a 32 kbit/s (G.721) e para os codecs ADPCM com quantização de árvore a 16 e 9.6 kbit/s ($M = 8, L = 16$), com (PON) e sem (SPON) ponderação de ruído, todos sem o emprego da técnica de pós-filtragem.

Na tabela 8.2 encontram-se as medidas subjetivas informais, em termos de MOS, ("Mean Opinion Score")¹ correspondentes aos codecs apresentados na tabela 8.1. Também são apresentados resultados adicionais para os codecs com pós-filtro (PF) e a combinação pós-filtro e ponderação de ruído (PF + PON). Para tanto, a audição dos arquivos de voz foi feita por 5 pessoas com reprodução através de aparelho telefônico. A sessão de avaliação era constituída de duas sub-sessões, independentes entre si, uma para o arquivo de sentenças femininas e outra para o arquivo de sentenças masculinas. Cada sub-sessão, por sua vez, era constituída de três experimentos, correspondentes aos níveis de entrada -12, -22 e -32 dB e que são descritos nas tabelas 8.3 e 8.4. Os valores de MOS apresentados na tabela 8.2 foram obtidos calculando-se, para cada experimento, a média das notas dadas

¹O MOS reflete a opinião do ouvinte sobre a qualidade geral dos arquivos ouvidos. Os valores de MOS são tais que $1 \leq \text{MOS} \leq 5$, onde, em termos de qualidade, têm-se: ruim (MOS=1), pobre (MOS=2), razoável (MOS=3), boa (MOS=4) e excelente (MOS=5).

CODEC	Nível de Entrada (dB)		
	-32	-22	-12
32 kbit/s (G.721)	19.93	17.16	11.75
16 kbit/s (SPON)	10.81	11.04	10.17
16 kbit/s (PON)	10.38	10.69	9.87
9.6 kbit/s (SPON)	5.09	5.27	4.79
9.6 kbit/s (PON)	4.91	4.95	4.53

Tabela 8.1: Desempenho objetivo em termos de S/R_{seg} .

CODEC	Nível de Entrada (dB)		
	-32	-22	-12
32 kbit/s (G.721)	4.5	4.2	3.5
16 kbit/s (SPON)	3.5	3.6	4.1
16 kbit/s (PON)	3.6	4.0	4.0
16 kbit/s (PF)	3.9	4.1	4.0
16 kbit/s (PF + PON)	4.0	3.5	4.3
9.6 kbit/s (SPON)	2.3	2.4	2.2
9.6 kbit/s (PON)	2.4	2.2	2.3
9.6 kbit/s (PF)	2.5	2.8	2.5
9.6 kbit/s (PF + PON)	2.4	2.4	2.5

Tabela 8.2: Desempenho subjetivo em termos de MOS.

pelos avaliadores.

Para a análise dos resultados obtidos, inicialmente torna-se importante fazer uma consideração a respeito dos níveis de entrada dos codecs testados. Conforme mencionado anteriormente, o sinal de entrada especificado para o codificador apresenta excursão em amplitude correspondente ao intervalo de -4096 a 4095. Verifica-se, portanto, que os níveis empregados na obtenção dos resultados expostos nas tabelas 8.1 e 8.2 não são compatíveis com este valor especificado.

De acordo com a tabela 8.1, comparando-se os codecs a 16 e 9.6 kbit/s, verifica-se o desempenho bastante superior do primeiro em relação ao segundo. Verifica-se também que a ponderação de ruído introduz uma pequena distorção no sinal de voz após seu processamento pelo codificador/decodificador. De acordo com a tabela 8.2, existe uma perda de qualidade subjetiva bastante significativa quando da redução da taxa de codificação de 16 para 9.6 kbit/s, o que é confir-

Experimento 1 -32 dB	Experimento 2 -22 dB	Experimento 3 -12 dB
16 kbit/s (SPON)	32 kbit/s	16 kbit/s (PF)
9.6 kbit/s (SPON)	16 kbit/s (PF+PON)	16 kbit/s (SPON)
9.6 kbit/s (PF)	9.6 kbit/s (PF+PON)	9.6 kbit/s (SPON)
16 kbit/s (PON)	9.6 kbit/s (PON)	9.6 kbit/s (PF)
9.6 kbit/s (PON)	16 kbit/s (PF)	9.6 kbit/s (PON)
9.6 kbit/s (PON+PF)	9.6 kbit/s (SPON)	16 kbit/s (PON)
16 kbit/s (PF)	9.6 kbit/s (PF)	16 kbit/s (PF+PON)
16 kbit/s (PON+PF)	16 kbit/s (PON)	9.6 kbit/s (PF+PON)
32 kbit/s	16 kbit/s (SPON)	32 kbit/s

Tabela 8.3: Experimentos 1,2 e 3 para arquivos de sentenças masculinas.

Experimento 1 -32 dB	Experimento 2 -22 dB	Experimento 3 -12 dB
16 kbit/s (PF)	32 kbit/s	16 kbit/s (SPON)
16 kbit/s (SPON)	16 kbit/s (PF+PON)	16 kbit/s (PF)
9.6 kbit/s (SPON)	9.6 kbit/s (PF+PON)	9.6 kbit/s (PF)
9.6 kbit/s (PF)	9.6 kbit/s (PON)	9.6 kbit/s (SPON)
9.6 kbit/s (PON)	16 kbit/s (PON)	16 kbit/s (PON)
16 kbit/s (PON)	9.6 kbit/s (SPON)	9.6 kbit/s (PON)
16 kbit/s (PON+PF)	9.6 kbit/s (PF)	9.6 kbit/s (PF+PON)
9.6 kbit/s (PON+PF)	16 kbit/s (PF)	16 kbit/s (PF+PON)
32 kbit/s	16 kbit/s (SPON)	32 kbit/s

Tabela 8.4: Experimentos 1,2 e 3 para arquivos de sentenças femininas.

mado pelas medidas objetivas apresentadas na tabela 8.1. Observa-se também o bom desempenho do codec ADPCM com quantização de árvore a 16 kbit/s que emprega conjuntamente as técnicas de pós-filtragem e ponderação de ruído.

Ainda referenciando-se à tabela 8.2, observa-se que alguns valores de MOS estão bastante distantes de um determinado valor esperado. Por exemplo, o valor de 3.5 para o codec a 32 kbit/s e nível de entrada -12 dB, bem como para o codec a 16 kbit/s (PF + PON) e nível de entrada -22 dB. No primeiro caso, deve-se considerar essencialmente a problemática de adequação do nível do sinal de entrada do codificador. Segundo a figura 2.1, onde encontra-se o diagrama em blocos do codificador ADPCM a 32 kbit/s (G.721), o sinal PCM a 64 kbit/s na entrada do codificador é gerado a partir de um sinal PCM linear de 13 bits de resolução. Este sinal PCM a 64 kbit/s é novamente convertido para PCM linear com resolução de 13 bits para ser processado pelo codificador. Entretanto, quando da realização dos testes, foi empregado um sinal de voz com níveis -12, -22 e -32 dB abaixo do nível de saturação de conversores A/D-D/A de 16 bits. Especificamente para o nível de entrada de -12 dB em discussão, tem-se uma excursão em amplitude no intervalo de aproximadamente -8231 a 8230, que é bastante superior à correspondente aos conversores de 13 bits, ou seja, no intervalo de -4096 a 4095, justificando-se, portanto, o valor discrepante. No segundo caso e também considerando-se todos os valores de MOS apresentados na tabela 8.2, verificou-se, quando da avaliação pelos ouvintes de cada condição dos experimentos, um desvio padrão muito grande, que pode ser justificado pelo número restrito de locutores e ouvintes empregados no teste, respectivamente 2 e 5.

De modo geral, apesar dos problemas citados, tais resultados confirmam os resultados obtidos nos capítulos anteriores, quando das medidas objetivas e subjetivas preliminares.

Capítulo 9

CONCLUSÕES

Este trabalho mostrou que um codec a 16 kbit/s com atraso e qualidade subjetiva compatíveis para aplicação telefônica pode ser implementado usando-se a estrutura do codec a 32 kbit/s padronizado pelo CCITT (Recomendação G.721), mediante a substituição do quantizador escalar por um quantizador de árvore e a incorporação de técnicas de redução da perceptibilidade do ruído de reconstrução, especificamente pós-filtragem e filtragem de ponderação de ruído. Entretanto, enfatiza-se que para aplicações do codec ADPCM com pós-filtragem em sistemas com ligações TANDEM, a distorção introduzida por tal técnica torna-se acumulativa e o desempenho de tais sistemas é bastante degradado.

A complexidade deste codec (transmissão e recepção), em termos de números de operações em ponto flutuante por segundo, foi estimada em aproximadamente 7.5×10^6 , de modo que, considerando-se a implementação “hardware” deste codec, verifica-se que, por exemplo, através do DSP (“Digital Signal Processor”) TMS320C30 da Texas, tal implementação torna-se factível. É importante ressaltar que no presente trabalho nenhuma tentativa de simplificação do algoritmo visando tal implementação foi realizada.

Adicionalmente, simulações a 9.6 kbit/s foram conduzidas com o objetivo de se ter uma idéia de uma taxa de bits possível de ser atingida sob a condição de que uma boa qualidade bem como inteligibilidade de voz fossem mantidas. À taxa de 9.6 kbit/s, verificou-se que uma boa inteligibilidade foi mantida, embora com um nível elevado de ruído granular. Entretanto, uma degradação significativa na

qualidade de voz foi bastante perceptível. Considerando-se este problema, estudos preliminares vislumbram que a otimização dos parâmetros do pós-filtro e do filtro de ponderação de ruído bem como a inclusão de um preditor de pitch à estrutura do codec permitirão uma melhoria de desempenho.

Referenciando-se aos testes subjetivos informais, ressalta-se que a qualidade de voz foi avaliada apenas em função da variação do nível do sinal de entrada do codificador, além do fato de ter-se usado apenas dois arquivos de voz. Tal fato motivou a elaboração de uma nova metodologia de testes, onde destacam-se os seguintes experimentos: pré-seleção de dois algoritmos a serem testados mais exaustivamente, testes para medir a qualidade de voz com a variação do nível de entrada, com taxa de erro e com transcódificações. A aplicação de tal metodologia encontra-se em andamento, sendo que está se estudando, entre vários tópicos, o número de ouvintes e o número de condições (por exemplo, sentenças masculinas e femininas, níveis de entrada) a serem considerados.

Apêndice I

QUANTIZAÇÃO PARA S/R ÓTIMA

Quando a densidade de probabilidade do sinal de entrada do quantizador é conhecida, torna-se possível escolher os níveis de quantização de forma a minimizar a variância do erro de quantização, maximizando assim a relação S/R. A variância do ruído de quantização é

$$\sigma_e = E[e^2(n)] = E[(\hat{x}(n) - x(n))^2], \quad (I.1)$$

onde $\hat{x}(n) = Q[x(n)]$ é o valor quantizado de $x(n)$.

Generalizando-se a característica do quantizador da figura I.1, têm-se M níveis de quantização denominados $\{\hat{x}_{-M/2}, \hat{x}_{-M/2+1}, \dots, \hat{x}_1, \dots, \hat{x}_{M/2}\}$, onde assumiu-se M par [6]. O nível de quantização associado ao intervalo $[x_{i-1}, x_i]$ é \hat{x}_i .

Para uma distribuição de amplitude simétrica e de média nula, define-se um ponto limitante central $x_0 = 0$, e, se a função densidade de probabilidade é diferente de zero para grandes amplitudes (por exemplo, densidades de Laplace e Gamma), atribuem-se às extremidades da faixa do quantizador os valores $\pm\infty$, isto é, $x_{\pm M/2} = \pm\infty$. Desta forma, pode-se escrever

$$\sigma_e^2 = \int e^2 p_e(e) de, \quad (I.2)$$

onde $p_e(e)$ é a função densidade de probabilidade do ruído de quantização.

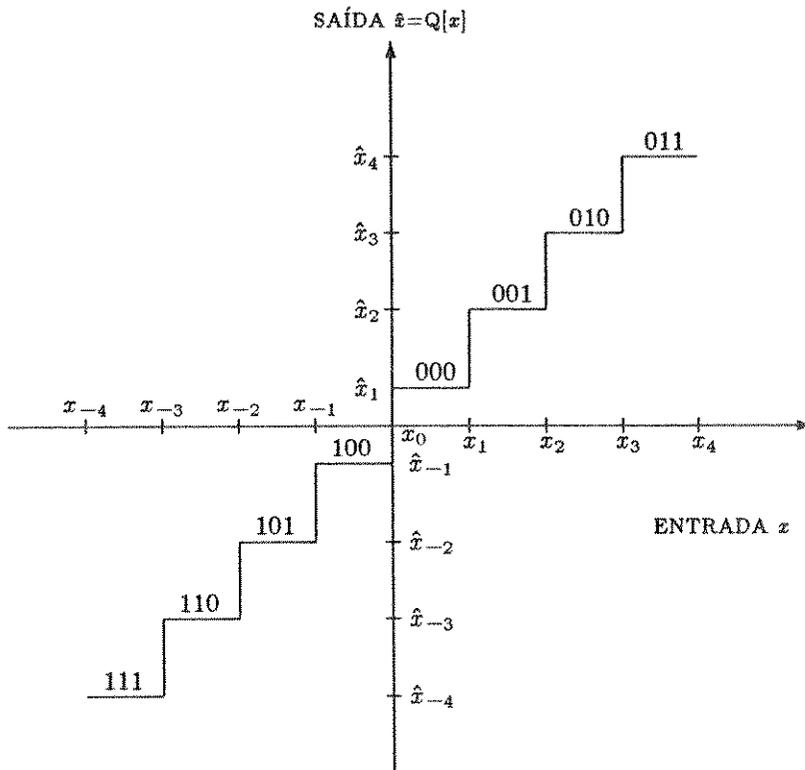


Figura I.1: Característica de entrada/saída de um quantizador de 3 bits.

Considerando-se os diferentes intervalos de quantização, pode-se escrever

$$\sigma_e^2 = \sum_{i=-M/2+1}^{M/2} \int_{x_{i-1}}^{x_i} (\hat{x}_i - x)^2 p_x(x) dx, \quad (I.3)$$

onde $p_x(x)$ é a função densidade de probabilidade do sinal $x(n)$.

Se $p_x(x) = p_x(-x)$, têm-se $\hat{x}_i = -\hat{x}_{-i}$ e $x_i = -x_{-i}$. Portanto,

$$\sigma_e^2 = 2 \sum_{i=1}^{M/2} \int_{x_{i-1}}^{x_i} (\hat{x}_i - x)^2 p_x(x) dx. \quad (I.4)$$

Deseja-se escolher o conjunto de parâmetros $\{x_i\}$ e $\{\hat{x}_i\}$ que minimizem σ_e^2 . Calculando-se $\partial\sigma_e^2/\partial\hat{x}_i = 0$ e $\partial\sigma_e^2/\partial x_i = 0$, obtêm-se as seguintes equações:

$$\int_{x_{i-1}}^{x_i} (\hat{x}_i - x) p_x(x) dx = 0, \quad i = 1, 2, \dots, M/2; \quad (I.5)$$

$$x_i = \frac{1}{2} (\hat{x}_i + \hat{x}_{i+1}), \quad i = 1, 2, \dots, M/2 - 1; \quad (I.6)$$

e, conforme suposições iniciais,

$$x_0 = 0; \quad (I.7)$$

$$x_{\pm M/2} = \pm\infty. \quad (I.8)$$

As equações (I.5) e (I.6) devem ser calculadas simultaneamente para $(M - 1)$ parâmetros não conhecidos do quantizador. Uma vez que estas equações geralmente são não lineares, um procedimento iterativo deve ser usado. Este procedimento foi explorado por Max [4].

De acordo com a referência [4], o quantizador de 16 níveis para sinais com função densidade de probabilidade gaussiana de desvio padrão unitário e de média zero, e que obedece às equações (I.5) e (I.6) é caracterizado pela figura I.2 e tabela I.1.

Experimentos relatados em [7] revelam que o quantizador ótimo para sinais gaussianos é uma boa aproximação para intervalos sonoros de voz, o mesmo não ocorrendo para transientes rápidos, intervalos não sonoros e para sinais de dados.

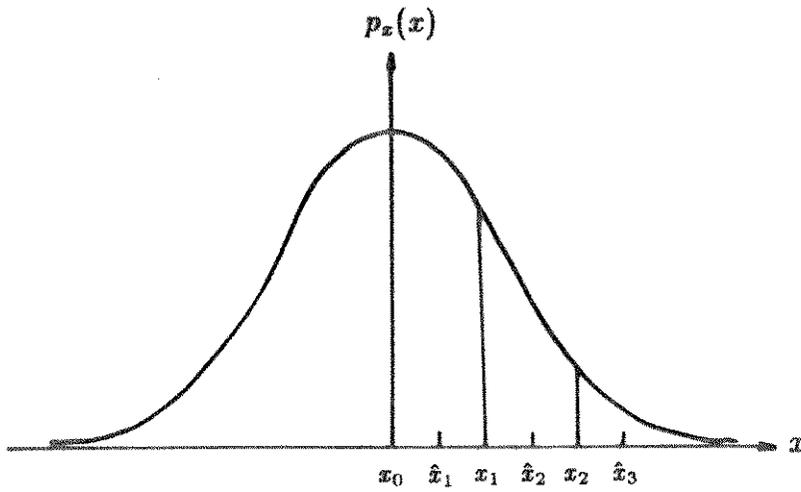


Figura I.2: Faixas de entrada e correspondentes níveis de saída para o quantizador ótimo (M par).

i	x_{i-1}	\hat{x}_i
1	0.0	0.1284
2	0.2582	0.3881
3	0.5224	0.6568
4	0.7996	0.9424
5	1.099	1.256
6	1.437	1.618
7	1.844	2.069
8	2.401	2.733

Tabela I.1: Parâmetros para o quantizador ótimo.

Apêndice II

ADAPTAÇÃO BIMODAL

II.1 Princípios da adaptação bimodal

A adaptação bimodal é uma técnica de adaptação do fator de escala $y(k)$. Deseja-se ajustar a velocidade de adaptação do quantizador para que este sempre opere dentro de sua faixa dinâmica, evitando-se, portanto, a ocorrência de distorção por sobrecarga, bem como de ruído granular. Na figura II.1 encontra-se um diagrama em blocos desta técnica.

De acordo com a técnica de adaptação bimodal, são gerados dois fatores de escala, y_u e y_l , que são combinados linearmente, constituindo assim um novo fator de escala y .

O fator y_u (y “unlocked”) é gerado a partir de uma adaptação do fator de escala y , como verificar-se-á posteriormente. O termo “unlocked” refere-se à necessidade de uma adaptação rápida (sinais de voz, onde o sinal diferença apresenta grandes flutuações de amplitude).

Também será verificado posteriormente que o fator de escala y_l (y “locked”) é obtido a partir de filtragem de y_u . O termo “locked” refere-se à necessidade de uma adaptação lenta (tons e dados na faixa de voz, onde o sinal diferença apresenta pequenas flutuações de amplitude). Segundo a referência [8], uma vez que tons e dados na faixa de voz produzem um sinal diferença com potência relativamente constante, um bom desempenho é alcançado com um quantizador fixo propriamente escalonado.

A combinação linear de y_u e y_l produz o fator de escala y :

$$y(k) = a_l(k)y_u(k-1) + (1 - a_l(k))y_l(k-1), \quad 0 \leq a_l \leq 1, \quad (\text{II.1})$$

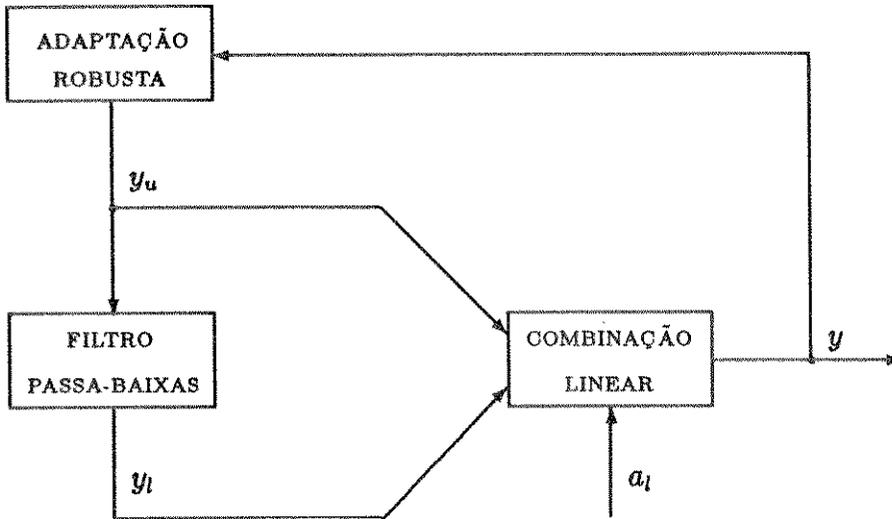


Figura II.1: Geração dos fatores de escala y_u e y_l .

onde a_l é o parâmetro de controle de velocidade e cuja geração será exposta posteriormente.

O fator de escala y atua no quantizador da seguinte forma: quando a envoltória do sinal diferença varia rapidamente, a_l se aproxima da unidade ($a_l \rightarrow 1$), tornando a adaptação do quantizador rápida ($y \rightarrow y_u$); caso contrário, quando a envoltória do sinal diferença é relativamente constante, a_l se aproxima de zero ($a_l \rightarrow 0$), tornando a adaptação do quantizador lenta ($y \rightarrow y_l$).

II.2 Geração dos fatores de escala y_u e y_l

II.2.1 Fator de escala y_u

O princípio de funcionamento de um quantizador adaptativo pode ser visto como uma forma de controle automático de ganho (CAG) com realimentação sobre o quantizador [9], como é mostrado na figura II.2 [9], onde tem-se uma entrada analógica Z multiplicada por $1/Y$ e então processada por um quantizador com ponto de sobrecarga 1 V. Se no conversor D/A a saída normalizada for multiplicada por Y , o efeito global desta multiplicação será o mesmo que processar Z através do quantizador da figura II.3 [9].

Se σ é o valor rms do sinal de entrada, o fator de carga é definido como

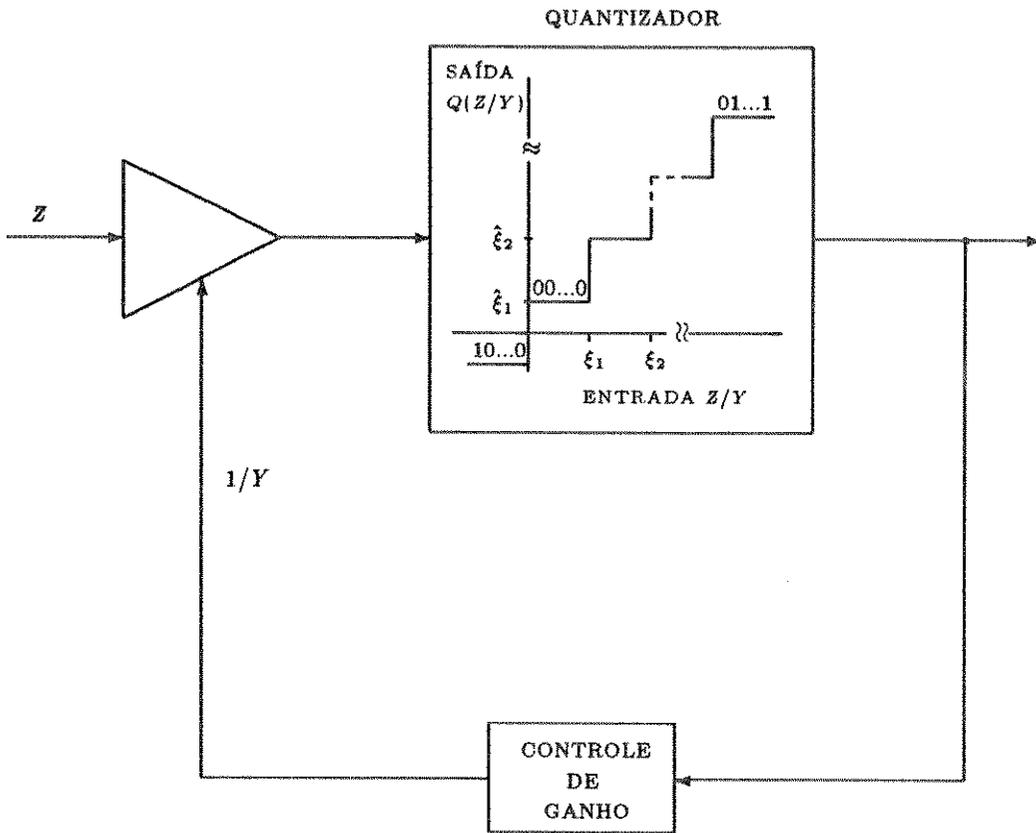


Figura II.2: Quantização adaptativa.

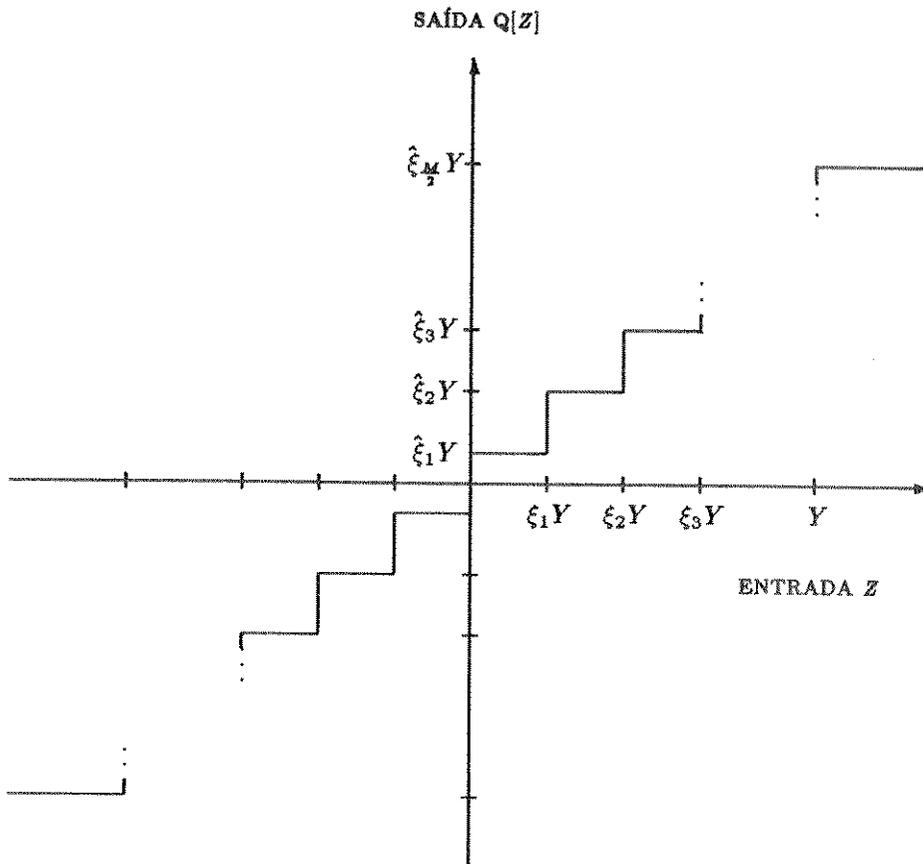


Figura II.3: Característica do quantizador.

FATOR DE CARGA (escala logarítmica)

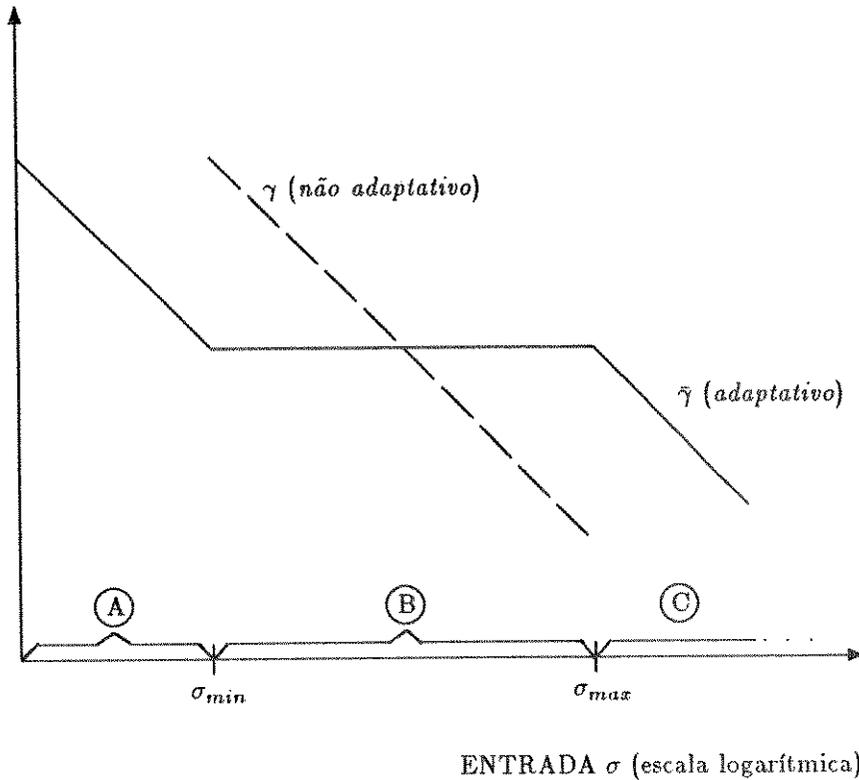


Figura II.4: Faixa dinâmica do quantizador.

sendo $\gamma = Y/\sigma$. Em um quantizador adaptativo, Y varia com o tempo de forma a reduzir a dependência do fator de carga γ em relação a σ . Uma vez que o sinal que controla Y contém informações estatísticas sobre σ , Y é necessariamente uma variável aleatória que flutua em torno de um valor médio \bar{Y} . O objetivo de projeto de um quantizador adaptativo é fazer com que o fator de carga $\bar{\gamma}$, definido por \bar{Y}/σ , permaneça independente de σ .

Na figura II.4 [9] são apresentadas curvas de variação do fator de escala em função da variância do sinal de entrada para um quantizador fixo e para um quantizador adaptativo. Através da figura verifica-se que o quantizador adaptativo apresenta melhor desempenho na região B, onde o fator de carga $\bar{\gamma} = \bar{Y}/\sigma$ é constante. As regiões A e C caracterizam o ruído de quantização granular e a distorção por sobrecarga, respectivamente.

• Regra de adaptação

A elaboração da presente regra de adaptação foi totalmente baseada na referência [9].

A condição para se encontrar $Z(k)$ dentro do i -ésimo intervalo de quantização positivo ou negativo é

$$Y(k)\xi_{i-1} \leq |Z(k)| < Y(k)\xi_i, \quad (\text{II.2})$$

onde $Y(k)$ é o valor de sobrecarga no instante k .

Neste caso, a saída do quantizador será $Q[Z(k)] = \pm Y(k)\hat{\xi}_i$ e a magnitude da palavra código será $|I(k)| = i - 1$.

Antes da próxima amostra $Z(k + 1)$ ser codificada, o ponto de sobrecarga varia da seguinte forma [10]:

$$Y(k + 1) = M(I(k))Y(k). \quad (\text{II.3})$$

Assim, considerando-se um quantizador com $2N$ níveis de saída, a ocorrência de uma palavra código com $|I(k)|$ próximo de 0 sugere que $Y(k)$ é muito grande e que deve ser diminuído. Reciprocamente, a ocorrência de $|I(k)|$ próximo de $N - 1$ sugere que o quantizador pode ser sobrecarregado e, portanto, $Y(k)$ deve ser aumentado. Desta forma, os N multiplicadores $M(I(k))$ devem satisfazer

$$\begin{aligned} M(0) < 1 \quad \text{e} \quad M(N - 1) > 1; \\ M(0) \leq M(1) \leq \dots \leq M(N - 1). \end{aligned} \quad (\text{II.4})$$

No caso de memória infinita, cada $Y(k)$ depende de toda a sequência de códigos passada, isto é,

$$Y(k + 1) = \left(\prod_{m=0}^k M(I(m)) \right) Y(0). \quad (\text{II.5})$$

A desvantagem da expressão acima é que o sistema torna-se bastante sensível a erros de transmissão. Isto pode ser explicado da seguinte forma: sejam \check{Y} e \check{I} as versões no receptor de Y e I , respectivamente. Assumindo-se que no instante $l < k$ tem-se $|I(l)| = i$, enquanto um erro de transmissão acarreta $|\check{I}(l)| = j$, se outros erros não ocorrerem,

$$\check{Y}(k + 1) = \frac{M(j)}{M(i)} Y(k + 1). \quad (\text{II.6})$$

Assim, cada erro de transmissão causa um erro multiplicativo entre o fator de adaptação do transmissor e do receptor que pode persistir indefinidamente.

Uma forma de minimizar o efeito de um erro de transmissão é através do seguinte procedimento de adaptação:

$$Y(k+1) = M(I(k))Y^\beta(k) = \left(\prod_{m=0}^k [M(I(m))]^{\beta^{k-m}} \right) Y^{\beta^{k+1}}(0), \quad (\text{II.7})$$

onde $0 < \beta < 1$.

Da mesma forma, se no instante l , $|I(l)| = i$ e, devido a um erro de transmissão, $|\tilde{I}(l)| = j$,

$$\tilde{Y}(k+1) = \left[\frac{M(j)}{M(i)} \right]^{\beta^{k-l}} Y(k+1). \quad (\text{II.8})$$

Verifica-se que, quando $\beta < 1$, o erro no fator de adaptação devido a cada erro de transmissão diminui exponencialmente, tornando o quantizador mais robusto em comparação com o procedimento caracterizado pela expressão (II.3), onde $\beta = 1$. Mas, analisando-se a figura II.5 [9], quando $\beta < 1$, $\bar{\gamma}$ varia com σ . Verifica-se também que conforme β se distancia de 1, $\bar{\gamma}$ torna-se mais sensível a variações em σ . Portanto, o valor de projeto de β é um compromisso entre a sensibilidade do fator de carga médio do quantizador $\bar{\gamma}$ a variações no valor rms de entrada σ e a sensibilidade do receptor a efeitos de erros de transmissão.

A figura II.1 mostra que y_u é obtido através da adaptação robusta de y . De fato, segundo a referência [8], o fator de escala "unlocked" Y_u é gerado a partir da expressão (II.7), isto é,

$$Y_u(k) = M(I(k))Y^\beta(k), \quad (\text{II.9})$$

e, no domínio logarítmico,

$$y_u(k) = \beta y(k) + \log_2 M(I(k)). \quad (\text{II.10})$$

Verifica-se que na expressão (II.10) tem-se $y_u(k)$ e não $y_u(k+1)$, uma vez que, segundo a expressão (II.1), se $a_l(k) = 1$, $y(k) = y_u(k-1)$.

II.2.2 Fator de escala y_l

O fator de escala y_l é gerado através de filtragem passa-baixas de y_u [8]:

$$y_l(k) = \alpha y_l(k-1) + (1-\alpha)y_u(k), \quad (\text{II.11})$$

onde $0 < \alpha < 1$. Sendo $\alpha < 1$, o sistema apresenta memória finita e, portanto, o efeito dos erros de transmissão se desvanece com o tempo.

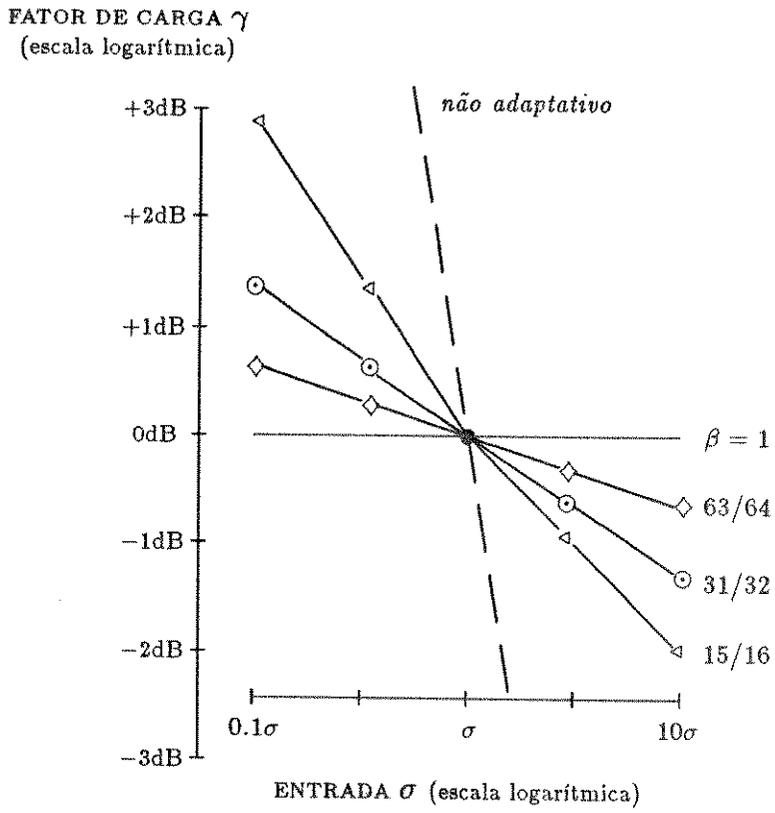


Figura II.5: Faixa dinâmica do quantizador robusto.

Apêndice III

CONSTANTE DE TEMPO DOS FILTROS DIGITAIS

Os filtros digitais que constituem o codificador/decodificador ADPCM da Recomendação G.721 são, em geral, da seguinte forma:

$$y(n) = \alpha_1 y(n-1) + \alpha_2 x(n). \quad (\text{III.1})$$

A resposta impulsiva do filtro da equação (III.1) é dada por

$$h(n) = \alpha_2 \alpha_1^n u(n), \quad (\text{III.2})$$

onde

$$u(n) = \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases} \quad (\text{III.3})$$

Para $0 < \alpha_1 < 1$, $h(n)$ tende exponencialmente a zero com uma constante de

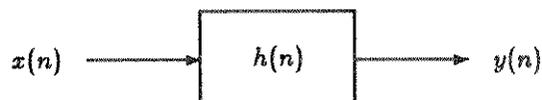


Figura III.1: Filtro digital de resposta impulsiva $h(n)$.

tempo τ dada por

$$e^{-nT/\tau} = \alpha_1^n, \quad (\text{III.4})$$

onde T é o período de amostragem. Assim,

$$\tau = \frac{-T}{\ln \alpha_1}. \quad (\text{III.5})$$

Apêndice IV

ESTRUTURA DO PREDITOR ADAPTATIVO

Várias estruturas constituídas de pólos e zeros foram estudadas [5,11] antes da obtenção da versão final do preditor adaptativo.

A estrutura “backward” foi preferida à estrutura “forward”, considerando-se que a primeira apresenta um atraso de codificação menor, bem como não necessita transmitir informação lateral (“side information”).

Inicialmente, foi proposta uma estrutura adaptativa constituída somente por pólos, de modo que a função de transferência do preditor era

$$P(z) = \frac{1}{1 - \sum_{i=1}^N a_i(k)z^{-i}}, \quad (\text{IV.1})$$

sendo $\{a_i(k), i = 1, \dots, N\}$ um conjunto de coeficientes no instante de amostragem k . Estes coeficientes são atualizados a cada nova amostra do sinal de entrada, de modo a minimizar a potência do sinal diferença $d(k)$.

Entretanto, de acordo com esta estrutura, quando ocorrer um erro de transmissão, as posições dos pólos poderão forçosamente se mover para fora da circunferência de raio unitário, e, conseqüentemente, o sistema torna-se instável.

Segundo a referência [11], uma alternativa para contornar o problema da estabilidade é a utilização de uma composição de pólos e zeros, de modo que a

nova função de transferência do preditor é

$$P(z) = \frac{1 + \sum_{i=1}^M b_i(k) z^{-i}}{1 - \sum_{i=1}^N a_i z^{-i}}, \quad (\text{IV.2})$$

onde os coeficientes a_i , que modelam os pólos, são fixos e os coeficientes $b_i(k)$, que modelam os zeros, variam de acordo com

$$b_i(k) = (1 - \delta_b) b_i(k-1) + \alpha_b d_q(k) d_q(k-i) / \sum_{i=1}^M (d_q(k-i))^2, \quad (\text{IV.3})$$

onde $\delta_b < 1$ e α_b é uma constante positiva escolhida experimentalmente.

Para implementação em tempo real, uma simplificação foi introduzida no método de adaptação dos coeficientes $b_i(k)$ [5]:

$$b_i(k) = (1 - \delta_b) b_i(k-1) + \alpha_b \text{sgn}[d_q(k)] \text{sgn}[d_q(k-i)], \quad (\text{IV.4})$$

Especificamente para o preditor da Recomendação G.721, são propostas uma estrutura de zeros de sexta ordem ($M = 6$), cujos coeficientes b_i são adaptados de acordo com a expressão (IV.4), e uma estrutura de pólos de segunda ordem ($N = 2$) com as seguintes equações de atualização dos coeficientes a_i ¹:

$$a_1(k) = (1 - \delta_{a_1}) a_1(k-1) + \alpha_{a_1} \text{sgn}[p(k)] \text{sgn}[p(k-1)] \quad (\text{IV.5})$$

e

$$a_2(k) = (1 - \delta_{a_2}) a_2(k-1) + \alpha_{a_2} \{ \text{sgn}[p(k)] \text{sgn}[p(k-2)] - 2a_1(k) \text{sgn}[p(k)] \text{sgn}[p(k-1)] \}, \quad (\text{IV.6})$$

onde $p(k)$ é a soma do sinal diferença quantizado $d_q(k)$ com o sinal reconstruído da estrutura de zeros $s_{ez}(k)$:

$$p(k) = d_q(k) + s_{ez}(k). \quad (\text{IV.7})$$

¹A definição desta estrutura de pólos adaptativos foi baseada na referência [5].

Apêndice V

COMPLEXIDADE EM CODIFICAÇÃO COM BUSCA MULTI-CAMINHO

A complexidade em codificação com busca multi-caminho C_t é constituída de duas componentes [16]: complexidade computacional C , que é definida como sendo o número de ramos que devem ser examinados para codificar uma amostra do sinal diferença, e memória M , que é definida como sendo o número total de amostras de um dado parâmetro que devem ser armazenadas em um determinado instante no processo de codificação. Deve-se salientar que na análise seguinte considerou-se o efeito de apenas um dos parâmetros envolvidos no processo de codificação¹, o sinal ADPCM $I(k)$, sendo que tal parâmetro está associado à maior complexidade C_t .

A seguir, são apresentadas as complexidades de dois esquemas ADPCM, um empregando busca exaustiva e outro empregando o algoritmo de busca (M, L) , sendo posteriormente feita uma comparação entre eles.

- Esquema ADPCM com busca multi-caminho exaustiva

Seja R o número de bits empregado na codificação de cada amostra do sinal diferença $d(k)$. Considerando-se que a busca envolve $L + 1$ amostras (janela

¹Conforme mencionado no Capítulo 4, podem ser citados, entre outros, os seguintes parâmetros responsáveis pela operação do codec: medida de distorção, sinal ADPCM codificado $I(k)$, sinal diferença quantizado $d_q(k)$, sinal estimado $s_e(k)$, sinal reconstruído $s_r(k)$, sinal fator de escala de adaptação $y(k)$ etc.

de observação), a complexidade computacional é dada pelo número de ramos contido nesta janela, ou seja,

$$C = 2^{R(L+1)},$$

que cresce exponencialmente com L . A memória é dada por

$$M = (L + 1)2^{R(L+1)} \tag{V.1}$$

e também cresce exponencialmente com L .

Como ilustração, na figura V.1 encontra-se a estrutura de árvore associada à palavra código $I(k)$ para $R = 1$ e $L = 2$.

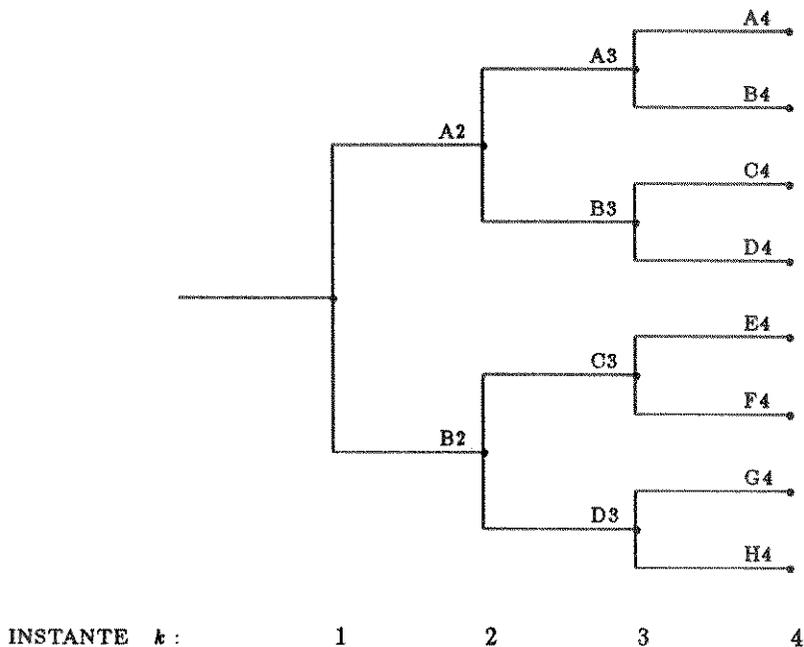


Figura V.1: Estrutura de árvore associada a $I(k)$ para $R = 1$ e $L = 2$.

De acordo com o algoritmo implementado, no instante $k = 4$ é decidido se o sinal ADPCM $I(k = 2) = A2$ ou $I(k = 2) = B2$ é transmitido, sendo que tal decisão é baseada na análise de caminhos futuros ao longo da árvore dentro da janela de observação de $L + 1 = 3$ amostras (instantes $k = 2, 3$ e 4). Assim, a complexidade, que é definida como o número de ramos contido

nesta janela, é 8 ($A4A3A2, B4A3A2, C4B3A2, \dots, G4D3B2, H4D3B2$). Ressalta-se que, para garantir a operacionalidade do algoritmo, para cada ramo, as 3 amostras do parâmetro $I(k)$ contidas na janela de observação devem ser armazenadas, de modo que neste caso a memória é dada por $3 \times 8 = 24$.

- Esquema ADPCM com busca multi-caminho não exaustiva

O codec ADPCM implementado enquadra-se neste caso, sendo que emprega como procedimento de busca multi-caminho o Algoritmo (M, L) , onde M é o número máximo de seqüências (ou caminhos na árvore) selecionadas na janela de observação de $L + 1$ amostras, sendo L definido como profundidade da árvore.

Para este esquema, a complexidade computacional é

$$C = M2^R, \quad (\text{V.2})$$

qualquer que seja a profundidade da árvore L , e a memória é

$$M \leq (L + 1)M2^R. \quad (\text{V.3})$$

Assim, a complexidade total para o codificador em questão é dada por

$$C_t \leq (L + 2)M2^R. \quad (\text{V.4})$$

Para o codificador operando à taxa de 16 kbit/s com amostragem a 8 kHz e $R = 2$ têm-se:

- Esquema ADPCM com busca multi-caminho exaustiva

$$C_t = (L + 2)2^{2(L+1)};$$

- Esquema ADPCM com busca multi-caminho não exaustiva

$$C_t \leq (L + 2)M2^2.$$

Comparando-se a complexidade total C_t dos dois esquemas apresentados, verifica-se que enquanto o procedimento com busca exaustiva apresenta complexidade que cresce exponencialmente com a janela de observação (ou seja, com a profundidade da árvore L), no procedimento de busca não exaustiva tal crescimento é função linear de L .

Apêndice VI

EFEITO DO PARÂMETRO α NA CONFORMAÇÃO DO ESPECTRO DO SINAL FILTRADO

Sabe-se que, para sinais de voz sonoros ou vocálicos, uma representação bastante satisfatória dos efeitos do trato vocal é o modelo “all-pole”¹ [6], caracterizado pela estrutura

$$H(z) = \frac{G}{1 - \sum_{i=1}^P a_i z^{-i}}, \quad (\text{VI.1})$$

onde G é um ganho e P , a ordem do modelo.

Uma vez que os coeficientes a_i são reais, os pólos de $H(z)$ ou são reais, ou ocorrem em pares complexos conjugados,

$$z_i, z_i^* = \rho_i e^{\pm j\omega_i}, \quad 0 \leq \rho_i \leq 1, \quad (\text{VI.2})$$

sendo que para cada par está associada uma formante. Na figura VI.1 é ilustrado o espectro $|H(e^{j\omega})|$, onde foram consideradas apenas três formantes com frequências centrais W_1, W_2 e W_3 .

Quando $\rho_i \rightarrow 1$, os picos nas regiões das formantes ficam mais acentuados e

¹As frequências das formantes correspondem às frequências de ressonância do trato vocal e estão associadas aos pólos do modelo.

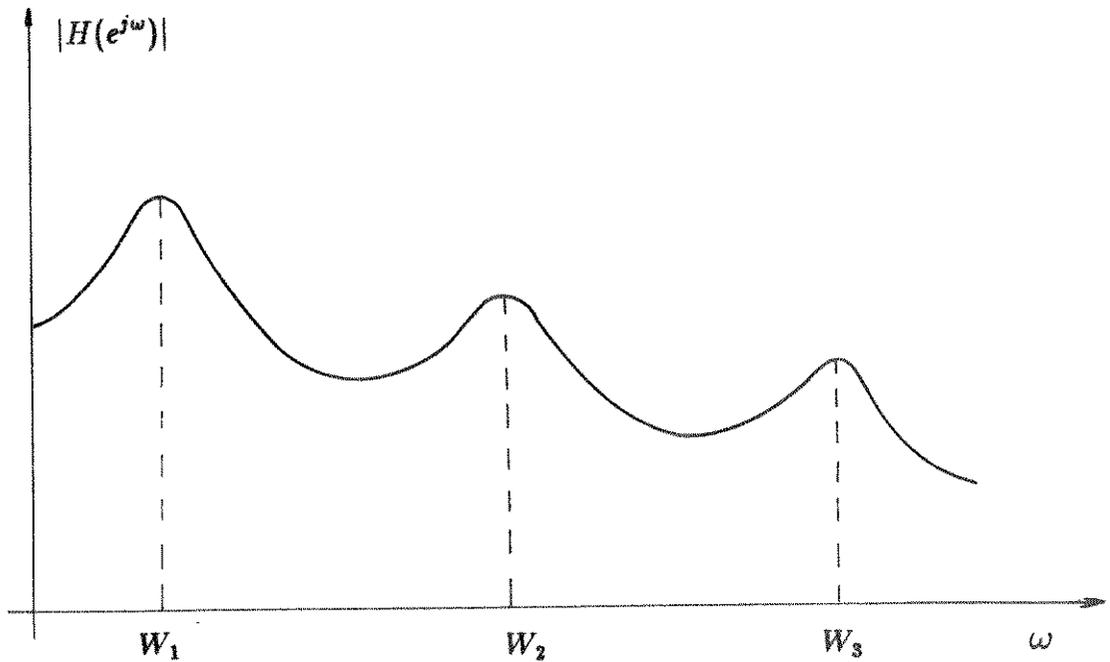


Figura VI.1: Representação das formantes W_1, W_2 e W_3 no espectro $|H(e^{j\omega})|$.

a largura de faixa das mesmas diminui. Fazendo-se

$$H'(z) = H(z/\alpha), \quad 0 \leq \alpha \leq 1, \quad (\text{VI.3})$$

os pólos de $H'(z)$ correspondem a

$$z'_i, z'^*_i = \alpha \rho_i e^{\pm j\omega_i}, \quad 0 \leq \rho_i \leq 1, \quad 0 \leq \alpha \leq 1, \quad (\text{VI.4})$$

de modo que para $\alpha \rightarrow 0$, os pólos se distanciam do círculo de raio unitário e a largura de faixa da região das formantes aumenta. Tal fato é ilustrado na figura VI.2.

APÊNDICE VI. EFEITO DO PARÂMETRO α NA CONFORMAÇÃO DO ESPECTRO DO SINAL FILTRADO

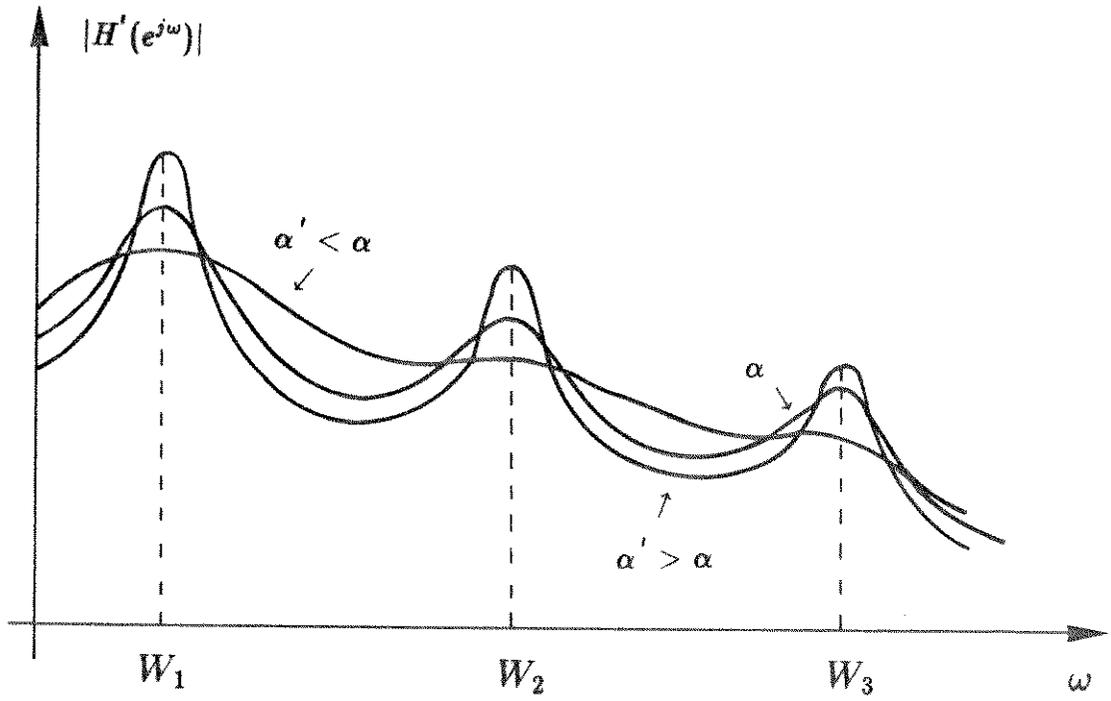


Figura VI.2: Espectro $|H'(e^{j\omega})|$ em função do parâmetro α .

Apêndice VII

MEDIDAS DE DESEMPENHO OBJETIVO

Sabe-se que o desempenho objetivo dos codecs simulados foi obtido através de medidas de *relação sinal/ruído* S/R e *relação sinal/ruído segmental* S/R_{seg} , que são definidas a seguir.

VII.1 Relação sinal/ruído S/R

A relação sinal/ruído no domínio logarítmico é definida como

$$[S/R]_{\log} = 10 \log_{10} \frac{P_s}{P_r} \quad \text{dB}, \quad (\text{VII.1})$$

onde P_s é a potência do sinal e P_r é a potência de ruído, ambas medidas sobre todas as amostras do arquivo de voz, ou seja,

$$P_s = \frac{1}{N_t} \sum_{k=1}^{N_t} s^2(k) \quad (\text{VII.2})$$

e

$$P_r = \frac{1}{N_t} \sum_{k=1}^{N_t} [s(k) - s_r(k)]^2, \quad (\text{VII.3})$$

onde N_t é o número total de amostras do arquivo de voz, $s(k)$ é o sinal de entrada do codificador e $s_r(k)$ é o sinal reconstruído na saída do decodificador.

VII.2 Relação sinal/ruído segmental S/R_{seg}

Para esta medida, os arquivos de voz são divididos em segmentos de 128 amostras. Considerando-se que a frequência de amostragem é 8000 Hz, estes segmentos apresentam duração de 16 ms. Em seguida, a relação sinal/ruído no domínio logarítmico de cada segmento i é calculada da seguinte forma:

$$[S/R(i)]_{log} = 10 \log_{10} \frac{P_s(i)}{P_r(i)} \quad \text{dB}, \quad (\text{VII.4})$$

onde $P_s(i)$ e $P_r(i)$, são, respectivamente, as potências de sinal e ruído correspondentes a cada segmento i de 128 amostras.

Posteriormente faz-se uma média de todas as $S/R(i)$ computadas e obtém-se a seguinte relação sinal/ruído segmental:

$$[S/R_{seg}]_{log} = \frac{1}{N_s} \sum_{i=1}^{N_s} [S/R(i)]_{log} \quad \text{dB}. \quad (\text{VII.5})$$

Referenciando-se à expressão (VII.5), N_s é o número de segmentos empregados no cálculo da S/R_{seg} . Uma vez que segmentos onde o silêncio é predominante podem afetar indevidamente esta medida, estabeleceu-se que não são considerados no cálculo da S/R_{seg} os segmentos cuja potência de sinal é mais de 50 dB inferior à máxima potência medida entre estes segmentos ($[P_s]_{max}$). Assim, N_s é definido como sendo o número de segmentos i com potência de sinal $P_s(i)$ tal que

$$[P_s]_{max} - P_s(i) < 50 \text{ dB}. \quad (\text{VII.6})$$

Referências Bibliográficas

- [1] Recomendação do CCITT G.721, "32 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)", Livro Vermelho, Outubro de 1984.
- [2] N. Benvenuto, G. Bertocci, "Enhancement in the 32 kbit/s ADPCM Algorithm for FSK Voice-Band Data Signals", IEEE Transactions on Communications, vol. 36, pag. 626-628, Maio de 1988.
- [3] Recomendação do CCITT G.711, "Pulse Code Modulation (PCM) of Voice Frequencies", Livro Vermelho, Outubro de 1984.
- [4] J. Max, "Quantizing for Minimum Distortion", IRE Transactions on Information Theory, vol. IT-6, pag. 16-21, Março de 1960.
- [5] D. Millar, P. Mermelstein, "Prevention of Predictor Mistracking in ADPCM Coders", IEEE ICC'84, Conference Proceedings, pag. 1508-1512, Junho de 1984.
- [6] L. R. Rabiner, R. W. Schafer, "Digital Processing of Speech Signals", pag. 120-126 e 191-199, Englewood Cliffs, Prentice Hall, 1978.
- [7] D. Cointot, "A 32 Kbit/sec ADPCM Coder Robust to Channel Errors", IEEE GLOBECOM'82, Conference Proceedings, pag. 964-967, Novembro-Dezembro de 1982.
- [8] D. W. Petr, "32 Kbps ADPCM-DLQ Coding for Network Applications", IEEE GLOBECOM'82, Conference Proceedings, pag. 239-243, Novembro-Dezembro de 1982.
- [9] D. J. Goodman, R. M. Wilkinson, "A Robust Adaptive Quantizer", IEEE Transactions on Communications, pag. 1362-1365, Novembro de 1975.
- [10] N. S. Jayant, "Adaptive Quantization with a one-word memory", Bell Syst. Tech. Journal, vol. 52, pag. 1119-1144, Setembro de 1973.
- [11] T. Nishitani, S. Aikoh, T. Araseki, K. Ozawa, R. Maruta, "32 Kbit/sec ADPCM Codec Using a Single Chip Signal Processor", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, pag. 960-963, Abril de 1982.

- [12] C. H. Belinazo, "Codificador/Decodificador ADPCM com taxa de bits variável - 40, 32, 24 e 16 Kbit/s", Relatório Técnico Parcial/ PDS - CPqD, Março de 1988.
- [13] J. D. Gibson, G. B. Haschke, "Adaptive Code Generators for Tree Coding of Speech", Proceedings of IEEE International Conference on Communications, pag. 1142-1146, Junho de 1987.
- [14] J. D. Gibson, G. B. Haschke, "Backward Adaptive Tree Coding of Speech at 16 kbps", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, pag. 251-254, Abril de 1988.
- [15] T. Berger, "Rate Distortion Theory. A Mathematical Basis for Data Compression", Englewood Cliffs, Prentice Hall, 1971.
- [16] N. S. Jayant, P. Noll, "Digital Coding of Waveforms-Principles and Applications to Speech and Video", pag. 440-442 e 443-444, Englewood Cliffs, Prentice Hall, 1984.
- [17] N. S. Jayant, S. A. Christensen, "Tree-Encoding of Speech Using the (M,L)-Algorithm and Adaptive Quantization", IEEE Transactions on Communications, vol.COM-26, pag. 1376-1379, Setembro de 1978.
- [18] J. B. Anderson, J. B. Bodie, "Tree Encoding of Speech", IEEE Transactions on Information Theory, vol. IT-21, pag. 379-387, Julho de 1975.
- [19] T. Svendsen, "Tree Encoding of LPC Residual", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, pag. 10.11.1-10.11.4, Abril de 1984.
- [20] Recomendação do CCITT G.114, "Mean One-Way Propagation Time", Livro Vermelho, Outubro de 1984.
- [21] Recomendação do CCITT G.161, "Echo-suppressors suitable for circuits having either short or long propagation times", Livro Vermelho, Outubro de 1984.
- [22] "Report on Question 27/Group XVIII Rapporteur's Group Meeting", Turin, 1-4 de Dezembro de 1987.
- [23] F. M. Ferreira, J. S. Yamamoto, F. Violaro, "Codec ADPCM com Quantização de Árvore", pag. 528-530, 7º Simpósio Brasileiro de Telecomunicações, Setembro de 1989.
- [24] CPqD-Telebrás, "16 kbit/s ADPCM with Tree Quantizer", Contribuição AH.89-D32 para o Grupo XV do CCITT, Julho de 1989.

- [25] N. S. Jayant, V. Ramamoorthy, "Adaptive Postfiltering of 16 kbit/s ADPCM Speech", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, pag. 829-832, Abril de 1986.
- [26] V. Ramamoorthy, N. S. Jayant, R. V. Cox, M. M. Sondhi, "Enhancement of ADPCM Speech Coding with Backward-Adaptive Algorithms for Postfiltering and Noise Feedback", IEEE Journal on Selected Areas in Communications, vol. 6, pag. 364-382, Fevereiro de 1988.
- [27] B. S. Atal, J. R. Remde, "A New Model of LPC Excitation for Producing Natural-Sounding Speech at Low Bit Rates", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, pag. 614-617, Abril de 1982.
- [28] B. S. Atal, M. R. Schroeder, "Predictive Coding of Speech Signals and Subjective Error Criteria", IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-27, pag. 247-254, Junho de 1979.
- [29] Simon Fraser University, University of California, Voicecraft Inc., "Description of the Consortium's Low Delay Vector Excitation Coder (LD-VXC), Version 2", Contribuição AH.89-D21 para o Grupo XV do CCITT, Junho de 1989.
- [30] F. M. Ferreira, J. S. Yamamoto, F. Violaro, "Enhanced ADPCM Tree Codec at 16 and 9.6 Kbit/s", Proceedings of the Fifth European Signal Processing Conference, vol. II, pag. 135-1338, Setembro de 1990.
- [31] J. D. Gibson, W. W. Chang, "Fractional Rate Multi-Tree Speech Coding", Department of Electrical Engineering, Texas A&M University.
- [32] Fábio Violaro, "Nova Versão do Sistema de Análise e Processamento Digital de Voz SAPDV-A", pag 50-53, 7º Simpósio Brasileiro de Telecomunicações, Setembro de 1989.