

Escolhendo um Sistema Sub-bandas para a Codificação de Imagens por um Critério Custo/Compactação

Ricardo L. de Queiroz

Dept. de Engenharia Elétrica
Universidade de Brasília
C.P.153041, 70910 Brasília, DF, BRASIL
Fone 061-3482328 Fax 061-2721053

Resumo

Bancos de filtros para codificação de imagens em sub-bandas são estudados. As tão comuns conexões de bancos de filtros de duas bandas são comparadas a bancos de M bandas, como as transformadas de blocos ou sobrepostas. É verificado, por um critério que pondera complexidade computacional e compactação espectral de energia, que estas transformadas provêm maior ganho aliado a uma menor complexidade. Tais resultados são comprovados em simulações.

1 Introdução

Sistemas ditos de sub-bandas (ou análise-síntese) são aqueles nos quais o sinal é decomposto em bandas do seu espectro, processado ou transmitido e, finalmente, recomposto a partir de suas bandas componentes. A decomposição é dita análise spectral e a recomposição é denominada como síntese. A maneira mais usual de se conseguir tal intento, se baseia na utilização de bancos de filtros (BFs) [1]. Cada filtro do banco se destina a selecionar uma banda do espectro do sinal. Na síntese, os filtros redelimitam as bandas que, somadas, reconstróem o sinal. Aqui nos ateremos ao estudo de filtros FIR, apenas. Sistemas ditos de dízimação máxima [1] são aqueles nos quais uma banda de faixa π/M é dízimada por M , de forma que o número total de amostras do sinal original se mantenha igual a soma das amostras de todas as sub-bandas. A Fig. 1 ilustra um BF de M bandas, de igual largura, com dízimação máxima, sendo $x(n)$ o sinal original, $\hat{x}(n)$ o reconstruído, $h_i(n)$ os filtros de análise e $g_i(n)$ os de síntese, para $i = 0, \dots, M-1$.

Os exemplos mais comuns de sistemas análise-síntese na codificação de sinais são os ditos por transformadas e por sub-bandas [2]. A transformada discreta de cossenos (DCT) [3], nada mais é que um sistema ortogonal de BF de M bandas, no qual o número de bandas é igual ao fator de dízimação e ao comprimento dos filtros. Desta forma, o sistema aparece como uma transformada em blocos.

2 Bancos de Duas Bandas

Por simplificação, um BF de duas bandas será denominado um BF2. Tais bancos possuem dois filtros: $h_0(n)$ e $h_1(n)$, para a análise, e $g_0(n)$ e $g_1(n)$ para a síntese. Um é o passa-baixa e o outro o passa-alta, ambos de comprimento L . Bancos QMF (Quadrature Mirror Filters) possuem a seguinte relação entre os filtros de análise [1, 4]

$$h_1(n) = (-1)^n h_0(L-1-n) \quad (1)$$

Bancos QMF para cancelamento de aliasing, com fase linear, foram projetados por Johnston e são amplamente utilizados. Tais filtros são listados em [1] para diversos valores de L . Utilizaremos dois destes filtros, denominados QMF-8A e QMF-16B. O grande problema destes bancos é o de não possuirem reconstrução perfeita [4], ou seja $\hat{x}(n) \neq x(n-d)$. Sendo $H_i(z)$ a transformada z de $h_i(n)$ e $P(z) = H_0(z)H_1(z)$, pode-se mostrar que soluções ortogonais, ou paraunitárias, para reconstrução perfeita são aquelas nas quais L é par e

$$g_0(n) = h_0(L-1-n) \quad (2)$$

$$g_1(n) = -(-1)^n h_0(n) \quad (3)$$

$$P(z) - P(-z) = az^{-L+1} \quad (4)$$

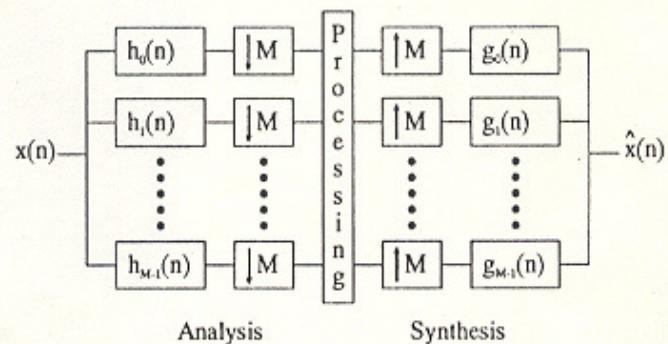


Fig. 1

Desta forma $h_i(n - 2l)$ (l inteiro), formariam bases ortogonais. Tais filtros são aqui denominados CQF (*Conjugate Quadrature Filters*) e podem ser implementados de maneira eficiente através de treliças [6]. Utilizaremos os filtros de [6], para $L = 8$ e $L = 16$, denominando-os CQF-8A e CQF-16B. A DCT, para $M = 2$, também foi testada como BF2 paraunitário. Nesta, para cada bloco de duas amostras, a soma é a saída passa-baixa e sua diferença a saída passa-alta. Por fim, cabe ressaltar que soluções bi-ortogonais existem [5]. Desta conceituação, em [5] são implementadas transformadas discretas *wavelet*.

3 Bancos de M Bandas

Transformadas de blocos são geralmente bancos de filtros com boa performance e baixa computação. Dentre as mais destacadas, as transformadas discretas de cosenos (DCT), Hadamard e Fourier foram amplamente utilizadas em codificação de imagens [2]. Além deste tipo, existem as transformadas sobrepostas (LTs), nas quais o comprimento dos filtros, L , é maior que M , de forma que as funções de base se extendem sobre blocos adjacentes. Sendo $L = 2KM$, adotaremos duas LTs para $K = 1$, denominadas MLT [7] e LOT [8]. Para $K = 2$, a ELT [9] será adotada. Além destas, a DCT, que se tornou um padrão mundial para codificação de imagens [3, 10], também será examinada. Tanto a DCT como a LOT são bastante adequadas para estes propósitos. Ambas possuem funções de base simétricas, ou, como um BF, filtros de fase linear. Mas a DCT possui filtros curtos, de forma que quando os coeficientes transformados são quantizados, efeitos-de-blocos aparecem na imagem. Estes efeitos são resultado de *aliasing* espúrio se apresentando como descontinuidades nas bordas de cada bloco que são visualmente incômodas [3]. As LTs atenuam em muito estes efeitos.

4 Partição Espectral

Um dos pontos mais importantes para a decisão do sistema de partição espectral a ser adotado, diz respeito à largura de faixa da banda de menor frequência. No caso de um sistema de M bandas, como na Fig. 1, esta largura seria π/M . Em [11] é sugerido, após várias simulações, que não haveriam muitas vantagens em se extender o valor de M além de 4 ou 8. De fato, em [3], o valor $M = 8$ é frequentemente citado como um dos mais adequados para imagens com resolução em torno de 512x512 amostras, bem como para as de 256x256 amostras. Aqui testaremos os valores $M = 4, 8, 16$, mas a maior atenção será dispensada ao caso $M = 8$.

Um BF de M bandas pode ser projetado pelo cascadingamento de N estágios de BF2s, se fizermos $M = 2^N$. Neste caso os bancos seriam conectados ao longo de uma árvore binária. Por outro lado, poderíamos apenas conectar BF2s

ao ramo passa-baixas de uma célula anterior de BF2, desta forma o espectro seria particionado em oitavas e não mais possuiria M bandas, mas $N + 1$. Além disto, as bandas não mais possuiriam larguras idênticas, mas a banda base possuiria largura π/M , com $M = 2^N$. A Fig. 2 ilustra a decomposição espectral bidimensional de uma imagem para o caso $M = 8$, tanto para partição em oitavas como uniforme.

Deve ser tomado muito cuidado com o BF2 que será cascadingado, em uma estrutura em árvore, pois geralmente tais bancos não são projetados obedecendo um critério de regularidade [5]. Esta regularidade é desejável por garantir que o longo cascadingamento das funções de base não resulte em um filtro de comportamento fractal. Adicionalmente, o cascadingamento de bons filtros, não regulares, pode gerar um filtro com pouca seletividade e lóbulos laterais de pouca atenuação [12].

5 Complexidade e Compactação

Para avaliar a complexidade de implementação de um BF, utilizaremos a unidade operações em ponto flutuante por amostra (FLOPS). Esta unidade engloba tanto adições como multiplicações, e se mostra conveniente, já que processadores de sinais modernos executam ambas operações com a mesma velocidade. Neste trabalho, a complexidade será medida apenas para a seção de análise, entretanto, como apenas nos referiremos a bancos paraunitários, a complexidade para a síntese seria idêntica. Além disto, nesta seção, nos referiremos a um processamento unidimensional. Com isto assumiremos que a imagem possa ser modelada como um sinal AR(1) separável [2], com correlação entre amostras $\rho = 0.95$, e que o processamento de separação das sub-bandas seja aplicado às linhas e, posteriormente, às colunas de uma imagem. Desta forma, cálculos para um sistema unidimensional seriam imediatamente extensíveis ao caso bidimensional.

Se C_2 for a complexidade em FLOPS para implementar um BF2, em uma estrutura em árvore, a complexidade total C para implementar um BF de N estágios, cuja banda base possuisse largura π/M , com $M = 2^N$, seria

$$C = NC_2 \quad (\text{M bandas}) \quad (5)$$

$$C = \frac{M-1}{M} C_2 \quad (\text{oitavas}) \quad (6)$$

A compactação de energia pode ser medida assumindo o modelo AR(1) ($\rho = 0.95$) citado, através do ganho de codificação [2], ou G_{TC} , que também pode ser visto como o ganho da transformada sobre o PCM. Sendo σ_i^2 a variância da i -ésima banda, tal ganho seria dado por

$$G_{TC} = 10 \log_{10} \frac{\frac{1}{M} \sum_{i=0}^{M-1} \sigma_i^2}{\left(\prod_{i=0}^{M-1} \sigma_i^2 \right)^{1/M}} \quad \text{dB} \quad (7)$$

A Tabela 1 lista valores de C e G_{TC} para vários BF2. A Tabela 2 reapresenta estes valores para bancos em árvore de M bandas e para $M = 4, 8, 16$. O mesmo se repete nas tabelas 3 e 4 para bancos com decomposição em oitavas e para bancos paralelos de M bandas, respectivamente.

Tabela 1 - Caso $M = 2$, complexidade, comprimento de filtro e ganho de codificação.

BF2	C	L	G_{TC}
DCT	3	2	5.05
QMF-8A	8	8	5.77
CQF-8A	9	8	5.84
QMF-16B	16	16	5.91
CQF-16B	17	16	5.92
IDEAL	∞	∞	5.94

Tabela 2 - Complexidade e ganho para BFs de M bandas, estruturados em árvores.

BF2	C	G_{TC}	C	G_{TC}	C	G_{TC}
	$M = 4$	$M = 8$	$M = 16$	$M = 4$	$M = 8$	$M = 16$
DCT	6	7.17	9	7.95	12	8.19
QMF-8A	16	8.31	24	9.32	32	9.67
CQF-8A	18	8.44	27	9.47	36	9.84
QMF-16B	32	8.52	48	9.56	64	9.92
CQF-16B	34	8.54	51	9.58	68	9.94

Tabela 3 - Complexidade e ganho para transformadas hierárquicas (em oitavas) de N estágios.

BF2	C	G_{TC}	C	G_{TC}	C	G_{TC}
	$M = 4$	$M = 8$	$M = 16$	$M = 4$	$M = 8$	$M = 16$
DCT	4.5	7.17	5.25	7.94	5.63	8.18
QMF-8A	12	8.29	14	9.23	15	9.54
CQF-8A	13.5	8.39	15.75	9.34	16.88	9.64
QMF-16B	24	8.49	28	9.45	30	9.76
CQF-16B	25.5	8.50	29.75	9.46	31.88	9.77

Tabela 4 - Complexidade e ganho para transformadas de M bandas.

BF	C	G_{TC}	C	G_{TC}	C	G_{TC}
	$M = 4$	$M = 8$	$M = 16$	$M = 4$	$M = 8$	$M = 16$
DCT	3.5	7.57	5.25	8.83	7.13	9.46
LOT	7	7.95	9.50	9.20	11.75	9.69
MLT	7	8.12	9	9.33	11	9.84
ELT	9	8.39	11	9.48	13	9.90
IDEAL	∞	8.56	∞	9.59	∞	9.94

6 Comparações

Das tabelas 1 a 4, pode-se extrair algumas informações úteis para a escolha de um banco de filtros adequado às necessidades de projeto. Comparando bancos de M bandas, apenas, pode-se inferir, das tabelas, que os hierárquicos pos-

uem uma complexidade bem superior aos paralelos. Por exemplo, para $M = 8$, a MLT e ELT (Tabela 4) possuem G_{TC} comparável à cascata de filtros QMF-8A e CQF-8A, respectivamente, com menos da metade da complexidade. Com a cascata de CQF-16B, obteríamos 0.4dB de ganho em relação a LOT ou MLT, mas com gastos computacionais quase 6 vezes maiores. O mesmo vale, em comparação a ELT, havendo um ganho de apenas 0.08dB e uma quadruplicação da complexidade. Para $M = 16$ as diferenças de ganho diminuem e as de complexidade se acentuam. Na realidade, devido a alta complexidade, estas conexões hierárquicas em BFs de M bandas são pouco utilizadas, geralmente apenas com $M = 2$ [15] ou $M = 4$ [11].

Com relação à decomposição do espectro em oitavas, a comparação das tabelas 3 e 4 nos revela uma maior proximidade entre os bancos. Contudo, para $M = 8$, a ELT na Tabela 4 possui maior G_{TC} que qualquer banco listado na Tabela 3 e a LOT possui G_{TC} comparável à cascata de QMF-8A, com menor complexidade. Outras comparações podem ser feitas, mas os bancos de filtros paralelos irão possuir uma relação complexidade-ganho igual ou superior aos hierárquicos. Além disto, já que bancos QMF são sujeitos a distorções, a conexão destes irá amplificá-las.

Outro fator que deve ser ponderado diz respeito ao comprimento dos filtros. Como iremos quantizar os coeficientes, bandas necessárias à reconstrução perfeita da imagem, na síntese, serão deletadas ou mal quantizadas. Face a uma borda, a deleção de algumas bandas acarretará o aparecimento de *ringing* em sua vizinhança. Quanto maior o filtro, maior a amplitude e alcance desta distorção. Por outro

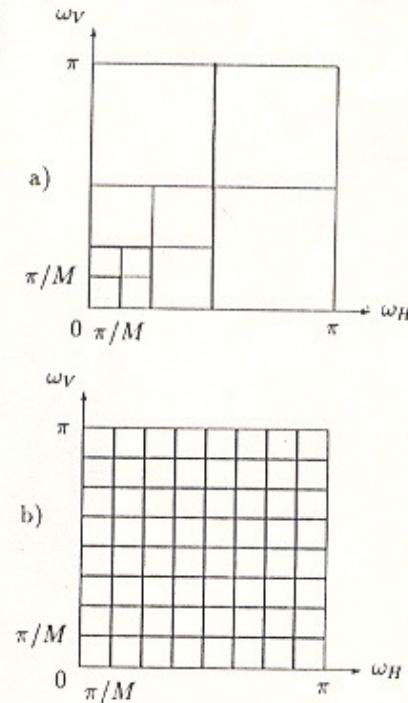


Fig. 2

lado, quando $L = M$ temos uma transformada de blocos e a distorção citada cede lugar a efeitos-de-bloco [3], como comentado em seção anterior. O valor $L = 2M$ parece ser um bom compromisso entre efeitos-de-bloco e *ringing*.

Foram feitas algumas simulações utilizando o codificador adaptativo de Chen e Smith para 8 classes [14], com quantizadores gaussianos otimizados e blocos de 8x8 amostras ($M = 8$). As transformações foram feitas através de processamento linha-coluna em imagens de 256x256 pixels. Nas simulações não foi introduzido nenhum codificador entrópico como o aritmético ou o de Huffman, já que se deseja fazer comparações e não atingir taxas específicas. As simulações incluem a LOT, DCT e as transformadas hierárquicas (em oitavas) utilizando filtros QMF. A Tabela 5 apresenta a SNR dos resultados, definida como a razão entre as variações do sinal e do erro de reconstrução. As taxas-de-bit são apresentadas em bits/ímagem (bpi) e incluem todo o *overhead* de informação necessário. Note que 64Kbpi corresponde a uma média de 1 bit/pixel. As Figuras 3(a), 3(b) e 3(c) mostram a imagem "Lena" codificada a 51Kbpi, utilizando a DCT, LOT e a cascata de QMF-16B, respectivamente.

7 Conclusão

A comparação isolada dos resultados da Tabela 5 nos levaria a concluir que, como as diferenças de SNR são pequenas, deveríamos apenas utilizar o BF de menor complexidade, que seria a DCT. Isto é verdade, em parte, mas o efeito-de-blocos é uma distorção bem visível, por possuir um padrão bem definido. Nas outras, o erro é melhor distribuído. A Fig. 4 apresenta as imagens-erro nos dois casos, para imagem "Girl" a 32Kbpi.

Desta forma, subjetivamente, seria mais interessante se adotar a LOT, que leva a um menor gasto computacional (afora a DCT) e é visualmente comparável ou superior aos outros métodos. Outra alternativa seria utilizar a ELT para $M = 8$, mas mantendo a taxa em um nível razoável, na qual pudesse se controlar o efeito de *ringing*. Estas simulações são deixadas para um trabalho futuro. O mesmo ocorre para sistemas hierárquicos utilizando os filtros 16B.

Por fim, gostaríamos de ressaltar que não é intenção deste trabalho indicar o melhor banco de filtros para codificação por transformadas/sub-bandas, mas apresentar sistemas de bom compromisso custo-desempenho. Tais sistemas, como a LOT, escolhida a partir de um critério teórico, se demonstra eficiente em simulações, mesmo comparada a sistemas mais complexos e, teoricamente, melhores. Tais bancos não demonstram grande melhoria do ponto de vista subjetivo que justifiquem o aumento substancial de complexidade computacional.



Tabela 5 - SNR(dB) resultante de simulações, utilizando a DCT, LOT e bancos hierárquicos baseados em filtros QMF. Imagem e taxa-de-bit são indicadas.

Taxa	DCT	LOT	QMF-8A	QMF-16B
LENA				
32Kbpi	21.8	22.0	21.5	21.8
51Kbpi	23.9	24.4	23.5	24.0
64Kbpi	25.4	25.6	24.8	25.4
128Kbpi	31.3	31.5	30.0	30.9
CAMERAMAN				
32Kbpi	23.0	22.9	22.2	22.6
51Kbpi	25.4	25.2	24.3	24.6
GIRL				
32Kbpi	23.3	23.5	23.0	23.5
64Kbpi	27.3	27.6	26.6	27.1

Agradecimentos : Dr. H. Malvar, Prof. E. Rubino e Ricardo Junqueira. Trabalho parcialmente suportado pelo CNPq. Processo 600.047/90-3.

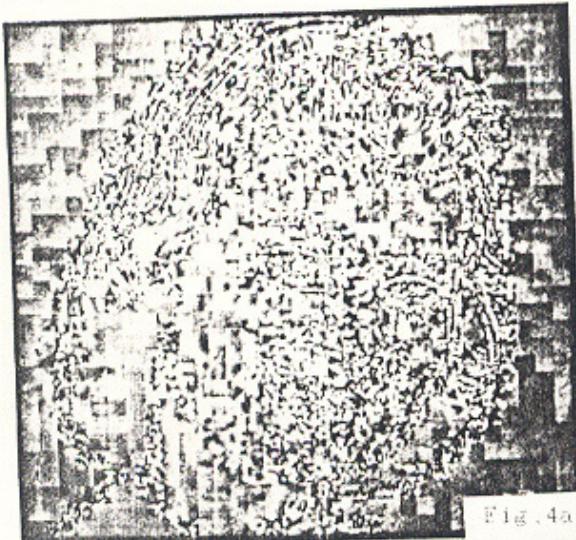


Fig. 4a

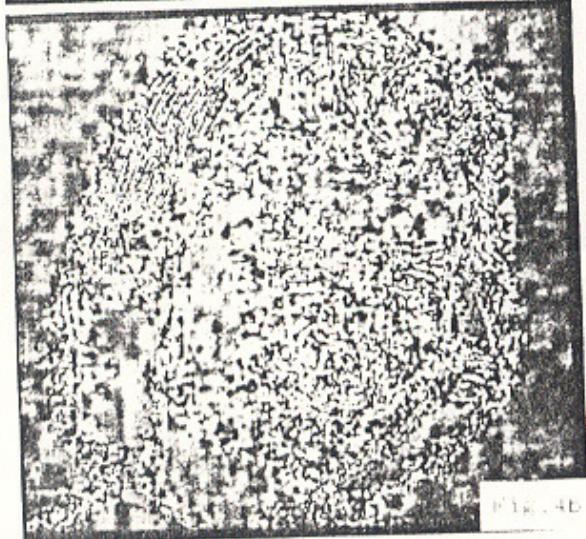


Fig. 4b

Referências

- [1] R. E. Crochiere, L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [2] N. S. Jayant, P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [3] K. R. Rao, P. Yip *Discrete Cosine Transform : Algorithms, Advantages, Applications* San Diego, CA : Academic Press, 1990.
- [4] M. J. Smith, T. P. Barnwell, 'Exact reconstruction techniques for tree-structured subband coders', *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-34, pp. 434-441, June 1986.
- [5] M. Vetterli, C. Herley, "Wavelets and filter banks: theory and design," Tech. Rep. 206/90/36, CTR, Columbia University, 1990.
- [6] P. P. Vaidyanathan, P. Hoang, "Lattice structures for optimal design and robust implementation of 2-channel PR-QMF banks," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-36, pp. 81-94, Jan. 1988.
- [7] H. S. Malvar, "Lapped transforms for efficient transform/subband coding," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-38, pp. 969-978, June 1990.
- [8] H. S. Malvar, D. H. Staelin, "The LOT: transform coding without blocking effects," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-37, pp. 553-559, Apr. 1989.
- [9] H. S. Malvar, "Modulated lapped transforms: properties, applications and fast algorithms," accepted for publication, *IEEE Trans. Signal Processing*, 1991.
- [10] G. K. Wallace, "The JPEG Still Picture Compression Standard," *Communications of the ACM*, Vol.34, No. 4, April 1991.
- [11] P. H. Westerink "Sub-band coding of Images," *PhD Thesis, Delft Technische Universiteit* The Netherlands, October 1989.
- [12] R. L. de Queiroz, "Complexity and selectivity of tree-structured filter banks," submetido para publicação, 1991.
- [13] R. L. de Queiroz, H. S. Malvar, "On the sub-optimality of hierarchical transforms," submetido para publicação, 1991.
- [14] "Adaptive coding of monochrome and color images," *IEEE Trans. Commun.*, COM-25, pp 1285-1292, Nov. 1977.
- [15] J. W. Woods, S. D. O'Neil, "Sub-band coding of images," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-34, pp. 1278-1288, Oct. 1986.