

## DIZIMAÇÃO E INTERPOLAÇÃO PROGRESSIVAS PARA CODIFICAÇÃO EM PIRÂMIDE

Ricardo L. de Queiroz e João B.T. Yabu-Uti

Depto. de Comunicações - Faculdade de Engenharia Elétrica

Universidade Estadual de Campinas - UNICAMP

C.P. 6101 - Campinas-SP - 13081

**RESUMO** - A codificação em pirâmide tem sido estudada ultimamente visando transmissão/codificação progressiva de uma imagem. Aqui propomos novo esquema de dizimação, adequado à interpolação através de filtros FIR bidimensionais. Estruturas polifásicas destes são apresentadas, bem como interpoladores bidimensionais baseados na Convolução Cúbica.

### 1. INTRODUÇÃO

A codificação em pirâmide consiste na representação da imagem original com diferentes resoluções. Com isto, é possível perfazer uma transmissão progressiva e, simultaneamente, reduzir a taxa de bits/pixel para codificação. Tal procedimento pode ser de grande utilidade para consulta a bancos de imagens via canais de faixa estreita ou sistemas de processamento lento, porém nada impede de aplicar esta codificação para simplesmente reduzir o montante de dados requeridos e melhor codificar a imagem.

Imaginaremos a imagem como uma matriz de  $N \times N$  pixels de forma que  $N$  seja da ordem de 512. Esta matriz, denotaremos por  $X(n_1, n_2)$  ou  $X(\underline{n})$ ,  $\underline{n} = (n_1, n_2)$ . Por todo o trabalho, utilizaremos notações familiares a [1] para representação de sinais e sistemas discretos multidimensionais.

### 2. PIRÂMIDE REDUZIDA

Se prefiltrarmos  $X(\underline{n})$  adequadamente, podemos representar esta versão filtrada, fielmente, com 1/4 das amostras originais, obtendo

$$X_{0F}(\underline{n}) = X(\underline{n}) \text{ filtrado} \quad (E1)$$

$$X_1(\underline{n}) = X_{0F}(2\underline{n}) \quad (E2)$$

$$L_0(\underline{n}) = X_0(\underline{n}) - X_{0F}(\underline{n}) \quad (E3)$$

$L_0(\underline{n})$  seria o primeiro nível da pirâmide e, repetindo o processo com  $X_1(\underline{n})$ , obteríamos novo nível e daí por diante. De forma geral

$$X_{i+1}(\underline{n}) = X_{iF}(2\underline{n}) \quad (E4)$$

$$L_i(\underline{n}) = X_i(\underline{n}) - X_{iF}(\underline{n}) \quad i=0,1,\dots,M-1 \quad (E5)$$

Este é o princípio das pirâmides Laplacianas ou Gaussianas [2] e neste referência o princípio pode ser melhor compreendido. Em [3], esforços são realizados de forma a definir pirâmides reduzidas, nas quais o número de amostras (nodos) representados em todos os níveis se soma a  $N^2$ , ou seja, o número de pixels originais. No primeiro caso, se  $M \gg 1$ , então o número de nodos é cerca de 1/3 maior que o original, ou

$$N_{\text{NODOS}} = \sum_{k=0}^{M-1} N^2/4^k = \left( N^2 + \frac{N^2}{3} \right) (1 - 1/4^{M-1}) \approx N^2 + N^2/3 \quad (E6)$$

Para reduzir esta total a  $N^2$ , basta não prefiltrar o sinal, sub-amostrá-lo e interpolá-lo, tomando  $L_k(\underline{n})$  como

$$L_k(\underline{n}) = X_k(\underline{n}) - X_{(k+1)}^{\text{INT}}(\underline{n}) \quad (E7)$$

Porém, 1/4 das amostras de  $L_k(\underline{n})$  devem ser descartadas pois, como não houve prefiltragem e a interpolação não deturpou as amostras originais, tais amostras serão sempre nulas. Daí, o número de nodos total se reduzirá à  $N^2$ . As referências [2] e [3] melhor explicam codificação em pirâmide e [4] e [5] tratam dos interpoladores para sinais unidimensionais, respeitando a restrição acima citada. A extensão bidimensional destes será feita posteriormente.

### 3. ESTRUTURA POLIFÁSICA BIDIMENSIONAL DO INTERPOLADOR

Na seção anterior, foi suposto que a sub-amostragem ou dizimação era retangular, como na Figura 1, e, não havendo pre-filtragem,  $X_{k+1}(\underline{n}) = X_k(\underline{Vn})$  e  $X_{k+1}(\underline{n}) = X(\underline{V}^k \underline{n})$ . Sendo  $\underline{V}$  uma matriz de sub-amostragem definida como

$$\underline{V} = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \quad (\text{no caso } Q_1=Q_2=2) \quad (\text{E8})$$

Um filtro FIR interpolador unidimensional pode ser dividido em seus sub-filtros polifásicos definidos por [4] e [5]

$$\begin{aligned} h(n) &= \text{interpolador FIR fase nula} \\ h_i(n) &= h(nQ + i) \quad i=0,1,\dots,Q-1 \end{aligned} \quad (\text{E9})$$

Em cada período de  $Q$  amostras, uma é a original,  $Q-1$  devem ser interpoladas e existem  $Q$  sub-filtros polifásicos, um para cada amostra a ser interpolada. Como uma é original e deve ser mantida,  $h_0(n) = \delta(n)$ , restando-nos projetar os  $Q-1$  restantes.

Para extensão ao caso bidimensional, seja (E8) e:

$$\begin{aligned} \underline{Q}^T &= (Q_1, Q_2) \\ \underline{0}^T &= (0, 0) \end{aligned}$$

Além disto, seja  $S$  o operador de superposição (para lidar com as diferentes taxas de amostragem entre  $h(n)$  e os sub-filtros) suplantando a somatória. Teremos:

$$h_{ij}(n_1, n_2) = h(Q_1 n_1 + i, Q_2 n_2 + j) \quad (\text{E10})$$

$$h(n_1, n_2) = \sum_{i=0}^{Q_1-1} \sum_{j=0}^{Q_2-1} h(Q_1 n_1 + i, Q_2 n_2 + j) \quad (\text{E11})$$

Em notação vetorial, sendo  $\underline{k}^T = (i, j)$ ,  $0 \leq i < Q_1$ ;  $0 \leq j < Q_2$ , temos

$$h(\underline{n}) = \sum_{\underline{k}=0}^Q h_{\underline{k}}(\underline{n}) = \sum_{\underline{k}=0}^Q h(\underline{V}\underline{n} + \underline{k}) \quad (\text{E12})$$

E, novamente,

$$h_0(\underline{n}) = \delta(\underline{n}) \quad (\text{E13})$$

O nosso caso de interesse, se restringe a  $Q_1 = Q_2 = 2$  e  $\underline{V} = 2\underline{I}$ , sendo  $\underline{I}$  a matriz identidade  $2 \times 2$ . O número de sub-filtros polifásicos, bem como o tamanho do período de sub-amostragem, é dado por  $D = |\det(\underline{V})| = Q_1 Q_2 = 4$ . Como conhecemos um destes por (E13), resta-nos 3 filtros a serem calculados. A Figura 2 ilustra um destes períodos e denota três amostras no plano por  $\underline{a}$ ,  $\underline{b}$  e  $\underline{c}$ , sendo que cada uma corresponde a um sub-filtro polifásico, sendo interpolada por este. Suponha que o interpolador seja  $7 \times 7$ , então a Figura 3 mostra as amostras de  $h(\underline{n})$  correspondentes a cada sub-filtro. Nesta figura, 0's e 1's são valores reais da amostra, já computando  $h_0(\underline{n})$ . A análise desta figura nos leva a concluir que, para cada amostra, devemos realizar uma filtragem com 12 ou 16 coeficientes de  $h(\underline{n})$ , ao invés dos 49 iniciais.

Para simplificar a notação, indicaremos os coeficientes de um filtro unidimensional por um pseudo-vetor e por uma pseudo-matriz o filtro 2D. Como  $h_0(n) = \delta(n)$ , então:

$$h_1(n) = [a_2 \ 0 \ a_1 \ 1 \ a_1 \ 0 \ a_2] \quad (\text{E14})$$

Para projetar o interpolador 2D, podemos utilizar Resposta Impulsiva Separável ou a Projeto pela Transformação de McClellan [4], como em [6]. Porém, o último, se aplicado a resposta impulsiva de (E14), formará um filtro FIR 2D que não atenderá (E13). Se tentarmos encontrar uma transformação com isopotenciais menos arredondadas, iremos nos aproximar de um filtro com Resposta Impulsiva Separável. Por isto simplificamos, e fizemos

$$h(n_1, n_2) = h_1(n_1)h_1(n_2) \quad (\text{E15})$$

$$h(n_1, n_2) = \begin{bmatrix} a^2 & 0 & a_1 a_2 & a_2 & a_1 a_2 & 0 & a_2^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_1 a_2 & 0 & a_1^2 & a_1 & a_1^2 & 0 & a_1 a_2 \\ a_2 & 0 & a_1 & 1 & a_1 & 0 & a_2 \\ a_1 a_2 & 0 & a_1^2 & a_1 & a_1^2 & 0 & a_1 a_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_2^2 & 0 & a_1 a_2 & a_2 & a_1 a_2 & 0 & a_2^2 \end{bmatrix} \quad (E16)$$

$$\underline{h}_a(n_1, n_2) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ a_2 & a_1 & a_1 & a_2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \underline{h}_c(n_1, n_2)^T \quad (E17)$$

$$\underline{h}_b(n_1, n_2) = \begin{bmatrix} a_2^2 & a_1 a_2 & a_1 a_2 & a_2^2 \\ a_1 a_2 & a_1^2 & a_1^2 & a_1 a_2 \\ a_1 a_2 & a_1^2 & a_1^2 & a_1 a_2 \\ a_2^2 & a_1 a_2 & a_1 a_2 & a_2^2 \end{bmatrix} \quad (E18)$$

Note que  $\underline{a}$  e  $\underline{c}$  seriam interpolados unidimensionalmente, mas iremos modificar a sub-amostragem de forma a sempre só utilizar  $\underline{h}_b(\underline{n})$ .

#### 4. DIZIM AÇÃO E INTERPOLAÇÃO PROGRESSIVAS

##### 4.1. Dizim ação

Nas seções anteriores consideramos, apenas, sub-amostragem retangular. Se dizirmos progressiva e sucessivamente com  $\underline{Q} = (2, 2)$  teríamos as seguintes matrizes de dizim ação (sub-amos

tagem) :

$$\begin{aligned} \underline{v}_0 &= \underline{I} & X(\underline{v}_0 \underline{n}) &= \text{imagem original} \\ \underline{v}_1 &= 2 \underline{I} \\ \underline{v}_k &= 2^k \underline{I} \end{aligned} \quad (\text{E19})$$

Em  $X(\underline{v}_k \underline{n})$ , a densidade de amostras coletados por originais seria

$$d_k = \frac{1}{D} = \frac{1}{2^{2k}} \quad (\text{E20})$$

Se, ao invés, fizéssemos

$$\underline{v}_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}; \quad \underline{v}_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}; \quad \underline{v}_3 = \begin{bmatrix} 2 & 2 \\ 2 & -2 \end{bmatrix}; \quad \underline{v}_4 = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}; \quad \dots$$

$$\underline{v}_{2k-1} = \begin{bmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & -2^{k-1} \end{bmatrix} \quad \underline{v}_{2k} = \begin{bmatrix} 2^k & 0 \\ 0 & 2^k \end{bmatrix} \quad (\text{E21})$$

Notariamos que

$$d_{2k-1} = \frac{1}{2} \cdot \frac{1}{2^{2(k-1)}}$$

$$d_{2k} = 1/2^{2k} \quad \text{ou}$$

$$d_k = 2^{-k} \quad (\text{E22})$$

Isto implica num maior número de passos para a formação da pirâmide, porém obteríamos duas vantagens básicas para redução da entropia da codificação. Para analisar a primeira vantagem, voltemos à figura 2. Note que com as dizimações sucessivas de (E21), reconstruiríamos o valor de  $\underline{b}$ , para em seguida interpolar  $\underline{a}$  e  $\underline{c}$ .

Seja  $\underline{v} = \underline{v}_1$  e  $e_{\underline{a}}$  o erro de interpolação de  $\underline{a}$ . Daí

$$H(e_{\underline{a}}|\underline{v}=\underline{v}_2) \geq H(e_{\underline{a}}|\underline{v}=\underline{v}_1) = H(e_{\underline{a}}|\underline{v}=\underline{v}_2; \underline{b}) \quad (E23)$$

com  $H(e)$  sendo a entropia do erro de interpolação.

$V_1$  configuraria dizimação hexagonal que permitiria uma interpolação mais eficiente para sinais circularmente limitados. O sinal, na realidade, não é limitado, mas a mesma afirmação vale para banda base circular, ou aproximadamente. Isto é tolerável, admitindo um modelo isotrópico para imagens.

Na Figura 4, pode-se ver a dizimação hexagonal para re-composição de  $\underline{a}$ , mas todos os resultados acima podem ser estendidos para  $\underline{c}$ . Além disso o raciocínio aqui aplicado também pode ser estendido para qualquer  $X(\underline{v}_{2k-1} \underline{n})$  e  $X(\underline{v}_{2k} \underline{n})$ .

#### 4.2. Interpolação

O segundo fator de redução de entropia, se baseia no fato que para aproximar  $X(\underline{v}_{2k-1} \underline{n})$  a partir de  $X(\underline{v}_{2k} \underline{n})$  só necessitamos utilizar  $h_{\underline{b}}(\underline{n})$ , que é um filtro polifásico que pondera amostras bidimensionalmente e é mais eficiente. Por outro lado, para aproximar  $X(\underline{v}_{2k} \underline{n})$  por  $X(\underline{v}_{2k+1} \underline{n})$ , também podemos utilizar apenas  $h_{\underline{b}}(\underline{n})$ . Voltemos a figura 4 para interpolar  $X(\underline{n})$  via  $X(\underline{v}_1 \underline{n})$ , e, no caso desejamos interpolar  $\underline{a}$ . Adotando novos eixos de coordenadas modificaríamos os índices para obter  $X(n'_1, n'_2) = X(\underline{n}')$ . A partir daí é fácil ver que novamente poderíamos aplicar um  $h_{\underline{b}}(\underline{n}')$ , desta vez ao longo das direções inclinadas.

Portanto,  $h_{\underline{b}}(\underline{n})$  seria o único interpolador utilizado na prática, já que  $h_0$ ,  $h_a$  e  $h_c$  seriam dispensados. Com isto teríamos interpolações eficientes, reduzindo a entropia dos erros que compõem a pirâmide.

#### 5. INTERPOLADORES PRÁTICOS

Como vimos, só necessitamos projetar  $h_{\underline{b}}(\underline{n})$ , que deve levar em conta os recursos computacionais disponíveis. Tomamos como exemplo um filtro  $h(\underline{n})$  de ordem  $7 \times 7$ , mas, para  $\underline{Q} = (2, 2)$  e in-

terpolando uma das 3 amostras restantes, podemos ter ordens  $2M+1 \times 2M+1$  ( $M$  ímpar) à escolha; gerados a partir de

$$h_1(n) = \left[ a_{\frac{M+1}{2}}, 0, \dots, a_2, 0, a_1, 1, a_1, 0, a_2, \dots, 0, a_{\frac{M+1}{2}} \right]$$

Tal interpolador unidimensional pode ter aproximação de ordem  $\frac{M+1}{2}$ . Em muitos casos é necessário reduzir  $M$  ou aplicar outro esquema. Para  $M=3$ , podemos utilizar uma extensão discreta do Núcleo de Convolução Cúbica de Keys [7] [5], neste

$$h_1(n) = [-1/16 \quad 0 \quad 9/16 \quad 1 \quad 9/16 \quad 0 \quad -1/16]$$

gerando

$$h_b(\underline{n}) = \begin{bmatrix} 1 & -9 & -9 & 1 \\ -9 & 81 & 81 & -9 \\ -9 & 81 & 81 & -9 \\ 1 & -9 & -9 & 1 \end{bmatrix} \times \frac{1}{256}$$

Se fizermos  $M=1$ , a melhor opção seria uma aproximação planar, ou

$$h_1(n) = [1/2 \quad 1 \quad 1/2] \quad h_b(\underline{n}) = \begin{bmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{bmatrix}$$

Neste último caso, a amostra a ser interpolada é, simplesmente, a média das amostras próximas a seu redor. Outra forma de interpolação atendendo (E13), mas fazendo interpolação apenas ao longo de duas direções ortogonais, pode levar a bons resultados mas utilizando 8 coeficientes ao invés de 16 (no caso  $M=3$ ). Utilizando  $M=3$  e a Convolução Cúbica, teríamos

$$h(n) = \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 & 0 & 0 & 0 \\ -1 & 0 & 9 & 16 & 9 & 0 & -1 \\ 0 & 0 & 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix} \times \frac{1}{16}$$

Este filtro deve ser utilizado nas sub-amostragens hexagonais nas quais o elemento central é desprezado, já que não interpolaremos amostras já existentes.

## 6. CONCLUSÃO

Propusemos, aqui, um método para implementação da codificação em pirâmide, baseando-nos na estrutura polifásica de um interpolador. Com isto formamos uma pirâmide reduzida ( $N^2$  nodos), com uma redução de entropia dada pela eficiência dos filtros FIR interpoladores em face da dizimação hexagonal. A principal idéia era a de conduzir o processo de forma a sempre recairmos neste procedimento de sub-amostragem e, por tal motivo, a compressão de dados na codificação será maior do que nos métodos convencionais que exploravam dizimações retangulares e interpolações menos trabalhadas. Resta-nos, por último, ressaltar que o método aqui descrito é apenas o primeiro passo na intenção de desenvolver uma codificação em pirâmide com alto desempenho.

## BIBLIOGRAFIA

- [1] D.E. Dudgeon, R.M. Merserean; "Multidimensional Digital Signal Processing", Englewood Cliffs, NJ, Prentice-Hall, 1983.
- [2] P.J. Burt, E.H. Adelson; "The Laplacian Pyramid as a Compact Image Code", IEEE Trans. on Comm., COM-31, pp. 532-540, April 1983.

- [3] L. Wang, M. Goldberg; "Reduced-Difference Pyramid", Optical Engineering, Vol. 28, 7, July 1989.
- [4] R.E. Crochiere, L.R. Rabiner; "Multirate Digital Signal Processing", Englewood Cliffs, NJ, Prentice-Hall, 1983.
- [5] R.L. Queiroz, J.B. Yabu-uti; "Técnicas de Projeto de Filtros FIR para Dizimação e Interpolação de Sinais Discretos, RT-178, Contrato 208/89, UNICAMP/TELEBRÁS, Setembro 1989.
- [6] R.J.P. de Figueiredo, T.C. Chen; "Image Decimation and Interpolation Techniques Based on Frequency Domain Analysis", IEEE Trans. on Comm., COM-32, April 1984.
- [7] R.G. Keys; "Cubic Convolution Interpolation for Digital Image Processing", IEEE Trans. on ASSP, ASSP-29, pp. 746-749, June 1983.

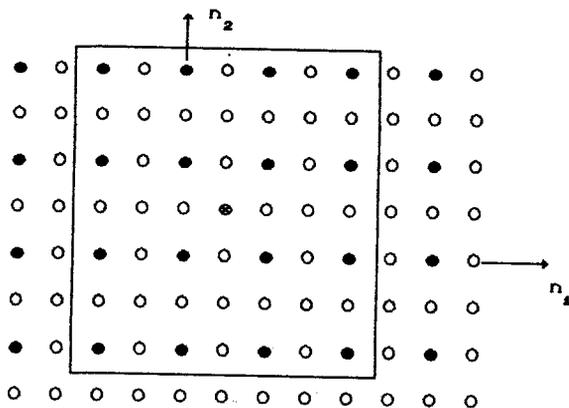


FIGURA 1 - Dizimação retangular, indicando região de suporte do filtro para interpolar amostra  $\bullet$ .

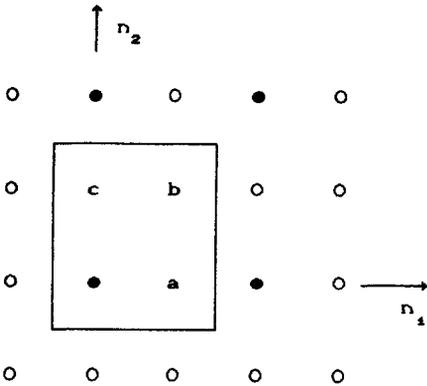


FIGURA 2 - Período de subamostragem, indicando amostras a serem interpoladas.

b	a	b	a	b	a	b
c	0	c	0	c	0	c
b	a	b	a	b	a	b
c	0	c	1	c	0	c
b	a	b	a	b	a	b
c	0	c	0	c	0	c
b	a	b	a	b	a	b

FIGURA 3 - Posição das amostras dos subfiltros de um interpolador 7x7 para dizimação retangular com  $Q=(2,2)$ .

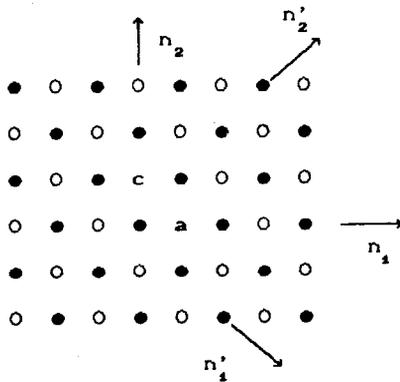


FIGURA 4 - Dizimação Hexagonal, destacando amostras faltantes da Dizimação Retangular e eixos auxiliares inclinados.