



Universidade Federal do Paraná  
Setor de Tecnologia  
Departamento de Engenharia Elétrica

## **Codificação de Áudio – Recomendação G.726**

Trabalho apresentado ao Prof<sup>o</sup> Marcelo de  
Oliveira Rosa.

Mestrando: Franco Brandelero

Curitiba  
Novembro de 2008

## 1. Introdução

A recomendação G.726 feita pelo CCITT em 1990, hoje conhecido como ITU-T, especifica a compressão do sinal de áudio para telefonia, usando-se a técnica ADPCM (Adaptive Differential Pulse Code Modulation). O uso da codificação G.726 permite que a taxa de transmissão de 64 kbps seja reduzida para 40, 32, 24 ou até 16 kbps.

## 2. Codificação

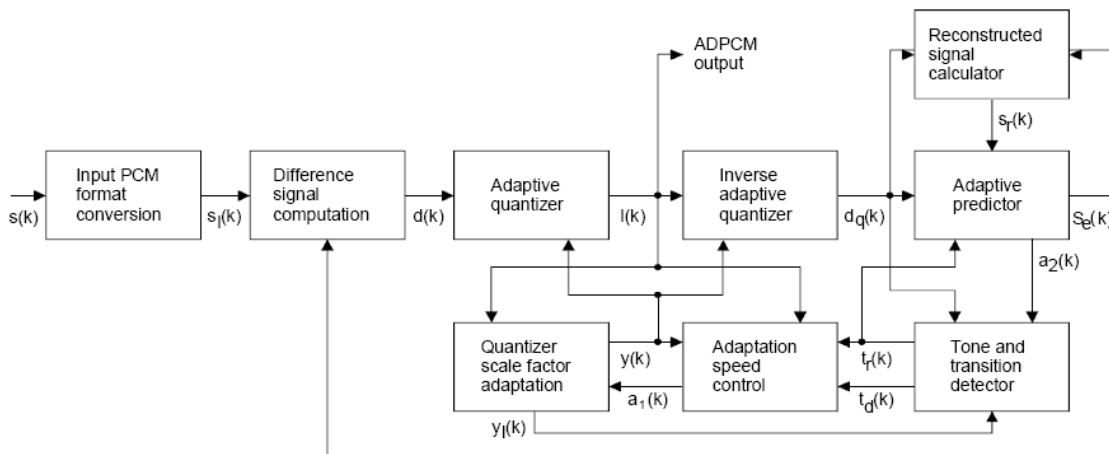


Figura 1 – Diagrama de Blocos do Codificador.

Primeiramente o sinal PCM lei-A ou lei- $\mu$  de entrada é convertido em um sinal PCM uniforme. A esse sinal é subtraído um sinal estimado e depois quantizado em uma quantidade menor de bits e enviado para transmissão.

Um quantizador inverso produz o sinal diferença novamente. O sinal estimado é adicionado ao sinal diferença para produzir a versão reconstruída do sinal de entrada. O sinal reconstruído e o sinal diferença são operados sobre um preditor adaptativo que produz o sinal de entrada estimado, assim completando o laço de realimentação.

Para cada variável descrita,  $k$ , é o índice de amostragem e as amostras são capturadas a cada  $125\mu\text{s}$ .

### 2.1. Conversor do formato PCM de entrada

Esse bloco converte o sinal  $s(k)$  de PCM lei-A ou lei- $\mu$  em um sinal PCM uniforme  $s_l(k)$ .

### 2.2. Cálculo do sinal Diferença

Este bloco calcula o sinal diferença  $d(k)$  do sinal PCM uniforme  $s_l(k)$  e do sinal estimado  $s_e(k)$ :

$$d(k) = s_l(k) - s_e(k)$$

### 2.3. Quantizador Adaptativo e Quantizador Adaptativo Inverso

Um quantizador adaptativo não-uniforme de 31, 15, 7 ou 4 níveis é usado para quantizar o sinal diferença  $d(k)$  operando a taxas de 40, 32, 24 ou 16 kbps, respectivamente.

Antes da quantização,  $d(k)$  é convertido para uma representação logarítmica de base 2 e então escalonado por  $y(k)$  que é computado pelo bloco de adaptação do fator escalar.

O quantizador adaptativo inverso produz  $d_q(k)$ , usando  $y(k)$  e os valores dados nas tabelas abaixo e então transformando o resultado a partir do domínio logarítmico.

#### 2.3.1. Operação à 40 kbps

Cinco bits são usados para especificar o nível quantizado (quatro para magnitude e um para sinal).

Entrada do quantizador $\log_2  d(k)  - y(k)$	$ I(k) $	Saída do quantizador inverso $\log_2  d_q(k)  - y(k)$
[4.31, $+\infty$ )	15	4.42
[4.12, 4.31)	14	4.21
[3.97, 4.12)	13	4.02
[3.70, 3.91)	12	3.81
[3.47, 3.70)	11	3.59
[3.22, 3.47)	10	3.35
[2.95, 3.22)	9	3.09
[2.64, 2.95)	8	2.8
[2.32, 2.64)	7	2.48
[1.95, 2.32)	6	2.14
[1.54, 1.95)	5	1.75
[1.08, 1.54)	4	1.32
[0.52, 1.08)	3	0.81
[-0.13, 0.52)	2	0.22
[-0.96, -0.13)	1	-0.52
$(-\infty, -0.96)$	0	$-\infty$

#### 2.3.2. Operação à 32 kbps

Quatro bits são usados para especificar o nível quantizado (três para magnitude e um para sinal).

Entrada do quantizador $\log_2  d(k)  - y(k)$	$ I(k) $	Saída do quantizador inverso $\log_2  d_q(k)  - y(k)$
[3.12, $+\infty$ )	7	3.32
[2.72, 3.12)	6	2.91
[2.34, 2.72)	5	2.52
[1.91, 2.34)	4	2.13
[1.38, 1.91)	3	1.66

[0.62, 1.38)	2	1.05
[-0.98, 0.62)	1	0.031
(-∞, -0.98)	0	-∞

### 2.3.3. Operação à 24 kbps

Três bits são usados para especificar o nível quantizado (dois para magnitude e um para sinal).

Entrada do quantizador $\log_2  d(k)  - y(k)$	$ I(k) $	Saída do quantizador inverso $\log_2  d_q(k)  - y(k)$
[2.58, +∞)	3	2.91
[1.70, 2.58)	2	2.13
[0.06, 1.70)	1	1.05
(-∞, 0.06)	0	-∞

### 2.3.4. Operação à 16 kbps

Dois bits são usados para especificar o nível quantizado (um para magnitude e um para sinal).

Entrada do quantizador $\log_2  d(k)  - y(k)$	$ I(k) $	Saída do quantizador inverso $\log_2  d_q(k)  - y(k)$
[2.04, +∞)	1	2.85
(-∞, -2,04)	0	0.91

## 2.4. Adaptação do fator de escala do Quantizador

O princípio básico usado para escalonar o quantizador é a adaptação bimodal, isto é, rápida para voz e lenta para dados e tons.

O fator rápido é calculado recursivamente:

$$y_u(k) = (1 - 2^{-5})y(k) + 2^{-5}W[I(k)]$$

Onde  $y_u(k)$  é limitado entre 1.06 e 10.00.

$W(I)$  é definido na tabela abaixo:

$ I(k) $	$W(I)$ à 40kbps	$W(I)$ à 32kbps	$W(I)$ à 24kbps	$W(I)$ à 16kbps
15	43.50	-	-	-
14	33.06	-	-	-
13	27.50	-	-	-
12	22.38	-	-	-
11	17.50	-	-	-
10	13.69	-	-	-
9	11.19	-	-	-

8	8.81	-	-	-
7	6.29	70.13	-	-
6	3.63	22.19	-	-
5	2.56	12.38	-	-
4	2.50	7.00	-	-
3	2.44	4.00	36.38	-
2	1.50	2.56	8.56	-
1	0.88	1.13	1.88	27.44
0	0.88	-0.75	-0.25	-1.38

O fator lento é derivado de  $y_u(k)$  através de um filtro passa-baixa:

$$y_l(k) = (1 - 2^{-6})y_l(k-1) + 2^{-6}y_u(k)$$

O fator rápido e lento são combinados para formar o fator de escala resultante:

$$y(k) = a_l(k)y_u(k-1) + [1 - a_l(k)]y_l(k-1)$$

Onde  $a_l(k)$ , o controle de velocidade de adaptação, é limitado entre 0 e 1.

## 2.5. Controle de velocidade de Adaptação

Esse parâmetro tende à unidade para sinais de voz e tende à zero para sinais de dados.

Dois medidas da magnitude média de  $I(k)$  são calculadas:

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

e

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

Onde  $F[I(k)]$  é definido como:

$ I(k) $	$F(I)$ à 40kpbs	$F(I)$ à 32kpbs	$F(I)$ à 24kpbs	$F(I)$ à 16kpbs
15	6	-	-	-
14	6	-	-	-
13	5	-	-	-
12	4	-	-	-
11	3	-	-	-
10	2	-	-	-
9	1	-	-	-
8	1	-	-	-
7	1	7	-	-
6	1	3	-	-
5	1	1	-	-
4	0	1	-	-
3	0	1	7	-
2	0	0	2	-

1	0	0	1	7
0	0	0	0	0

Usando essas duas médias, a variável  $a_p(k)$  é definida:

$$a_p(k) = \begin{cases} (1-2^{-4})a_p(k-1) + 2^{-3}, & \text{se } |d_{ms}(k) - d_{ml}| \geq 2^{-3}d_{ml}(k) \\ (1-2^{-4})a_p(k-1) + 2^{-3}, & \text{se } y(k) < 3 \\ (1-2^{-4})a_p(k-1) + 2^{-3}, & \text{se } t_d(k) = 1 \\ 1, & \text{se } t_r(k) = 1 \\ (1-2^{-4})a_p(k-1), & \text{outro} \end{cases}$$

E então definimos controle de velocidade de adaptação como:

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

Este limite assimétrico tende a eliminar transições prematuras de sinais pulsados de entrada como transmissão de dados.

## 2.6. Preditor Adaptativo e Cálculo do sinal reconstruído

O sinal estimado é dado por:

$$s_e(k) = \sum_{i=1}^2 a_i(k-1)s_r(k-i) + s_{ez}(k)$$

Onde

$$s_{ez}(k) = \sum_{i=1}^6 b_i(k-1)d_z(k-i)$$

E o sinal reconstruído é definido como:

$$s_r(k-i) = s_e(k-i) + d_q(k-i)$$

Onde

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

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

$$p(k) = d_q(k) + s_{ez}(k)$$

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

e  $\text{sgn}[0]=1$ , exceto quando  $i=0$  e  $p(k-i)=0$  e  $\text{sgn}[p(k-i)]=0$ .

$$b_i(k) = (1 - 2^{-8})b_i(k-1) + 2^{-7} \operatorname{sgn}[d_q(k-i)]$$

## 2.7. Detector de transição e Tom

Para melhorar a performance de sinais FSK originados por modems, uma detecção em dois passos é definida:

$$t_d(k) = \begin{cases} 1, & a_2(k) < -0.71875 \\ 0, & \text{outro} \end{cases}$$

$$t_r(k) = \begin{cases} 1, & a_2(k) < -0.71875 \text{ e } |d_q(k)| > 24 \cdot 2^{y_1(k)} \\ 0, & \text{outro} \end{cases}$$

## 3. Decodificação

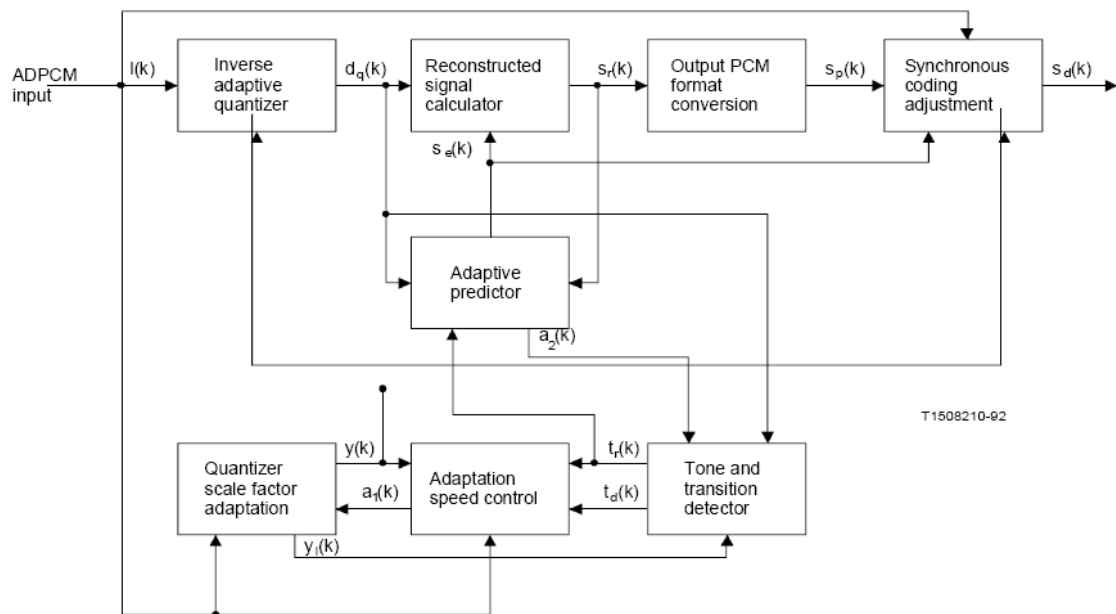


Figura 2 - Diagrama de Blocos do Decodificador.

Com exceção do bloco de ajuste de código Síncrono, todos os outros blocos já foram descritos no codificador.

### 3.1. Ajuste de código Síncrono

Este bloco previne a acumulação de distorção que ocorre em codificação síncrona de tandem (ADPCM-PCM-ADPCM).

Se o codificador e decodificador têm diferentes condições iniciais, então o sincronismo pode levar tempo. Se esta propriedade está perturbada, ela pode ser recobrada para os sinais cujo nível espectral ocupa a maioria da banda entre 200 e 3400 Hz (ex: voz).

Isto é conseguido, primeiro convertendo-se o sinal lei A ou  $\mu$  em um sinal PCM uniforme  $s_{tx}(k)$  e então calculando-se o sinal diferença  $d_x(k)$ :

$$d_x(k) = s_{tx}(k) - s_e(k)$$

O sinal diferença é então comparado ao intervalo de decisão do quantizador ADPCM.

O sinal  $s_d(k)$  é então definido abaixo:

$$s_d(k) = \begin{cases} s_p^+(k), & d_x(k) < \text{Limite Inferior do Intervalo} \\ s_p^-(k), & d_x(k) \geq \text{Limite Superior do Intervalo} \\ s_p(k), & \text{Outro Caso} \end{cases}$$

#### 4. Conclusão

O G.726 possui uma qualidade praticamente idêntica ao G.711, porém com menos consumo de banda. O G.726, perdeu espaço durante os anos 90, devido a sua incapacidade de trabalhar com sinais de modem e de fax maiores que 12 kbps, porém ele desperta o interesse de quem implementa sistemas de telefonia e transmissão de voz por exigir pouco processamento.