

TRANSMISSÃO DE IMAGENS COM CODIFICAÇÃO DPCM EM PIRÂMIDE

RICARDO L. DE QUEIROZ

JOÃO B. T. YABU-UTI

DEPARTAMENTO DE COMUNICAÇÕES - FACULDADE DE ENGENHARIA ELÉTRICA

UNIVERSIDADE ESTADUAL DE CAMPINAS

CP 6101 - CAMPINAS - SP 13081

Resumo :

A codificação em pirâmide se baseia na representação da imagem original com diferentes resoluções, imaginando cada destas imagens como dispostas em uma pirâmide. A base da pirâmide é a imagem original, o segundo nível tem metade da resolução, o terceiro $1/4$ e assim por diante. No trabalho, nos limitaremos a 2 e 3 níveis. A novidade aparece quando codificamos o terceiro nível com DPCM preditivo e os restantes com um esquema interpolativo, formando um sistema híbrido. Usaremos interpoladores baseados na Convolução Cúbica, comentaremos vantagens e desvantagens e simularemos o método comparando-o ao DPCM convencional.

codificação diferenciada em frequência, como em codificação Sub-Bandas (1) e pode ser estendido para um sistema de transmissão de imagens com multiresoluções.

Neste trabalho buscaremos a simplicidade de implementação de qualquer sistema envolvido, por tal nos limitaremos ao caso unidimensional, embora possa ser facilmente estendido ao caso bidimensional. Em todo o trabalho, o sinal unidimensional compreendido como o sinal de vídeo é obtido pelo rastreamento da imagem linha a linha em dois campos entrelaçados, como em TV convencional. Este sinal (representativo do quadro bidimensional) é monocromático e está convenientemente digitalizado com 8 bits por amostra a uma taxa de amostragem (f_s) em torno de 10 M samples/segundo, limitando a faixa do sinal a 5MHz.

A seguir definiremos o conceito de Codificação em Pirâmide, aplicaremos DPCM à tal codificação e descreveremos o método como um todo.

1) INTRODUÇÃO

A motivação básica de qualquer estudo de codificação de imagens digitalizadas é a intenção de comprimir os dados para economia de memória ou de taxa de transmissão. Este trabalho tenta modificar o sistema convencional DPCM (1) (Differential Pulse Code Modulation) para comprimir dados e lidar com a falta de tempo para processamento entre duas amostras de um sinal de vídeo, que em tempo real gira em torno de 100ns. Além disto, o sistema proposto permite uma

2) CODIFICAÇÃO EM PIRÂMIDE

A codificação em pirâmide segue dos trabalhos de Burt & Adelson (2) e de Tanimoto (3), mas pode ser bem explicado em (4). Em todos estes trabalhos, a principal finalidade era a transmissão progressiva da imagem a ser codificada.

Seja a imagem original um array de $N \times N$ pixels, sendo $N=2^i$, i inteiro. Em nossa aplicação voltaremos ao caso

unidimensional e utilizaremos a sequência $x(n)$, obtida pela amostragem ao longo das linhas da imagem. Se filtramos este sinal com um Passa-Baixas digital com corte em $\pi f_s/2$ o sinal resultante da filtragem terá metade da banda de frequência original e podemos reamostrá-lo com $f_s/2$. Seja $x_1(n)$ esta versão passa-baixas e $x'_1(n)$ a aproximação de $x(n)$ por inserção de zeros e interpolação. Claramente $x'_1(n)$ é uma versão "borrada" da imagem original. Considere a imagem representada por $x_1(n)$ e, agora, novamente, filtre e reamostré-a obtendo $x_2(n)$. Novamente interpole $x_2(n)$ para encontrar $x'_2(n)$. Definimos, então,

$$\begin{aligned}x_0(n) &= x(n) \\ L_0(n) &= x_0(n) - x'_1(n) \\ L_1(n) &= x_1(n) - x'_2(n)\end{aligned}$$

Sendo $L_0(n)$ diferença entre o sinal original e a sua aproximação, a partir do sinal subamostrado. Em geral

$$L_k(n) = x_k(n) - x'_{k+1}(n) \quad k=0,1,\dots,M-1 \quad [E1]$$

é a diferença entre a imagem com $1/2^k$ da resolução original e sua aproximação. Se a imagem original contiver $2^M = N \times N$ pixels, então [E1] e os processos de prefiltragem e interpolação podem ser repetidos M vezes a partir de $x(n)$ obtendo uma imagem $x_M(n)$ que possui apenas uma amostra e corresponde grosseiramente à média de intensidade da imagem.

Para recuperar a imagem original a partir de $x_M(n)$ e dos $L_k(n)$, basta-nos proceder inversamente.

$$L_M(n) = x_M(n)$$

$$x_{M-1}(n) = x_M(n) + L_{M-1}(n) \quad [E2a]$$

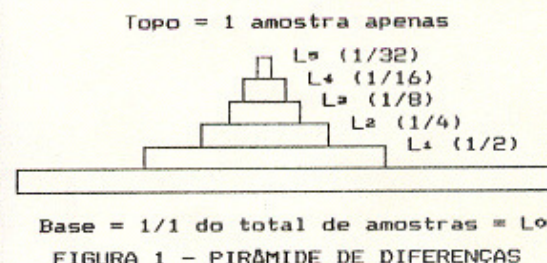
$$x_{M-2}(n) = x'_{M-1}(n) + L_{M-2}(n) \quad [E2b]$$

$$x_{k-1}(n) = x'_k(n) + L_{k-1}(n) \quad [E2c]$$

$$x(n) = x_0(n) = x'_1(n) + L_0(n) \quad [E2d]$$

Constrói-se a pirâmide colocando os níveis $L_k(n)$ amontoados, como na figura 1. Como os níveis são compostos por diferenças, chamamos a pirâmide como Pirâmide de Diferenças.

Neste esquema, pode-se dizer que L_M é codificado via PCM (1) (Pulse Code Modulation) e os restantes através de DPCM, só que estes são preditos pela interpolação do nível superior. Podemos claramente notar que envolveremos muito mais diferenças que pixels originais, aproximadamente o dobro ou $(1-1/2^{M-1}) \times N \times N$ diferenças redundantes quando comparadas ao DPCM comum. Por isto são propostos esquemas de Pirâmides Reduzidas, nas quais o número de amostras não se altera (4). A seguir iremos incorporar o DPCM extrapolativo convencional à codificação em pirâmide.



3) PIRÂMIDE REDUZIDA

Iremos, em diante, utilizar apenas os três níveis inferiores da Pirâmide, L_0 , L_1 e L_2 . Utilizando as equações E1 e E2 podemos construir a Pirâmide e recuperar o sinal original, mas como codificar L_2 , o topo da pirâmide? Lembre-se que neste caso $L_2(n) = x_2(n)$. Este trabalho tem como idéia principal codificar L_2 com um DPCM convencional extrapolativo. L_1 e L_0 são por definição diferenças e podemos considerá-los DPCM com predição interpolativa, formando um sistema híbrido extra/interpolativo. Este processo lembra, mas sendo estruturalmente diferente, o DPCM Interpolativo de Sethia & Anderson (5) e de Handa & Tanaka (6). Vamos supor, por enquanto, o DPCM de L_2 como uma caixa preta perfeita e iremos para as equações.

$$L_2(n) = x_2(n) = x_{1/4}(n) \quad [E3a]$$

$$x_1(n) = x_2'(n) + L_1(n) \quad [E3b]$$

$$x(n) = x_1'(n) + L_0(n) \quad [E3c]$$

$$x_2'(n) = \text{INTERP} [\text{DPCM} x_{1/4}(n)] \quad [E4]$$

A equação E4 também vale para $x_1(n)$. Porém anteriormente supomos que o sinal era prefiltrado antes de sofrer subamostragem, mas podemos considerar o sinal sem prefiltragem permitindo aliasing do espectro. Neste caso, ainda assim podemos recuperar completamente a imagem original, formando uma Pirâmide Reduzida. Para isto note que $x_1(n)$ contém metade das amostras originais de $x(n)$ e $x_2(n)$ metade das de $x_1(n)$ e $1/4$ do original. Portanto nas aproximações, como em E4, não necessitamos interpolar todas as amostras, mas metade, pois a outra metade já é perfeitamente conhecida e nenhum refinamento, dado pelos L_i , precisa ser somado a estas amostras. Definiremos os sinais $x_i'(n)$ como divididos em $x_i(n)$ e $x_i''(n)$, sendo que cada destes sinais tem a mesma frequência de amostragem, mas com metade das amostras nulas, nas posições das amostras do outro. Logo

$$x_1'(n) = x_1(n) + x_1''(n) \quad [E5a]$$

$$x_2'(n) = x_2(n) + x_2''(n) \quad [E5b]$$

$$L_0(n) = \text{DEC}_2(x(n) - x_1''(n)) \quad [E5c]$$

$$L_1(n) = \text{DEC}_2(x_1(n) - x_2''(n)) \quad [E5d]$$

Sendo $\text{DEC}_2(y(n))$ o sinal $y(n)$ dizimado, ou subamostrado por 2. Para recuperar o sinal definamos $\text{INT}_2(y(n))$ como a superamostragem e interpolação de $y(n)$ por um fator 2. Temos L_0 , L_1 e $x_2(n)$. Após codificar perfeitamente $x_2(n)$ via DPCM e L_0 e L_1 em uma codificação PCM auxiliar, temos :

i) Ache $\text{INT}_2(x_2(n))$ e some L_1 às amostras resultantes não coincidentes com as de $x_2(n)$, obtendo $x_1(n)$.

ii) Ache $\text{INT}_2(x_1(n))$ e some L_0 às amostras resultantes não coincidentes com as de $x_1(n)$, obtendo $x(n)$.

Podemos exemplificar o processo na Figura 2. É apresentado $x(n)$ e suas versões dizimadas $x_1(n)$ e $x_2(n)$. Interpolando x_2 , tiramos as diferenças entre x_1 e as amostras interpoladas. Estas diferenças formam L_1 que estão apenas representadas pelos seus sinais, sem magnitude. Repete-se o processo para obter L_0 a partir de x_1 e $x(n)$.

A pirâmide formada é reduzida e a ilustramos na figura 3. Indicamos em cada nível o número de amostras envolvidas e podemos notar que estes números se somam à unidade.

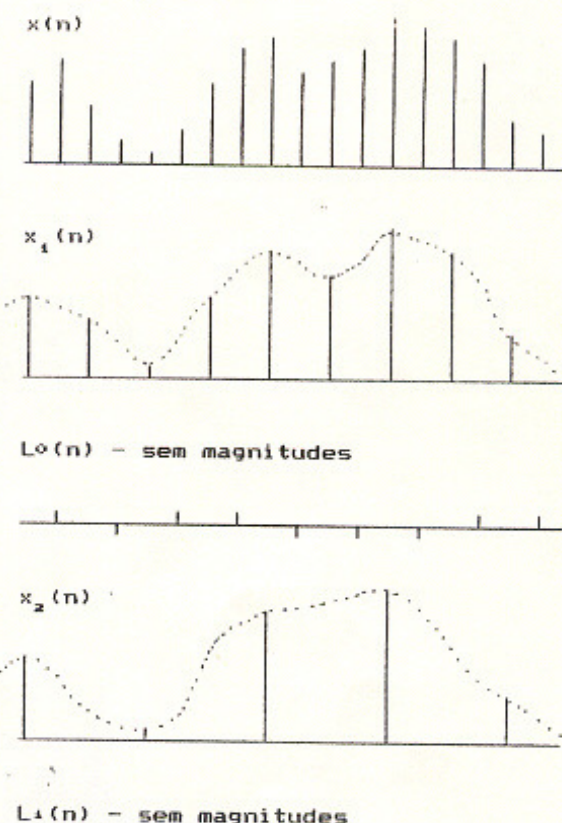


FIGURA 2 - Exemplo.

4) DPCM EM PIRÂMIDE

A partir dos resultados da seção precedente, formar o DPCM em Pirâmide é apenas uma questão de especificar a codificação DPCM do nível L_2 , os quantizadores de L_0 e L_1 , os interpoladores e desenhar o diagrama final do sistema. Na Figura 4 identificamos os blocos de um DPCM (1)

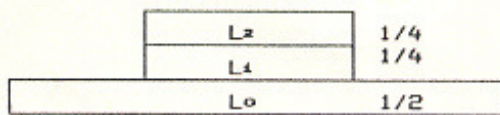


FIGURA 3 - Pirâmide de Diferenças Reduzida.

com entrada $x(n)$ codificando os erros $e(n)$ e recuperando, via decodificação local, $\hat{x}(n)$.

Aqui utilizaremos um esquema de predição adaptativo bidimensional (7) e, por simplificação, quantizadores fixos. As amostras utilizadas são as de marcas a, b e c como na Figura 5.

Por fim, a Figura 6 apresenta o diagrama do DPCM em pirâmide, sendo que $h(n)$ são os filtros não causais que perfazem a interpolação e os dispositivos DEMUX e MUX são responsáveis pela dizimação e recomposição do sinal como na Figura 2. O DEMUX demultiplexa o sinal em dois com metade da amostragem inicial formando os sinais x_1 e, digamos, x_{1a} a partir de x e x_2 e x_{2a} a partir de x_1 . O MUX reintegra o sinal no receptor. É previsto que um multiplexador de linha irá atuar nos 3 níveis, mas para

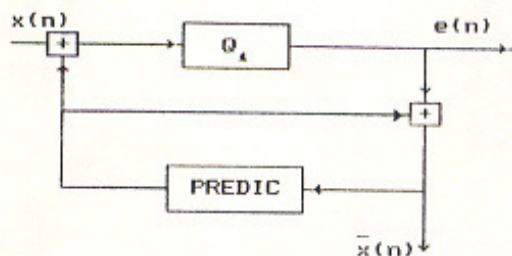
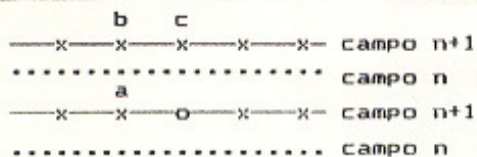


FIGURA 4 - Diagrama DPCM



o - amostra a ser predita

FIGURA 5 - Amostras envolvidas na predição

as necessidades deste trabalho, supomos que os 3 níveis são enviados paralelamente no canal.

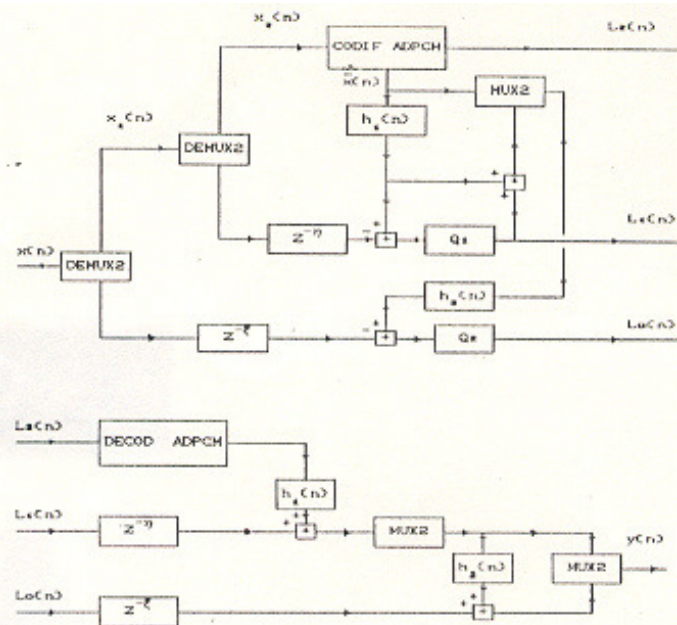


Figura 6 - Codificador e Decodificador do DPCM em Pirâmide.

5) INTERPOLADORES E PREFILTROS

Como já comentado, não é prevista nenhuma forma de pré-filtragem para limitar a faixa do sinal a ser dizimado, pois L_2 seria um sinal deturpado pela filtragem. Contudo as interpolações para estimação das amostras da figura original seriam bem mais precisas. Se a fidelidade da codificação tolerar esta pequena degradação, tal estratégia deve ser aplicada.

O interpolador aqui utilizado é um filtro digital FIR bastante simples e baseado na convolução cúbica de Keys (8). Para formar este núcleo de convolução para interpolação, foram atendidas condições de continuidade do sinal interpolado e de suas derivadas, como em Splines, mas atendendo uma forma polifásica adequada (9) para que amostras originais não sejam alteradas pela filtragem. Desta forma, a resposta impulsiva discreta para um filtro desta natureza para dobrar a amostragem do sinal (interp. com fator 2) é dada por (10) :

$$\begin{aligned} h(0) &= 1 \quad h(1) = 9/16 \quad h(2) = 0 \quad h(3) = -1/16 \\ h(n) &= 0 \text{ para } |n| > 3 \end{aligned} \quad [E6]$$

Se necessário prefiltrar o sinal, devemos utilizar um filtro ótimo, dadas as estatísticas do sinal de imagem (o qual tem como bom modelo um processo Gauss-Markov com $\rho=0.95$) e a densidade espectral de potência do Pós-filtro interpolador. Isto pode facilmente ser feito como em (11).

6) VANTAGENS x DESVANTAGENS

As principais vantagens desta codificação sobre o DPCM são:

a) Maior compressão de dados devido ao fato de podermos quantizar diferenciadamente os três níveis, otimizando o processo. Isto porque uma predição interpolativa é mais eficiente que uma extrapolativa, para mesmos graus de complexidade. Além disto os

três níveis carregam informação de bandas de frequência distintas, permitindo codificação mais grosseira dos níveis inferiores, sem interferir na qualidade subjetiva da imagem resultante.

b) Aumento do intervalo de processamento entre amostras (1+4 em L2 e L1 e 1+2 em Lo). Este intervalo é fator crítico em CODECs de vídeo amostrado a 10MHz.

c) Possibilidade de transmissão progressiva/condicional de L1 e Lo.

E as principais desvantagens são: necessidade de adição de hardware ao DPCM convencional (embora mais lento) e necessidade de sincronização das transmissões dos 3 níveis. Estes são os preços pagos pela melhoria de desempenho.

7) RESULTADOS

Foram realizados levantamentos e simulações em quatro imagens. Nestes primeiros passos foram medidas entropias de ordem zero dos erros de predição para estimar a taxa de bits/pixel necessárias a uma reconstrução sem erros. Em primeiro lugar, medimos as entropias das seguintes quantidades:

$$H_0 = H[x(n)]$$

$$H_1 = H[x(n) - x(n-1)]$$

$$H_4 = H[x(n) - x(n-4)]$$

$$H_{L1} = H[L_1]$$

$$H_{Lo} = H[Lo]$$

Com isto queríamos verificar o quanto o ganho de entropia dos erros de interpolação compensaria o maior entropia do erro de predição de L2 em relação ao do DPCM aplicado à imagem original. Com isto a predição da amostra atual seria a amostra anterior, $H_4 = H[L_2]$ e G seria o ganho em entropia do DPCM em pirâmide, como na TABELA I.

$$G = \frac{H_4 + H_{L1} + 2H_{Lo}}{4} - H_1$$

TABELA I - SEM PREFILTRAGEM

IMAGEM	H ₀	H ₁	H ₄	H _{Lo}	H _{L1}	G
1	7.33	4.61	5.97	3.49	4.79	0.18
2	6.72	3.84	5.05	2.93	3.99	0.12
3	6.97	3.40	4.93	2.16	2.81	0.38
4	6.86	3.56	4.90	2.59	3.56	0.15

TABELA II - COM PREFILTRAGENS

IMAGEM	H ₄	H _{Lo}	H _{L1}	G
1	5.84	2.91	4.45	0.40
2	4.94	2.28	2.22	0.73
3	4.86	1.50	2.42	0.67
4	4.78	1.98	3.19	0.42

Pelas tabelas podemos ver que o ganho médio esteve em torno de 0.2 b/pel sem prefiltragem e de 0.5 com prefiltragem adequada. De forma similar foram refeitas as tabelas para um preditor como o da Figura 5, resultando em um ganho em entropia de 0.4 b/pel.

O tipo de levantamento realizado serve apenas como referência, pois alguns destes erros podem ser codificados de forma mais grosseira que os outros. Em uma segunda fase são projetados os quantizadores e codificações da imagem são feitas comparando DPCM x DPCM em Pirâmide. Na

Figura 7 temos a comparação do detalhe do cabelo da imagem "Zelda" para 4 codificações diferentes. Em todas, o DPCM utilizado foi o mesmo: adaptativo LMS (APB) bidimensional com quantizador fixo de 33 níveis. Os segmentos de imagem são:

- Superior esquerdo: Imagem original.
- Superior direito: DPCM em pirâmide, com Q_1 de 7 níveis e Q_2 de 11 níveis. SNR = 40.3 dB e taxa média de 1.45 b/pel.
- Inferior esquerdo: DPCM em pirâmide, com Q_1 de 9 níveis e Q_2 de 11 níveis. SNR = 43.3 dB e taxa média de 1.85 b/pel.
- Inferior direito: DPCM com SNR = 48 dB e taxa média de 2.4 b/pel.

9) CONCLUSÃO

Aqui procuramos encarar o método proposto como uma alternativa ao DPCM e por isso sempre comparamos o desempenho dos dois sistemas. Nestas condições, o esquema em pirâmide se mostra uma grande alternativa combinando as codificações em Sub-Bandas, Pirâmide e DPCM, atingindo uma performance superior. Deve ficar bem claro que a performance aqui atingida não é a melhor permissível (longe disto), mas apenas deve ser encarada com relação à performance do ADPCM. Melhorias podem ser realizadas através de transmissão condicional dos níveis inferiores, interpolação adaptativa etc., mas a aplicação do DPCM à codificação em pirâmide sem dúvida revela-se como método atrativo dentro do Processamento e Codificação de Imagens.

9) REFERÊNCIAS

(1) N. S. Jayant, P. Noll, *Digital Coding of Waveforms*, Englewood Cliffs, NJ, Prentice-Hall, 1984.
 (2) P. J. Burt, E. H. Adelson, *The Laplacian Pyramid as Compact Image Code*, IEEE Trans. Commun., COM-31, pp 532-540, April 1983.



FIGURA 7

(3) S. L. Tanimoto, *Image Transmission with Gross Information First*, Comp. Graph and Image Proc. 9, pp 72-76, 1979.
 (4) L. Wang, M. Goldberg, *Reduced-difference Pyramid*, Optical Engineering, Vol 28 #7, July 1989.
 (5) M. Sethia, J. B. Anderson, *Interpolative DPCM*, Trans. Commun., COM 32, pp 710-736, June 1984.
 (6) S. Handa, H. Tanaka, *Two-D Interpolative DPCM*, Electr. and Commun. in Japan, Part 1, vol 71, #5, 1988.
 (7) M. Hadhoud, D. Thomas, *Two-D Adaptive LMS Algorithm*, IEEE Trans. on Circuits and Systems, CAS-35, May 1988.
 (8) R. G. Keys, *Cubic Convolution Interpolation For Digital Image Processing*, IEEE Trans. on ASSP, ASSP-29, pp 746-749, June 1983.
 (9) R. E. Crochiere, L. R. Rabiner, *Multirate Digital Signal Processing*, Englewood Cliffs, NJ, Prentice-Hall, 1983.
 (10) R. L. Queiroz, J. B. Yabu-uti, *Técnicas de Projeto de Filtros FIR para Dízimação e Interpolação de Sinais Discretos*, RT-178, Contrato 208/89, UNICAMP/TELEBRÁS.
 (11) H. S. Malvar, D. H. Stealin, *Optimal Pre and Postfilters for Decimation and Interpolation of Random Signals*, IEEE Trans. on Communications, pp 67-73, January 1988.