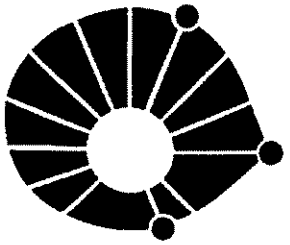
defendida por Rubem Toledo Bergamo e aprovada pela Comissão Julgada em 29 09 00
Orientador

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e Computação
Departamento de Comunicações

UNICAMP



UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Análise da Compactação de Energia na Codificação de Imagens por Sub-Bandas, Utilizando Banco de Filtros Lineares e Não-Lineares

Rubem Toledo Bergamo

Orientador: Prof. Dr. João Baptista T. Yabu-uti

Banca examinadora:

Prof. Dr. João Baptista T. Yabu-uti
DECOM/FEEC/UNICAMP

Prof. Dr. Rui Seara
UFSC – Florianópolis, SC

Prof. Dr. Max Henrique M. Costa
DECOM/FEEC/UNICAMP

Prof. Dr. Yuzo Iano
DECOM/FEEC/UNICAMP

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação (FEEC) da Universidade Estadual de Campinas (UNICAMP), como parte dos requisitos exigidos para obtenção do título de Mestre em Engenharia Elétrica.

Área de Concentração:
Telecomunicações e Telemática

Dissertação de Mestrado
Campinas – SP – Brasil
Setembro, 2000

UNICAMP

UNIDADE BC
N.º CHAMADA
T/UNICAMP
B452a
V. Ex.
TOMBO BC/ 43800
PROC. 16-392/01
C D
PREÇO R$ 11,00
DATA 14/02/01
N.º CPD

CM-00154316-2

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

B452a

Bergamo, Rubem Toledo
Análise da compactação de energia na codificação de imagens por sub-bandas, utilizando banco de filtros lineares e não-lineares / Rubem Toledo Bergamo.-- Campinas, SP: [s.n.], 2000.

Orientador: João Baptista T. Yabu-uti
Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Processamento de imagens – Técnicas digitais. 2. Compressão de dados (Telecomunicações). 3. Teoria da codificação. I. Yabu-uti, João Baptista T. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Ao meu amigo, muita vezes pai e avô Toledo, *in memoriam*, pela presença e incentivo constante na minha vida acadêmica. Muito do que sou hoje devo a você. Meu muito obrigado onde quer que você esteja.

BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

AGRADECIMENTOS

Eis aqui a árdua tarefa de agradecer a todos que, de uma forma direta ou indireta, contribuíram e conviveram comigo nestes anos de vida acadêmica

Agradeço ao meu orientador Dr. João Baptista T. Yabu-uti, por sua competência e dedicação que foram importantes para o êxito do meu trabalho. Agradeço especialmente o apoio que me deu durante esses anos.

Meu agradecimento ao Professor Doutor Max Henrique Machado Costa por sua ajuda e dedicação, fossem em aulas ou conversas que foram essenciais para que esse sonho fosse realizado. Agradeço também ao Prof. Dr. Yuzo Iano as conversas que tivemos em momentos difíceis que foram fundamentais para a realização deste sonho. Aos professores Dr. Dalton Soares Arantes e Dr. Fábio Violaro pela clareza e simpatia que transmitiram-me seus conhecimentos e sempre estiveram a disposição para me ajudar.

Meu agradecimento e Escola Técnica Federal de Santa Catarina - UNED - São José, em especial ao grupo de Telecomunicações, pelo apoio e incentivo que me deram para que eu pudesse realizar este trabalho e a todos que de uma forma ou outra acreditaram em mim.

Meu agradecimento em especial ao amigo e companheiro Felipe Antônio Oliveira de Souza pelo apoio e diálogos que tivemos durante esses anos de convívio, que foram fundamentais para o êxito do meu trabalho. Aos amigos que lá fiz, Érico, Ernesto, Ruyther, Bruno, Cláudio, Edson, Rômulo, Fabbryccio, Carlos Henrique, Franciraldo, Raul, Vitor, Mehdi, Vicente, Wilson e entre outros que conviveram comigo.

Aos funcionários do DECOM, Lúcia e Ricardo.

Agradeço à Fundação Coordenação de Aperfeiçoamento do Pessoal de Nível Superior – CAPES- pelo suporte financeiro concedido a este trabalho desde o início.

A minha filha, Gabriella, uma das principais razão deste meu sonho, agradeço a sua compreensão e espera. Para ela, todo meu amor.

Ao meu pai, Rubem, meu maior incentivador, sempre me apoiou nas horas mais difíceis da minha vida. À minha mãe Maria José, que sempre sofreu comigo nos momentos mais difíceis e, principalmente pelas palavras de ânimo e carinho que sempre dispôs.

As minhas avós Cléo e Nenzinha, por palavras de carinho e sensibilidade. Aos meus avôs Toledo(*in memorian*) e Albino(*in memorian*) onde quer que estejam, obrigado por tudo.

A nova família que me acolheu de braços abertos em Campinas, Seu Napoleão, Dona Edneia, Diego, Richard, Cláudia, Frederico e Guilherme.

Agradeço muitíssimo a minha mulher, Carla, o apoio e carinho que me dedicou durante todos esses anos, sem os quais teria sucumbido nos momentos difíceis. A tarefa de terminar este trabalho seria muito mais árdua sem a sua presença. A ela meu eterno amor.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

RESUMO

Este trabalho avalia sistemas baseados em transformadas de multi-resolução ou decomposição em sub-bandas na aplicação em codificação de sinais e imagens. Uma importante propriedade que estas formas de representação exibem é a compactação de energia no domínio dos coeficientes resultantes da transformada. Este trabalho investiga esta propriedade através de uma medida empírica definida como ganho de compactação (GC). Em muitos casos esta medida é igual ao ganho de codificação em sub-bandas em relação à codificação PCM. A medida proposta apresenta a vantagem de poder ser calculada exclusivamente a partir das variâncias das sub-bandas geradas. Segundo esta medida, a eficiência de compactação de energia de diversos bancos de filtros lineares e não-lineares é estudada para várias imagens de teste. Estimam-se as complexidades computacionais dos bancos de filtros que tiveram melhor desempenho quanto à compactação de energia e implementa-se um sistema básico de compressão de imagens para avaliar o desempenho dos bancos de filtros através da PSNR versus taxa de compressão e da qualidade visual.

ABSTRACT

This study evaluates some systems based on multiresolution transforms or subband decomposition in image coding. An important characteristic presented by those forms of representation is the compaction of energy in the transform domain. This work investigates such a characteristic using of an empirical measure, witch is defined as compaction gain (CG). In many cases, this measure is equal to the subband coding gain in relation to PCM coding. The proposed measure is advantageous since it may be calculated exclusively through the variance of the subbands. According to that measure, the efficiency in energy compaction of several linear and non-linear filter banks is analyzed for a number of test images. Computational complexity is evaluated for the filter banks which achieved the best performances regarding energy compaction, and a basic system for image compaction has been implemented in order to evaluate the performance of the filter banks concerning PSNR versus compression rate and visual quality.

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO 1

CAPÍTULO 2 – DECOMPOSIÇÃO EM SUB-BANDAS 5

2.1 TÉCNICAS DE COMPRESSÃO DE IMAGEM 5

2.2 BANCO DE FILTROS DE SINAIS DISCRETOS NO TEMPO 8

2.2.1 Dizimação 8

2.2.2 Interpolação 11

2.3 DECOMPOSIÇÃO DE SINAIS EM MULTIBANDAS 12

2.3.1 Decomposição em Duas Bandas Uniformes 15

2.3.2 Banco de Filtros QMF 18

2.4 DECOMPOSIÇÃO LINEAR DE MULTI-RESOLUÇÃO DE IMAGENS 20

2.5 DECOMPOSIÇÃO NÃO LINEAR EM SUB-BANDAS 24

2.5.1 Decomposição Morfológica com Reconstrução Perfeita 25

2.5.2 Filtros de Mediana 29

2.5.3 Decomposição/ Reconstrução Não Linear para Imagens 30

CAPÍTULO 3 – COMPACTAÇÃO DE ENERGIA E COMPRESSÃO DE IMAGENS USANDO BANCO DE FILTROS 34

3.1 INTRODUÇÃO 34

3.2 CRITÉRIOS PARA SELEÇÃO DOS BANCOS DE FILTROS 35

3.2.1 Fase Linear Generalizada 35

3.2.2 Comprimento dos Filtros 36

3.2.3 Ganho de Compressão 37

3.3 ANÁLISE DO GANHO DO COMPACTAÇÃO DE ENERGIA 38

3.3.1 Ganho de Codificação para Sub-bandas (G_{SBC}) 38

3.3.2 Estimativa do Ganho de Compactação de Energia 40

3.3 ESQUEMAS DE COMPRESSÃO DE IMAGENS VIA CODIFICAÇÃO DE SUB-BANDAS 42

3.3.1 Quantização das Sub-Bandas....45
3.3.2 Quantização com Alocação de Bits....48
3.3.3 Quantização com Ponderação das Sub-Bandas....50
3.3.4 Codificação de Entropia....51

CAPÍTULO 4 – DESCRIÇÃO DO SISTEMA IMPLEMENTADO NAS SIMULAÇÕES....54

4.1 SISTEMAS IMPLEMENTADOS PARA SIMULAÇÃO....54
4.1.1 Implementação dos Métodos da Convolução Circular e da Extensão Simétrica....56
4.1.2 Os Quantizadores Implementados....58
4.2 COMPLEXIDADE COMPUTACIONAL....60
4.3 MEDIDAS DE GANHO E CRITÉRIO DE QUALIDADE VISUAL OBJETIVA....61

CAPÍTULO 5 – APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS....63

5.1 MEDIDAS DE GANHO DE COMPACTAÇÃO (GC_N)....64
5.2 SIMULAÇÕES COM COMPRESSÃO DE IMAGENS VIA BANCO DE FILTROS....71
5.3 REDUÇÃO DA COMPLEXIDADE COMPUTACIONAL....75

CAPÍTULO 6 – CONSIDERAÇÕES FINAIS E SUGESTÕES DE TRABALHOS....76

APÊNDICES DE IMAGENS....79

REFERÊNCIAS BIBLIOGRÁFICAS....86

CAPÍTULO 1

INTRODUÇÃO

Durante as últimas décadas, vem crescendo o interesse em pesquisar métodos de codificação de imagens que apresentem altas taxas de compressão. A compressão de imagem é essencial à maioria das aplicações que possibilitam a transmissão de imagem através da Internet, a videoconferência, o videofone, a televisão de alta definição, os vídeos domésticos, a televigilância, os bancos interativos de imagens médicas e comerciais e as imagens obtidas por satélites e sondas espaciais. Toda a informação contida numa determinada imagem deve ser convenientemente comprimida para tornar eficiente a sua transmissão remota, bem como para não sobrecarregar a capacidade dos dispositivos de armazenamento.

Um dos métodos mais utilizados na compressão de imagens é a codificação por transformada, cujo objetivo principal é a redução da correlação existente entre os elementos de uma determinada imagem, procurando concentrar a maior quantidade de energia possível no menor número de coeficientes, de modo que estes possam ser quantizados separadamente. No entanto, esse método de compressão introduz distorções na imagem original quando se exige altas taxas de compressão.

Dentre essas distorções, existem duas que são mais freqüentes nos codificadores de imagem. A primeira é chamada de efeito de bloco (*blocking effect*) e surge quando o

processamento da imagem é feito por blocos, como por exemplo nos padrões, JPEG (*Joint Photographic Experts Group*) e MPEG (*Motion Pictures Expert Group*) . Esses formatos são ambos baseados na DCT (*Discrete Cosine Transform*) e constituem padrão para imagem estática e vídeo, respectivamente. Para fatores de compressão entre 30 a 40 vezes, a transformada utilizada pelo JPEG, por exemplo, introduz um forte efeito de bloco na imagem. Esse efeito aumenta para se ter taxas de transmissão menores.

Infelizmente, o olho humano é muito sensível a este tipo de distorção. Por isso, codificadores que utilizam a DCT não são apropriados para ambientes em que se requer altas taxas de compressão e se têm baixas taxas de transmissão.

O segundo tipo de distorção é devido ao fenômeno de Gibbs dos filtros lineares, também conhecido como "*ringing effect*". Este efeito, que aparece como contornos nas bordas, é mais visível ao redor das imagens de alto contraste e pode ser também muito desagradável. Para se reduzir o efeito de *ringing*, um banco de filtros lineares apropriado pode ser utilizado [EGGE95b], mas não é possível achar um banco de filtros lineares que não tenha nenhum tipo de efeito *ringing*. Recentemente, são alvo de intensas pesquisas os métodos baseados na decomposição de sinais em bandas de freqüências disjuntas, usando bancos de filtros e, em particular, a decomposição multi-resolução de sinais baseada na teoria das funções *wavelets*.

A análise em multi-resolução ou decomposição em sub-bandas, implementada através de bancos de filtros, constitui uma poderosa ferramenta na codificação de sinais. A idéia consiste em representar um sinal discreto através de uma série de sinais elementares, constituídos por um sinal de aproximação e uma série de sinais de detalhamento. Estes sinais possuem conteúdos em freqüências aproximadamente disjuntas e por isso são chamados sinais de sub-bandas, ou simplesmente sub-bandas do sinal original.

Geram-se estes sinais através da passagem do sinal original por um banco ou árvore de filtros. É conseguida uma codificação eficiente alocando-se resoluções diferentes na representação ou quantização destas sub-bandas de acordo com os parâmetros de importância exibidos por cada uma delas. Na ausência de quantização, uma condição desejável ao banco de filtros é que a síntese do sinal a partir das sub-bandas produza reconstrução perfeita. Com isso, a realização de uma quantização criteriosa sobre os coeficientes da decomposição provê mínima ou nenhuma distorção perceptível do sinal original.

O presente trabalho define uma medida de mérito de decomposição com a qual se analisa a eficiência da representação em sub-bandas de sinais bidimensionais ou imagens, em termos de relações entre as variâncias (energias) dos coeficientes utilizados na representação da imagem original. São estudados os casos de bancos lineares e não-lineares [EGGE95a]. Após a verificação desta medida de compactação de energia, foram implementados algoritmos com o objetivo de simular sistemas de compressão para avaliar o desempenho destes bancos de filtros no que se refere a PSNR (*Peak Signal to Noise Ratio*) e qualidade visual objetiva para determinadas taxas de compressão.

O trabalho, além deste capítulo introdutório, apresenta os seguintes capítulos. O Capítulo 2, contém uma revisão dos conceitos relativos à decomposição em sub-bandas através de bancos de filtros lineares unidimensionais baseados na teoria de filtros ortogonais e não-ortogonais, com atenção especial aos filtros QMF (*Quadrature Mirror Filter*). É vista também a decomposição em sub-bandas com banco de filtros não-lineares e uma apresentação dos filtros medianos que serão usados neste tipo de decomposição e o filtro MED 2. Também se discute neste capítulo uma extensão para o caso bidimensional destes bancos de filtros para os casos lineares e não-lineares.

O Capítulo 3 apresenta os critérios utilizados para a escolha de um bom filtro para compressão de imagens aplicada a uma estrutura de banco de filtros. É discutida a importância das estatísticas de cada sub-banda em relação à operação de quantização e se apresenta uma medida de compactação de energia relacionada à decomposição em sub-bandas, definida como Ganho de Compactação (GC_N). Também é apresentado um esquema básico de compressão de imagens utilizando-se a decomposição de multi-resoluções na aplicação de compressão de imagens. Nesse capítulo, ainda se apresentam os quantizadores utilizados nas simulações, o algoritmo de alocação de bits com ponderação nas sub-bandas e a codificação de entropia.

O Capítulo 4 apresenta e descreve o modelo de compressão básico utilizado nas simulações, os métodos de extensões periódicas utilizados, os quantizadores utilizados, a alocação de bits de acordo com sua importância e comentários sobre estes procedimentos.

O Capítulo 5 contém os gráficos e as tabelas comparativas baseados na medida de Ganho de Compactação (GC_N) para as diversas decomposições realizadas. São mostrados os gráficos referentes a PSNR (*Peak Signal to Noise Ratio*) *versus* bits por pixel (bpp) e ainda PSNR *versus* entropia do sinal quantizado, fazendo-se análises e comentários sobre os resultados obtidos.

Por fim, o Capítulo 6 apresenta as principais conclusões e implicações do trabalho e oferece sugestões para trabalhos futuros.

CAPÍTULO 2

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

DECOMPOSIÇÃO EM SUB-BANDAS

Os métodos mais utilizados na compressão, processamento e codificação de imagens, adotam transformadas lineares. A transformada de sub-bandas é uma subclasse da transformada linear que oferece propriedades bastante úteis para as aplicações anteriormente descritas.

Neste capítulo, será introduzida, de forma sucinta, a teoria de banco de filtros lineares e não-lineares. Começa-se pela descrição sumária de algumas técnicas de compressão de imagens. Em seguida, apresenta-se um estudo de banco de filtros, quer no domínio da freqüência, quer no domínio discreto, com particular ênfase nos bancos de filtros com reconstrução perfeita. Em seguida, mostram-se as características dos filtros lineares espelho em quadratura (QMF – *Quadrature Mirror Filter*), os filtros medianos, e as decomposições em multi-resoluções com bancos de filtros lineares e não-lineares.

2.1 TÉCNICAS DE COMPRESSÃO DE IMAGEM

As especificidades de cada aplicação e, em particular, a taxa de compressão média que é necessária atingir, que é função do formato inicial das imagens, da capacidade de uso dos

meios e da qualidade visual que as imagens devem possuir após o processo de compressão, são fatores que condicionam a seleção das técnicas de compressão a ser usada em cada caso. As técnicas de compressão disponíveis são baseadas na redução (idealmente na eliminação) das interdependências e das redundâncias entre amostras, quer espaciais, quer temporais, de forma a que se possa diminuir a quantidade de informação que é necessário transmitir ou armazenar.

Na Fig. 2.1, apresenta-se, em diagrama de blocos, um sistema genérico de compressão de imagens. Inicialmente, as imagens são digitalizadas de acordo com um formato (normalizado) – por exemplo, a norma CCIR 601 [CCIR82]. Se considerada esta norma, cada elemento da imagem ou amostra é codificado em PCM (*Pulse Code Modulation*) com uma resolução de 8 bits.

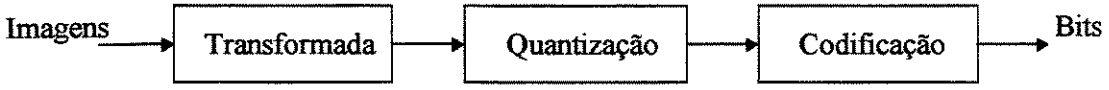


Figura 2.1 – Modelo simplificado de um sistema de compressão de imagens.

Sabe-se que imagens representativas de cenas naturais (não sintéticas) contêm um elevado nível de redundância e correlação estatística entre amostras vizinhas. De forma que se possam eliminar e reduzir estes efeitos e obter taxas de compressão consideráveis, é conveniente que se passe a representação original das amostras em PCM para um espaço de representação mais adequado do ponto de vista da compressão. Esta operação está simbolizada na Fig. 2.1 pelo bloco designado por *Transformada.*

A mudança do espaço de representação das amostras permite diminuir a entropia dos dados a codificar e definir critérios de importância visual, que são bastantes relevantes na fase de quantização. As principais técnicas de transformações são as técnicas preditivas e as transformadas lineares.

No grupo das técnicas preditivas, enquadram-se as técnicas baseadas na teoria de predição linear. A idéia principal consiste em predizer os valores das amostras seguindo uma codificação do erro de predição. Trata-se do que se denomina por DPCM (*Differential Pulse Code Modulation*). Os preditores podem ser fixos ou adaptativos, e a predição pode se realizar

quer no domínio espacial quer no domínio temporal. A predição temporal tem grande interesse e maior aplicabilidade quando se quer comprimir seqüências de imagens animadas.

Já o grupo das transformadas lineares envolve as técnicas baseadas na utilização de transformadas lineares inversíveis, que possuem a capacidade de concentrar a energia das amostras num conjunto reduzido de coeficientes e, simultaneamente, de diminuir ou eliminar as dependências estatísticas entre as amostras. A transformada de Karhunen-Loeve é ótima, segundo os critérios considerados, mas sua aplicação prática é bastante complexa, pois a transformada depende das características estatísticas dos sinais [LIM89]. Como os sinais são, geralmente, não estacionários, não é possível desenvolver algoritmos rápidos de cálculo para a transformada de Karhunen-Loeve.

Uma transformada linear muito utilizada e popular é a Transformada Discreta do Coseno (DCT). Esta desempenha um papel fundamental para as recomendações atuais de compressão de imagens estáticas e animadas: H.261, JPEG e MPEG [NETR88 e EBRA98]. A DCT usada nesses padrões é um caso particular das transformadas trigonométricas e corresponde à transformada de Fourier discreta no tempo com extensão simétrica de meia amostra para o sinal.

Quando se combinam num sistema de compressão as técnicas preditivas e as transformadas lineares, é comum denominá-lo de sistema híbrido. Para que se possam explorar as características dos dados transformados é necessário combinar esta técnica com outras, definidas na Fig. 2.1 como a *Quantização* e a *Codificação*.

A quantização tem por objetivo limitar o universo da representação dos coeficientes gerados pelas transformadas a um conjunto reduzido de possibilidades. Os coeficientes podem ser quantizados individualmente ou em grupos, o que geralmente se define, respectivamente, por quantização escalar ou vetorial [GERS92]. A operação de quantização torna o processo de compressão irreversível, isto é, as imagens reconstruídas são numericamente diferentes das imagens originais devido à introdução do erro de quantização. Com o objetivo de atenuar os efeitos indesejáveis visíveis desse erro e proceder a uma distribuição eficiente da taxa de transmissão disponível, é comum que sejam usados diferentes níveis de quantização em função da importância visual de cada coeficiente ou grupo de coeficientes.

Após o processo de quantização, pode-se ainda explorar a redundância estatística presente nos dados, usando as técnicas de codificação por entropia [JAYA84]. Essa operação é representada na Fig. 2.1 pelo bloco designado *Codificação*.

2.2 BANCO DE FILTROS DE SINAIS DISCRETOS NO TEMPO

Os métodos de compressão baseados em decomposições lineares de sinais, também denominados de métodos das transformadas, têm como principal objetivo a concentração de energia dos sinais num conjunto limitado de coeficientes e o ordenamento ou a hierarquização dos diferentes coeficientes ou bandas, de acordo com sua importância relativa. Neste grupo, enquadram-se as técnicas baseadas na divisão dos sinais em bandas de freqüência disjuntas, usando um banco de filtros com capacidade de reconstrução perfeita.

Os primeiros trabalhos nesta área foram realizados por Crochiere, no âmbito da codificação de sinais de áudio [CROC76]. Mais tarde, Woods dedicou-se à compressão de imagens e vídeo [WOOD86].

Após estes trabalhos iniciais, tornou-se de grande interesse a aplicação destas técnicas e, em particular, o desenvolvimento de novos projetos de bancos de filtros [VAID87]. Neste campo, assume particular relevância o trabalho de Smith [SMIT86], no qual o autor apresenta uma família de bancos de filtros que permitem eliminar simultaneamente os vários tipos de distorção: amplitude, fase e sobreposição dos espectros (também denominada de *aliasing*). Estes banco de filtros possuem a capacidade de reconstruir o sinal original de forma perfeita.

De modo a facilitar a compressão dos principais aspectos da teoria de bancos de filtros, será efetuada uma breve revisão preliminar de algumas noções básicas de processamento digital de sinais. Para começar, apresentam-se dois componentes básicos de um sistema de decomposição utilizando banco de filtros, que são a dizimação (*downsampling*) e a interpolação (*upsampling*).

2.2.1 Dizimação

Na Fig. 2.2, apresenta-se um sistema de dizimação por um fator M de sinais discretos no tempo.

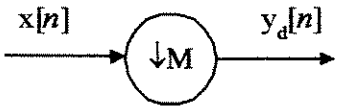


Figura 2.2 – Dizimação por um fator M.

Inicialmente, podem-se analisar as propriedades do sistema dizimador representado na Fig. 2.2. A relação entre a entrada e a saída no domínio discreto no tempo é dada pela seguinte equação:

$$y_d[n] = x[nM], \text{ com } M \in Z. \tag{2.1}$$

O sinal de saída, $y_d[n]$, é o resultado da amostragem do sinal de entrada $x[n]$ com um período igual a M. Isto é, em cada conjunto de M amostras, é retida uma. Este processo é geralmente denominado por subamostragem, uma vez que o número de amostras do sinal de saída é menor.

Para se determinar a relação equivalente a (2.1) no domínio da freqüência, considere-se a formação de um sinal auxiliar, $x'[n]$, com o mesma taxa de $x[n]$, definido por,

$$x'[n] = x[n].\sum_{k=0}^{+\infty}\delta[n - kM]. \tag{2.2}$$

Ou seja, é o resultado do produto do sinal original por um trem de impulsos periódicos. Como o trem de impulsos é periódico, então pode ser expresso a partir da série de Fourier discreta no tempo [OPPE89], dada por:

$$\sum_{k=0}^{+\infty}\delta[n - kM] = \frac{1}{M}\sum_{k=0}^{M-1}e^{j2\pi nk/M} \quad , \tag{2.3}$$

o que permite rescrever a equação (2.2) da seguinte forma:

$$x'[n] = \frac{1}{M}\sum_{k=0}^{M-1}x[n]e^{j2\pi nk/M} \quad . \tag{2.4}$$

Se for aplicada a transformada **Z** à equação (2.4), obtém-se, após algumas manipulações, a seguinte expressão,

$$X'(z) = \frac{1}{M}\sum_{k=0}^{M-1} X(zW^k), \tag{2.5}$$

onde $W = e^{-j2\pi/M}$ e $z = e^{j\omega}$. Neste momento, é possível determinar a transformada **Z** da saída, dada por:

$$Y_d(z) = \sum_{k=-\infty}^{+\infty} x'[kM]\, z^{-k} = \sum_{k=-\infty}^{+\infty} x'[k](z^{1/M})^{-k}, \tag{2.6}$$

ou

$$Y_d(z) = \mathrm{X}'(z^{1/M}). \tag{2.7}$$

Substituindo (2.5) em (2.7), obtém-se finalmente a transformada **Z** da saída do dizimador por M, dada por,

$$Y_d(z) = \frac{1}{\mathrm{M}}\sum_{k=0}^{M-1} X(z^{1/M}W^k). \tag{2.8}$$

Uma vez conhecida a transformada **Z** de (2.1), pode-se saber a transformada de Fourier. Para isto, basta calcular (2.8) ao longo da circunferência de raio unitário, isto é,

$$Y_d(e^{j\omega}) = \mathrm{Y_d}(z)\Big|_{z=e^{j\omega}} = \frac{1}{\mathrm{M}}\sum_{k=0}^{M-1} X(e^{j(\omega-2\pi k)/M}). \tag{2.9}$$

A expressão (2.9) corresponde à soma de M réplicas de X(ω/M) (expansão na freqüência) espaçadas de 2π/M rad. Como no caso do teorema de Nyquist para sinais contínuos no tempo, se o sinal não for de banda limitada, verificar-se-á um fenômeno destrutivo por perda de informação contida em x[*n*], denominado sobreposição de espectro (*aliasing*). Esta sobreposição dos espectros pode ser vista na Fig. 2.3.

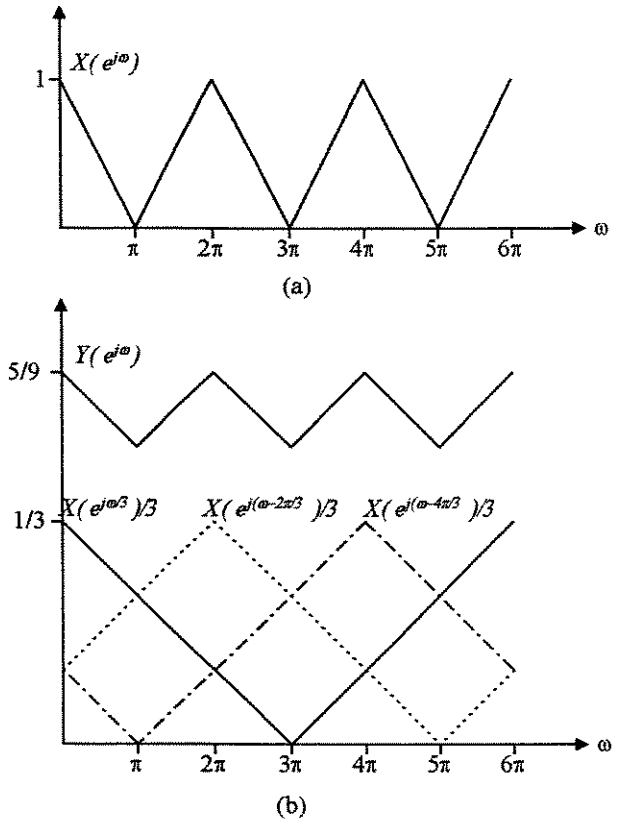


Figura 2.3 – Dizimando um sinal $X(e^{j\omega})$ no domínio da freqüência por um fator M = 3. (a) espectro original. (b) As três réplicas do sinal após a dizimação e a soma deles para formar $Y(e^{j\omega})$.

2.2 Interpolação

Analogamente, pode-se realizar um estudo similar para o interpolador representado pela Fig. 2.4.

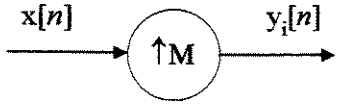


Figura 2.4 – Interpolação por um fator M.

A ação do interpolador no sistema resulta num aumento da variância do sinal de entrada, por inserção de amostras nulas, ou seja, amostras de valores zeros. A relação entre a entrada e a saída da função interpoladora no domínio discreto do tempo, é definida por

$$y_i[n] = \begin{cases} x[n/M], & n = kM, k \in \mathrm{Z} \\ 0 & \text{caso contrário.} \end{cases} \tag{2.10}$$

No domínio da transformada de Fourier, tem-se

$$\mathrm{Y}(e^{j\omega}) = \mathrm{X}(e^{jM\omega}). \tag{2.11}$$

Aplicando a transformada **Z** na expressão (2.10), facilmente fica prova-se o resultado da equação (2.12):

$$Y_e(z) = \sum_{k=-\infty}^{+\infty} x[k/M] z^{-k} = \sum_{k=-\infty}^{+\infty} x[k] (z^M)^{-k} = X(z^M). \tag{2.12}$$

A expressão (2.12) significa, em termos de freqüência, uma compressão por M do espectro do sinal de entrada, seguido da formação de M réplicas ou imagens do mesmo espectro, devido à periodicidade de $\mathrm{X}(\omega)$. A este sistema normalmente está associado um filtro interpolador, cuja função é eliminar as imagens do espectro indesejáveis.

2.3 DECOMPOSIÇÃO DE SINAIS EM MULTIBANDAS

Na Fig. 2.5, está representado um sistema genérico de processamento de multibandas de sinais discretos no tempo, $\mathrm{x}[n]$, baseado em bancos de filtros. Geralmente os sinais de entrada são o resultado da amostragem de um sinal qualquer contínuo no tempo. Será considerado, nesta seção, unicamente um sinal unidimensional. Uma abordagem semelhante para sinais não-separáveis de duas dimensões é perfeitamente possível também.

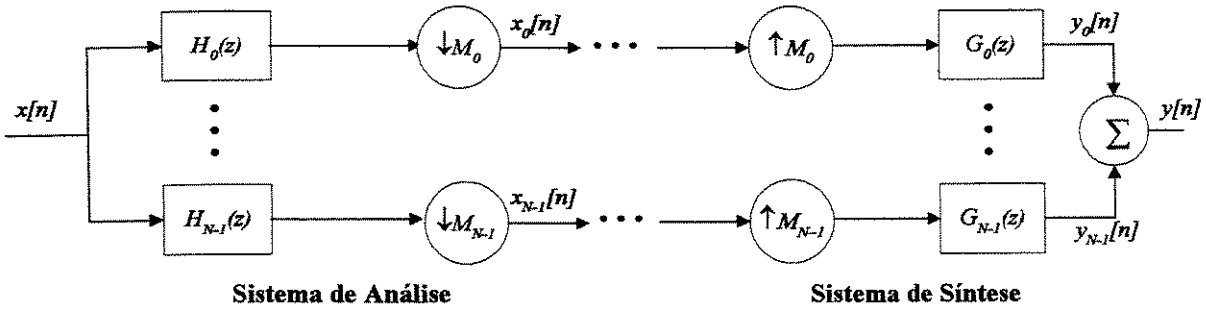


Figura 2.5 – Sistema genérico de decomposição de sinais em multibandas baseado em bancos de filtros.

No sistema de análise, o sinal é filtrado pelos filtros de análise, $H_i(z)$, dizimado e separado em N bandas de freqüência disjuntas, $x_i[n]$. A largura de banda de cada filtro é π/Mi, onde M*i* é o fator de dizimação. Os fatores de dizimação são escolhidos de forma que o sistema seja criticamente amostrado, isto é que a soma das taxas de dados das sub-bandas seja igual a do sinal original:

$$\sum_{i=0}^{N-1} \frac{1}{M_i} = 1. \tag{2.13}$$

O sistema de síntese tem como objetivo reconstruir o sinal original a partir da informação contida em cada sub-banda, $x_i[n]$. Os sistemas interpoladores recuperam a taxa do sinal original por inserção de amostras nulas. Após esta operação, são usados os filtros de síntese, $G_i(z)$, de forma a eliminar as imagens do espectro que são indesejáveis e a sobreposição dos espectros (*aliasing*) devido à dizimação. Os sinais resultantes, $y_i[n]$, são combinados por adição, de forma que o sinal resultante, $y[n]$, seja uma réplica fiel (eventualmente atrasada) do sinal de entrada, $x[n]$.

Dependendo do tipo de aplicação em que o sistema da Fig. 2.5 é utilizado, podem-se encontrar diversos sistemas de processamento entre a análise e a síntese, como, por exemplo, compressão e codificação das sub-bandas, canais de transmissão de dados e sistemas de armazenamento. Quando os sistemas entre a análise e a síntese não introduzem erros nas sub-bandas, $x_i[n]$, é possível, por escolha adequada dos filtros, reconstruir fielmente o sinal de

entrada, mesmo quando os filtros não possuem características em freqüências ideais [SMIT86].

Nesta seção, mostrar-se-á as relações que os filtros de análise e síntese devem satisfazer para que se possa construir um sistema com reconstrução perfeita. Com este objetivo, serão escritas, no domínio da transformada **Z**, as equações de análise para cada sub-banda, que integram simultaneamente o efeito da filtragem e da dizimação, de acordo com (2.8), ou seja,

$$X_k(z) = \frac{1}{M_k} \sum_{n=0}^{M_k-1} H_k(z^{1/M_k} W_k^n) X(z^{1/M_k} W_k^n), \tag{2.14}$$

para $k = 0,1,\ldots\ldots$, N-1 e $W_k = e^{-j2\pi/Mk}$.

Analogamente, para o sistema de síntese, pode-se escrever, de acordo com (2.12) as seguintes equações:

$$Y_k(z) = \mathrm{X_k}(z^{M_k}) G_k(z), \tag{2.15}$$

para $k = 0,1, \ldots$, N-1. Atendendo à linearidade da transformada **Z**, pode-se escrever

$$Y(z) = \sum_{k=0}^{N-1} Y_k(z). \tag{2.16}$$

Substituindo (2.14) e (2.15) em (2.16), pode-se, finalmente, escrever

$$\begin{aligned} Y(z) = &\left[\sum_{k=o}^{N-1} \frac{1}{M_k} H_k(z) G_k(z) \right] \mathrm{X(z)} + \\ &+ \sum_{\mathrm{k=o}}^{\mathrm{N\text{-}1}} \sum_{n=1}^{M_k-1} \frac{1}{M_k} H_k(zW_k^n) X(zW_k^n) G_k(z) \ . \end{aligned} \tag{2.17}$$

A primeira prioridade de qualquer método de projeto de banco de filtros é garantir que a segunda parcela da equação (2.17) seja nula, independentemente do sinal de entrada, isto é, que o conjunto de filtros escolhidos elimine o componente responsável pela sobreposição dos

espectros (*aliasing*). Admitindo que tal solução exista, resta garantir que a função transferência do sistema, seja unicamente um atraso [VAID87].

2.3.1 Decomposição em Duas Bandas Uniformes

Um caso particular do sistema representado na Fig. 2.5, muito discutido e de grande interesse prático, é a decomposição em duas bandas uniformes [SMIT86], [VAID87] e [WOOD86] , cujo diagrama está representado na Fig. 2.6.

O sistema de análise é constituído por um filtro passa-baixas e por um filtro passa-altas, com largura de banda idealmente igual a $\pi/2$ rad e dizimadores por um fator 2. O sistema de síntese é formado por interpoladores com fator 2 e filtros de reconstrução com características semelhantes.

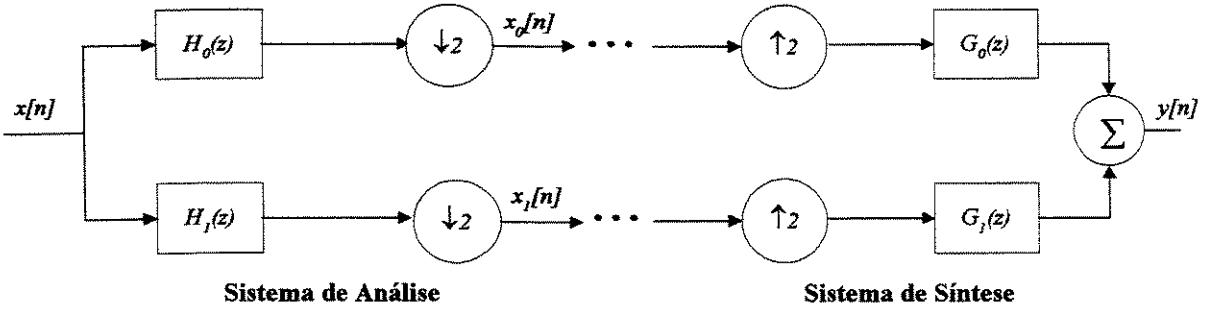


Figura 2.6 – Decomposição em banco de filtros em duas bandas.

Como se disse anteriormente, nesta seção, o foco de interesse é o estudo de um banco de filtros de duas bandas uniformes. Neste caso, a equação (2.17) reduz-se a:

$$\begin{aligned} Y(z) = &\frac{1}{2}\left[H_0(z)G_0(z) + H_1(z)G_1(z)\right]X(z) + \\ &+\frac{1}{2}\left[H_0(-z)G_0(z) + H_1(-z)G_1(z)\right]X(-z), \end{aligned} \tag{2.18}$$

assim, é possível expressar Y(z) da seguinte forma:

$$Y(z) = T(z)X(z) + A(z)X(-z), \tag{2.19}$$

onde $T(z)$ é a função transferência do sistema e $A(z)$ representa o componente de sobreposição do espectro.

Para que a saída seja uma réplica fiel da entrada (reconstrução perfeita), é necessário que se verifiquem as seguintes condições [VAID87]:

$$A(z) = 0, \ \forall z \ \text{ e } \ T(z) = cz^{-D}, \tag{2.20}$$

onde c é uma constante e D é o atraso do sistema. Como forma de se eliminar o problema de sobreposição dos espectros, $A(z)$, os filtros de síntese devem satisfazer as seguintes relações:

$$G_0(z) = H_1(-z) \quad \text{e} \quad G_1(z) = -H_0(-z), \tag{2.21}$$

ou seja, os filtros de síntese passam a ser definidos a partir dos filtros de análise, o que reduz o problema inicial do projeto completo de um banco de filtros. Agora a procura se resume a um filtro passa-baixas e um passa-altas para os quais a função transferência,

$$T(z) = \frac{1}{2}\left[H_0(z)H_1(-z) - H_1(z)H_0(-z)\right], \tag{2.22}$$

satisfaça a condição imposta em (2.20).

Pode-se ainda reduzir a complexidade do problema admitindo que os filtros $H_0(z)$ e $H_1(z)$ podem se relacionar da seguinte forma:

$$H_1(\mathrm{z}) = H_0(-z), \tag{2.23}$$

o que no domínio discreto corresponde a considerar,

$$\mathrm{h}_1[\mathrm{n}] = (-1)^{\mathrm{n}}\,\mathrm{h}_0[\mathrm{n}]. \tag{2.24}$$

e assim, o problema de projeto do banco de filtros da Fig. 2.6 se reduz a achar um filtro passa-baixas, $H(z) = H_0(z)$, de resposta ao impulso infinita ou finita, para qual a equação

$$T(z)=\frac{1}{2}\left[H^2(z)-H^2(-z)\right], \tag{2.25}$$

satisfaça a condição (2.20).

Neste trabalho, interessa unicamente um banco de filtros de resposta finita ao impulso. Neste caso, é relativamente simples verificar que (2.25) nunca será um monômio e, portanto, não é possível satisfazer plenamente a condição (2.20). No entanto, podem-se construir soluções de reconstrução perfeita baseadas em filtros de fase linear. A família de filtros que possibilita esta solução será vista na próxima seção e é chamada de filtros espelho em quadratura, ou simplesmente QMF.

Para satisfazer as condições de reconstrução perfeita (2.20), pode-se seguir uma estratégia diferente. De forma a eliminar o componente de sobreposição de espectros, continua-se a considerar a opção (2.21). Mas, em contrapartida, admite-se uma nova forma de relacionar $H_0(z)$ e $H_1(z)$, do tipo

$$H_1(z)=-z^{-(K-1)}H_0(-z^{-1}), \tag{2.26}$$

ou no domínio discreto,

$$h_1[n]=(-1)^{n-1}h_0[K-n-1], \tag{2.27}$$

onde K é um número par que representa o comprimento do filtro. As equações (2.26) e (2.27) partem do pressuposto que se está interessado numa solução baseada em filtros de resposta ao impulso finita e causal. Nesse caso, a função de transferência do sistema é dada por

$$T(z)=\frac{1}{2}z^{-(K-1)}\left[H_0(z)H_0(z^{-1})+H_0(-z)H_0(-z^{-1})\right]. \tag{2.28}$$

A família de soluções obtida de acordo com (2.28) é denominada de filtros em quadratura conjugada ou CQF (*Conjugate Quadrature Filter*) [SMIT86] e [VAID93]. Analogamente, pode-se projetar um banco de filtros como o da Fig. 2.6, tendo um filtro passa-baixas, $H(z) = H_0(z)$, que satisfaça a equação:

$$H(z)H(z^{-1}) + H(-z)H(-z^{-1}) = 2\ . \tag{2.29}$$

Recorrendo às propriedades da transformada **Z** e, em particular, à transformada da autocorrelação [AKAN92] de uma seqüência dada por

$$r[n] = h[n] * h[-n] \xleftrightarrow{z} R(z) = H(z)H(z^{-1}), \tag{2.30}$$

pode-se reescrever (2.29) na forma

$$R(z) + R(-z) = 2, \tag{2.31}$$

onde $R(z)$ é uma função densidade espectral de potência complexa. Desta forma, uma das maneiras de se projetar banco de filtros ortogonais é achar a função $R(z)$, na qual é válida (2.31) [VETT95].

2.3.2 Banco de filtros QMF

O banco de filtros espelho em quadratura ou QMF (*Quadrature Mirror Filter*) é assim chamado pelo fato de as respostas em freqüências dos filtros passa-baixas e passa-altas serem simétricas em torno de $\pi/2$ (Fig. 2.7), ou seja, um quarto da freqüência de Nyquist. Os filtros QMF são capazes de fazer o cancelamento da sobreposição dos espectros devido à dizimação.

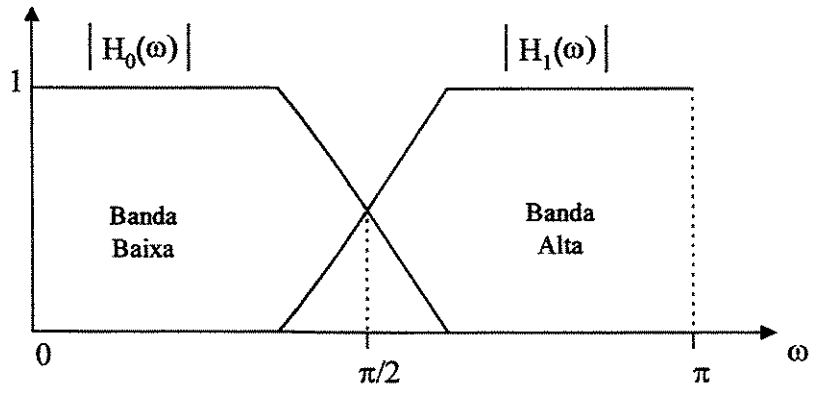


Figura 2.7 – Resposta em freqüência de filtros QMF.

A solução proposta na seção anterior por (2.23) define que, para se ter um banco de filtros como o da Fig. 2.7, basta ter um único filtro passa-baixas, $h_0[n]$. Em outras palavras, o problema de projeto é reduzido se o filtro passa-baixas escolhido tiver sua transformada de Fourier $H(\omega)$ que satisfaça a seguinte condição [SIMO91]:

$$\frac{1}{2}\left[H(\omega)H(-\omega)+H(-\omega+\pi)H(\omega+\pi)\right]=1 \tag{2.32}$$

ou

$$|H(\omega)|^2 + |H(\omega+\pi|^2 = 2\,. \tag{2.33}$$

O filtro passa-altas de análise e os filtros de síntese são formados simplesmente modulando $h_0[n]$ por $e^{j\pi n}$ ou, equivalentemente, por $(-1)^n$. Resumindo, o conjunto de filtros de análise/síntese, segundo [SMIT91], é definido como sendo:

$$G_0(e^{j\omega}) = H_0(e^{j\omega}) \tag{2.34}$$

$$\mathrm{H_1(e^{j\omega}) = H_0(e^{j(\omega+\pi)})} \tag{2.35}$$

$$G_1(e^{j\omega}) = -G_0(e^{j(\omega+\pi)})\,. \tag{2.36}$$

Por suas boas propriedades de reconstrução e facilidade de projeto, os conjuntos de filtros QMF têm sido largamente utilizado no contexto de decomposição em banco de filtros com reconstrução perfeita. Quando esses filtros são aplicados a uma estrutura como a da Fig. 2.6, eles apresentam um grau de distorção insignificante. Como o interesse, neste trabalho, são os sinais com duração finita, os filtros QMFs utilizados nas simulações serão filtros de resposta ao impulso finita (FIR – *Finite Impulse Response*) e de fase linear. Os filtros lineares utilizados neste trabalho foram retirados do estudo realizado por Villasenor [VILL95], que analisou o desempenho de diversos pares H_0 e H_1 com relação à eficiência na compressão de imagens. Utilizaram-se os filtros 1 e 4 do referido trabalho, denominados aqui *QMF(9x7)* e *QMF(5x3)*.

2.4 DECOMPOSIÇÃO LINEAR EM MULTI-RESOLUÇÃO DE IMAGENS

A decomposição em multi-resolução de imagens é um tema de investigação importante, quer no domínio teórico quer no domínio prático das aplicações. O conceito de resolução de um sinal está associado ao conteúdo ou riqueza espectral de representação. Nos sistemas de compressão de imagens baseadas em banco de filtros, a concentração de energia dos elementos da imagem num conjunto reduzido de coeficientes e a exploração das redundâncias espaciais são realizadas a partir da decomposição em bandas de freqüência, segundo uma topologia pré-definida.

No entanto, é sempre possível achar uma decomposição em sub-bandas segundo o objetivo que se quer. Geralmente, o objetivo é baseado em minimizar em conjunto a taxa de transmissão e a distorção. No caso deste trabalho, o intuito é analisar a capacidade dos métodos de decomposição em sub-bandas em compactar energia no domínio dos coeficientes e verificar suas imagens reconstruídas em termos de PSNR e qualidade visual objetiva. Estudos teóricos sobre a capacidade de concentração de energia dos métodos de decomposição em sub-bandas mostram que a decomposição em multi-resolução tem desempenho similar e uma menor complexidade computacional, quando comparada com a decomposição uniforme em N bandas [AKAN92]. Dentre todas as topologias possíveis de divisão em sub-bandas, neste trabalho será a utilizada a decomposição linear e não-linear [EGGE95a] de multi-resolução, baseada em banco de filtro com reconstrução perfeita.

As imagens são sinais bidimensionais, e os componentes do sistema de processamento, ou seja, os filtros, os dizimadores e os interpoladores, devem ter, em princípio, a mesma dimensão. Na literatura, pode-se encontrar um conjunto extenso de exemplos de banco de filtros não-separáveis [LIM89].

Neste trabalho, optar-se-á por uma configuração de decomposição do tipo separável, isto é, baseada em banco de filtros de duas sub-bandas uniformes para sinais unidimensionais. Esta opção justifica-se pela reduzida complexidade computacional do método e a enorme variedade de banco de filtros disponíveis.

Na Fig. 2.8, apresenta-se a decomposição de uma imagem em quatro bandas uniformes, usando um banco de filtros separáveis. Em primeiro lugar, as linhas da imagem são filtradas com os filtros passa-baixas e passa-altas e dizimadas por um fator dois, dando origem a duas subimagens intermediárias. Em seguida, as colunas das imagens intermediárias são filtradas por filtros passa-baixas e passa-altas e dizimadas por um fator dois, dando origem a quatro subimagens finais ou sub-bandas. Obviamente é possível trocar a ordem pela qual são processadas as linhas e as colunas da imagem.

Essa decomposição gera sub-bandas com diferentes níveis de resolução. A decomposição segue sobre a sub-banda de aproximação gerada em baixas freqüências. A quantidade de sub-bandas geradas é $3N+1$, onde N é o número de estágios.

Se for repetida de forma sistemática a decomposição sobre a sub-banda de aproximação, obtém-se uma decomposição de imagem do tipo multi-resolução ou hierárquica. Na Fig. 2.9, apresentam-se exemplos da divisão do espectro espacial de freqüência, segundo a decomposição multi-resolução.

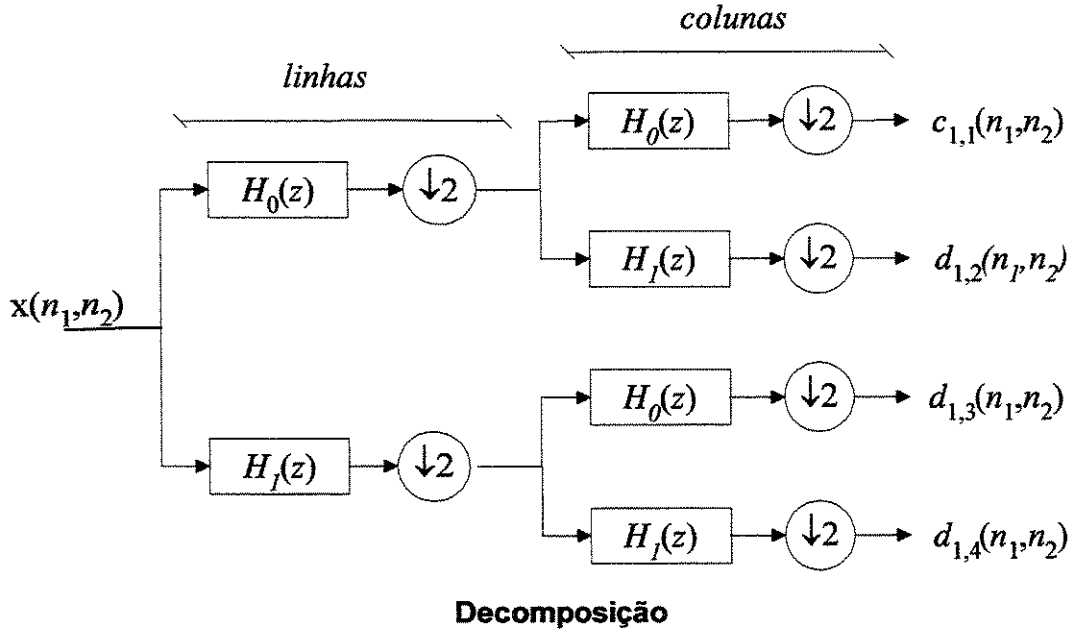


Figura 2.8 – Decomposição linear em multi-resolução de sub-bandas, por um banco de filtros separáveis. Esta decomposição está limitada a 1 nível de profundidade. As bandas estão referenciadas segundo a seguinte nomenclatura: $c_{1,1}$ como sendo a sub-banda de aproximação (baixas freqüências) e $d_{1,2}$, $d_{1,3}$ e $d_{1,4}$ como sendo as sub-bandas de detalhamento (altas freqüências).

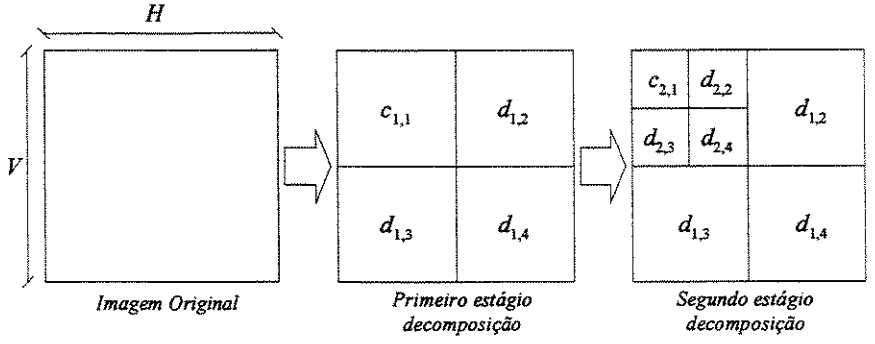


Figura 2.9 – Decomposição de multi-resolução de dois estágios.

O processo de reconstrução da imagem a partir da informação das sub-bandas segue o caminho inverso da decomposição mostrada na Fig. 2.8. As bandas $c_{1,1}$, $d_{1,2}$, $d_{1,3}$ e $d_{1,4}$ são interpoladas, filtradas e combinadas para formar e dar origem ao sinal $y[n_1, n_2]$. Consegue-se uma representação eficiente, principalmente pela quantização apropriada das $3N$ sub-bandas de altas freqüências (ou detalhamento), $d_{i,j}$, $1 \leq i \leq N$, $2 \leq j \leq 4$, visto que estas sub-bandas retêm a maior parte dos coeficientes da decomposição. Os coeficientes presentes nestas sub-bandas

tendem a se distribuir em torno de zero, seguindo uma distribuição gaussiana generalizada [ANTO91].

A Fig. 2.10 mostra a decomposição multi-resolução da imagem *Lena* para três estágios. A sub-banda de aproximação $c_{3,1}$ da Fig. 2.10, é uma versão suavizada e dizimada da imagem original e exibe uma alta correlação entre os coeficientes, além de concentrar a maior energia (variância) entre todas as sub-bandas. No entanto, ela possui 4^3 menos amostras do que a imagem original e não acarreta perda de eficiência na compressão. Para casos típicos de níveis de profundidade, ou estágios de decomposição, de 3 ou 4, esta sub-banda representa 1,6% e 0,4%, respectivamente, do total dos coeficientes a serem codificados. Devido ao seu tamanho, nenhuma técnica de compressão pode ser aplicada a ela, tal como a DPCM [JAYA84], por exemplo.

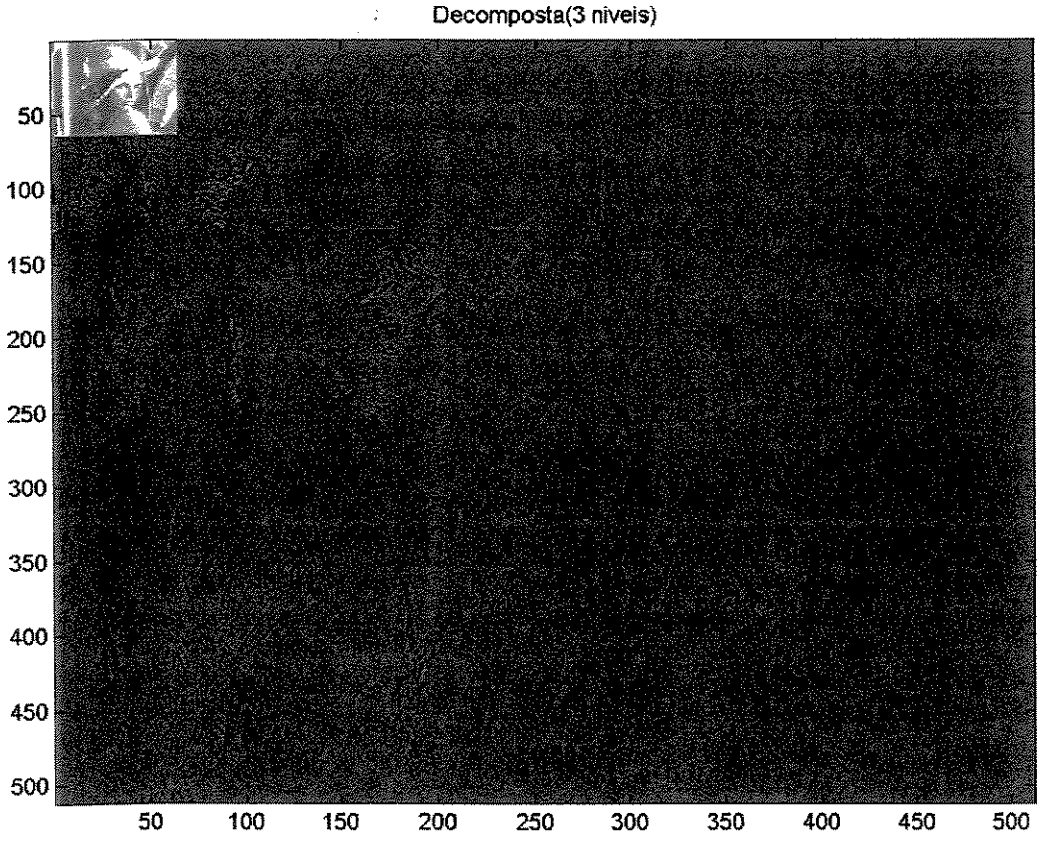


Figura 2.10 – Decomposição multi-resolução em 3 níveis de profundidade da imagem *Lena(512X512)*.

Conforme ilustra a Fig. 2.10, as sub-bandas de detalhamento $d_{i,j}$, $1 \leq i \leq N$, $2 \leq j \leq 4$, apresentam regiões de coeficientes aproximadamente nulos correspondentes às regiões de

baixa atividade espacial na imagem original. Estas regiões podem ser eficientemente codificadas através de técnicas de varredura e de codificação de entropia. As regiões de textura ou de bordas da imagem original causam o surgimento de coeficientes de alta magnitude nas regiões correspondentes das sub-bandas $d_{i,j}$. Esses coeficientes possuem alta correlação com os coeficientes da sub-banda de mesma orientação em níveis diferentes.

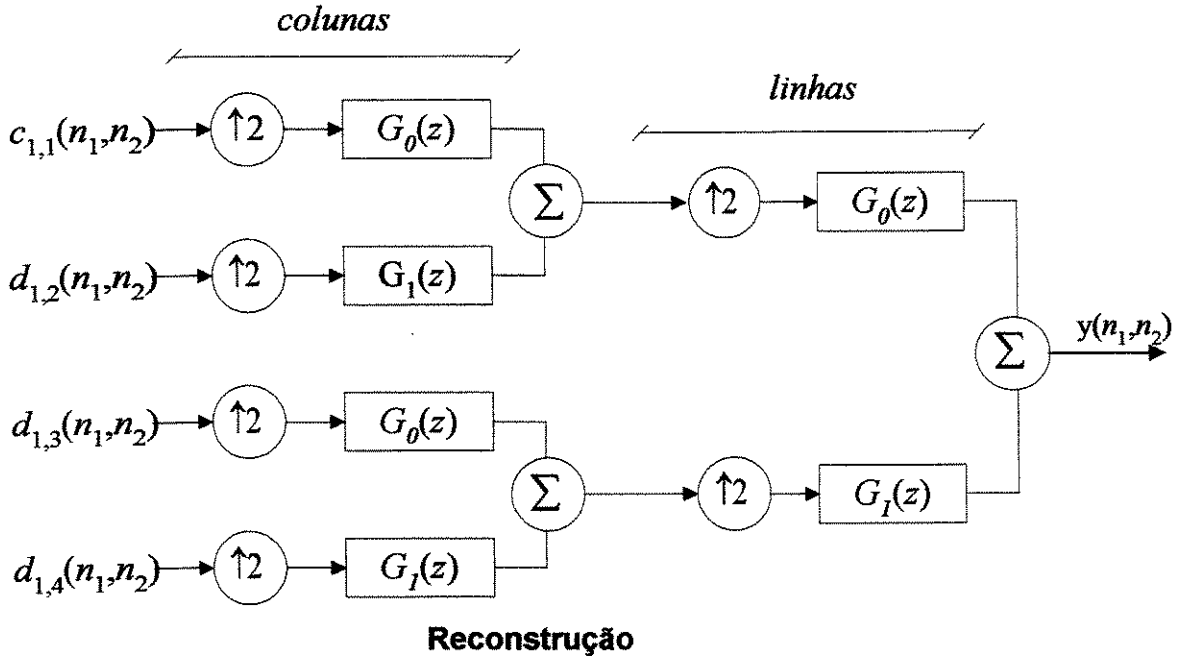


Figura 2.11 – Reconstrução da imagem referente ao último estágio da síntese bidimensional, que correspondente ao primeiro estágio de decomposição em multi-resolução .

O processo de reconstrução da imagem a partir das sub-bandas segue o caminho inverso da decomposição. As sub-bandas $c_{1,1}$, $d_{1,2}$, $d_{1,3}$ e $d_{1,4}$ são interpoladas, filtradas e combinadas de forma a dar origem ao sinal $y(n_1, n_2)$. Na Fig. 2.11, apresenta-se um sistema de decomposição correspondente a um nível de decomposição.

2.5 DECOMPOSIÇÃO NÃO-LINEAR EM SUB-BANDAS

Até o momento, apresentou-se a decomposição em sub-bandas que utiliza um banco de filtros lineares. Porém, existe a possibilidade de se ter decomposição em sub-bandas utilizando-se de filtros não-lineares [EGGE95a]. No início deste trabalho, foi mencionado que, na compressão de imagens por transformada, existiam dois tipos básicos de distorções: o

efeito de bloco e o efeito de *ringing*. No caso da decomposição em sub-bandas, o primeiro tipo de distorção não ocorre, porém o segundo tipo de distorção, o *ringing*, é impossível de se eliminar na decomposição linear em sub-bandas. No entanto, existe a possibilidade de minimizá-lo através de um projeto apropriado de banco de filtros [EGGE95b]. Para evitar estas duas distorções, foi proposta por Egger uma decomposição morfológica em sub-bandas [EGGE95a]. Este padrão de decomposição morfológica é a análise de multi-resoluções descritas por [ZHOU92] e por [SALE92]. Uma outra aproximação é a proposta por [TOET89], que é baseada no mesmo princípio da análise de multi-resolução, com a diferença de introduzir um fator dizimador após os filtros morfológicos.

Nesta seção, será apresentada a decomposição não-linear em sub-bandas, usando amostragem crítica através de uma banco de filtros não-lineares.

2.5.1 Decomposição Morfológica com Reconstrução Perfeita

A maior vantagem, em relação à decomposição linear, da proposta de representação da imagem em sub-bandas através de um banco de filtros morfológicos, é a supressão do efeito de *ringing* e a compactação de energia no domínio dos coeficientes. No entanto, imagens nas quais se tenha textura são melhor representadas por bancos de filtros lineares. Para evitar este problema, [EGGE95a] propõe uma decomposição adaptativa por meio da qual as regiões de textura são processadas por filtros lineares e as outras por filtros morfológicos (não-lineares).

A apresentação abaixo segue de perto o desenvolvimento de Egger [EGGE95a], que propõe uma decomposição em sub-bandas usando filtros morfológicos. Para se ter uma ferramenta que permita uma generalização de um banco linear para um não-linear, utiliza-se o conceito de filtro *half-band*. A definição é a seguinte:

Definição 1: Um filtro linear é chamado *half-band* se todas as amostras ímpares da resposta impulsiva f(k) são zero, exceto a amostra f(δ), isto é:

$$f(2k+1)=\begin{cases}1/2 & 2k+1=\delta \\ 0 & \text{caso contrário}\end{cases} \tag{2.37}$$

onde $k = 0, 1, 2, ..., \frac{L-1}{2}$ e L é o comprimento do filtro f(k).

Na Fig. 2.12 é ilustrado um exemplo de um filtro *half-band*, o qual supõe-se que δ é igual a cinco.

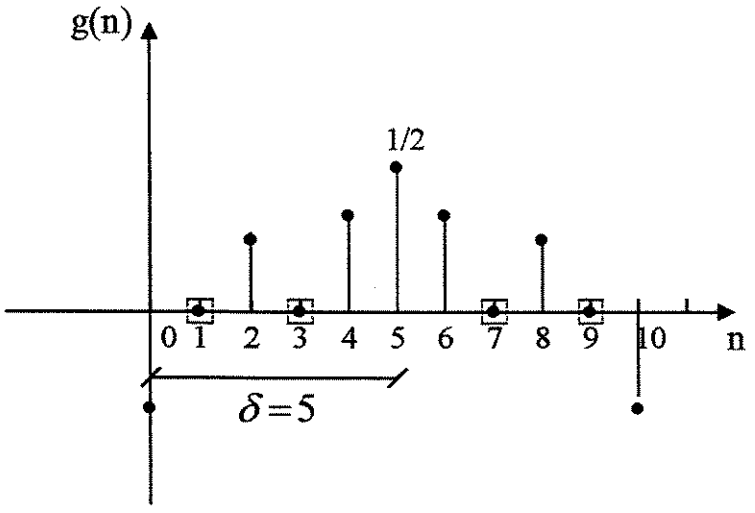


Figura 2.12 – Exemplo de um filtro *half-band*. Onde δ é igual a cinco, as amostras pares tem valores arbitrários e as amostras ímpares são nulas

Considere-se agora que um sinal x(n) é dizimado e interpolado por um fator 2, para se obter y(n), que é então filtrado por um filtro *half-band* com resposta impulsiva f(n), tendo como saída u(n), como mostra a Fig. 2.13.

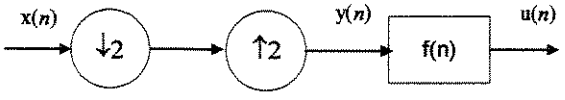


Figura 2.13 – Um sinal é dizimado e interpolado por um fator 2 e depois filtrado por um filtro *half-band*.

Nesta figura, definiu-se u(n) como a convolução de y(n) com f(n); então, toda amostra alternada de u(n) é dada por;

$$u(2n) = \frac{1}{2} x(2n - \delta), \tag{2.38}$$

isto é, a entrada é apenas atrasada de δ e atenuada por um fator de 1/2 para estas amostras [EGGE95a]. Note-se que (2.38) implica em (2.37). Desta forma pode-se generalizar o conceito de um filtro *half-band* para o caso não-linear.

Definição 2: Um filtro não-linear é chamado de filtro *half-band* se toda a amostra alternada do sinal y(k) é igual à amostra correspondente do sinal original, exceto por uma constante multiplicativa. Seja Fy(k) a saída de um filtro não-linear *half-band*. Assim

$$Fy(2k) = c.x(2k - \delta), \tag{2.39}$$

onde c é uma constante arbitrária.

Obtém-se a reconstrução perfeita com a estrutura de bancos de filtros da Fig. 2.6 se a convolução de $h_0(k)$ e $g_0(k)$ for um filtro *half-band* e os filtros respeitarem as condições vistas anteriormente em (2.21). Agora, supondo-se que $H_0(z) = 1$, a reconstrução perfeita é atingida se $G_1(z) = -1$, $G_0(z)$ é um filtro *half-band* e $H_1(z) = G_0(-z)$. Este banco de filtros é mostrado na Fig. 2.14.

Considere-se a substituição do filtro $G_0(z)$ por um filtro generalizado de *half-band* denotado por F(•). Não se pode descrever este filtro por sua transformada **Z**, visto que se trata de um filtro não-linear.

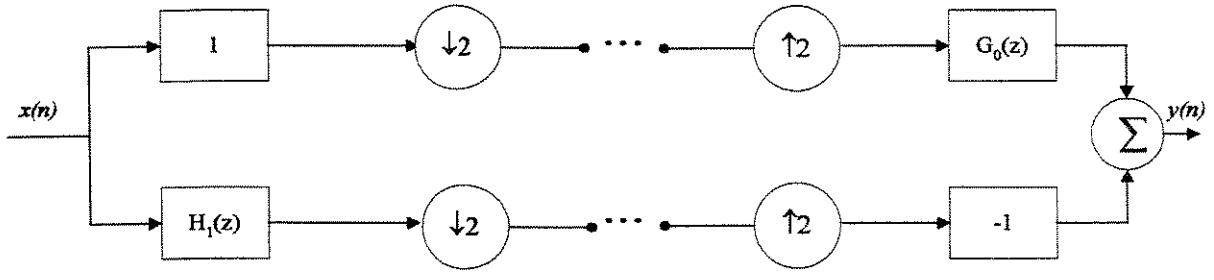


Figura 2.14 – Estrutura de Banco de Filtros com Reconstrução Perfeita se $G_0(z)$ é um filtro *half-band* e $H_1(z) = G_0(-z)$.

No entanto, a condição de biortogonalidade estabelece que $H_1(z) = G_0(-z)$ para reconstrução perfeita. Dessa forma deseja-se achar um filtro passa-altas de análise correspondente para o caso não-linear. Segundo [EGGE95A], este filtro pode ser obtido da seguinte forma: Negando-se a variável z, tem-se o seguinte efeito na resposta impulsiva:

$$Z^{-1}\{G(-z)\}=(-1)^k g(k). \tag{2.40}$$

Como toda amostra ímpar é zero, negando-se a variável z somente se produzirá efeito na amostra central. Sendo assim, para um filtro de fase-zero tem-se

$$G(-z)=1-G(z). \tag{2.41}$$

Esta situação pode ser generalizada para o caso não-linear. Denotando por A(•) o filtro não-linear de análise passa-altas, tem-se

$$A = I - F, \tag{2.42}$$

onde *I* é o operador identidade. Seguindo a mesma estrutura, tem-se um banco de filtros não-lineares que provê reconstrução perfeita como mostrado na Fig. 2.15.

Egger [EGGE95a] mostra que a decomposição não-linear com reconstrução perfeita pode ser realizada por bancos de filtros morfológicos. Os filtros estudados no trabalho de Egger foram o filtro de dilatação, o filtro de fechamento (*closing*) e filtro de mediana.

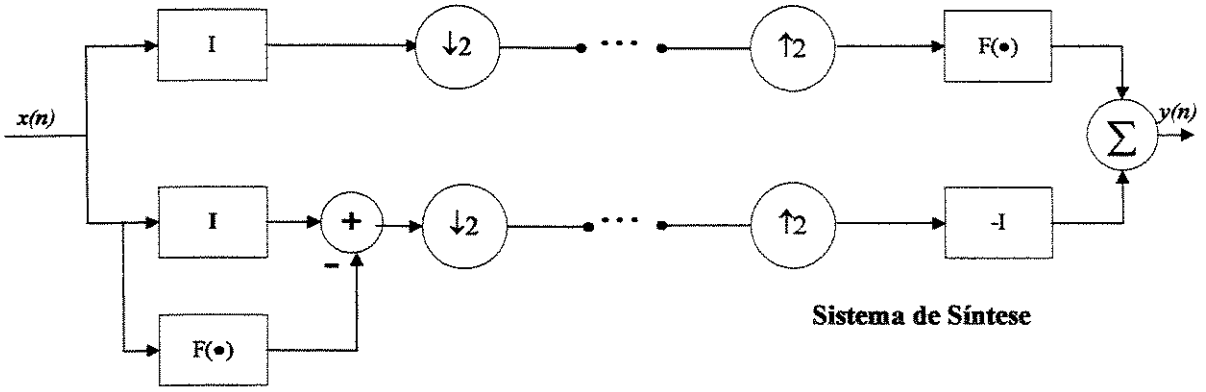


Figura 2.15 – Banco de filtros não-lineares com reconstrução perfeita.

Segundo o estudo, o filtro de mediana obteve desempenho superior aos outros dois. Assim, este filtro, apresentado a seguir, foi utilizado para a decomposição não-linear neste trabalho.

2.5.2 Filtros de mediana

A filtragem mediana é uma técnica de processamento de sinal não-linear bastante útil na supressão de ruídos em imagens e também na preservação das bordas. A filtragem mediana calcula valores de pixels da imagem de saída com base na vizinhança dos correspondentes pixels da imagem de entrada. Normalmente, a filtragem mediana, para o caso bidimensional, consiste de uma janela (mxn) que varre a imagem com um número ímpar de pixels. A Fig. 2.16 ilustra uma forma de filtragem do filtro de mediana utilizando uma janela, ou região de suporte, 3x3.

No exemplo da Fig. 2.16, se o pixel de valor 60 fosse um pico de ruído branco na imagem de entrada, ele seria eliminado após a filtragem. No caso de se usar um número par de pixels na janela, ou na região de suporte do filtro, o valor da mediana será a média aritmética dos dois valores medianos da seqüência de pixels ordenados.

Neste trabalho, são feitos testes com três regiões de suporte para os filtros utilizados neste tipo de decomposição. Devido à estrutura de decomposição não-linear ter uma particularidade especial, que será descrita na próxima seção, no que se refere à decomposição de sinais bidimensionais, os filtros usados têm as regiões de suporte apresentadas na Fig. 2.17. Dos três filtros apresentados, o MED 2 tem uma particularidade especial. Como se tem apenas duas amostras, o valor do píxel reconstruído é a média entre as duas amostras. Isso faz com que este filtro possua uma característica linear, apesar de ele estar sendo utilizado na mesma estrutura de decomposição na qual os outros dois filtros não-lineares são colocados. Os filtros MED 4 e MED 6 foram propostos por Egger [EGGE95a], o filtro MED 2 é o filtro linear que será utilizado neste trabalho. É interessante ressaltar que o filtro MED 2 é uma degeneração dos filtros MED 4 e MED6. E ele será chamado de MED 2 apenas por uma questão de nomenclatura, sendo que ele não é definido como filtro mediano, mas sim como um filtro de média.

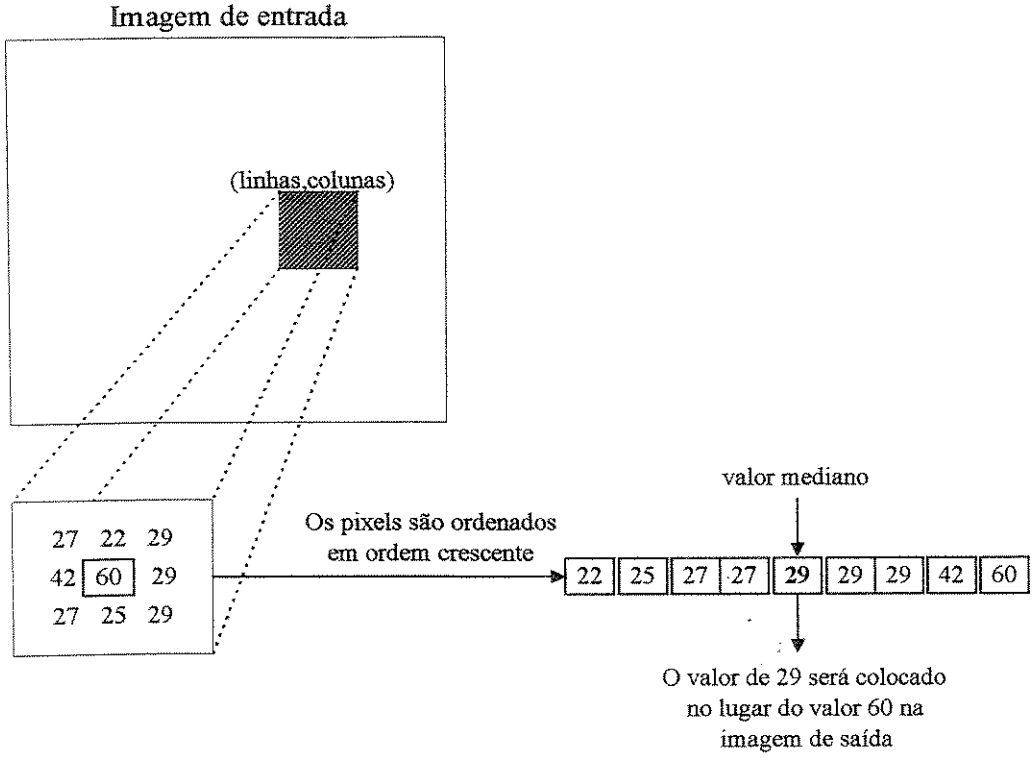


Figura 2.16 – Exemplo ilustrativo de uma filtragem mediana (3x3) sobre o pixel desejado.

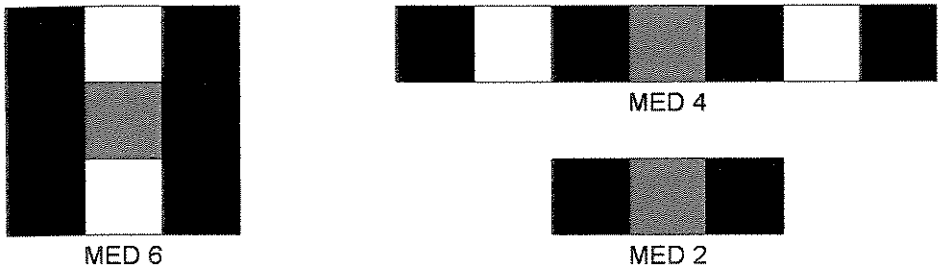


Figura 2.17 – As amostras em preto são as entradas do filtro, as amostras em branco são zeros e a amostra em cinza é aquela a ser reconstruída.

2.5.3 Decomposição/Reconstrução não-linear para imagens

Na estrutura de decomposição apresentada na Fig. 2.15, apresenta-se a decomposição com reconstrução perfeita para um sinal unidimensional. Agora, deseja-se ter uma decomposição que seja compatível com um sinal bidimensional. Para isso, será apresentada a estrutura de decomposição em sub-bandas utilizada nas simulações. O intuito é decompor esta

estrutura não-linear como a decomposição utilizada na Fig. 2.8. Assim, a estrutura de decomposição da Fig. 2.15 leva a uma estrutura como ilustra a Fig. 2.18.

Para a estrutura de decomposição não-linear em sub-bandas, a sub-banda de aproximação é apenas uma versão dizimada da imagem original. Por exemplo, para o primeiro estágio, a sub-banda de aproximação é obtida dizimando-se primeiro as colunas e depois as linhas. Com isso, na sub-banda de aproximação, os coeficientes serão os mesmos coeficientes da imagem original que não sofreram dizimação. A sub-banda de detalhamento $d_{1,3}$ sofre uma dizimação nas colunas e uma filtragem nas linhas. A sub-banda de detalhamento $d_{1,2}$ sofre uma filtragem nas colunas e uma dizimação nas linhas. Por fim, a sub-banda de detalhamento $d_{1,4}$ corresponde a filtrar nas colunas e depois nas linhas, ou seja, esta sub-banda sofre duas etapas de filtragem para o primeiro estágio.

É importante notar que os filtros da Fig. 2.17 têm regiões de suporte que filtram apenas nas colunas. Após a primeira dizimação na parte de baixa e a filtragem na parte de alta, faz-se a transposta da imagem resultante para serem dizimadas ou filtradas as linhas. Uma outra característica importante deste tipo de decomposição é que a dizimação na decomposição é feita nas colunas pares da parte de baixa e nas colunas ímpares da parte de alta. Este procedimento poder ser alternado, mas nunca dizimando-se as mesmas colunas na parte de baixa e de alta. Isto é uma condição para a reconstrução perfeita do esquema apresentado na Fig. 2.18.

Na síntese, verifica-se que apenas as amostras que foram dizimadas na análise serão interpoladas com zeros. Na parte de baixa da síntese, apenas as amostras que foram interpoladas serão filtradas, ou seja, os coeficientes que não sofreram dizimação não serão filtrados, pois correspondem às próprias amostras da imagem original.

As sub-bandas estão referenciadas segundo a nomenclatura utilizada na decomposição linear da Fig. 2.8, ou seja, a estrutura apresentada na Fig. 2.18 representa o primeiro estágio de decomposição de uma imagem.

Nesta estrutura de decomposição, o bloco representado por $\boxed{I - F(\bullet)}$ corresponde à estrutura apresentada na Fig. 2.19.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

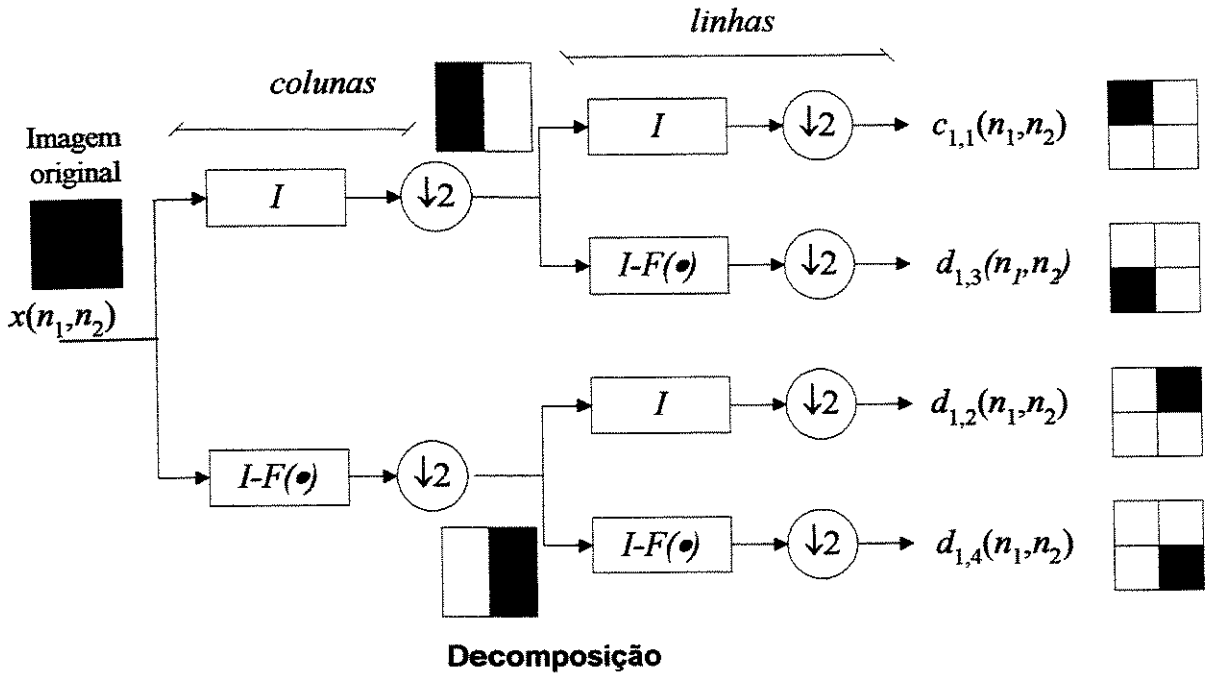


Figura 2.18 – Decomposição não-linear de multi-resolução de sub-bandas por um banco de filtros não-lineares. Esta decomposição está limitada a 1 nível de profundidade.

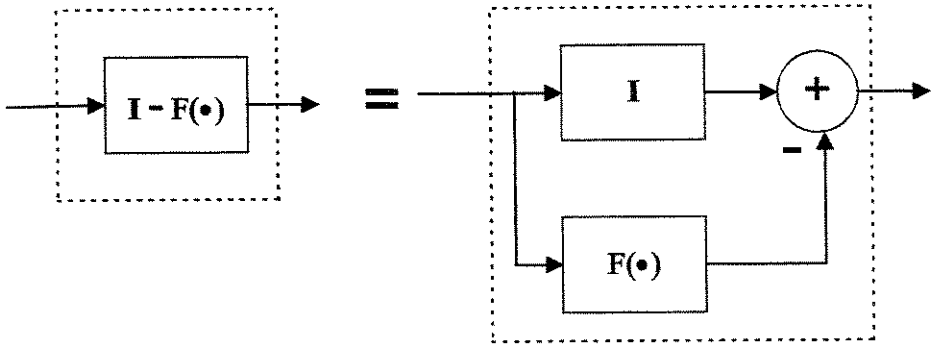


Figura 2.19 – Bloco correspondente à diferença entre a imagem de entrada e a mesma filtrada pelo filtro mediano.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

A estrutura de reconstrução da imagem segue a mesma estrutura do caso linear e é ilustrada pela Fig. 2.20.

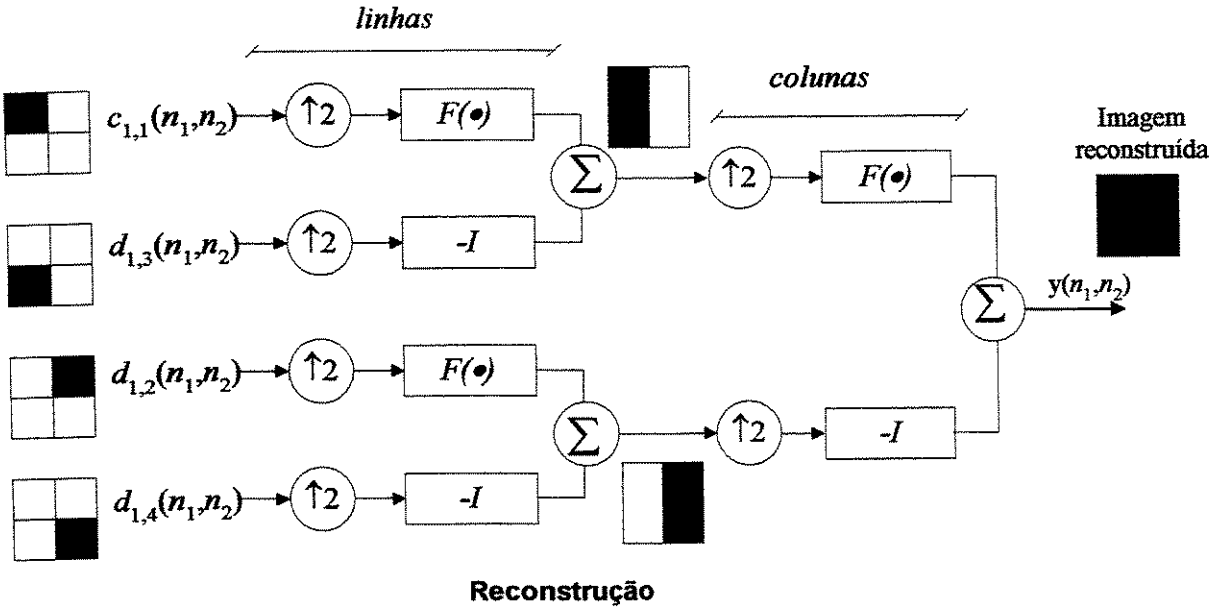


Figura 2.20 – Reconstrução da imagem referente ao último estágio da síntese bidimensional, que correspondente ao primeiro estágio de decomposição não-linear de multi-resolução.

CAPÍTULO 3

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

COMPACTAÇÃO DE ENERGIA E COMPRESSÃO DE IMAGENS USANDO BANCO DE FILTROS

Este capítulo é dedicado à apresentação de uma análise feita sobre os coeficientes das sub-bandas, a qual diz respeito à compactação de energia no domínio dos coeficientes resultantes da transformada. É investigada esta propriedade através de uma medida definida como Ganho de Compactação (GC). Segundo essa medida, a eficiência da compactação de energia de diversos bancos de filtros lineares e não-lineares é estudada para várias imagens de teste.

É apresentado um esquema básico de compressão de imagens utilizando-se a decomposição de multi-resoluções na aplicação de compressão de imagens. São também apresentados os quantizadores utilizados nas simulações, bem como o algoritmo de alocação de bits com ponderação nas sub-bandas.

3.1 INTRODUÇÃO

O enorme volume de informação gerado por uma seqüência de imagens, ou mesmo por uma imagem fixa, pressupõe a necessidade da utilização de métodos de compressão eficientes, de

forma a tornar possível o desenvolvimento de aplicações nos domínios da transmissão e do armazenamento digital da imagem. Ao longo das últimas décadas, um extenso e variado leque de técnicas de compressão têm sido estudados, porém, de forma geral, as soluções desenvolvidas são sempre a redução das dependências estatísticas entre os elementos da imagem e a compactação de energia em um reduzido número de coeficientes.

Os métodos de compressão mais utilizados são baseados no uso das transformadas lineares unitárias, entre as quais a transformada discreta do coseno (DCT) e a decomposição em sub-bandas ocupam um lugar de destaque.

Este trabalho objetiva a aplicação e o desenvolvimento de técnicas de compactação de energia e compressão de imagens a partir da decomposição de sinais em sub-bandas e, em particular, da decomposição em multi-resoluções com banco de filtros lineares e não-lineares.

3.2 CRITÉRIOS PARA A SELEÇÃO DOS BANCOS DE FILTROS

Num sistema de compressão de imagens baseado em bancos de filtros, é natural esperar que o êxito global do sistema dependa, de alguma forma, das características do banco de filtros.

Nesta seção, são examinadas as características dos bancos de filtros com capacidade de reconstrução perfeita. As características que mais influenciam no desempenho de um sistema de compressão de imagem são fase linear, tamanho dos filtros e ganho de compressão.

3.2.1 Fase linear generalizada

Na área do processamento digital de imagem e, em particular, no domínio da compressão de imagens fixas, são utilizados normalmente filtros digitais de fase linear, com a justificativa de que estes minimizam os efeitos visuais indesejáveis, em torno de transições abruptas ou contornos das imagens [LIM89].

Em sistemas de compressão baseados em bancos de filtros com capacidade de reconstrução perfeita, este argumento só tem sentido quando se quantizam grosseiramente várias sub-bandas, de forma a obter elevadas taxas de compressão. Quando as taxas de compressão são baixas, o processo de quantização introduz pequenas perdas ou distorções. Neste caso, é difícil

distinguir os méritos de um banco de filtros com fase linear em relação a um banco de filtros de fase não-linear [QUEI92].

No Capítulo 2, foi apresentado os bancos de filtros com capacidade de reconstrução perfeita. A ortogonalidade e a fase linear são duas características incompatíveis, exceto no caso do banco de Haar, descrito em [VETT95]. No caso dos banco de filtros biortogonais, abdica-se da ortogonalidade para se obter soluções de fase linear.

Na seleção de bancos de filtros biortogonais, tendo em vista a sua aplicação na compressão de imagens, devem-se procurar filtros de soluções regulares com critérios mais gerais, como o comprimento, a resposta ao impulso e ao degrau e a ortogonalidade do filtro. Na literatura, existem várias formas de se selecionar bancos de filtros biortogonais como as descritas por [VILL95] e [EGGE95b].

3.2.2 Comprimento dos Filtros

Como já foi mencionado, neste trabalho, tem-se interesse apenas pelos filtros de resposta ao impulso finita. Geralmente os filtros com respostas longas correspondem a soluções com melhores características, quer no nível da sua resposta em freqüência, quer em relação a outras funções importantes para cada aplicação. Por exemplo, em sistemas de compressão de sinais de voz, são utilizados filtros com 32 coeficientes não nulos, devido à necessidade de uma grande seletividade entre os canais [VETT95].

No âmbito da compressão de imagens, deve-se recorrer a bancos de filtros com comprimentos mais modestos, de forma a que se possam construir sistemas com capacidade de processamento em tempo real e a uma menor exigência em termos da resposta em freqüência dos filtros.

Quando são usados bancos de filtros biortogonais assimétricos [EGGE95b], isto é, formados por dois filtros com comprimentos desiguais, a primeira dificuldade que se apresenta é saber qual deles se deve utilizar na síntese. Uma aproximação a este problema pode ser realizada em função dos respectivos comprimentos.

Geralmente, num sistema de compressão, as bandas de alta freqüência são grosseiramente quantizadas ou mesmo eliminadas, porque a qualidade dos sinais reconstruídos depende, em primeiro lugar, das características do filtro passa-baixas. Pelo que foi exposto, deve-se escolher

para o filtro passa-baixas da síntese aquele com o menor número de coeficientes não nulos. Esta opção oferece a vantagem de que o filtro passa-altas de análise passa a ter automaticamente o mesmo número de coeficientes. Um filtro de análise passa-altas de menor comprimento vai gerar um maior número de amostras nulas e, portanto, vai reduzir a variância e a entropia das respectivas bandas [EGGE95b].

3.2.3 Ganho de compressão

O objetivo fundamental de qualquer transformação linear de sinais é permitir transferir a sua análise e processamento para um espaço de representação mais adequado e eficiente, do ponto de vista da aplicação que se quer. No domínio da compressão de sinais, significa a possibilidade de se concentrar a energia dos sinais num conjunto reduzido de amostras ou coeficientes não correlacionados.

No caso particular dos métodos baseados na decomposição de sinais em bandas de freqüências ou simplesmente sub-bandas, visa-se saber quais os benefícios que podem ser obtidos em termos da compressibilidade dos dados em relação à representação original em código PCM. A capacidade de compressão dos métodos baseados em transformadas lineares é normalmente medida a partir do chamado ganho de compressão [JAYA84].

Para um sistema ortogonal genérico de decomposição em M bandas não uniformes e para um modelo de quantização de grande qualidade, ou seja, pequenas perdas e erro de quantização do tipo ruído branco de baixa energia, Pearlman [PEAR91] definiu o ganho de compressão teórico máximo, GCT, dado pela expressão

$$GCT = \frac{\sigma_x^2}{\prod_{i=0}^{M-1} (M_i \sigma_{x_i}^2)^{1/M_i}}, \tag{3.1}$$

com $\sum_{i=0}^{M-1} 1/M_i = 1$. Isto é, o ganho de compressão teórico máximo é dado pela relação entre a variância do sinal de entrada e a média geométrica ponderada das variâncias das M bandas. O limite inferior da expressão (3.1) é 1, no caso do sinal de entrada ser um ruído branco. Para sinais correlacionados, o ganho aumenta com o número de sub-bandas, mas satura rapidamente, pois, ao

passo que os estágios de decomposição vão aumentando, a parcela de contribuição das sub-bandas fica cada vez menor. Conseqüentemente, o ganho de compressão tende a se estabilizar com o aumento dos estágios.

A expressão (3.1) não contempla duas situações importantes: a utilização de bancos de filtros não ortogonais e a decomposição de sinais discretos no tempo de duração finita. Esta abordagem é feita na próxima seção.

3.3 ANÁLISE DO GANHO DO COMPACTAÇÃO DE ENERGIA

A propriedade de preservação de energia é muito usada pelas transformadas ortogonais na compressão de imagens, pois o erro de reconstrução pode ser calculado pela somatória dos erros das sub-bandas. Esta propriedade de preservação de energia faz com que as transformadas ortogonais sejam bastante atrativas. No entanto, para o caso de bancos de filtros de multi-resoluções, a ortogonalidade é incompatível com a linearidade de fase, que é bastante relevante para sistemas de compressão de imagem. O uso de bancos de filtros biortogonais na codificação de imagens é então requerido, porque evita distorção de fase nas sub-bandas e também porque é interessante para a implementação do algoritmo EZW (*Embedded Zerotree Wavelet*), desenvolvido por Shapiro [SHAP93], que será visto adiante.

A decomposição em sub-bandas tem uma grande vantagem em relação aos outros esquemas de codificação de imagens, no que se refere à propriedade de compactação de energia. Katto e Yasuda [KATT91] definem uma medida de ganho G_{SBC} (*Gain Subband Code*) que não depende da ortogonalidade dos filtros e pode ser aplicada em todas as transformadas lineares, incluindo DPCM e as transformadas ortonormais. A formulação desta medida encontra-se na próxima seção.

3.3.1 Ganho de Codificação para Sub-bandas (G_{SBC})

Katto e Yassuda definiram alguns parâmetros para se calcular o ganho de codificação para sub-bandas. A formulação apresentada em [KATT91] está relacionada para um sinal unidimensional, mas pode ser estendida para o caso bidimensional. Sejam os parâmetros A_i e B_i, dados pelas seguintes equações:

$$\sigma_i^2 = A_i.\sigma_x^2 \tag{3.2}$$

e

$$\sigma_r^2 = \sum_{i=1}^{M} B_i.\sigma_{q,i}^2 \,, \tag{3.3}$$

onde σ_i^2 é a variância da sub-banda *i-ésima*, σ_x^2 é a variância da imagem de entrada, σ_r^2 é a variância do erro da imagem reconstruída, $\sigma_{q,i}^2$ é a variância do erro de quantização da sub-banda *i-ésima* e M é o número de sub-bandas. Também se define α_i como o fator inverso de sub-amostragem para as sub-bandas *i-ésima*, dado por

$$\sum_{i=1}^{M} \alpha_i = 1. \tag{3.4}$$

Sob a suposição de erro de quantização descorrelatado, B_i é definido como

$$B_i = \alpha_i \sum_{n=0}^{L_i-1} g_i^2[n] \,, \tag{3.5}$$

onde L_i é o número de coeficientes do filtro de síntese e $g_i[n]$, a resposta impulsiva do filtro de síntese . A fórmula geral do G_{SBC} definida em [KATT91] é dada por

$$G_{SBC} = \frac{1}{\prod_{i=1}^{M} \left(\frac{A_i.B_i}{\alpha_i} \right)^{\alpha_i}}, \tag{3.6}$$

a qual pode ser expressa como

$$G_{SBC} = \frac{\sigma_x^2}{\prod_{i=1}^{M} \left(\sigma_i^2 \sum_{n=0}^{L_i-1} g_i^2[n] \right)^{\alpha_i}}. \tag{3.7}$$

3.3.2 Estimativa do Ganho de Compactação de Energia

Um dos objetivos deste trabalho é analisar a capacidade de compactação de energia obtida pela decomposição em sub-bandas, através de uma medida da distribuição das variâncias correspondentes às várias sub-bandas geradas. A análise agora será feita para um caso bidimensional, onde i representa o estágio de decomposição ou a profundidade da decomposição e j a posição da sub-bandas. A motivação deste estudo é encontrada em [JAYA84] e se deve ao fato de que o erro de quantização para cada sub-banda é proporcional à sua variância, isto é,

$$\sigma_{q,i,j}^2 = \varepsilon^2 2^{-2R(i,j)} \sigma_{i,j}^2, \tag{3.8}$$

onde $\sigma^2_{q,i,j}$ é a variância do erro de quantização para a sub-banda $d_{i,j}$ (ou $c_{i,j}$) que estão ilustradas na Fig.2.9. Isto é, a sub-banda na posição j ($j \in \{1,2,3,4\}$) e profundidade i ($i \in \{1, \ldots, N\}$), $\sigma^2_{i,j}$ é a variância da mesma sub-banda, $R(i,j)$ é o número de bits por amostra alocados para a quantização desta sub-banda e ε é a constante de desempenho do quantizador.

Para estimar esta capacidade de compactação, utiliza-se, neste trabalho, uma medida empírica. Esta medida, denominada de Ganho de Compactação para o N-ésimo estágio de decomposição (GC_N), é definida pela razão entre a média aritmética ponderada e a média geométrica ponderada das variâncias normalizadas das sub-bandas geradas [BERG00], ou seja,

$$GC_N = \frac{MA_{\alpha_{i,j}}\left(\dfrac{\sigma_{i,j}^2 B_{i,j}^{'}}{\alpha_{i,j}}\right)}{MG_{\alpha_{i,j}}\left(\dfrac{\sigma_{i,j}^2 B_{i,j}^{'}}{\alpha_{i,j}}\right)}, \tag{3.9}$$

onde $\sigma_{i,j}^2$ denota a variância da sub-banda i,j, $B^{'}_{i,j}$ é o fator de normalização dessas variâncias e $\alpha_{i,j}$ é o fator inverso de sub-amostragem, que relaciona o número de coeficientes da banda i,j e o número total de coeficientes. Observa-se que $\alpha_{i,j} = 1/4^i$, $i = 1, \ldots, N$, $j = 1, \ldots, 4$. (O índice j varia de 2 a 4 para valores de i entre 1 e N-1. Quando $i = N$, j assume valores de 1 a 4.) Os fatores $\alpha_{i,j}$ caracterizam uma distribuição de probabilidade, pois são não-negativos e têm soma igual a um, como ilustra a equação (3.4). É em relação a esta distribuição que as médias aritmética (MA) e geométrica (MG) de (3.9) são calculadas. Portanto, estas médias podem ser expressas por

$$MA_{\alpha_{i,j}}\left(\frac{\sigma_{i,j}^2 B_{i,j}^{'}}{\alpha_{i,j}}\right)=\sum_{i,j}\sigma_{i,j}^2 B_{i,j}^{'} \tag{3.10}$$

e

$$MG_{\alpha_{i,j}}\left(\frac{\sigma_{i,j}^2 B_{i,j}^{'}}{\alpha_{i,j}}\right)=\prod_{i,j}\left(\frac{\sigma_{i,j}^2 B_{i,j}^{'}}{\alpha_{i,j}}\right)^{\alpha_{i,j}}. \tag{3.11}$$

Os fatores de normalização $B_{i,j}^{'}$ são usados para ajustar as variâncias das sub-bandas, de modo que a média aritmética expressa em (3.10) seja idêntica a σ_x^2, a variância da imagem original. Por exemplo, para as sub-bandas produzidas no primeiro estágio, normalizam-se as variâncias $\sigma_{1,1}^2, \sigma_{1,2}^2, \sigma_{1,3}^2$ e $\sigma_{1,4}^2$ com

$$B_{1,j}^{'}=\frac{\sigma_x^2}{(\sigma_{1,1}^2+\sigma_{1,2}^2+\sigma_{1,3}^2+\sigma_{1,4}^2)},$$

sendo que i =1, j =1, ..., 4, isto é, $B_{1,j}^{'}$, para o primeiro estágio, é a razão entre a variância original e a soma das variâncias das sub-bandas geradas. No segundo estágio, as variâncias correspondentes às sub-bandas de baixas freqüências, $\sigma_{2,1}^2$, $\sigma_{2,2}^2$, $\sigma_{2,3}^2$ e $\sigma_{2,4}^2$ produzidas por duas etapas de filtragem, são normalizadas com

$$B_{2,j}^{'}=\frac{B_{1,j}^{'}\sigma_{1,1}^2}{(\sigma_{2,1}^2+\sigma_{2,2}^2+\sigma_{2,3}^2+\sigma_{2,4}^2)},$$

sendo que agora i =2, j =1, ..., 4. Isto é, no segundo estágio, $B_{2,j}^{'}$ é o produto do fator de normalização da sub-banda de baixas freqüências (do primeiro estágio) pela razão entre a variância da sub-banda de baixas freqüências (do primeiro estágio) e a soma das variâncias das sub-bandas descendentes. Este processo é repetido para cada estágio da decomposição, gerando uma seqüência normalizada de variâncias para todas as sub-bandas, de modo análogo às variâncias dos coeficientes de uma transformada unitária. Observa-se que o critério de escolha

dos fatores de normalização descrito acima faz com que o numerador de (3.9) seja, em todos os estágios, idêntico à variância da imagem original.

Esta medida de compactação é igual ao ganho de codificação de sub-banda, G_{SBC}, proposto por Katto e Yasuda [KATT91], quando os fatores de normalização $B'_{i,j}$ introduzidos acima são idênticos aos coeficientes B_i, que relacionam as variâncias dos erros de quantização $\sigma^2_{q,i}$ dos coeficientes das sub-bandas *i-ésimas* e a variância do erro de reconstrução da imagem, σ^2_r, através da expressão (3.3). Tais condições são satisfeitas por bancos de filtros ortogonais com erros de quantização descorrelatados. Conjectura-se que estas condições sejam também satisfeitas em muitos casos de bancos de filtros não ortogonais. Essa relação entre a medida empírica expressa por (3.9) e a medida de Katto e Yasuda é importante, pois pode-se mostrar que a medida de Katto e Yasuda corresponde ao ganho de codificação em sub-bandas em relação à codificação PCM quando os coeficientes das sub-bandas são submetidos à quantização fina com alocação de bits ótima [KATT91].

A vantagem da utilização de (3.9) em relação à expressão de Katto e Yasuda consiste em que esta última depende dos coeficientes dos filtros de síntese e, portanto, não se aplica facilmente a bancos de filtros não-lineares. Por outro lado, a expressão (3.9) depende exclusivamente das variâncias das sub-bandas geradas pela decomposição.

3.3 ESQUEMAS DE COMPRESSÃO DE IMAGENS VIA CODIFICAÇÃO DE SUB-BANDAS

As propriedades exibidas pelos coeficientes das sub-bandas ensejaram esquemas bem sucedidos de compressão de imagens. Os primeiros trabalhos exploravam as propriedades dos coeficientes de cada sub-banda através de técnicas de codificação como o DPCM, técnicas de varredura juntamente com codificação *run-length* e codificação de entropia [WOOD86] e [GHAR88].

Estes primeiros estudos deixaram de lado uma importante propriedade que é a correlação entre sub-bandas. Olhando a Fig. 2.10, fica claro que existe uma dependência entre as bandas de detalhes. Esta observação levou ao desenvolvimento de uma forma de codificação especialmente projetada para a decomposição de multi-resoluções de transformadas implementadas por banco de filtros. Ela é baseada numa estrutura de dados chamada de árvores de zeros (*zero tree*), que é

análoga à varredura em zig-zag e ao símbolo de fim de bloco usados na codificação via DCT (JPEG). A idéia é definir uma árvore de símbolos nulos ou zeros que começa em uma raiz na decomposição de multi-resoluções que também é nula. Para que se entenda melhor isso, é apresentado na Fig. 3.1 o ordenamento das sub-bandas em três níveis de profundidade.

Neste ordenamento das sub-bandas, é possível dizer que as bandas de detalhamento 2, 5 e 8 são altamente correlacionadas [SHAP93], uma vez que a sub-banda 2 é uma aproximação grosseira da sub-banda 5, e esta, por sua vez, é uma aproximação grosseira da sub-banda 8. O mesmo é válido para as sub-bandas 3, 6 e 9 e as sub-bandas 4, 7 e 10. Suponha-se agora que o pixel no canto superior esquerdo da sub-banda 2 da Fig. 3.2 seja nulo.

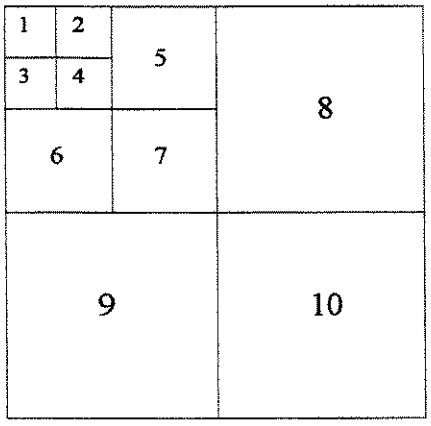


Figura 3.1: Sub-bandas resultantes de uma decomposição em três níveis de profundidade através de banco de filtros

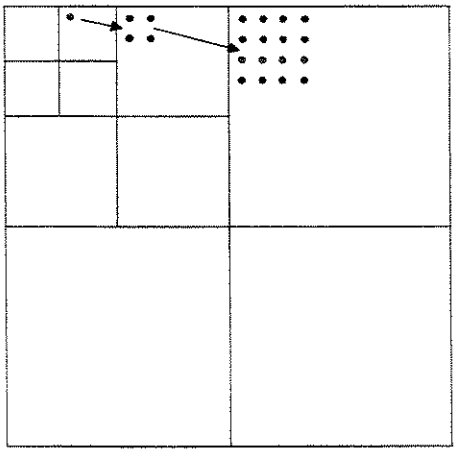


Figura 3.2: Estrutura de árvores de zeros em uma decomposição de imagem através de banco de filtros.

Olhando-se para a estrutura da Fig.3.2, há uma grande probabilidade de que os pixels do quadrado 2x2 da sub-banda 5 também sejam nulos. Da mesma forma, os pixels do quadrado 4x4 da sub-banda 8 provavelmente também serão nulos. Essas regiões quadradas são denominadas regiões espacialmente correlacionadas.

Esta importante característica foi utilizada por Shapiro [SHAP93] no desenvolvimento de um algoritmo denominado EZW (*Embedded Zerotree Wavelet*). Este é um algoritmo para compressão de imagens cuja principal característica é o fato de os bits serem codificados pela sua ordem de importância visual. Esse algoritmo oferece vantagens com relação aos anteriormente citados, pois nele o codificador pode interromper a codificação quando uma desejada taxa de compressão for atingida. Similarmente, o decodificador pode interromper a decodificação a uma determinada taxa desejada. Este método de compressão produz excelentes resultados sem requerer conhecimento a priori da imagem ou de tabelas de codificação previamente armazenadas. Neste trabalho, não será usado o algoritmo EZW. No entanto, os resultados podem ser melhorados se o algoritmo for utilizado [SAID96].

A decomposição de sinais em várias resoluções foi estudada ao longo dos últimos anos em diversos campos de aplicação, entre os quais a compressão de imagem. A decomposição por si só não comprime a informação, só a organiza de uma forma mais favorável. Quando a taxa de informação é menor do que a entropia das bandas, é necessário quantizar as sub-bandas e codificar por entropia os dados resultantes, como se mostra na Fig. 3.3. A quantização é um processo irreversível. Por isso, deve-se explorar ao máximo as características particulares de cada sub-banda, como, por exemplo, a concentração de energia que cada uma delas exibe, de forma a minimizar a distorção e melhorar a qualidade visual das imagens reconstruídas. Esta tarefa é realizada pelo controlador da taxa de transmissão, que aloca os bits nas sub-bandas de acordo com a capacidade de compactação de energia de cada uma delas.

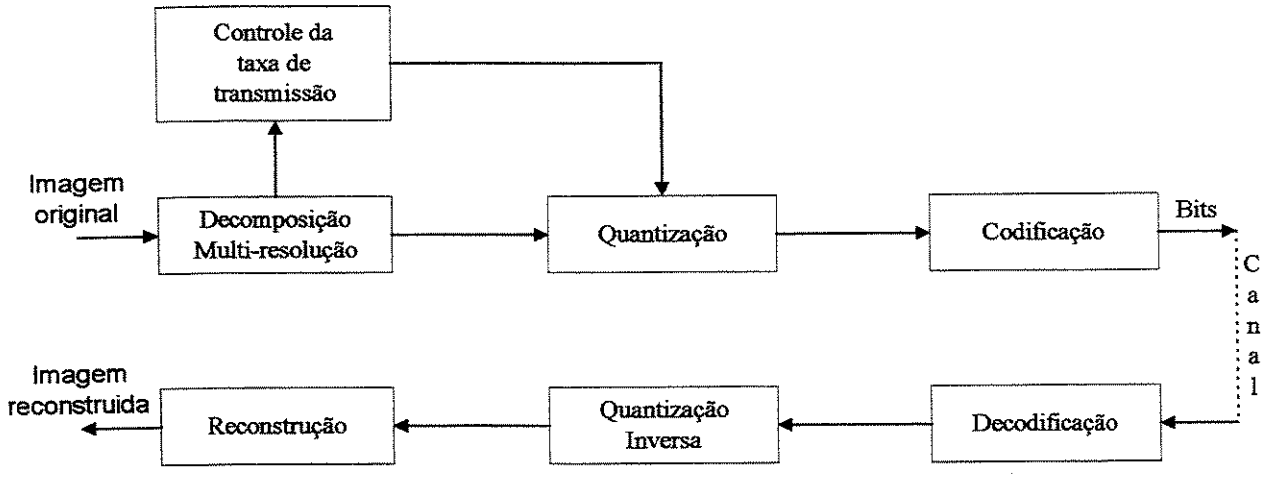


Figura 3.3: Diagrama de um sistema genérico de compressão de imagens baseado na Decomposição em Multi-resolução de imagem.

O esquema de compressão proposto na Fig. 3.3 mostra, em diagrama de blocos, um sistema genérico de compressão de imagens constituído por uma etapa de quantização e por um codificador de entropia. Esses dois blocos são discutidos nas próximas seções.

3.3.1 Quantização das Sub-bandas

O resultado da decomposição em bandas de freqüência, seja das imagens originais ou ainda das imagens do erro de predição, para o caso de vídeo, são conjuntos finitos de coeficientes ou amostras, cujos domínios são subconjuntos de **N**. Outra abordagem possível é considerar que as amostras passaram por um processo de pré-quantização ou arredondamento, pois os domínios são subconjuntos de **Z**.

Seja x uma quantidade escalar contínua representando a intensidade do píxel. Para representar x com um número finito de bits, um número finito de níveis de quantização pode ser utilizado. Assume-se que um total de L níveis são utilizados para representar x. O processo de atribuição de um dos L níveis a cada valor de x é denominado quantização da amplitude, ou simplesmente, quantização [JAYA84]. De forma grosseira, pode-se dizer que, se cada escalar é quantizado de forma independente ou individualmente, o procedimento é denominado de quantização escalar. Se dois ou mais escalares formam um grupo ou um vetor e são quantizados conjuntamente, o procedimento é chamado de quantização vetorial.

Em sistemas de compressão de imagens cujo objetivo principal seja comprimir de forma significativa não é possível representar completamente todo o domínio de coeficientes, isto é, atribuir a cada valor um símbolo diferente. É necessário, portanto, limitar o conjunto de símbolos com que se trabalha, mesmo que isso implique a introdução de alguma distorção nos sinais reconstruídos.

Neste estudo, trabalha-se unicamente com quantizadores escalares no processo de quantização das sub-bandas, devido à sua menor complexidade computacional. Na Fig. 3.4, estão representadas as características de dois quantizadores típicos.

Um quantizador escalar é constituído por dois sistemas. Em primeiro lugar, pelo sistema direto que mapea a seqüência de valores de entrada, $x[i]$, na seqüência de símbolos, $q[i]$, pertencentes ao conjunto, $\{s_1, s_2, \cdots, s_L\}$. Esta operação é realizada, pela divisão do eixo real em L intervalos de decisão adjacentes,

$$]l_0,l_1], \;]l_1,l_2], \; \ldots,]l_{L-2},l_{L-1}], \;]l_{L-1},l_L[, \tag{3.12}$$

com $l_0 = -\infty$ e $l_L = +\infty$. A um valor de $x[i]$ que pertença ao intervalo $]l_{j-1}, l_j]$, é atribuído o símbolo s_j. Em segundo lugar, o quantizador escalar é constituído pelo sistema inverso que mapea a seqüência de símbolos $q[i]$, na seqüência de saída $\hat{x}[i]$, cujos valores pertencem ao conjunto de reconstrução, $\{r_1, r_2, \cdots, r_L\}$.

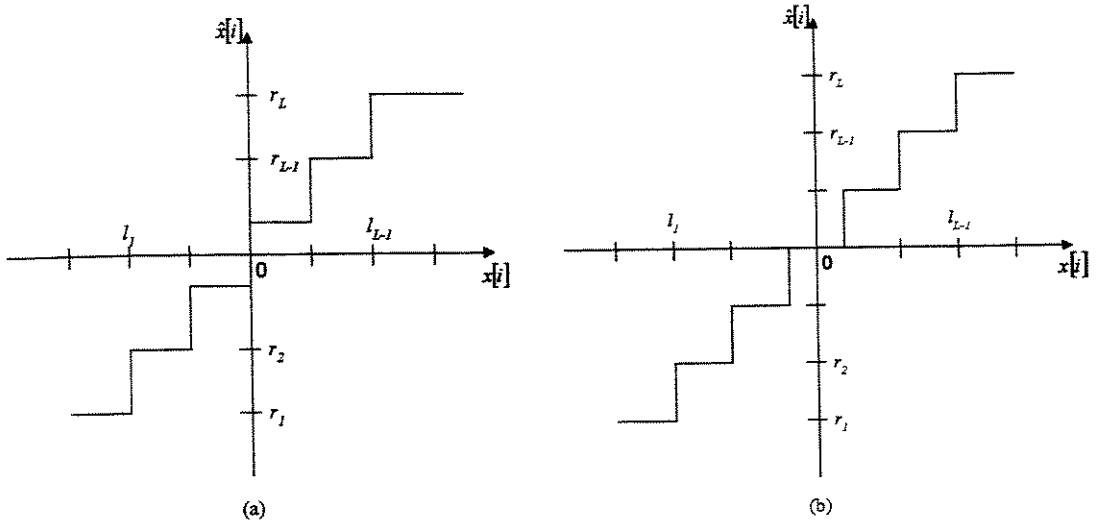


Figura 3.4: Tipos de quantizadores escalares genéricos:
(a) Com limiar de decisão em zero (*midrise*).
(b) com nível de reconstrução em zero (*midtread*).

Um tipo de quantizador com grande interesse prático é aquele em que os níveis de decisão e de reconstrução são igualmente espaçados. Esse quantizador denomina-se quantizador uniforme. Neste trabalho, serão utilizados quantizadores uniformes, como ilustra a Fig. 3.4(b). Os quantizadores da Fig.3.4(a) não são os mais indicados para a quantização de sub-bandas resultantes da decomposição da imagem, pois não possuem um nível de reconstrução nulo quando, em geral, a maioria dos coeficientes são nulos, como é o caso das sub-bandas de detalhamento. Além disso, seu desempenho em termos de entropia é também menor [JAYA84].

Em qualquer quantizador, estão sempre associados dois aspectos importantes. O primeiro desses aspectos é que a seqüência de símbolos, $q[i]$, é uma fonte de informação discreta, cuja redundância estatística deve ser eliminada ou reduzida, através da utilização de técnicas de codificação por entropia. O segundo aspecto é que a quantização é uma operação irreversível. Isto é, as seqüências de entrada e de saída não são iguais, devido à introdução do erro de quantização, que é definido por

$$e_Q[i] = x[i] - \hat{x}[i], \tag{3.13}$$

onde x[i] e $\hat{x}[i]$ são, respectivamente, a entrada e a saída do quantizador. O valor e_Q^2 pode ser interpretado como um caso especial da medida de distorção $d(x-\hat{x})$, que é uma medida de distância entre x[i] e $\hat{x}[i]$.

Os níveis de representação e decisão são freqüentemente determinados através da minimização de algum critério de erro baseado em $d(x,\hat{x})$. Um exemplo é a média da distorção, definida por:

$$D = E[d(x,\hat{x})] = \int_{x=-\infty}^{\infty} d(x,\hat{x}).p_X(x).dx \ , \tag{3.14}$$

onde $p_X(x)$ é a função densidade de probabilidade de x. Seja $d(x,\hat{x}) = |x-\hat{x}|^2$. O erro médio quadrático (MSE – *Mean Square Error*) da saída do quantizador é definido pela seguinte expressão:

$$D = E\left(|x-\hat{x}|^2\right) = \sum_{i=1}^{L-1} \int_{s_{i-1}}^{s_i} (x-r_i).p_X(x).dx. \qquad (3.15)$$

Como exemplo, tem-se a quantização uniforme com L intervalos, onde $s_i - s_{i-1} = \Delta$, sendo que Δ é o tamanho do passo, ou seja, o espaçamento entre dois níveis de representação, e $p_X(x)$ é uma entrada com distribuição uniforme. Este é o caso de quantização ótima, no qual a quantidade (3.13) é minimizada. Para esta situação, o MSE vale

$$D = \frac{\Delta^2}{12}. \qquad (3.16)$$

Quando a distribuição $p_X(x)$ não é uniforme, a quantização uniforme não é ótima, e para ocorrer a minimização de (3.13), o quantizador deverá ser não uniforme e satisfazer certos critérios estudados por Lloyde e Max, descritos em [JAYA84].

3.3.2 Quantização com Alocação de Bits

A quantização com alocação de bits utiliza a estatística das sub-bandas como base para o processo de quantização. A cada sub-banda da decomposição, é alocada uma certa quantidade de bits para sua representação, de acordo com suas características estatísticas e a taxa de bits total que se deseja obter.

O processo de alocação ótima de bits é realizado de acordo com um critério de minimização [JAIN88] previamente escolhido. Existe um algoritmo de alocação de bits que permite fazer uma distribuição de bits entre as sub-bandas, de acordo com o valor das suas variâncias [STRA96]. Descreve-se sumariamente este algoritmo. Sejam:

- N o número total de bits na imagem original;
- M o número total de sub-bandas;
- R a taxa de bits desejada;
- $\alpha_k = N_k/N$ o tamanho relativo da k-ésima sub-banda;
- $b = (b_1,....,b_N)$ o vetor contendo o número de bits alocados a cada sub-banda;
- ε_k o índice de desempenho do quantizador;

- σ_k^2 a variância da k-ésima sub-banda;
- ω_k o fator de ponderação da sub-banda (maiores detalhes serão dados na próxima seção; por enquanto, assume-se que $\omega_k = 1$, $\forall k$.).

Define-se o erro de quantização como:

$$D(b) = \sum_{k=1}^{M} \alpha_k . \omega_k . \varepsilon_k^2 . 2^{-2b_k} . \sigma_k^2 \tag{3.17}$$

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

e a taxa de bits como:

$$R(b) = \sum_{k=1}^{M} \alpha_k . b_k . \tag{3.18}$$

O problema da alocação de bits consiste em minimizar $D(b)$, mantendo fixa a taxa de bits total $R(b)=R_C$. Isso pode ser resolvido utilizando um multiplicador de Lagrange λ [JAIN88].

$$\min\{D(b) + \lambda . R(b)\} = \min\left\{\sum_{k=1}^{M} \alpha_k . (\omega_k . \varepsilon_k^2 . 2^{-2b_k} . \sigma_k^2 + \lambda . b_k)\right\}. \tag{3.19}$$

Assumindo que $\varepsilon_k = \varepsilon$ e definindo $\tilde{\lambda} = \lambda / \varepsilon^2$, a expressão (3.19) se reduz a:

$$\min\left\{\sum_{k=1}^{M} \alpha_k . \varepsilon^2 (\omega_k . 2^{-2b_k} . \sigma_k^2 + \tilde{\lambda} . b_k\right\}. \tag{3.20}$$

Diferenciando (3.20) em relação a b_k, obtém-se a seguinte solução [STRA96].

$$b_k = \frac{1}{2} \log_2 \left(\frac{(2 \log_2 2) . \omega_k . \sigma_k^2}{\tilde{\lambda}} \right), \tag{3.21}$$

Substituindo (3.21) em (3.18), obtém-se:

$$R_C = \sum_{k=1}^{M} \alpha_k . b_k = \frac{1}{2} \sum_{k=1}^{M} \alpha_k . \log_2 \left(\frac{(2 \log_e 2) . \omega_k . \sigma_k^2}{\tilde{\lambda}} \right), \tag{3.22}$$

onde $\tilde{\lambda}$ é definido por:

$$\tilde{\lambda} = 2^{\left\{\sum_{k=1}^{N} \alpha_k . \log_2\left((2\log_e 2).\omega_k . \sigma_k^2\right) - 2R_C\right\}}. \qquad (3.23)$$

Primeiramente, o valor de $\tilde{\lambda}$ é calculado utilizando-se as variâncias σ_k^2 das sub-bandas. Em seguida, o número de bits alocado a cada sub-banda é calculado através da expressão (3.21). Este algoritmo deve ser repetido cada vez que uma imagem nova for codificada.

3.3.3 Quantização com Ponderação das Sub-bandas`

A quantização com ponderação nas sub-bandas é um aperfeiçoamento da quantização com alocação de bits. É notório que uma minimização do erro médio quadrático não garante resultados visualmente ótimos [JAIN88]. Os olhos humanos são menos sensíveis a perdas em altas freqüências do que a perdas em médias e baixas freqüências, fator que não foi considerado na seção anterior. O desempenho do quantizador pode, então, melhorar se forem introduzidos aspectos do sistema visual humano no modelo de alocação de bits. Isto é feito através da ponderação do ruído de quantização de cada sub-banda de acordo com a sensibilidade do sistema visual humano. Apresenta-se aqui um esquema simples e eficaz para ponderação do ruído de quantização. Neste, mais bits são alocados às bandas de baixas e médias freqüências [STRA96]. O algoritmo de alocação de bits é o mesmo utilizado na seção anterior, porém os valores atribuídos aos pesos de ponderação das sub-bandas assumem valores diferentes de 1. O peso ω_k, correspondente à k-ésima sub-banda, assume um valor igual à a^j, que irá depender do conteúdo em freqüências, ou seja, do índice desta sub-banda. Define-se a como um número real e maior que 1.

Na Fig. 3.5, é apresentado um modelo de distribuição dos pesos de ω_k para uma decomposição de multi-resoluções com três camadas, ou seja, três estágios.

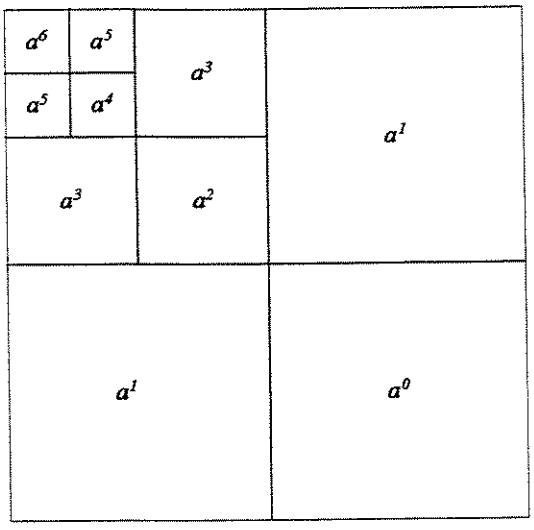


Figura 3.5: Modelo de distribuição dos pesos de ponderação para as sub-bandas de uma decomposição em multi-resoluções com nível de profundidade igual a 3.

Seguindo a mesma notação da Fig. (2.11), para este caso, o peso de ponderação da primeira banda, ou banda de aproximação $c_{1,3}$, é igual a a^6, o das bandas $d_{2,3}$ e $d_{3,3}$ é igual a a^5, o da banda $d_{4,3}$, é igual a a^4, e assim por diante. Para se calcular o número de bits alocados a cada sub-banda, substituem-se os valores de ω_k nas expressões (3.21), (3.22) e (3.23) e se utiliza o algoritmo apresentado na seção anterior. O valor de a utilizado na simulação foi de $a=2$. Através das simulações, pôde-se concluir que, ao passo que o valor de a cresce, a taxa de bits cresce também. No esquema da seção anterior, a alocação de bits era feita de forma uniforme entre todas as sub-bandas, independentemente do conteúdo de freqüência e da sua importância visual. Neste esquema, mais bits são alocados às médias e baixas freqüências, que são visualmente mais importantes. A ponderação garante que as sub-bandas sejam comprimidas levando em conta sua importância visual.

3.3.4 Codificação de Entropia

A última etapa num esquema de compressão é a codificação de entropia. Nesta etapa, é eliminada a redundância estatística da imagem ou vídeo. Este processo é reversível e não há necessidade de aproximações. Na etapa anterior à codificação, mais especificamente no quantizador, o problema foi minimizar a distorção média do sinal para um dado número fixo de

níveis de quantização ou minimizar o número de níveis para uma certa distorção média, dada por (3.14).

Nesta etapa, em aplicações práticas de compressão de imagens, o interesse maior é minimizar o número de bits e, para tal, associam-se palavras código apropriadas aos níveis de quantização gerados em instante anterior ao sistema de codificação. É muito importante salientar que os problemas de quantização e de associação de palavras-códigos estão intimamente ligados. Se os níveis de quantização não são igualmente prováveis, associá-los a palavras-códigos de comprimento variável, de acordo com as probabilidades de ocorrência de cada nível, permite uma grande redução da taxa média de bits, em comparação à associação de palavras-códigos de comprimento fixo. Esse fato não se verifica quando os níveis de quantização não apresentam uma probabilidade de ocorrência aproximadamente igual. Portanto, o que é feito no estágio de quantização afeta posteriormente o estágio de associação de palavras-códigos aos níveis.

No processamento de sinais de vídeo, por exemplo, as quantidades a serem processadas apresentam funções densidades de probabilidade não-uniformes, e o fato de associá-las a palavras-códigos de comprimento variável pode significar grande redução na taxa média de bits do sistema de codificação. Contudo, deve-se ressaltar que, em virtude das imprecisões nas estatísticas estimadas ou nas medidas de distorção assumidas, os resultados teóricos de vários métodos de quantização e de associação de palavras-códigos devem ser considerados apenas como orientação na decisão de qual método usar em uma aplicação de codificação de imagem.

Para que o decodificador (ou receptor) seja capaz de identificar unicamente os níveis de quantização, a cada nível deve ser associada uma palavra-código diferente. Além disso, os níveis são transmitidos usualmente em série, e as palavras-códigos devem ser projetadas de tal modo que possam ser identificadas quando recebidas em série. Um código com essas características é dito ser unicamente decodificável [YABU91].

No caso de se associarem palavras-códigos de comprimento variável para reduzir a taxa média de bits, está se supondo que algumas mensagens apresentam probabilidade de ocorrência maior que as outras. A essas mensagens com maiores probabilidades de ocorrência são associadas palavras-códigos mais curtas, e às mensagens com probabilidades de ocorrência menores são associadas palavras-códigos mais longas (mais bits/mensagem). No caso de se associarem aos níveis palavras-códigos de comprimento fixo, a taxa média de bits é o próprio número de bits requeridos para codificar uma mensagem. A entropia [H] do sinal é definida como

$$H = -\sum_i p_i . \log_2 (p_i), \tag{3.13}$$

onde p_i = probabilidade de ocorrência da mensagem ou nível " i ".

Desde que $\sum_i p_i = 1$, pode-se mostrar que $0 \leq H \leq log_2 L$, onde L é o número de níveis de quantização. Se as probabilidades variam durante a codificação, uma associação de palavras-códigos mais sofisticada pode ser empregada para se obterem melhores resultados, como a codificação aritmética [SAID96] e [WITT87]. É importante ressaltar que, como a codificação por entropia não apresenta perda de informação, podem-se associar duas ou mais formas desta codificação após a quantização.

Do ponto de vista da teoria de informação, a entropia é teoricamente a taxa média de bits mínima possível que é requerida para codificar uma mensagem, ou é a quantidade média mínima de informação que uma mensagem contém. Assim, no projeto de palavras-códigos, aquele que der a taxa média de bits mais próxima da entropia é o melhor.

CAPÍTULO 4

DESCRIÇÃO DO SISTEMA IMPLEMENTADO NAS SIMULAÇÕES

Neste capítulo, é apresentado o sistema de compressão que foi implementado, baseado na decomposição multi-resolução de imagens. Descrevem-se as simulações realizadas e se procuram avaliar os bancos de filtros lineares e não-lineares, no que diz respeito a sua capacidade de compactação de energia no domínio da transformada. São descritos sumariamente os métodos da convolução circular e da extensão simétrica, que foram utilizados nas simulações, quantização e alocação de bits para os coeficientes. São apresentadas as medidas de distorção e detalhamento na compressão das imagens.

4.1 SISTEMAS IMPLEMENTADOS PARA SIMULAÇÃO

O sistema simulado neste trabalho é mostrado no digrama da Fig. 4.1. O objetivo é avaliar e verificar as vantagens de cada banco de filtro. Em um primeiro momento, é avaliada a capacidade de cada banco de filtro no que se refere à capacidade de compactação de energia no domínio da transformada dos coeficientes em sub-bandas, ou seja, na decomposição. A seguir, serão implementados no sistema de compressão os bancos de filtros que tiveram melhor desempenho, no que se refere ao Ganho de Compactação proposto. Os bancos de filtros serão

aplicados ao sistema de compressão apresentado na Fig. 4.1 para se verificar sua PSNR e a qualidade visual das imagens reconstruídas.

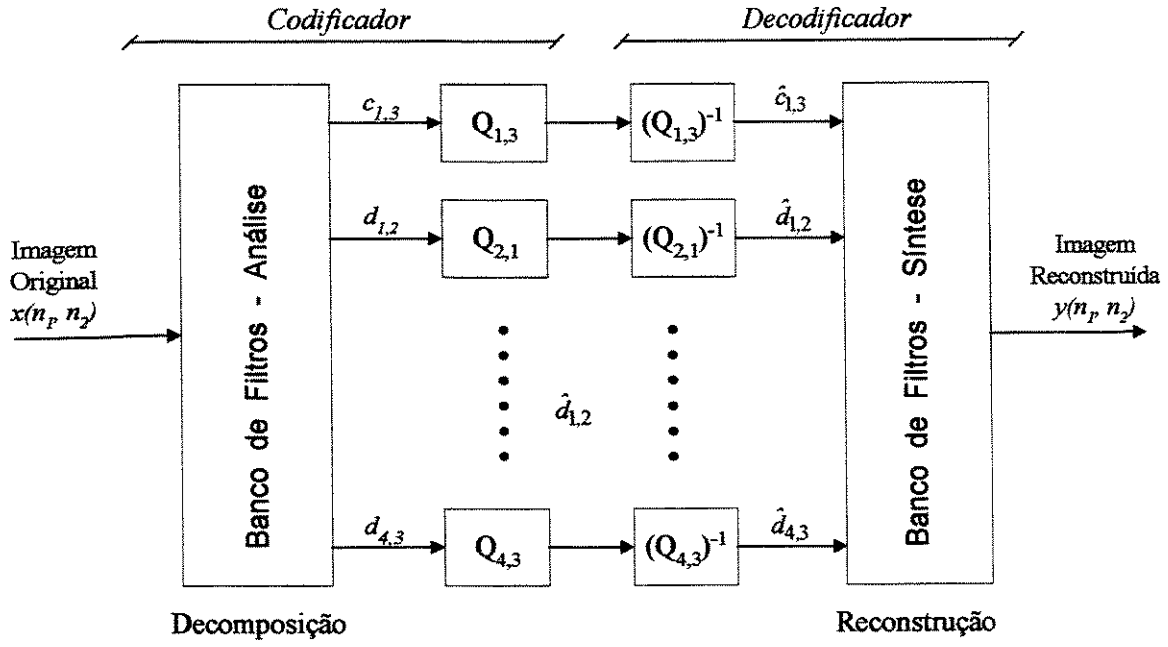


Figura 4.1: Sistema de compressão implementado para simulação com os bancos de filtros

O primeiro bloco é o estágio de decomposição (bidimensional) da imagem. Foi implementado, por extensão separável (seção 2.4), em três níveis de decomposição, gerando um total de 10 sub-bandas. Foram utilizados dois pares de filtros biortogonais de comprimento ímpar, cujos coeficientes são mostrados na Tabela 4.1. Somente são mostrados os coeficientes centrais e os da direita, visto que cada filtro é simétrico. O conjunto *QMF(9x7)* apresenta os filtros FIR $\bar{h}$ e h com comprimentos $p = 9$ e $q = 7$, respectivamente. Já o conjunto *QMF(5x3)* possui $p = 5$ e $q = 3$. Como dito anteriormente, na Seção (2.3.2), esses filtros estão apresentados no trabalho de Villasenor [VILL95], onde se avaliou a regularidade das funções de escala *wavelet* associada ao par de filtros, a resposta impulsiva e o degrau da eficiência na codificação a uma taxa de 16:1 de imagens teste, utilizando três níveis de profundidade.

Tabela 4.1 – Coeficientes dos pares de filtros utilizados na simulação, com o coeficiente central e as amostras da direita.

	Filtro *QMF 9x7*	**Filtro *QMF 5X3***
h_0	0.852699; 0.377402; -0.110624; -0.023849; 0.037828	0.750; 0.250; -0.125
h_1	0.788486; -0.418092; -0.040689; 0.064359	0.500; 0.250

Os blocos a seguir são os quantizadores, denominados $Q_{1,3}$ e $Q_{i,j}$, com $2 \leq i \leq 4$ e $1 \leq j \leq 3$. Cada sub-banda é quantizada individualmente e os coeficientes pertencentes a ela são mapeados em índices inteiros. O próximo passo na compressão seria a codificação de entropia. No entanto, esta codificação não foi implementada. Portanto, a forma de avaliar o desempenho dos bancos de filtros será apenas na presença de quantização. Além da medida de taxa de bits, também é calculada a entropia dos coeficientes quantizados, para se estimar o ganho que se obteria com a codificação de entropia.

O decodificador compreende o mapeamento dos índices de quantização dos coeficientes de cada sub-banda em valores próximos aos originais e a síntese a partir das sub-bandas. Utilizou-se a notação $(Q_{1,3})^{-1}$ e $(Q_{i,j})^{-1}$ para o processo de quantização inversa, porém é sabido que a quantização é um processo irreversível. Mas é mantida esta notação de forma a facilitar a compreensão. E, por fim, tem-se a interpolação e o banco de filtros de síntese. Esta implementações foram realizadas em programações no MATLAB 5.3® e em linguagem C++, em plataforma Windows NT 95®.

4.1.2 Implementação dos Métodos da Convolução Circular e da Extensão Simétrica

Um dos problemas relativos ao uso de bancos de filtros para codificação de dados ou sinais de duração finita, como vetores de linhas ou colunas em uma imagem digital, é o manejo de condições de contorno nas bordas do vetor. Se forem aplicados diretamente os filtros nas imagens, ou seja, convoluindo linearmente as linhas e as colunas, estar-se-á incrementando um número maior de pixels na imagem original. Em outras palavras, o problema é que, convoluindo

linearmente uma imagem (VxH) com um filtro (pxq), a imagem resultante terá um tamanho de (V+p-1) x (H+q-1). Esta propriedade da convolução linear é contrária ao objetivo da compressão de dados. E para reduzir este impacto, torna-se necessário definir métodos nos quais este efeito seja minimizado e a imagem original não sofra uma expansão. As transformadas que conservam o número de amostras entre o sinal discreto ou o vetor original e sua representação no domínio da transformada são chamadas não expansivas. Um exemplo é a transformada discreta do coseno (DCT). A transformada implementada com banco de filtros mostra-se, então, uma transformada expansiva, já que as convoluções do sinal de entrada com os filtros aumentam o comprimento do sinal. Para minimizar o aumento de pixels, as bordas das sub-bandas podem ser retiradas ou espacialmente truncadas, para que se tenha uma fração proporcional ao tamanho da imagem original. No entanto, estas soluções introduzem descontinuidades e estas adicionam variâncias nas sub-bandas de detalhamento ou altas freqüências, reduzindo o ganho de codificação.

Existem soluções para este problema que permitem reconstrução exata e não aumentam o número total de pixels da imagem original. Podem-se encontrar algumas destas técnicas em [KARL89]. Caso certas restrições relacionadas ao sinal de entrada e ao banco de filtros utilizados sejam satisfeitas, a transformada implementada através de banco de filtros pode-se tornar uma transformada não expansiva. Brislawn [BRIS95] verificou que, se o comprimento N_0 de uma entrada unidimensional for par e se os filtros exibirem simetria (fase linear) e possuírem comprimentos menores do que N_0, então a decomposição em multi-resoluções por banco de filtros poderá ser implementada, sem causar aumento no tamanho do vetor de entrada, através de convoluções circulares entre uma extensão simétrica do sinal de entrada e os filtros do banco. Felizmente, este é o caso dos banco de filtros biortogonais. As descontinuidades introduzidas pela implementação circular são eliminadas porque se utiliza uma extensão par do sinal de entrada. Isto dobra o comprimento do sinal de entrada, aparentemente dobrando o número de coeficientes resultantes. A chave para se fazer este procedimento de uma forma não expansiva está em se usar filtros de fase linear que produzam bandas simétricas.

Detalhes do procedimento e derivações algébricas do método da extensão simétrica e da convolução circular podem ser encontradas em [SMIT91] e [BRIS95]. Em [SOUZ00], é possível encontrar uma descrição completa do método da extensão simétrica para fins de implementação de algoritmo.

Neste trabalho, os métodos da convolução circular e o método da extensão simétrica foram implementados nas decomposições, ou análise, para analisar a capacidade de compactação de energia das sub-bandas do domínio da transformada. E na implementação do modelo de compressão da Fig. 4.1, é utilizado o método da extensão simétrica, por apresentar um ligeiro ganho em relação ao método da convolução circular.

4.1.3 Os Quantizadores Implementados

Neste trabalho, foram utilizados dois tipos de quantizadores uniformes para a quantização dos coeficientes das sub-bandas. Para a sub-banda de baixas freqüências, $c_{1,3}$, os coeficientes foram quantizados com um número par de intervalos dado por $N=2^R$, onde R é o número de bits alocados a esta sub-banda. Os limitantes de sobrecarga, x_0 e x_{N-1} são a menor e a maior ocorrência em $c_{1,3}$, respectivamente. Portanto, a codificação é realizada, então, da seguinte forma :

$$i = \lfloor (x - x_0)/\Delta + 0.5 \rfloor , \tag{4.1}$$

onde $\lfloor \bullet \rfloor$ denota o maior inteiro menor ou igual ao argumento. O índice i pode assumir os valores 0, 1, 2, ..., N-1. A decodificação é realizada fazendo-se

$$\hat{x}(i) = \Delta i + x_0 . \tag{4.2}$$

Os quantizadores para os coeficientes das sub-bandas de detalhes $d_{i,j}$, $1 \le i \le 3$ e $2 \le j \le 4$, são simétricos em torno da origem e apresentam um número ímpar de intervalos. Isto se deve à distribuição dos coeficientes apresentar um pico em zero e decair rapidamente em ambos os lados, sendo aproximada por uma função densidade de probalidade gaussiana generalizada [ANTO92]. Para cada sub-banda $d_{i,j}$, o passo Δ é dado por $2max/(L+2)$ onde max é o maior valor absoluto entre os coeficientes pertencentes à sub-banda e L é o número de intervalos fora o intervalo central atribuído à quantização de $d_{i,j}$. A codificação é dada por

$$i = sinal(x) \left\lfloor \frac{|x|}{\Delta} \right\rfloor , \tag{4.3}$$

onde $sinal(x)=1$ se $x>0$, -1 se $x<0$ e 0 se $x=0$. A decodificação se dá com,

$$\hat{x}(i) = i\Delta + sinal(x).\frac{\Delta}{2} . \qquad (4.4)$$

O intervalo central em torno do zero é chamado zona morta e possui comprimento 2Δ. Esta zona morta ajuda a eliminar o ruído presente no sinal, suavizando a representação, [GHAR88] e facilita a tarefa de um possível codificador de entropia.

Resta então a tarefa de estabelecer a resolução com a qual as 10 sub-bandas serão quantizadas. Para tanto, deve-se medir a importância que cada sub-banda tem na representação da imagem original. Esta medida pode ser quantificada pela área relativa da sub-banda (o número de coeficientes que ela possui) e por sua energia (variância). No Capítulo 3, foi descrito um algoritmo de alocação de bits de acordo com estes dois critérios. Sejam b_1, b_2,....., b_{10} os números de bits alocados à quantização das sub-bandas de acordo com sua importância (ver Fig. 3.1). Calculam-se os valores de b_k por (3.9), onde σ_k^2 é a variância da sub-banda k e ω_k um fator de ponderação que privilegia as sub-bandas que se encontram em maior profundidade. Dessa forma $\omega_1 = a^6$, $\omega_2 = \omega_3 = a^5$, $\omega_4 = a^4$, $\omega_5 = \omega_6 = a^3$, $\omega_7 = a^2$, $\omega_8 = \omega_9 = a$ e $\omega_{10} = 1$, com $a \geq 1$. Utilizou-se $a = 2$ nas simulações.

O parâmetro λ é um multiplicador de Lagrange utilizado na minimização da função erro em (3.5), que é calculado por (3.11), onde α_k é o número de coeficientes da k-ésima sub-banda, normalizado pelo número total de pixels da imagem, e $R_C = \sum_{k=1}^{10} \alpha_k b_k$ é a taxa média almejada em bits por píxel após a quantização de todos os coeficientes em todas as sub-bandas k=1,2,...10. É importante notar que este algoritmo gera $b_k's$ não inteiros, devendo, portanto, haver arredondamento. Além disto, a quantização dos coeficientes das bandas de detalhes requer um número ímpar de níveis, pois é utilizada a zona morta em torno da origem. Portanto para as sub-bandas $2 \leq k \leq 10$, são alocados $b_k + 1$ bits na quantização. Nas simulações, os valores de R_C foram variando entre 0 e 8, a um passo de 0,5.

A quantização das sub-bandas geradas pelo filtro MED 2 tem uma particularidade especial no que se refere à quantização da sub-banda de aproximação. Pelo fato de a sub-banda de aproximação ser uma versão dizimada da imagem original, ela consiste em coeficientes

originais que não sofreram dizimação. Por essa razão, nas simulações, manteve-se a precisão de 8 bits nesta sub-banda, independentemente do valor de R_C.

4.2 COMPLEXIDADE COMPUTACIONAL

Um fator relevante a qualquer algoritmo de compressão de imagem é a complexidade computacional dos codificadores e decodificadores. A possibilidade de utilização destes algoritmos em sistemas de compressão de vídeo em tempo real, onde a taxa temporal de quadros é alta, pode tornar-se inviável caso as máquinas de operação do codificador e decodificador possuam baixo poder computacional.

Neste trabalho, foi comparada a complexidade computacional dos bancos de filtros. Compararam-se apenas três bancos de filtros: QMF(9x7), QMF(5x3) e MED 2. Estes bancos de filtros foram os que apresentaram melhor compactação de energia no domínio da transformada.

O cálculo do número de operações dos filtros QMF's foi baseado nas estruturas de análise e síntese das Figs. 2.8 e 2.11. Na decomposição de uma imagem de V linhas e H colunas em i níveis utilizando um par de filtros biortogonais de comprimentos p e q, são necessárias $2VH(p+q).(1-4^{-i})$ multiplicações e $(4/3).VH(p+q-2).(1-4^{-i})$ adições. Para a reconstrução, ou seja, a decodificação, há um aumento no número de operações devido à interpolação. São necessárias $4.VH(p+q).(1-4^{-i})$ multiplicações e $4.VH(p+q-1).(1-4^{-i})$ adições.

Para o filtro MED 2 aplicado às estruturas de análise e síntese das Figs. 2.18 e 2.20, o número de operações é significativamente menor do que os QMF's. Isto é devido a sua simplicidade e tamanho. Para a decomposição da imagem de V linhas e H colunas, são necessárias $(4/3).VH(1-4^{-i})$ multiplicações e $(8/3).VH(1-4^{-i})$ adições. Para a reconstrução, são necessários os mesmos números de operações que na decomposição.

Para se ter uma idéia do número de operações do banco de filtros MED 2 em relação ao banco de filtros QMF, foi definido um fator de redução de operações, que é definido como:

$$\text{Fator de redução de adições na análise} = \frac{2}{(p+q-2)} \times 100\% \qquad (4.5)$$

$$\text{Fator de redução de multiplicações na análise} = \frac{1}{(p+q)} \times 100\% \qquad (4.6)$$

$$\text{Fator de redução de adições na síntese} = \frac{2}{3(p+q-1)} \times 100\% \qquad (4.7)$$

$$\text{Fator de redução de multiplicações na síntese} = \frac{1}{3(p+q)} \times 100\% \qquad (4.8)$$

A tabela com os valores de percentagens é mostrada no próximo capítulo de resultados.

4.3 MEDIDAS DE GANHO E CRITÉRIO DE QUALIDADE VISUAL OBJETIVA

Nesta seção, são definidas as medidas utilizadas para avaliar o grau de compactação de energia para cada banco de filtros na decomposição e o conteúdo de informação ou detalhamento contido nas imagens reconstruídas para estimar o desempenho do sistema de compressão simulado.

Para avaliar o desempenho dos bancos de filtros com relação à compactação de energia no domínio da transformada, é definida uma fórmula geral baseada na proposta feita na seção (3.3.2), na qual o ganho de compactação de energia (GC_N) é dado pela equação (3.9).

Uma outra medida para avaliar o desempenho do sistema de compressão de imagem estudado foi a mais usual, que é o erro quadrático médio (MSE – *Mean Square Error*), calculado como

$$MSE = \frac{1}{VH} \sum_{i=0}^{V-1} \sum_{j=0}^{H-1} \left(x(i,j) - y(i,j)\right)^2 , \qquad (4.10)$$

onde $x(i, j)$ é a amostra da imagem original e $y(i, j)$ é a amostra reconstruída, ambas com dimensões VxH. O critério de qualidade visual objetiva é a relação sinal ruído de pico (PSNR – *Peak Signal Noise Ratio*), medida em decibéis, dada por

$$PSNR_{dB} = 10 \log_{10} \frac{255^2 . VH}{\sum_{i=0}^{V-1} \sum_{j=0}^{H-1} \left(x(i,j) - y(i,j)\right)^2} , \qquad (4.11)$$

onde 255^2 é o maior valor de pico a pico que o sinal de entrada pode assumir (tipicamente 255 para imagens em 8 bits). As medidas de distorção baseadas em erros quadráticos oferecem

limitações quando tomadas para medir a qualidade de sinais codificados como imagens, já que elas não correlacionam 100% com as degradações percebidas pelo olho humano.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

CAPÍTULO 5

APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Neste capítulo, são descritas as simulações realizadas e seus resultados no contexto dos capítulos anteriores. Primeiramente, foram realizados testes a fim de comprovar a eficiência dos bancos de filtros no que se refere à compactação de energia na decomposição. Depois foi analisado o desempenho dos bancos de filtros, que obtiveram maior ganho de compactação, em um sistema básico de compressão de imagem, no qual se avaliou seu desempenho por PSNR em relação a taxa de compressão. Para finalizar, fez-se uma análise da complexidade computacional sob os filtros que foram utilizados no sistema de compressão implementado.

As imagens de teste utilizadas foram *Lena*, com dimensões V=512 e H=512, *Salesman*, com dimensões V=352 e H=288, *Sail* e *Monarch*, ambas com dimensões V=512 e H=768. A imagem *Salesman* foi retirada atraves de uma seqüência de quadros, sendo que, o primeiro quadro da seqüência foi utilizado no trabalho para as simulações. Também se utilizou uma imagem de desenho, *Tina*, com dimensões V=256 e H=160. Todas as imagens foram obtidas via Internet, no endereço http://links.uwaterloo.ca/bragzone.base.html, e a imagem de desenho Tina foi conseguida no site http://www.monica.com.br .

5.1 MEDIDAS DE GANHO DE COMPACTAÇÃO (GC_N)

Analisou-se a compactação de energia dos bancos de filtros no domínio da transformada. Para isso, foram utilizados os filtros QMF's descritos na Tabela 4.1 do capítulo anterior e os filtros da Fig. 2.17. Os filtros QMF's foram testados em duas situações. Primeiro, utilizou-se a convolução circular e, em seguida, o método da extensão simétrica. Os filtros QMF's foram utilizados no esquema de decomposição da Fig. 2.9 e o filtro MED 2 foi utilizado no esquema de decomposição da Fig. 2.18. O número máximo de estágios foi feito igual a quatro, pois, a partir do quarto estágio, não se têm ganhos significativos de compactação de energia, como pode ser verificado na análise dos gráficos de ganho.

O método da extensão simétrica oferece um ligeiro ganho em relação à convolução circular [SMIT91]. Porém, no caso de imagens que possuem uma grande quantidade de bordas (componentes de altas freqüências) não se tem um aumento significativo. A complexidade computacional dos dois métodos é idêntica.

Nas Tabelas 5.1 a 5.5, apresentam-se os ganhos de compactação de cada imagem testada com os respectivos gráficos de ganho de compactação em função do número de estágios da decomposição. São mostradas as imagens originais com suas respectivas tabelas de ganho de compactação e gráficos de GC_N x Número de Estágios.

Figura 5.1 – Imagem original *Lena* (512x512).

Na Fig. 5.1, temos a imagem original de *Lena* e a Tabela 5.1 demonstra os resultados referentes aos Ganhos de Compactação (GC_N) dos bancos de filtros utilizados na decomposição dessa imagem. Na Fig. 5.2, apresenta-se o gráfico referente à Tabela 5.1.

Tabela 5.1 – Ganho de compactação de cada filtro para a imagem *Lena*

Filtro / Estágio	1	2	3	4
QMF(5/3)CIR	7,1156	9,0900	9,3650	9,4044
QMF(5/3)EXT SIM	7,5597	9,7102	10,004	10,0523
QMF(9/7)CIR	6,2149	8,2846	8,6876	8,7498
QMF(9/7)EXT SIM	6,7299	9,0546	9,4953	9,5714
MED 2	7,3173	9,4003	9,6682	9,6907
MED 4	4,4292	5,4133	5,5479	5,5693
MED 6	4,9973	6,2471	6,4223	6,4359

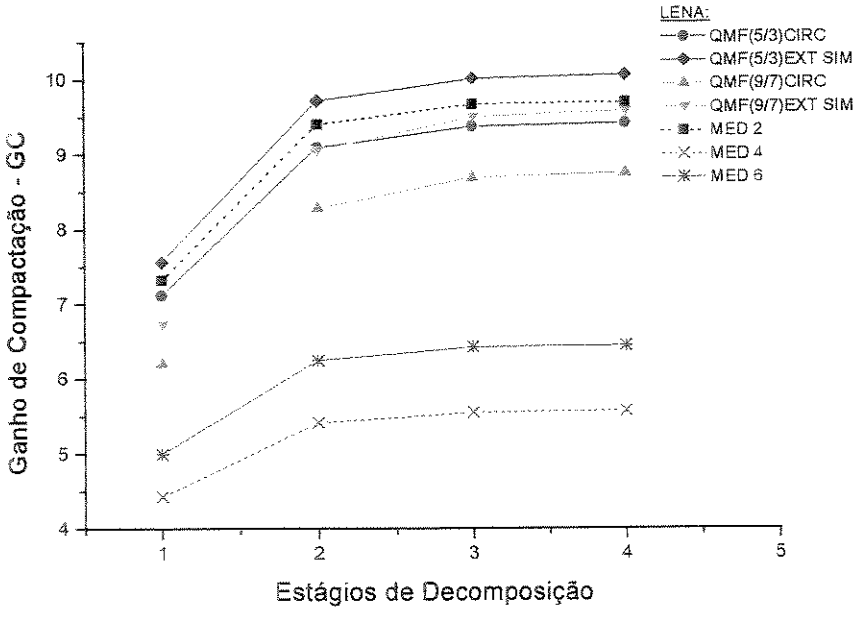


Figura 5.2 – Gráfico de GC x Estágios de Decomposição para a imagem *Lena*.

Na Fig. 5.3, é apresentada a imagem original *Salesman* e a Tabela 5.2 demonstra os resultados referentes aos Ganhos de Compactação (GC_N). Na Fig. 5.4, apresenta-se o gráfico referente à Tabela 5.2.

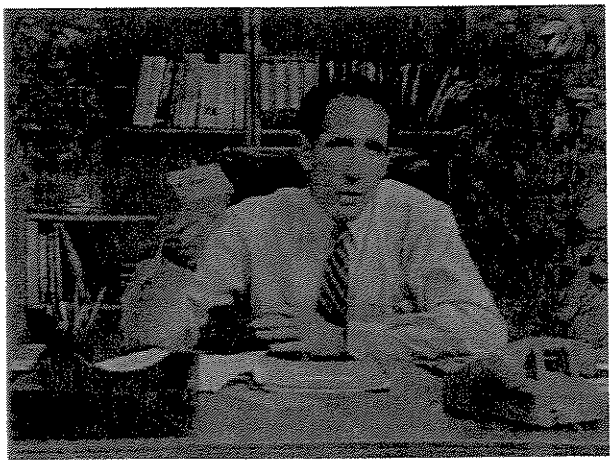


Figura 5.3 – Imagem original *Salesman* (352x288).

Tabela 5.2 – Ganho de compactação de cada filtro para a imagem *Salesman*

Filtro / Estágio	1	2	3	4
QMF(5/3)CIR	3,8835	4,4507	4,5131	4,5200
QMF(5/3)EXT SIM	4,1311	4,7409	4,8085	4,8176
QMF(9/7)CIR	3,2075	3,7056	3,7950	3,8102
QMF(9/7)EXT SIM	3,4005	3,9505	4,0493	4,0645
MED 2	3,9001	4,4805	4,5509	4,5603
MED 4	2,6032	2,8953	2,9321	2,9345
MED 6	2,0819	2,2495	2,2761	2,2779

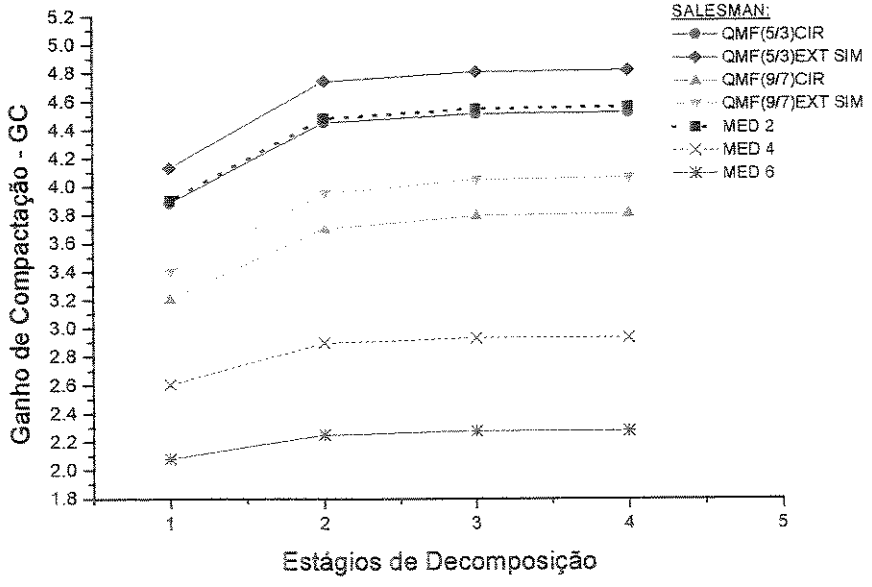


Figura 5.4 – Gráfico de GC x Estágios de Decomposição para a imagem *Salesman*.

Na Fig. 5.5, apresenta-se a imagem original *Sail* e a Tabela 5.3 demonstra os resultados referentes aos Ganhos de Compactação (GC_N). Na Fig. 5.6, apresenta-se o gráfico referente à Tabela 5.3.

Figura 5.5 – Imagem original *Sail* (512x768).

Tabela 5.3 – Ganho de compactação de cada filtro para a imagem *Sail*

Filtro / Estágio	1	2	3	4
QMF(5/3)CIR	2,5002	2,6242	2,6393	2,6405
QMF(5/3)EXT SIM	2,5092	2,6320	2,6477	2,6491
QMF(9/7)CIR	2,0471	2,1831	2,2090	2,2125
QMF(9/7)EXT SIM	2,0565	2,1917	2,2204	2,2238
MED 2	2,4672	2,5881	2,5998	2,5981
MED 4	1,6598	1,7313	1,7381	1,7387
MED 6	1,8742	1,9742	1,9843	1,9853

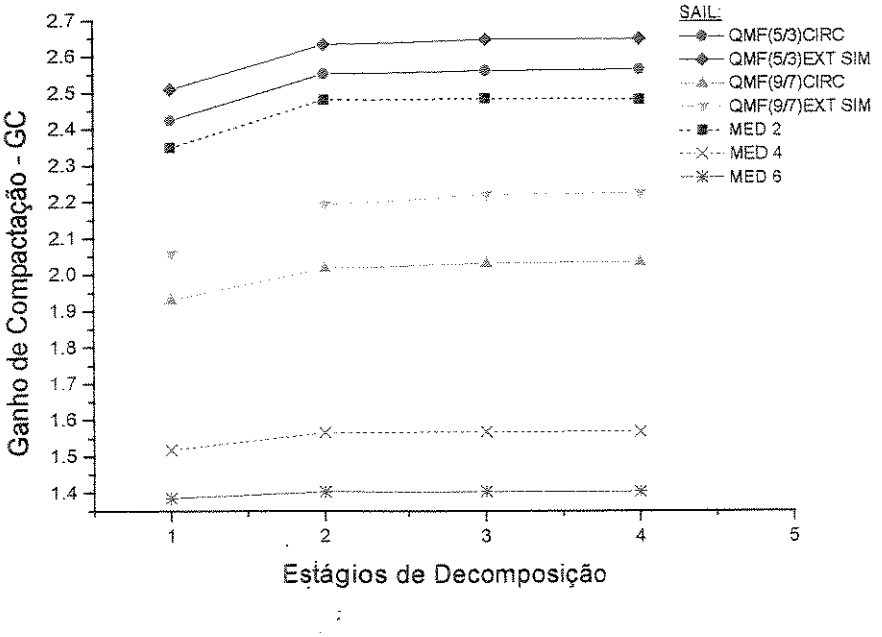


Figura 5.6 – Gráfico de GC x Estágios de Decomposição para a imagem *Sail*

Na Fig. 5.7, é apresentada a imagem original *Monarch* e a Tabela 5.4 demonstra os resultados referentes aos Ganhos de Compactação (GC_N). Na Fig. 5.8, apresenta-se o gráfico referente à Tabela 5.4.

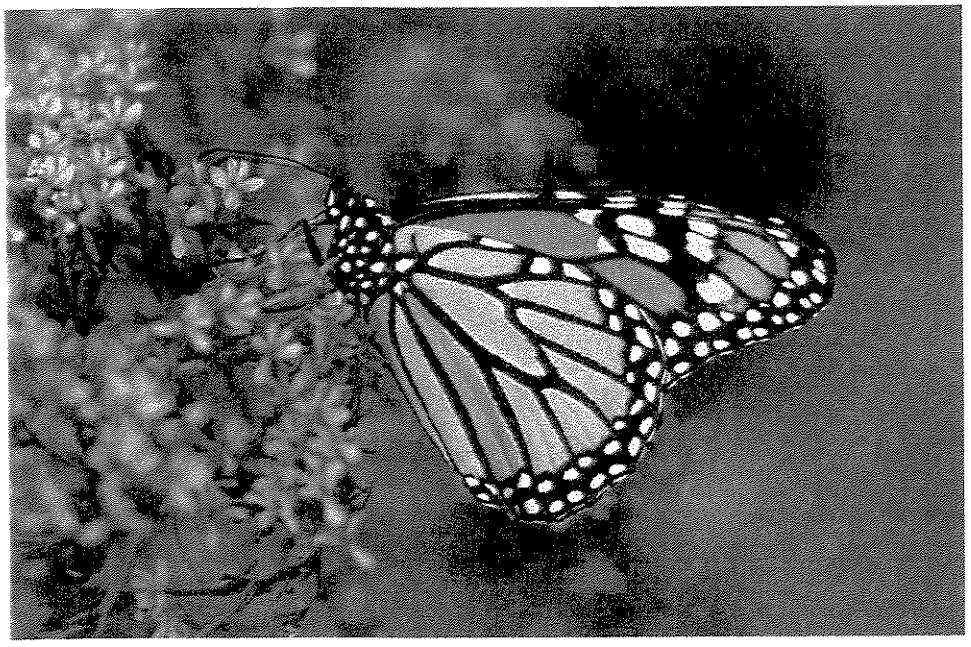


Figura 5.7 – Imagem original *Monarch* (512x768).

Tabela 5.4 – Ganho de compactação de cada filtro para a imagem *Monarch*

Filtro / Estágio	1	2	3	4
QMF(5/3)CIR	4,5152	5,0914	5,1315	5,1336
QMF(5/3)EXT SIM	4,5302	5,1043	5,1425	5,1443
QMF(9/7)CIR	3,7409	4,2833	4,3361	4,3392
QMF(9/7)EXT SIM	3,7734	4,3301	4,3810	4,3845
MED 2	4,4889	5,0402	5,0784	5,0767
MED 4	2,7880	3,0050	3,0093	3,0096
MED 6	3,6726	4,1391	4,1631	4,1635

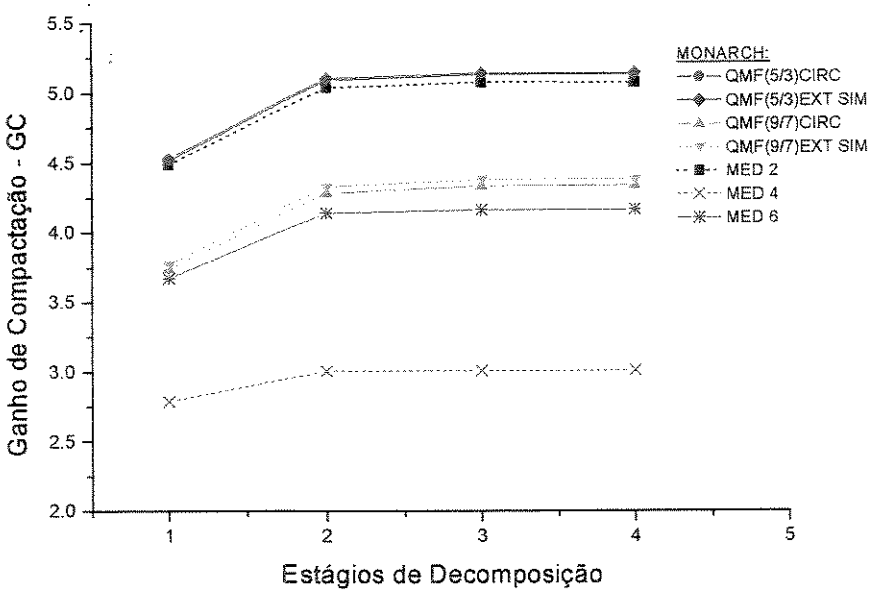


Figura 5.8 – Gráfico de GC x Estágios de Decomposição para imagem *Monarch*

Na Fig. 5.9, é apresentada a imagem original *Tina* e a Tabela 5.5 demonstra os resultados referentes aos Ganhos de Compactação (GC_N). Na Fig. 5.10, apresenta-se o gráfico referente à Tabela 5.5.

Figura 5.9 – Imagem original *Tina* (256x160)

Tabela 5.5 – Ganho de compactação de cada filtro para a imagem *Tina*

Filtro / Estágio	1	2	3	4
QMF(5/3)CIR	1,5854	1,6063	1,6133	1,6143
QMF(5/3)EXT SIM	1,5844	1,6001	1,6067	1,6080
QMF(9/7)CIR	1,3184	1,3482	1,3614	1,3633
QMF(9/7)EXT SIM	1,3193	1,3462	1,3560	1,3581
MED 2	1,5453	1,5601	1,5622	1,5586
MED 4	1,1978	1,210	1,2117	1,2117
MED 6	1,2374	1,2507	1,2530	1,2532

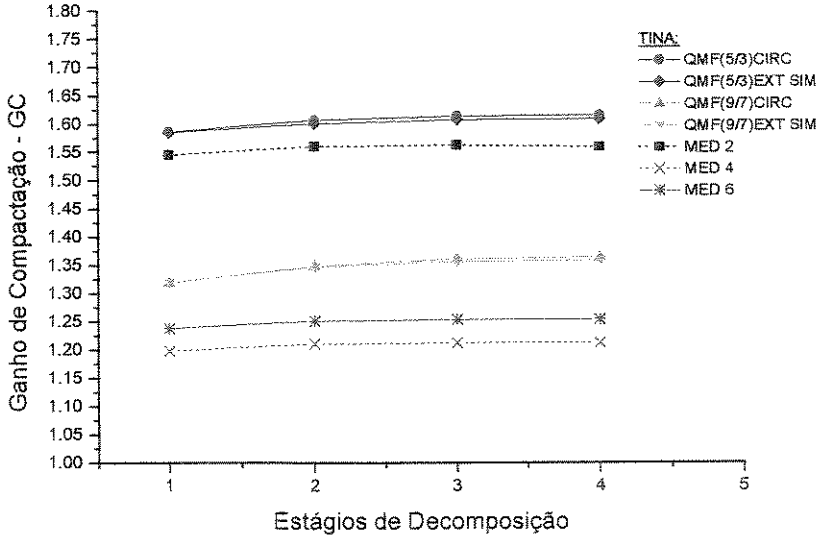


Figura 5.10 – Gráfico de GC x Estágios de Decomposição para a imagem *Tina*

Nos gráficos de ganho de compactação *versus* estágios de decomposição, verifica-se que o filtro QMF (5x3) utilizando extensão simétrica na decomposição, foi o banco de filtro que obteve o melhores resultados, no entanto o banco MED 2 obteve resultados muito próximos de ganho de compactação.

No caso da imagem *Sail* não se verifica diferenças significativas de ganho com relação aos bancos QMF's que utilizaram convolução circular e extensão simétrica. Esta imagem possui uma grande quantidade de bordas (componentes de altas freqüências), e portanto, não há grandes diferenças em valores de ganho de um método em relação ao outro [SMIT91].

5.2 SIMULAÇÕES COM COMPRESSÃO DE IMAGENS VIA BANCO DE FILTROS

São apresentados agora os gráficos de desempenho *versus* taxa de compressão para o sistema simulado de acordo com a Fig. 4.1. É utilizada uma medida objetiva de desempenho, a razão sinal ruído de pico, PSNR. As medidas de taxa de compressão referem-se às taxas alcançadas após a quantização em bits por pixel e à entropia dos coeficientes quantizados.

Nas Figuras 5.11 a 5.15, são mostrados os gráficos para as cinco imagens teste. A decomposição em sub-bandas é realizada em três níveis, utilizando apenas três bancos de filtros: QMF 9x7, QMF 5x3 e MED 2. Esses três bancos de filtros foram os que apresentaram melhor compactação de energia na decomposição. Os filtros QMF's foram implementados com o método da extensão simétrica, pois este teve um desempenho um pouco melhor que o método da convolução circular. Os parâmetros de alocação de bits R_C variam entre 0 e 8, a passos de 0,5. Nos gráficos, os quadrados representam o banco de filtros QMF 9x7, os triângulos representam o banco de filtros QMF 5x3 e os círculos representa o banco de filtros MED 2.

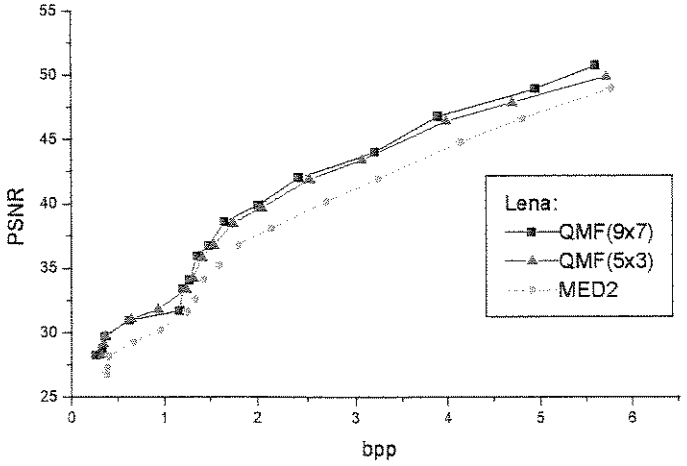


Figura 5.11 – Curva de PSNR *versus* bits por pixel da imagem *Lena*.

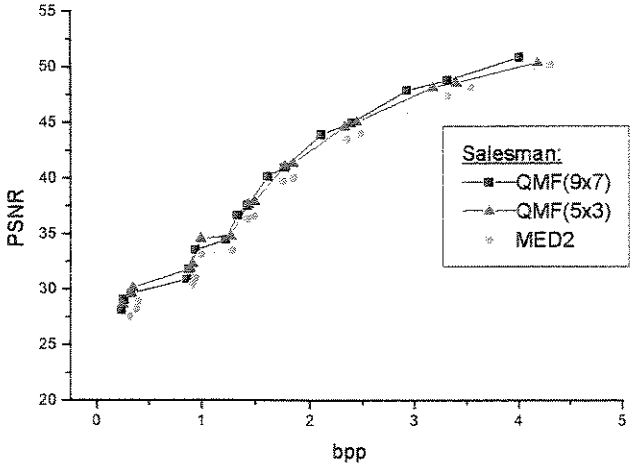


Figura 5.12 – Curva de PSNR *versus* bits por pixel da imagem *Salesman*.

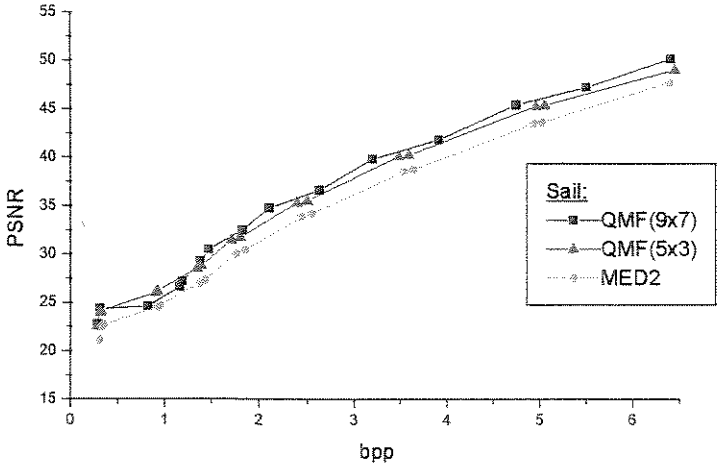


Figura 5.13 – Curva de PSNR *versus* bits por pixel da imagem *Sail*.

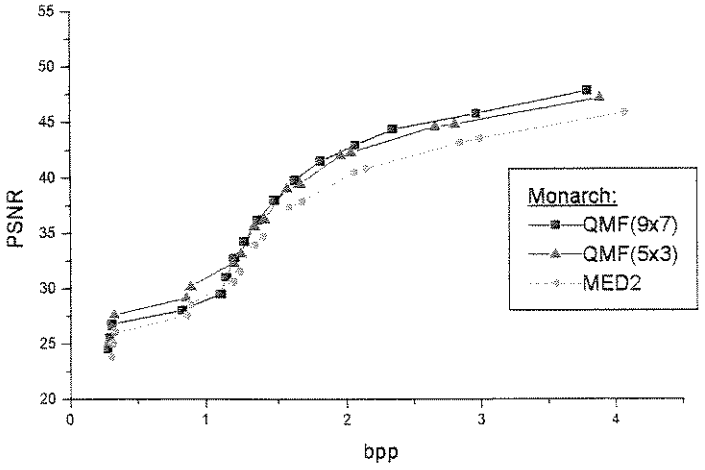


Figura 5.14 – Curva de PSNR *versus* bits por pixel da imagem *Sail*.

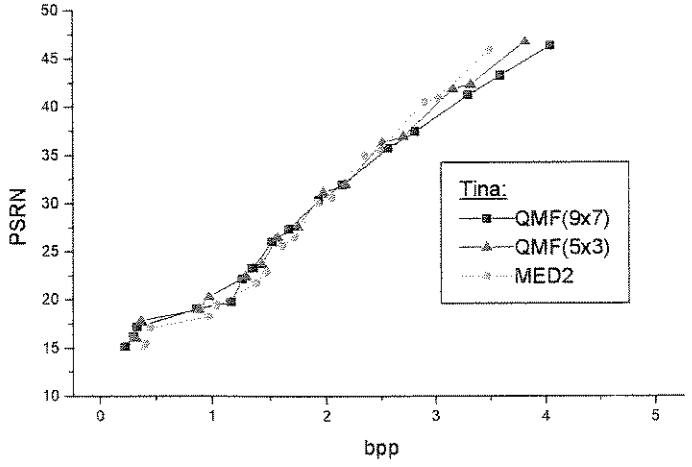
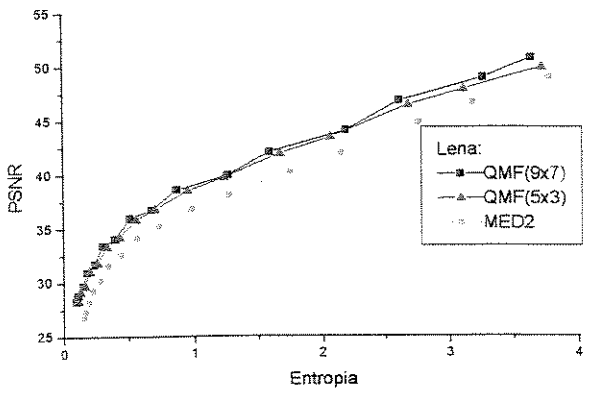
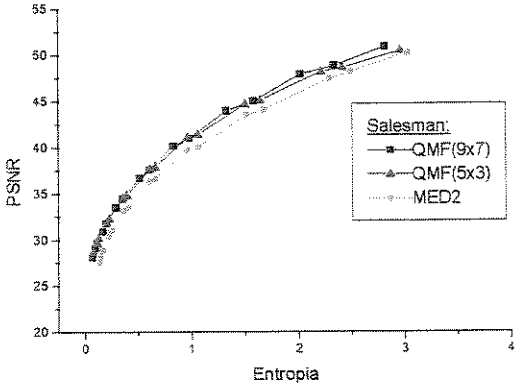


Figura 5.15 – Curva de PSNR *versus* bits por pixel da imagem *Sail*

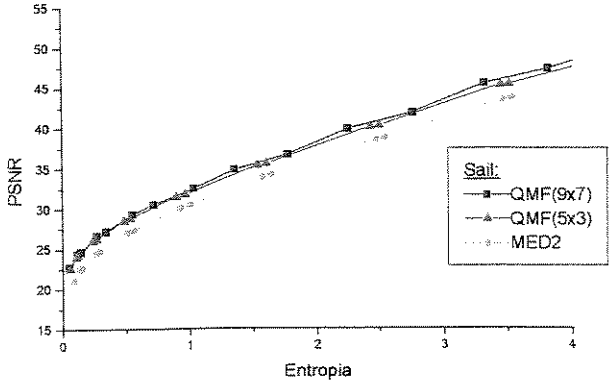
Nas Figuras 5.16a-e são mostrados os gráficos de PSNR *versus* entropia das cinco imagens de teste. É importante ressaltar que a entropia dos coeficientes quantizados funciona como uma medida de taxa de compressão, conforme a Seção 3.3.4. Na verdade, ela é uma estimativa do que seria alcançado com uma codificação por entropia de comprimento variável, como a codificação de Huffman, por exemplo.



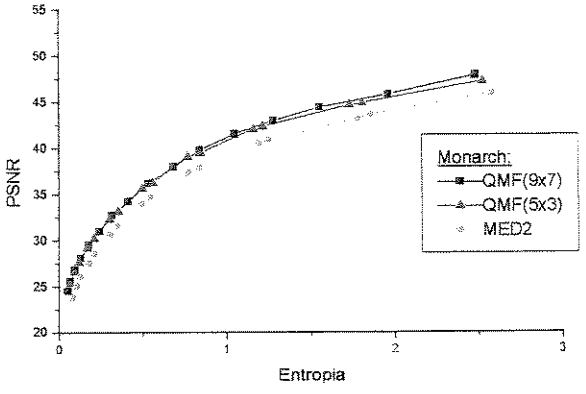
(a)



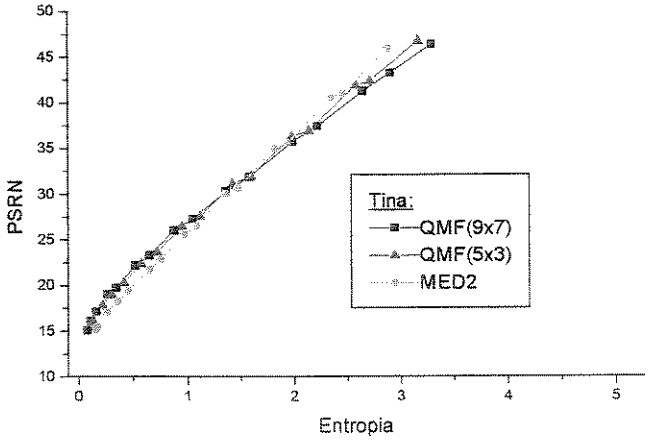
(b)



(c)



(d)



(e)

Figura 5.15 – Curvas de PSNR versus entropia para as imagens (a) *Lena*, (b) *Salesman*, (c) *Sail*, (d) *Monarch*, (e) *Tina*.

Observa-se pelos gráficos que as curvas PSNR *versus* entropia, se mostram bastante parecidas com os gráficos de PSNR *versus* bpp, isto demonstra que em termos de desempenho, não ocorreram mudanças significativas, apenas uma diminuição do PSNR proporcional para cada banco.

5.3 REDUÇÃO DA COMPLEXIDADE COMPUTACIONAL

Outro aspecto analisado foi a complexidade computacional das decomposições e reconstruções que são mostrados na Tabela 5.6. O filtro MED 2 obteve um melhor desempenho de compactação em relação aos dois outros filtros medianos. Por isso, a análise de complexidade computacional é feita comparando-se o número de adições e multiplicações realizadas pelo filtro MED 2 em relação aos números referentes aos filtros QMF. Observa-se que os números de adições e multiplicações do MED 2 são significativamente menores, tanto na decomposição quanto na reconstrução da imagem.

A forma de cálculo para obter os valores de adições e multiplicações, encontra-se na Seção 4.2 do capítulo anterior.

Tabela 5.6 – Tabela referente à porcentagem de operações de somas e multiplicações que o banco de filtros MED 2 realiza em relação aos bancos com filtros QMF.

	+ Análise	x Análise	+ Síntese	x Síntese
QMF(9/7)	14,3%	6,3%	4,4%	2,1%
QMF(5/3)	33,3%	12,5%	9,5%	4,2%

+ Análise = número de adições na análise (decomposição);

x Análise = número de multiplicações na análise;

+ Síntese = número de adições na síntese (reconstrução);

x Síntese = número de multiplicações na síntese.

CAPÍTULO 6

CONSIDERAÇÕES FINAIS E SUGESTÕES DE TRABALHOS

Neste capítulo, faz-se uma revisão geral do trabalho realizado e, simultaneamente, apontam-se algumas direções de investigação futura. O trabalho descrito ao longo desta dissertação centrou-se no estudo e na aplicação das teorias dos bancos de filtros lineares e não-lineares com capacidade de reconstrução perfeita. Visava-se, especialmente, analisar a compactação de energia no domínio da transformada e o desempenho destes bancos de filtros aplicados a um sistema básico de compressão de imagem.

Foi proposta neste estudo uma medida empírica de compactação de energia dos coeficientes de decomposições em sub-bandas. Em muitos casos, esta medida é igual à medida de Katto e Yasuda [KATT91], que representa o ganho de codificação de sub-bandas em relação à codificação PCM. A medida empírica apresenta a vantagem de poder ser aplicada a bancos de filtros não-lineares. Esta vantagem da utilização da medida empírica de ganho de compactação proposta neste trabalho em relação à expressão de Katto e Yasuda, ocorre porque esta última depende dos coeficientes dos filtros de síntese e, portanto, não se aplica facilmente a bancos de filtros não-lineares. Por outro lado, a medida proposta no trabalho depende exclusivamente das variâncias das sub-bandas geradas pela decomposição e independe dos coeficientes dos banco de

filtros utilizados. Assim, estimam-se as variâncias das sub-bandas e calcula-se o ganho de compactação para o banco de filtros em questão, seja este banco linear ou não-linear.

Utilizou-se uma estrutura de decomposição e reconstrução perfeita com bancos de filtros lineares e não-lineares. O banco de filtros MED 2, apesar de possuir características lineares, é uma extensão dos filtros não-lineares MED 4 e MED 6 e foi implementado na mesma estrutura de decomposição que a dos filtros não-lineares. O banco MED 2, devido à sua simplicidade e fácil implementação, apresenta uma complexidade computacional muito menor do que a dos filtros QMF, como se ilustrou na Tabela 5.6.

Com relação ao ganho de compactação de energia, o banco de filtros QMF (5/3) com extensão simétrica apresentou os melhores resultados. Não obstante, o banco MED 2 apresentou ganhos muito próximos em quase todos os estágios das imagens testadas, apesar de sua reduzida complexidade computacional. Observou-se também que os bancos de filtros de mediana utilizados, MED 4 e MED 6, não apresentaram desempenhos satisfatórios em comparação aos bancos de filtros QMF. É importante ressaltar que

Após a análise de compactação de energia, os bancos que ofereceram um melhor desempenho (MED 2, QMF 5x3 e QMF 9x7) foram implementados em um sistema básico de compressão. Verificou-se que as imagens reconstruídas com os filtros QMF's tiveram um ligeiro ganho em relação ao MED 2. Porém, no caso da imagem *Tina*, sem textura, apenas com bordas e regiões homogêneas, o desempenho do banco MED 2 na região de quantização fina foi superior a dos QMF's. E na região com quantização grosseira, o desempenho foi muito próximo. Por exemplo, para esta imagem (Apêndice, Fig. A.5(a) e (b)), pôde-se verificar que a distorção por efeito de *ringing* nas imagens reconstruídas com os bancos QMF's foi visível. Também se pôde verificar isso nas Figuras A.4(c) e (d) (Apêndice). Já as imagens reconstruídas com o banco MED 2 não apresentaram a distorção de *ringing*.

Portanto, o filtro MED 2 demonstra ser uma ferramenta de grande potencial nas aplicações de compressão de imagens, principalmente levando-se em conta o fato de ele possuir uma complexidade computacional significativamente menor que os bancos QMF's.

Como sugestão para trabalhos futuros, o banco de filtros MED 2 pode ter resultados melhores quando utilizado em imagens com pouca incidência de textura que é o caso dos desenhos animados, apesar deste tipo de imagem não ser natural, bons resultados podem ser conseguidos com reduzida complexidade computacional. Uma outra sugestão seria utilizar o

banco de filtro não-lineares (neste caso pode-se incluir o MED 2 que é linear mas aplicado a uma estrutura não-linear , para utilizar na predição de movimento na codificação de vídeo, pois, conjectura-se que as imagens diferenças resultante da subtração do quadro original e seu predito, possuem características de alta freqüência (detalhes) e os filtros lineares usuais não são adequados para filtrar imagens com estas características. No caso do filtro MED 2, (que é linear, mas aplicado a uma estrutura não-linear), trabalhando com duas amostras, pode-se prover resultados interessantes.

APÊNDICE DE IMAGENS

Neste Apêndice, procura-se mostrar, com mais detalhes, as imagens tratadas neste trabalho. No entanto, essas imagens impressas podem ter sofrido pequenas distorções decorrentes do processo de impressão.

Fig. A.1 (a) – Codificação por sub-bandas com filtro QMF 5x3 a 1.30 bpp, Entropia=0,42 e PSNR=34,18.

Fig. A.1 (b) – Codificação por sub-bandas com filtro MED 2 a 1.41 bpp, Entropia=0,56 e PSNR=34,14.

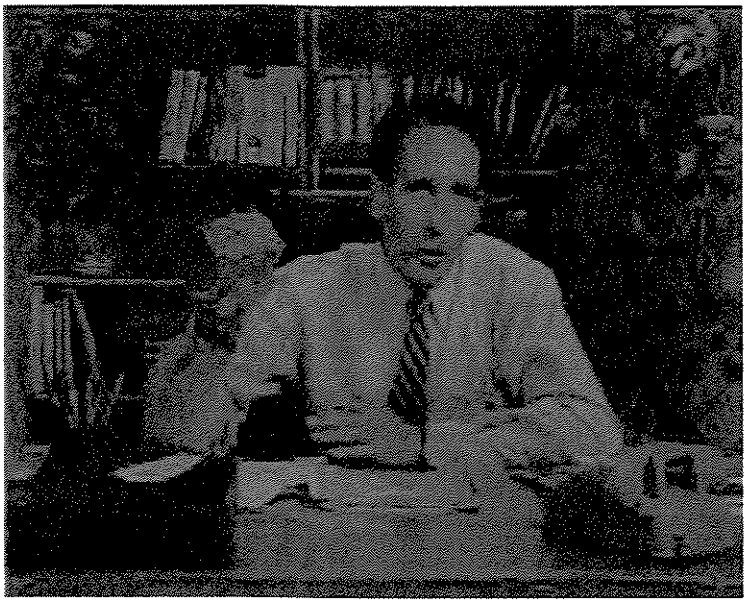


Fig. A.2 (a) – Codificação por sub-bandas com filtro QMF 9x7 a 1.01 bpp, Entropia=0,39 e PSNR=31,62.

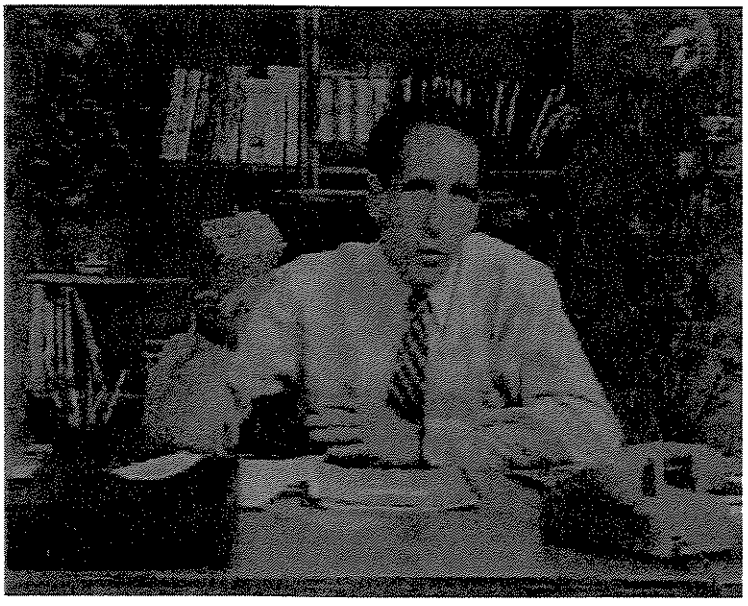


Fig. A.2 (b) – Codificação por sub-bandas com filtro MED 2 a 1.08 bpp, Entropia=0,48 e PSNR=31,22.

Fig. A.3 (a) – Codificação por sub-bandas com filtro QMF 5x3 a 1.33 bpp, Entropia=0,48 e PSNR=28,4.

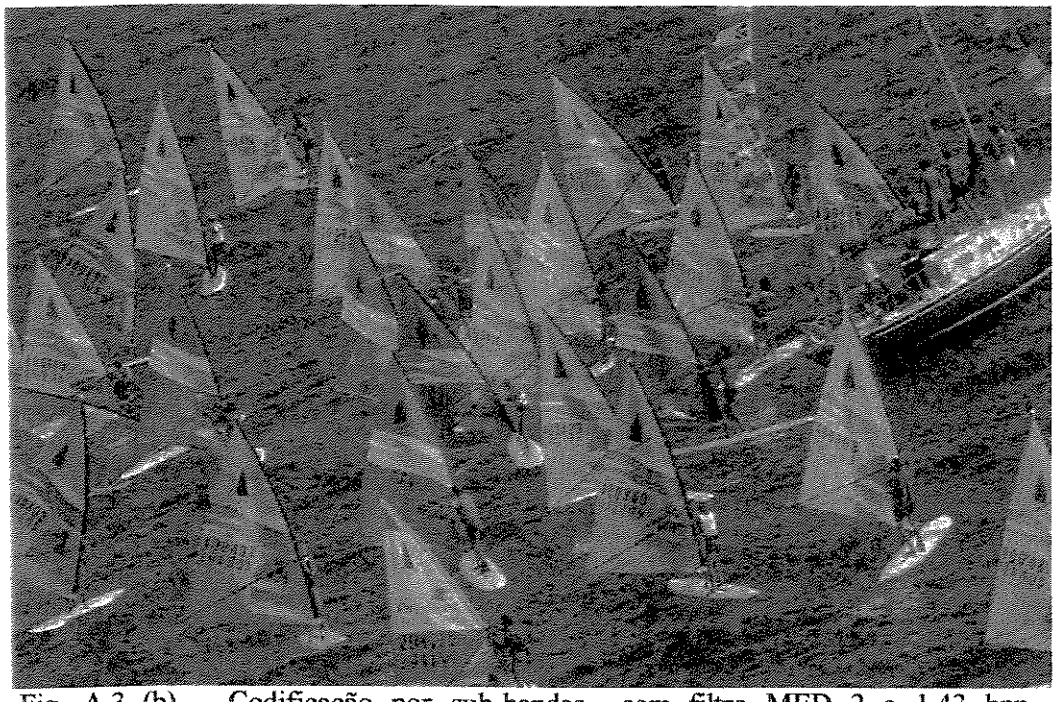


Fig. A.3 (b) – Codificação por sub-bandas com filtro MED 2 a 1.43 bpp, Entropia=0,57 e PSNR=27,34.

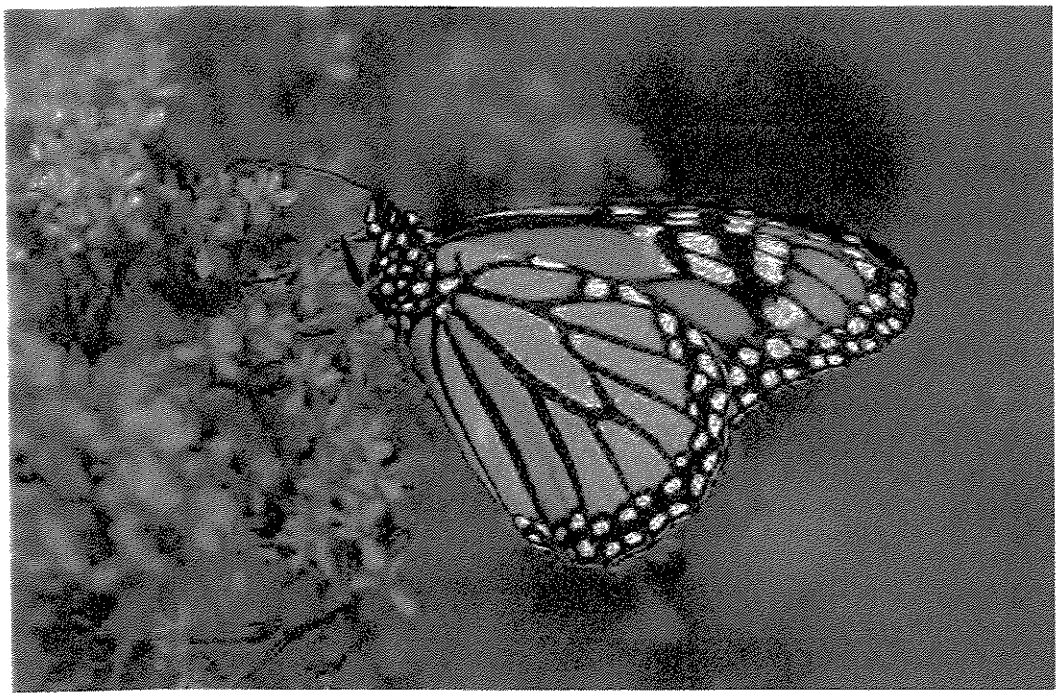


Fig. A.4 (a) – Codificação por sub-bandas com filtro QMF 9x7 a 1.12 bpp, Entropia=0,19 e PSNR=29,49.

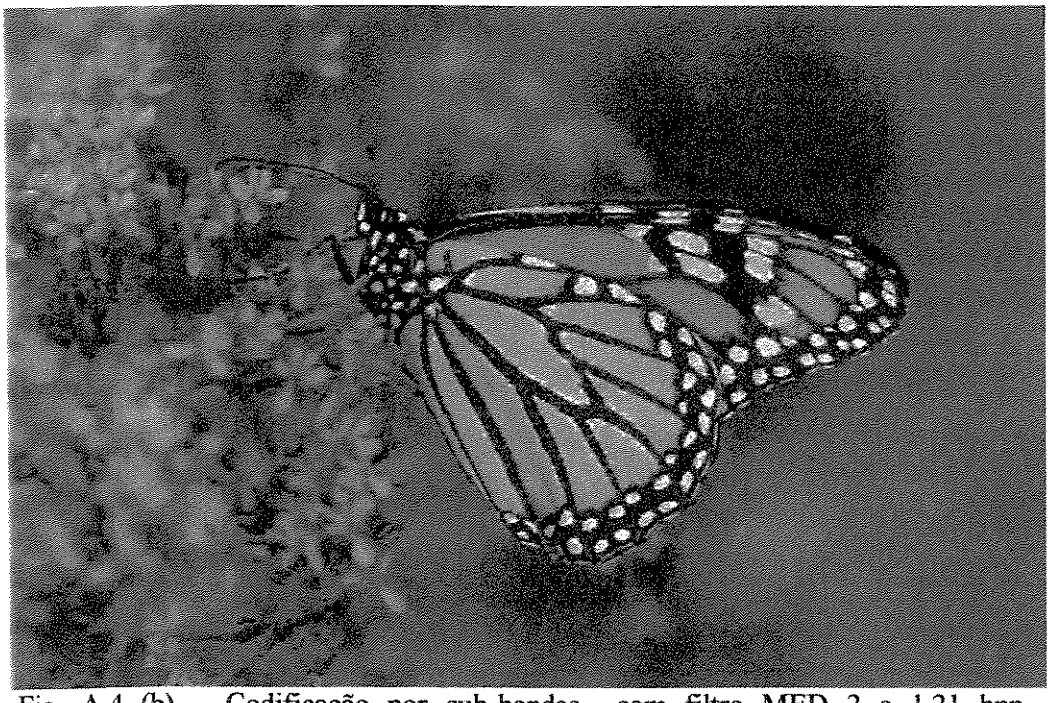


Fig. A.4 (b) – Codificação por sub-bandas com filtro MED 2 a 1.21 bpp, Entropia=0,33 e PSNR=30,66.

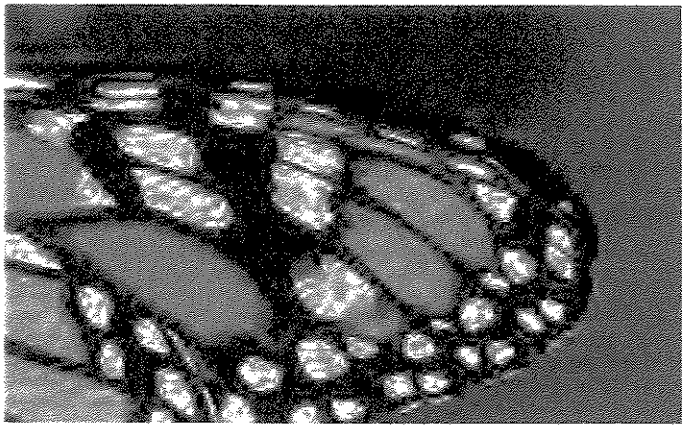


Fig A.4(c) Detalhe da Figura A.4(a)

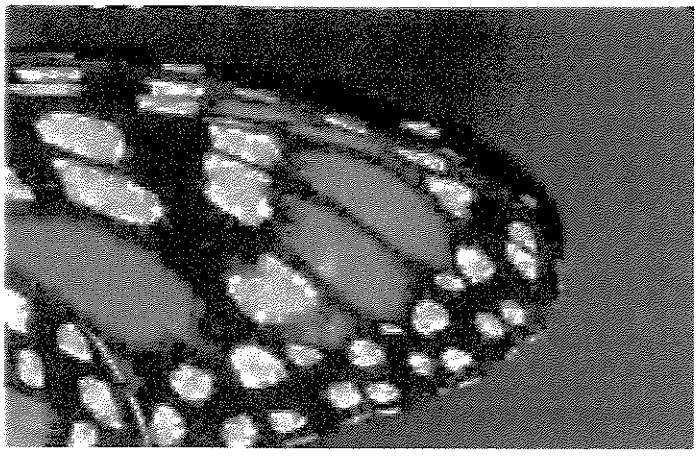


Fig A.4(d) Detalhe da Figura A.4(b)

Fig. A.5 (a) – Codificação por sub-bandas com filtro QMF 9x7 a 1.16 bpp, Entropia=0,37 e PSNR=20,3.

Fig. A.5 (b) – Codificação por sub-bandas com filtro MED 2 a 1.04 bpp, Entropia=0,46 e PSNR=19,86.

BIBLIOGRAFIA

[AKAN92] AKANSU, Ali, HADDAD Richard, Multiresolution Signal Decomposition: transforms, Subbands and Wavelets, Academic Press, 1992.

[ANTO92] ANTONINI Marc, DAUBECHIES Ingrid, "Image coding using wavelet transform", IEEE trans. on Signal Processing, vol. 1, nº 2, pp. 205-220, abril de 1992

[BERG00] BERGAMO Rubem T., SOUZA Felipe A. O., COSTA Max H. M., YABU-UTI João B. T., "Análise da Compactação de Energia na Representação de Imagens por Decomposição em Sub-bandas". Anais do XVII Simpósio Brasileiro de Telecomunicações. Gramado, RS, setembro de 2000.

[BRIS95] BRISLAWN C. M., "Preservation of subband symetry in multirate signal coding", IEEE Trans. on Signal Processing, vol. 43, nº 12, pp. 1-4, dezembro de 1995.

[CCIR82] INTERNATIONAL RADIO CONSULTATIVE COMMITTEE, Encoding parameters of digital television for studios, Recommendation 601-1, 1982.

[CROC76] CROCHIERE R., WEBER S., FLANAGAN F., "Digital coding of speech in subbands", Bell Systems Tech. J.,pp.1069-1085, 1976.

[EBRA98] EBRAHIMI T., KUNT M., "Visual data compression for multimedia applications", Proccedings of the IEEE, vol. 86, nº 6, pp.1109-1125, junho 1998.

EGGE95a] EGGER O., LI W., KUNT M., "High compression image coding using na adaptive morphological subband decomposition", *Proceedings of the IEEE*, vol.83, pp.272-287, fevereiro de 1995.

[EGGE95b] EGGER O., LI W., "Sub-band coding of images using asymmetrical filter banks", *IEEE Trans. on Image Processing*, vol. 4, nº 4, pp. 478-485, abril de 1995.

[GERS92] GERSHO Allen, GRAY Robert, *Vector quantization and signal compression*, Kluwer Academic, 1992.

[GHAR88] GHARAVI H., TABATABAI A., "Subband coding of monochrome and color images", *IEEE Trans. on Circuits and Systems*, vol. 35, nº 2, pp. 207-214, fevereiro de 1988.

[JAIN88] JAIN A. K., *Fundamental of Digital image Processing*, Englewood Cliffs, NJ: Prentice Hall, 1988.

[JAYA84] JAYANT, N. S., NOLL Peter. Digital coding of waveforms: Principles and applications to speech and vídeo, Prentice-Hall, 1984.

[KARL89] KARLSSON, G., VETTERLI, M. "Extension of finite length signals for subband coding", *Signal Processing*, nº 17, pp. 161-168, 1989.

[KATT91] KATTO J., YASUDA Y., "Performance evaluation of subband coding and optimization of its filter coefficients", *SPIE Proceedings of Visual Comunications and Image Processing*, vol. 1605, pp. 95-106, novembro de 1991

[LIM89] LIM, J. S. *Two-Dimensional Signal and Image Processing*, Englewood Cliffs, NJ: Prentice Hall, 1989.

[NETR88] NETRAVALI A. N., HASKEL B. G., *Digital Pictures: Representation and Compression Aplications of Communications Theory*. In: R. W. Lucky, Series Editor. New York: Plenum, 1988.

[OPPE89] OPPENHEIN, Alan, SCHAFER, Ronald. *Digital Signal Processing,* Prentice hall, 1989.

[PEAR91] PEARLMAN William. "Performance bounds for subband coding", *Subband Image Coding*, editado por John Woods, Kluwer Academic Publishers, 1991.

[QUEI92] QUEIROZ, Ricardo, MALVAR, Henrique. " On the asynptotic performance of hierarchical transforms" . *IEEE Trans. On Circuits and Systens for Video Technolog*, vol. 40, nº 10, pp. 2620-2622, outubro de 1992.

[SAID96] SAID, A., PEARLMAN,W. "A new, fast and efficient image codec based on set partitioning in hierarchical trees", *IEEE Trans. on Circuits and Systens for Video Technolog*, vol. 6, nº 3, pp. 243-250, junho de 1996.

[SALE92] SALEMBIER, P., KUNT, M. "Size-sensitive multiresolution decomposition of images with rank order besed filters", *Signal Processing*, vol. 27, pp. 205-241, 1992.

[SHAP93] SHAPIRO,J. M. "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans. on Signal Processing*, vol. 41, nº 12, dezembro de 1993.

[SIMO91] SIMONCELLI, E. P., ADELSON, E. H. "Subband Transforms", *Subband Image Coding*, editado por John Woods, Kluwer Academic Publishers, 1991.

[SMIT86] SMITH, M., BARNWELL, T. "Exact reconstruction techniques for tree estrutured subband coders", *IEEE Trans. on ASSP*, vol. 34, nº 3, junho de 1986.

[SMIT91] SMITH, M. "IIR analysis/synthesis Systems", *Subband Image Coding*, editado por John Woods, Kluwer Academic Publishers, 1991.

[SOUZ00] SOUZA, Felipe. A. O." Um estudo sobre a técnica de subamostragem Quincunx e sua aplicação na codificação de imagens", Dissertação de Mestrado, Universidade Estadual de Campinas, junho de 2000.

[STRA96] STRANG, G., NGUYEN, T. *Wavelets and bank Filters*, Wellesley-Cambridge Press, 1996.

[TOET89] TOET, A. "A morphological pyramid image decomposition," *Patt. Recog. Letter*, vol. 9, pp. 255-261, 1989.

[VAID87] VAIDYANATHAN, P. P. "Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques", *IEEE ASSP Magazine*, julho de 1987.

[VETT95] VETTERLI, M., KOVACEVIC, J.. *Wavelets and subband coding*, Prentice Hall, 1995.

[VILL95] VILLASENOR, John, BELZER, Benjamin, LIAO, Judy. "Wavelet filter evaluation for image compression", *IEEE Trans. on Image Processing*, vol. 4, nº 8, pp. 1053-1060, agosto de 1995.

[WOOD86] WOODS, J., O'NEIL, S. " Subband coding of images", *IEEE Trans. on ASSP*, vol. 34, nº 5, outubro de 1986.

[YABU91] YABU-UTI, J. B. T. "Técnicas de Codificação do sinal de televisão PAL-M", Relatório Técnico – Contrato Telebrás (387/90). Publicação FEE 40/91, dezembro de 1991.

[ZHOU92] ZHOUAND, Z., VENETSANOPOULOS, A. N. "Morphological methods in image coding", *Proc. Int. Conf. Acoust., Speech and Signal Processing*, vol. 3, pp. 481-484, março de 1992.