

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA ELÉTRICA

DEPARTAMENTO DE COMUNICAÇÕES

**COMPRESSÃO ESTATÍSTICA DE
IMAGENS ESTÁTICAS USANDO
TRANSFORMADA COSSENO
DISCRETA EM DUAS DIMENSÕES
COM ALGORITMO DE BLOCO
VARIÁVEL ADAPTATIVO**

autor: Guillermo Ramón Fernández Segovia

orientador: Luiz César Martini

Este exemplar corresponde a edição final da tese
defendida por GUILLERMO R. FERNÁNDEZ
SEGÓVIA aprovada pela Comissão
Julgadora em 29 / 10 / 93.

Luiz César Martini
Orientador

Dissertação submetida à Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas, para preenchimento dos pré-requisitos parciais para obtenção do Título de Doutor em Ciências.

outubro 1993

Resumo

O objetivo desta dissertação é apresentar um método eficiente de compressão de imagem aplicado na codificação digital dos sinais na forma de componentes Y , C_R , C_B , usando um algoritmo adaptativo de bloco variável aplicado à Transformada DCT em duas dimensões. Visam-se os defeitos usando o algoritmo de codificação por DCT convencional que trata todas as áreas segmentadas em blocos fixos não sobrepostos em forma indiscriminada produzindo uma qualidade não uniforme para diferentes conteúdos de imagem. Emprega-se uma recente técnica de segmentação por árvore. Os resultados são promissores para transmissão de imagens paradas com o melhor compromisso entre a qualidade de imagem e a codificação eficiente.

Abstract

This work intends to show an efficient coding method of images using transform coding DCT-2D as a key element for image data compression. Conventional DCT coding algorithms, treat all the image areas indiscriminately, causing nonuniform image quality for various image contents, consequently this motivates works on DCT schemes adaptive to the images contents so that a better tradeoff between bit rate and image quality can be achieved. In this work, a DCT-2D algorithm with adaptive zone and quantization is combined with variable block size segmentation to further improve the coding performance. A recent segmentation technique is used. The results are given to illustrate the superiority of this adaptive DCT algorithm with segmented blocks and this algorithm poses itself as a promising compression method to deal with transmission of good quality still images.

Agradecimentos

Ao Prof. Dr. Luiz César Martini, pela amizade e orientação deste trabalho.

Aos Professores Yabu-Uti, Yuzo, Afonso da equipe PDS/DECOM/FEE e ao Professor Michel DECOM/FEE pelas consultas e apoio.

Aos Engenheiros Hélio Marcos Machado Graciosa e Carlos Antonio Taube pelo apoio no convenio UCV-Telebrás.

Aos Professores da FEE, com os quais tive relacionamento.

Aos Colegas do curso de Doutorado e Amigos, Osvaldo, Sílvia, Isamara, Andréa, Nina, Marcelo, Sindi e Cabral pelo apoio e amizade.

Aos Colegas e Funcionários da Telebrás e da Unicamp, pelo apoio.

Aos Amigos, novos e antigos, pela amizade.

Aos Funcionários da FEE, pelo auxílio.

Ao CNPQ pelo suporte financeiro.

Ao Amigo César Augusto Rosales Córdova pela atenciosa colaboração técnica e amizade. “Hermano, que en tu paz celestial sepas que tus amigos te recordaran por siempre”.

E finalmente,

Graças a Deus.

Dedico este trabalho a minha família, como uma pequena compensação pelo tempo que deixamos de passar juntos durante a preparação e escrita desta Tese:
Minha esposa, Ximena e meus filhos, Ximenita, Sebastián e Paulinho.

Meus Pais pelo apoio espiritual.

Conteúdo

RESUMO	i
ABSTRACT	ii
AGRADECIMENTOS	iii
CONTEÚDO	v
1 CONSIDERAÇÕES GERAIS	1
1.1 Introdução	2
1.2 Motivação e Avaliação do Trabalho	2
1.3 Contribuição do Trabalho	5
1.4 Descrição dos Capítulos	5
2 POSSÍVEIS SOLUÇÕES DE CODIFICAÇÃO EFICIENTE	7
2.1 Introdução.	8
2.2 Codificação Preditiva	9
2.3 Codificação por Transformada	11
2.4 Codificação Híbrida	11
3 CONCEITOS BÁSICOS DO SINAL DE VÍDEO E DA CODIFICAÇÃO POR TRANSFORMADAS	13

3.1	Largura de Faixa e Capacidade de Canal.	14
3.2	Informação fonte	14
3.3	Futuro Sistema Multi-mídia e o Formato Intermediário Comum	16
3.4	Defeitos na Informação de Fonte	18
3.5	Representação Espectral da Imagem	19
3.5.1	Série de Fourier	20
3.5.2	Imagens de Faixa Limitada - Transformada de Fourier Discreta - Teorema da Amostragem Bidimensional	22
3.5.3	Espectro de Frequência de uma imagem amostrada	24
3.5.4	Espectro de Frequência para a Transformada Discreta Cosseno	29
3.6	Introdução aos Conceitos Básicos de Codificação por Transformadas	30
3.7	Fatores que decidem a Escolha da Transformada	31
3.8	Propriedades do uso de Transformadas: a Transformada Ótima e Imagem Markoviana.	32
3.9	A Transformada Cosseno Discreta	37
3.10	Implementação da DCT	38
3.11	Implementação da IDCT e Análise do Não Casamento	41
3.12	Introdução à Transformação de Imagem por Blocos	49
3.13	Tamanho do Bloco	50
3.14	Características no Domínio da Transformada	51
3.15	Quantização e Codificação dos Coeficientes da Transformada	52
3.15.1	Introdução	52
3.15.2	Quantização, Distorção e Entropia	53
3.15.3	Minimização da distorção δ tendo como limitante a Entropia H.	56
3.15.4	Minimização da Entropia H tendo como limitante a distorção δ	58
3.16	Codificação dos valores Quantizados	61
3.16.1	Codificação Zonal	61

3.16.2	Codificação por Limiar	62
3.16.3	Classes de Varredura	62
3.17	Distribuição de Bits e Codificação Entrópica	63
4	SISTEMA PROPOSTO, SIMULAÇÕES E AVALIAÇÕES	64
4.1	Introdução.	65
4.2	Generalidades do Sistema Empregado	68
4.2.1	Introdução	68
4.2.2	Defeitos do Sistema Convencional	70
4.3	Imagens Padrão e Ferramentas de Avaliação.	74
4.3.1	Introdução - Imagens Digitais Padrão	74
4.3.2	Imagens de Teste	75
4.3.3	Pacotes de software desenvolvidos para avaliação do sistema empregado.	77
4.4	Avaliações para o Sistema Convencional de Bloco Fixo.	80
4.4.1	Introdução	80
4.4.2	Características no domínio da transformada	81
4.4.3	Resultados	85
4.4.4	Determinação dos Intervalos Adequados dos Parâmetros Estatísticos de Controle para as Imagens de Teste	89
4.4.5	Resumo dos Parâmetros Estatísticos de Controle para as Imagens de Teste.	122
4.5	Algoritmo Adaptativo Empregado	123
4.5.1	Introdução	123
4.5.2	Diagrama em blocos adaptativo simplificado	124
4.5.3	Representação por Regiões Estruturadas	128
4.5.4	Indexação codificada de Árvore de Quarto Grau.	129
4.5.5	Seleção de Coeficientes e Quantização por Blocos Segmentados	132

4.5.6	Plano de distribuição de bits e sequencia de varredura em zig-zag. . .	138
4.5.7	Avaliações Finais do Algoritmo Empregado na Recuperação das imagens de Teste	141
5	CONCLUSÕES	179
a	GLOSSÁRIO DE TERMOS USADOS	186
b	LISTA DE TRABALHOS PRODUZIDOS E RELACIONADOS.	188
c	LISTAS DE INTERESSE.	191
	LISTA DE FIGURAS	192
	LISTA DE TABELAS	199
d	APLICAÇÃO DOS PROGRAMAS	201
e	CABEÇALHO DE SINCRONISMO DE QUADRO	203
f	IMAGENS “.gif” E “.ras”	205
g	ARREDONDAMENTO E EXATIDÃO DA IDCT.	207
	BIBLIOGRAFIA	209

Capítulo 1

CONSIDERAÇÕES GERAIS

1.1 Introdução

A grande limitação da transmissão analógica ou digital da informação de vídeo é a grande capacidade de canal requerida para se recuperar uma imagem de alto padrão de qualidade. A largura de faixa ocupada pelo sinal de vídeo composto em banda base, por exemplo no sistema PAL-M, é de 4.2 Mhz. Considerando-se, uma frequência de amostragem Super-Nyquist de 10 Mhz e uma quantização uniforme correspondente a 8 bits/amostra [2], a qual permite obter uma qualidade de imagem com degradação imperceptível de acordo com a escala de 5 graus para a medida subjetiva do CCIR [2] é necessário uma taxa da ordem de 85 Mb/s para transmissão digital (incluindo-se os pulsos de sincronismo e serviços) e de acordo com a Recomendação CCIR 601 sobe para 166 Mb/s para o caso de sinal de vídeo em componentes e no caso de armazenar um “vídeo clips” de 10 segundos submetidos a digitalização desde uma transmissão de TV “broadcasting” requer mais de 200 Mbytes [3]. Isto mostra que o processamento do sinal de TV por meio digital será vantajoso caso se tenha uma técnica de codificação eficiente, ou seja, uma técnica que de alguma forma remova as redundâncias e irrelevâncias contidas no sinal de TV (Compressão de imagem) para minimizar a taxa de bits de transmissão ou espaço de memória ocupado.

1.2 Motivação e Avaliação do Trabalho

Neste trabalho apresenta-se um estudo teórico de compressão de imagens paradas baseado na codificação digital dos sinais de TV na forma de componentes Y, C_R , C_B , usando um algoritmo adaptativo de bloco variável à Transformada DCT em duas dimensões.

Propõe-se aqui uma estrutura de codificação eficiente e algoritmo base diferente do sugerido pela CCIR e JPEG, visando o avanço da tecnologia de componentes RISC (processamento paralelo), a compressão de imagem e a redução de taxa.

Foram realizados testes subjetivos e objetivos com a visualização das imagens padrão fornecidas pela SMPTE (Society of Motion Pictures and Television Engineers) através de pacotes de avaliação desenvolvidos pelo autor e pelos sistemas de análise de imagens: KHOROS(The KHOROS Group. University of New Mexico. E.U.A.); ALV(Phill Everson,

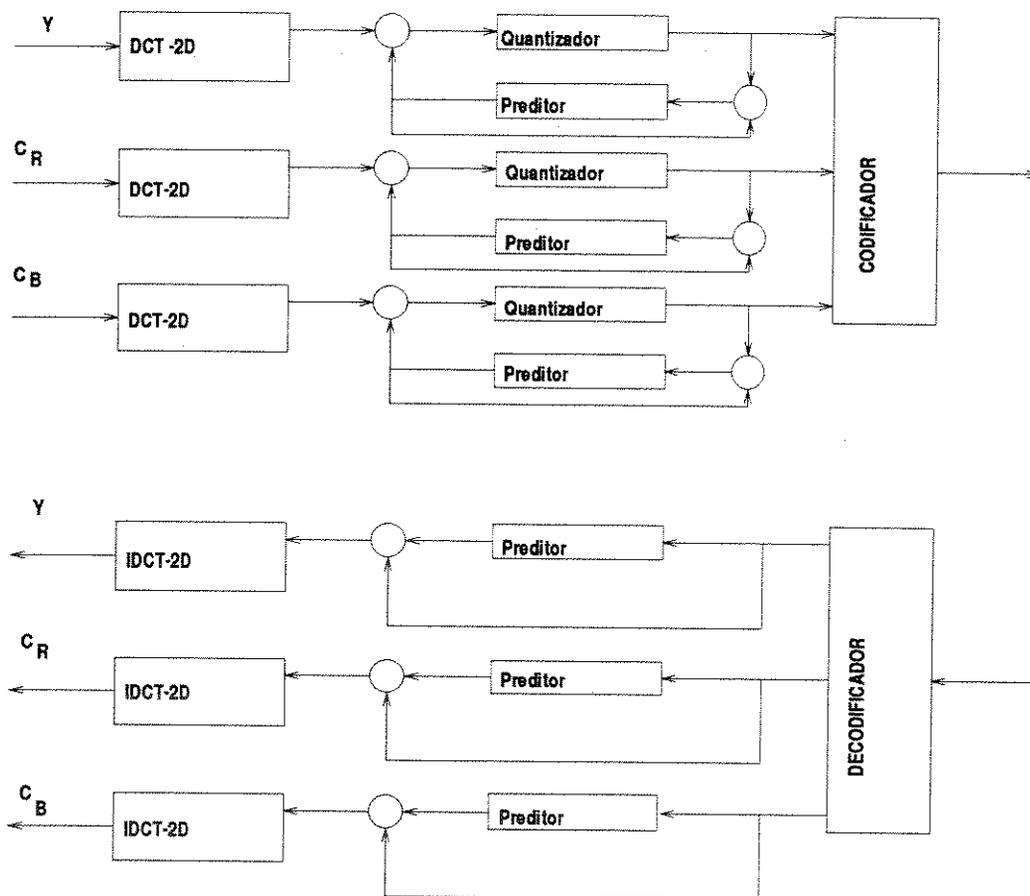


Figura 1.1: Codificação por Transformada DCT do sinal de vídeo em componentes

University of Bristol.U.K.) e XV (Version 2.21.John Bradley, University of Pennsylvania, 1992), presentes no ambiente público das "Workstations SUN".

Dentro das avaliações de caracter objetivo dos pacotes aqui desenvolvidos, salienta-se as medidas de relação Sinal/Ruído, Entropia Acumulada e características estatísticas de ocorrência de amostras (histogramas) para as imagens padrão e para sinais de teste determinísticos usados nas avaliações do sinais de TV. Também usa-se de preferência as imagens padrão em componentes obtidas da referência [1] e principalmente para avaliar o codificador, utiliza-se o sinal de Luminância Y por ser mais crítico quanto a informação de vídeo.

A codificação ficou baseada no desenvolvimento de pacotes de "software" criados para simular só a parte fundamental do sistema de codificação e algoritmo base, tais como: o uso da Transformada Discreta Cosseno (DCT) em duas dimensões com segmentação de blocos fixo e variável, técnicas adaptativas de seleção de coeficientes e técnicas de quantização que permitirão a recuperação integral da imagem livre de interferências perceptíveis inerentes de um processamento com perdas na compressão com precisão de codificação de 2 a 3 bits/E.I.. As outras partes do esquema completo de codificação são mencionadas mas não avaliadas, porque outros autores têm trabalhado exaustivamente em tais matérias (DPCM, Codificação entrópica, memórias elásticas, detecção e correção de erros, multiplexador, etc..) [1], [2].

Este trabalho apresenta resultados do uso de intensos recursos computacionais empregando "workstations SUN" da UNIX de tipo Sparc-2, trabalhando com duas áreas de disco perfazendo em conjunto uma área ocupada de 100 Mbytes comprimidos, associada ao uso de fita magnética e minidiscos de 3 1/2 polegadas de alta e baixa densidade.

A maioria dos programas de simulação foram criados usando linguagem Fortran 77 e alguns em linguagem C.

Alguns cálculos estatísticos empregados para avaliar as imagens atingem mais de 300 horas de uso de "workstations" sendo necessário separar no tempo os cálculos para permitir compartilhar a "workstation" com outros usuários do domínio DECOM. O tratamento dos programas criados e empregados não são objeto de análise deste trabalho porque constituem um desenvolvimento longo e paralelo cujo objetivo é fornecer os resultados para análise.

A avaliação primária subjetiva dos dados foi feita transferindo as imagens para minidiscos, visualizada na mesma “workstation” e testada através do uso de um sistema de aquisição de imagens.

1.3 Contribuição do Trabalho

Este trabalho demonstra que um algoritmo de estrutura de bloco variável adaptativo com a “Média” e o “Desvio Padrão” diferente das normas recentes da JPEG e futuras da MPEG no uso da codificação DCT também pode ser usada para a recuperação integral do sinal usando compressão com perdas com precisão de codificação de 2 a 3 bits/EI. Logo, este trabalho está dirigido a usar como algoritmo fundamental, a Transformada DCT em duas dimensões submetida a uma segmentação de bloco variável adaptativa e com futura aplicação num sistema híbrido. O maior processamento requerido é justificado com o avanço da tecnologia VLSI em “chips” do tipo “RISC” de processamento paralelo.

1.4 Descrição dos Capítulos

Este trabalho é composto por 5 Capítulos e 7 Apêndices.

No Capítulo 1, foi introduzido, de forma bastante geral, os objetivos e a motivação que originaram o desenvolvimento da dissertação.

No Capítulo 2, são apresentadas em forma geral, as codificações eficientes relacionadas.

No Capítulo 3, é apresentada toda a Teoria básica relacionada ao sinal de vídeo, formato intermediário comum de sistemas de TV e a Teoria básica de codificação usando transformadas e a transformada DCT-2D.

No Capítulo 4 é apresentado o Sistema Proposto, as simulações e as respectivas avaliações.

No Capítulo 5, são feitas algumas conclusões e sugestões sobre trabalhos futuros que podem ser realizados para complementar o algoritmo proposto como sistema.

Nos apêndices são apresentados alguns desenvolvimentos de interesse secundário e alguns trabalhos parciais apresentados em Congressos e originados desta dissertação.

Capítulo 2

POSSÍVEIS SOLUÇÕES DE CODIFICAÇÃO EFICIENTE

2.1 Introdução.

A codificação eficiente da imagem permite transmitir uma imagem de fidelidade aceitável, com um número de bits por segundo o menor possível. Uma possível forma de codificação universal da imagem, consiste em usar o sinal analógico de banda base de vídeo constituído pelo sinal de Luminância (Y) mais dois sinais Diferenças de cor (C'_R e C'_B). Esta conformação recebe o nome de sinal de vídeo em componentes de primeira transformação, representa o melhor compromisso com todos os Sistemas de Televisão à cores atuais (SECAM, NTSC, PAL) e favorece certa compatibilidade, por isso o termo universal. A base teórica deste fato é desenvolvida pela referência [1]. Outras codificações eficientes diferentes, têm sido estudadas por muitos autores, e um resumo interessante é apresentado por Netravali e Limb [5].

Três técnicas de codificação são a base deste trabalho como possíveis soluções, “codificação preditiva” (MCPD e MCPDA) que trata com dados no domínio espacial, “codificação por transformada DCT” que em forma geral trata com dados no domínio da frequência e a combinação de ambas que define a “codificação híbrida” onde procura-se combinar as características mais atrativas das duas codificações anteriores.

Os dois tipos de codificação mencionados acima, são estatísticos por natureza, pois ambos operam em virtude das propriedades estatísticas das imagens e não do ponto de vista subjetivo do observador (ou seja, nesse caso frisa-se o mais importante da imagem para o olho humano). Portanto se as propriedades estatísticas são “adequadas”, a eficiência da codificação deverá ser satisfatória, e, para assegurar o termo “adequadas” precisa-se de um sistema adaptado às propriedades locais da imagem, caso contrário espera-se resultados pobres em regiões onde as características saem significativamente da média estatística (caso das bordas).

Essas técnicas estatísticas de compressão são baseadas na “correlação entre elementos adjacentes” no domínio espacial, onde por exemplo, o valor da magnitude de um elemento pode ser estimado dos valores dos elementos prévios adjacentes já codificados.

É preciso frisar que “altos valores de correlação” entre elementos adjacentes estão em geral associados a “lenta variação no tempo” ou distribuição espectral de potência por natureza passa-baixa implicando numa codificação com pequena capacidade de canal de

transmissão. Felizmente em imagens ainda ativas (conteúdo razoável de detalhe espacial) o coeficiente de correlação inter-elementos é relativamente alto, causando uma taxa de compressão na codificação por transformada que pode chegar a 8:1 [7] para uma imagem de baixa resolução, ou seja, se a imagem original é codificada em 8 bit/E.I. poderá ser reduzida à 1 bit/E.I.. Agora, se esta for originalmente de 16 bit/E.I., chegaríamos a 2 bit/E.I., dando uma imagem mais aceitável. Não obstante a qualidade de imagem final, ou seja, o número de bit/E.I. é o verdadeiro indicador da qualidade da imagem fornecida pelo sistema e não a taxa de compressão.

2.2 Codificação Preditiva

É um método alternativo de compressão de dados de imagem que concorre com a codificação por transformada.

Dois fatos geram este método: primeiro, devido a que o sistema PCM apenas transmite as amplitudes quantizadas de cada EI, não aproveitando nenhuma das propriedades de uma imagem real, e segundo, devido a que os EI adjacentes dessa imagem possuem uma forte correlação tanto espacial como temporal. Este método vai explorar a correlação dos EI usando um sistema PCM de um sinal diferença. Em um sistema básico de codificação preditiva, uma predição aproximada da amostra a ser transmitida é feita da informação codificada prévia que tem sido transmitida, ver figura (2.1). O erro (ou sinal diferença) resultante da subtração do valor de predição com o valor atual do EI, é logo quantizado em uma sequência de "L" valores de amplitude discretos. Estes níveis são representados como palavras binárias de comprimento fixo ou variável e logo enviados ao canal transmissor.

Em resumo a codificação preditiva tem três componentes básicos: 1) Preditor, 2) Quantizador e 3) Distribuidor de código. Dependendo do número de níveis do quantizador, se faz distinção entre a modulação Delta (MD) que tem apenas $L=2$ níveis e Modulação por Codificação de Pulso Diferencial (MCPD), onde L é maior a dois níveis. O uso de MD requer para conseguir uma qualidade de imagem adequada que a amostragem seja muito maior que a frequência de Nyquist, por isso não tem sido muito usado em codificação de

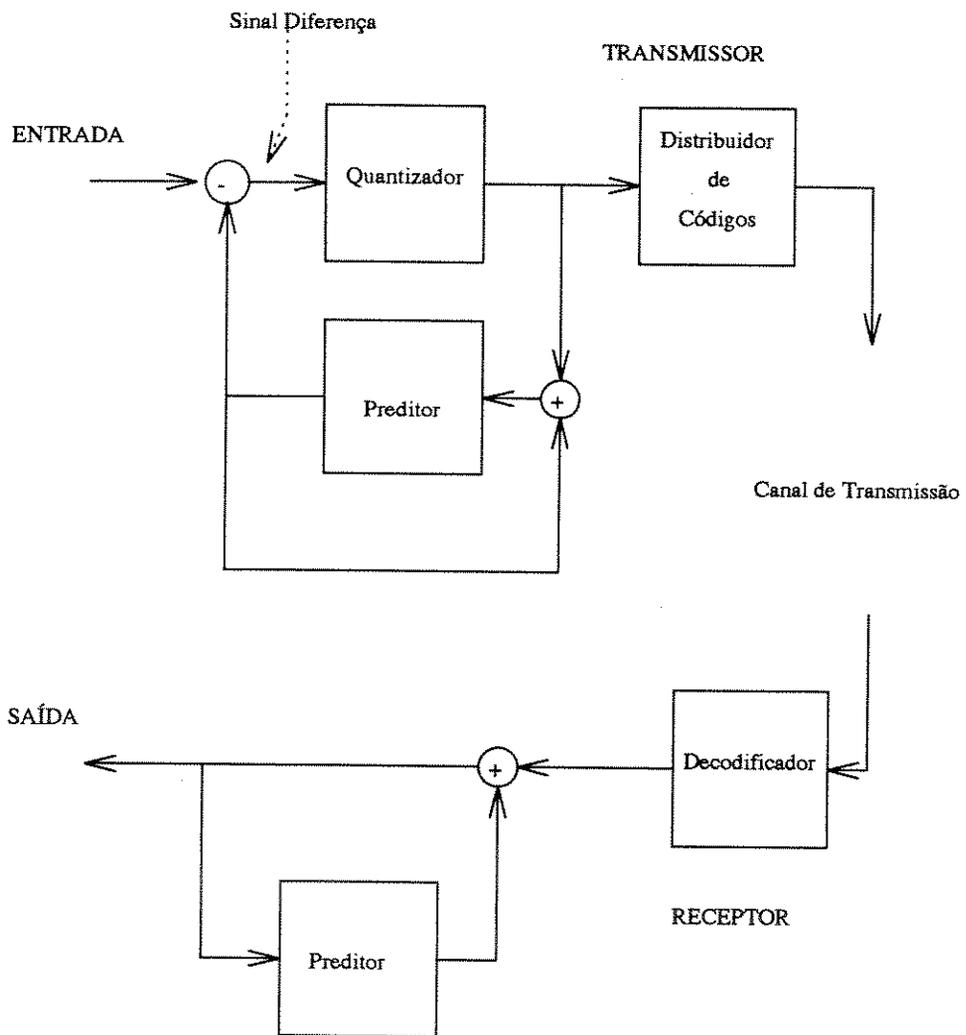


Figura 2.1: Codificação Preditiva básica

imagem e prefere-se o uso do MCPD e MCPD adaptativo por serem mais eficientes.

2.3 Codificação por Transformada

O mecanismo de operação é mais complexo do ponto de vista de cálculos, visto que está baseado na decomposição do sinal de entrada temporal ou espacial no domínio do espectro de frequência. Uma das transformadas clássicas usada foi a Transformada de Fourier em codificação de imagem, mas hoje tem sido mudada por outras transformadas que têm a vantagem de alta eficiência na codificação com manipulação baseada só em números reais.

Para uma imagem com alta correlação entre elementos, muitas componentes espectrais de alta ordem serão de magnitudes pequenas e podem ser codificadas com poucos bits ou desprezadas totalmente. Uma operação por transformada pode ser vista como uma operação que tenta originar uma sequência de coeficientes espectrais que são aproximadamente não correlatos, reduzindo ao máximo a redundância de informação.

Codificação preditiva é mais complexa que transformada. Transformada pode ser mais eficiente e simples, mas requer muitos cálculos e é menos sensível aos erros que a codificação preditiva, além do mais, os erros que influem na eficiência do sistema, manifestam-se de uma maneira totalmente diferente à codificação preditiva. A comparação direta é difícil.

2.4 Codificação Híbrida

Usa-se com o objetivo de se conseguir o melhor compromisso entre a taxa de compressão e a complexidade do sistema, posto que são reduzidos os requerimentos de memória de armazenamento e a complexidade de cálculo.

Com o avanço da tecnologia VLSI e RISC [27] a Transformada Rápida em duas dimensões é disponível e neste caso a operação preditiva pode fazer-se entre os coeficientes transformados de uma mesma localização, mas situados em quadros diferentes para permitir a codificação híbrida de imagens em movimento.

Em resumo, é possível variações no processo, por exemplo, fazer primeiro a codificação preditiva (MCPD) do que fazer a Transformada Rápida em duas dimensões, mais em geral, as características de um ou outro sistema de processamento estão em compromisso com os dois métodos básicos mencionados, e a mínima taxa de codificação, não é tão baixa como aquela da codificação por transformada pura, mas a implementação é viável. O uso de técnicas adaptativas, também é preciso para atingir valores finais de compressão de 1 bit/E.I. com qualidade de reconstrução de imagem aceitável para videofone, não obstante, para qualidade de "Broadcasting" requer-se valores superiores a 2 bit/EI.

Capítulo 3

CONCEITOS BÁSICOS DO SINAL DE VÍDEO E DA CODIFICAÇÃO POR TRANSFORMADAS

3.1 Largura de Faixa e Capacidade de Canal.

É preciso suficiente Capacidade de Canal para transmitir os dados digitais de vídeo, onde existem limitações fundamentais.

Os sistemas de comunicações tem uma Largura de Faixa “W” que limita a velocidade de variação do sinal. Shannon estabeleceu que a velocidade de uma transmissão de informação livre de erro não pode exceder à Capacidade de Canal,

$$c \leq W \log\left(1 + \frac{P_b}{P_n}\right) \quad (3.1)$$

Onde P_b é potência transmitida média e P_n é potência de ruído média. Esta relação é conhecida como lei de Hartley-Shannon (HS) e estabelece o limite superior para a eficiência de um sistema de comunicações tendo conhecidos a Largura de Faixa “W” e a relação de Potência de Sinal/Ruído. A lei HS especifica a Largura de Faixa ótima versus intercâmbio de potência e sugere a possibilidade de compressão de Largura de Faixa.

$$c = W \log\left(1 + \frac{P_b}{\mu W}\right) \quad (3.2)$$

Onde, $\mu W = P_n$ significando que a Potência de ruído é proporcional à Largura de Faixa.

A relação (3.2) é um compromisso atual entre o sinal e o ruído para o caso de distorção.

3.2 Informação fonte

Não obstante que os custos dos sistemas de armazenamento digital e de processamento rápido digital estão sendo cada vez menores pelo avanço das tecnologias VLSI e RISC, ainda o investimento é alto. Agora é viável armazenar e processar de um até dois quadros de uma cena de TV. O formato de informação de fonte deve tomar em conta as normas de TV tais como PAL e NTSC. A recomendação 601 da CCIR (3.1) estabelece os

	PAL	NTSC
Parâmetros	625/50	525/60
Frequência de amostragem	13,5Mhz(Y) 6,75Mhz(U,V)	13,5Mhz(Y) 6,75Mhz(U,V)
Estrutura de amostragem	ortogonal	ortogonal
Número de linhas ativas	576	480
Número de E.I. ativos/linha		
Luminância(Y)	720 E.I./linha	720 E.I./linha
Crominância(U)	360 E.I./linha	360 E.I./linha
Crominância(V)	360 E.I./linha	360 E.I./linha
Frequência de Campo	50	59.91
Quantização	8 bit PCM 220 níveis de quantização preto=16 e branco=235	8 bit PCM 220 níveis de quantização preto=16 e branco=235

Tabela 3.1: Parâmetros CCIR 601

parametros básicos do formato de imagem em componentes que normaliza a Capacidade de Canal para transmitir a informação fonte nos sistemas PAL e NTSC.

A luminância (Y) é submetida a amostragem em 13.5 Mhz e as componentes de crominância (C_R e C_B) em 6.75 Mhz, ou seja o número de E.I. da crominância é a metade do número de E.I. da luminância. O número de bits em um quadro para uso em estúdio de TV pode ser expresso como,

$$N_b = N(x) \cdot N(y) \cdot bpp \quad (3.3)$$

Onde,

N_b = No. de bits por quadro

bpp = No. de bits por E.I.

$N(x)$ =No. de amostras ativas horizontais

$N(y)$ =No. de linhas ativas verticais

Logo, o número de bits por quadro de informação de fonte para os sistemas PAL e NTSC, usando a equação (3.3) e a tabela (3.1) para luminância e crominância são :

$$N_{bPAL} = \underbrace{N_{bL}}_{\text{Luminância}} + 2 \cdot \underbrace{N_{bC}}_{\text{Crominância}} \quad (3.4)$$

$$N_{bPAL} = 6.635.520 \quad \text{bits/quadro} \quad (3.5)$$

$$N_{bNTSC} = \begin{array}{ccc} N_{bL} & + & 2 \cdot N_{bC} \\ \text{Luminância} & & \text{Crominância} \end{array} \quad (3.6)$$

$$N_{bNTSC} = 5.529.600 \quad \text{bits/quadro} \quad (3.7)$$

Se, as frequências de quadro (F_q) são 25 Hz e 30 Hz aproximadamente para PAL e NTSC, a Capacidade de Canal necessária para transportar a informação de fonte segundo a recomendação CCIR 601 é :

$$C_t = N_b \cdot F_q \quad (3.8)$$

Portanto,

$$C_{PAL} = 166 \text{Mbit/s} = C_{NTSC} \quad (3.9)$$

3.3 Futuro Sistema Multi-mídia e o Formato Intermediário Comum

A necessidade de compartilhar os tele-serviços de comunicações entre os diferentes países que tem normas diferentes de sistemas de TV leva a criar um formato intermediário comum (CIF) que permite convergir os parâmetros fundamentais do sinal de entrada, tais como: resolução espacial e resolução temporal a uma norma única. Este formato poderá fornecer uma melhor conexão para a futura codificação de vídeo.

O CIF nasceu quando resolveu-se compatibilizar as normas de TV inter-regionais na Europa [4], veja figura (3.1).

Os parâmetros básicos para o formato CIF comparado com os sistemas PAL e NTSC são mostrados na tabela (3.2).

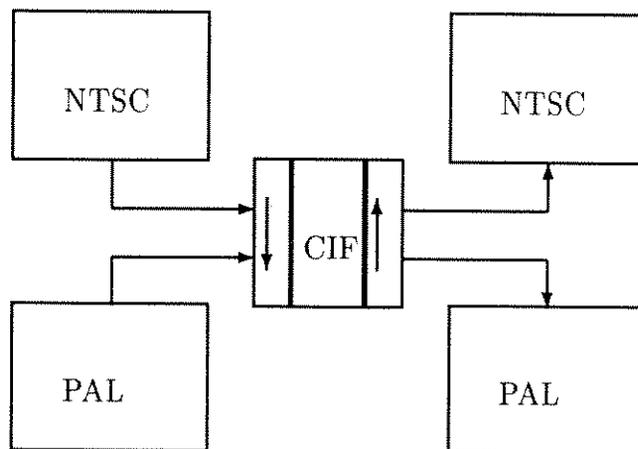


Figura 3.1: Princípio de Formato Intermediário Comum que permite a conexão direta entre países com Sistemas NTSC e PAL. Conversão para CIF é possível com filtragem e amostragem apropriados.

	PAL	CIF	NTSC
Luminância			
No. de EI ativos	720	360(352)	720
Linhas ativas	576	288	480
Crominância			
No. de EI ativos	360	180(176)	360
Linhas ativas	576	144	480
Frequência de quadro	25	29.97	29.97
Entrelaçado	2:1	1:1	2:1

Tabela 3.2: Comparação do formato CIF com NTSC e PAL

3.4 Defeitos na Informação de Fonte

Não é possível representar uma imagem sintética $I(x,y)$ de tamanho finito com absoluta fidelidade devido ao número finito de bits resultantes da quantização.

Na digitalização de um sinal analógico de imagem o conjunto de pontos de amostragem (x,y) pode ser dispostos em “arrays” retangulares que dependendo da frequência de amostragem (f_A) vão definir a Densidade de Amostragem e o tipo de mosaico. Este último pode ser : Mosaico Ortogonal, Mosaico Quincux na linha e Mosaico Quincux no campo [6].

Os pontos de amostras, ou Elementos de Imagem chamados em forma resumida “EI”, podem se referir a amostras fixas no plano de imagem. O termo $EI(x,y)$ e o valor de $I(x,y)$ são equivalentemente usados neste trabalho.

A resolução espacial pode ser especificada de várias formas: a mínima distância determinável entre dois pontos ou linhas, a frequência máxima da informação de vídeo do ponto de vista espectral ou o número máximo de linhas contidas dentro de uma imagem.

O fenômeno de “pré-aliasing” na informação de fonte aparece quando acontecem transições randômicas muito rápidas, por exemplo o ruído nas fontes de imagem pode ser amostrado por casualidade; uma forma de evitar este fenômeno é recolher uma média ponderada da vizinhança do ponto de amostragem, para este propósito usa-se uma função de ponderação de Gauss de tipo:

$$h_c(u, v) = K \cdot e^{-a \cdot u^2 - b \cdot v^2} \quad |u|, |v| < \delta \quad (3.10)$$

Onde K é uma constante de normalização dependendo da aplicação. Usando esta função Gaussiana é possível tirar a média ao filtrar a imagem,

$$\bar{I}(x, y) = \int_{y-\delta}^{y+\delta} \int_{x-\delta}^{x+\delta} I(w, z) \cdot h_c(x-w, y-z) dw dz \quad (3.11)$$

Neste caso, a intensidade em cada ponto é substituída pela média ponderada da vizinhança do ponto. Este processo é chamado pré-filtragem, devido a que a imagem é filtrada antes de ser amostrada. $\bar{I}(x, y)$ é a imagem filtrada, e $h_c(u, v)$ é a função de pré-filtragem da câmara.

Em resumo, a quantidade de informação visual é reduzida de três maneiras em uma representação espacial de imagem quantizada:

- Primeiro apenas um número finito de amostras de intensidade são representadas.
- Segundo, transições rápidas de intensidade são removidas por pré-filtragem.
- Terceiro os valores de intensidade são quantizados com uma exatidão finita.

Agora o fenômeno de “pós-aliasing” tem lugar na recuperação da imagem, pois a partir do sinal quantizado ($\bar{I}(x, y)$) da imagem pré-filtrada, deve-se tentar recuperar a réplica da imagem original aproximada $\tilde{I}(x, y)$. A dificuldade é decidir que valor de intensidade de luz deve-se transferir aos pontos (x, y) que não são pontos de amostragem como acontece na recuperação por “sample and hold”, estes defeitos de “pós-aliasing” apresentam-se principalmente nas bordas de imagem e podem ser reduzidas usando interpolação entre os EI para se obter transições de intensidade relativamente suaves, por exemplo, se do ponto de vista analítico, $I_s(x, y)$ representa a imagem submetida a amostragem por funções Delta Dirac,

$$I_s(x, y) = \sum_{k=1}^{N_s} \bar{I}(x_k, y_k) \cdot \delta(x - x_k) \cdot \delta(y - y_k) \quad (3.12)$$

A imagem reconstruída é dada por,

$$\tilde{I}(x, y) = \iint I_s(w, z) \cdot h_d(x - w, y - z) dw dz \quad (3.13)$$

Onde, $h_d(u, v)$ = função linear de interpolação.

Em resumo a representação de imagem espacial submetida a amostragem e quantização precisam de “pré-filtragem” prévia à amostragem e “pós-filtragem” na recuperação de imagem tal como mostra-se na fig.(3.2). Portanto inadequada pré-filtragem produz “pre-aliasing”, enquanto que inadequada pós-filtragem produz “pós-aliasing”.

3.5 Representação Espectral da Imagem

Amostragem espacial e quantização da intensidade de imagem não é o único método para representar informação visual usando uma quantidade finita de dados. Se

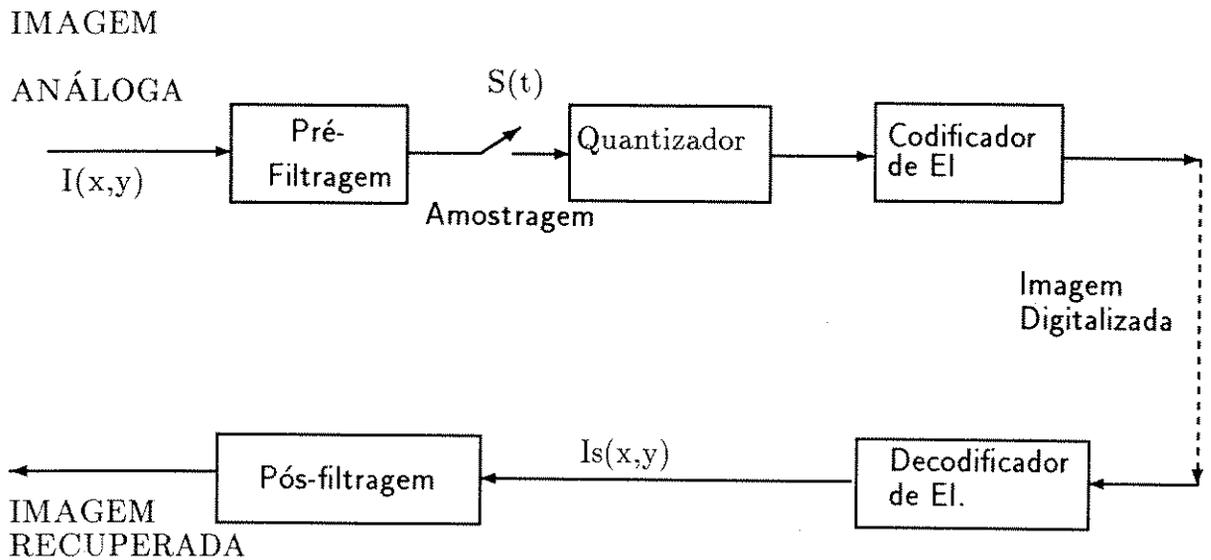


Figura 3.2: Processamento básico da imagem espacial submetida a amostragem e quantização

a imagem original $I(x,y)$ é transformada antes de submeter à amostragem e quantização pode-se conseguir uma codificação mais conveniente.

3.5.1 Série de Fourier

Supondo que os limites da imagem são:

$$0 \leq x \leq X \quad (3.14)$$

$$0 \leq y \leq Y \quad (3.15)$$

E para (x,y) define-se a função exponencial complexa que gera um conjunto ortonormal chamado de “base de imagens”.

$$\Phi_{mn}(x, y) = \frac{1}{\sqrt{XY}} \exp\left\{(2\pi \cdot \sqrt{-1}) \cdot \left[\frac{nx}{X} + \frac{my}{Y}\right]\right\} \quad (3.16)$$

Logo, para toda imagem razoável é possível na prática escrever sua representação

por Série de Fourier:

$$I(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} C_{mn} \cdot \Phi_{m,n}(x, y) \quad (3.17)$$

Onde, seus coeficientes:

$$C_{mn} = \int_0^X \int_0^Y I(x, y) \cdot \Phi'_{m,n}(x, y) dy dx \quad (3.18)$$

e,

$\Phi'_{m,n}$ = função exponencial complexa conjugada.

A “base de imagens” é usada (veja 3.16) para representar um conjunto de imagens canônicas que quando combinadas apropriadamente produzem qualquer imagem $I(x,y)$ de interesse prático.

Neste trabalho é utilizado o conjunto ortonormal de funções ou “base de imagens” representado pela Série Cosseno:

$$\Phi_{m,n}(x, y) = \frac{K_m \cdot K_n}{\sqrt{XY}} \cos(\pi n \frac{x}{X}) \cdot \cos(\pi m \frac{y}{Y}) \quad |m, n \geq 0 \quad (3.19)$$

Onde,

$$K_l = \begin{cases} 1 & \text{if } l = 0 \\ \sqrt{2} & \text{if } l > 0 \end{cases}$$

Representando $I(x,y)$ em termos de “base de imagens” exponenciais de Fourier como na equação 3.16, conduz à decomposição da imagem em componentes de frequência espaciais (C_{mn}), ou seja, é a representação espectral da imagem. Para cada “base de imagens” $\Phi_{mn}(x, y)$ as frequências espaciais nas direções x e y são respectivamente:

$$f_x = \frac{n}{X} \quad (\text{ciclos por unidade de comprimento}) \quad (3.20)$$

$$f_y = \frac{m}{Y} \quad (\text{ciclos por unidade de comprimento})$$

A frequência espacial total do “pattern” periódico bidimensional é definido como:

$$f_s = \sqrt{f_x^2 + f_y^2} \quad (3.21)$$

Em geral, as componentes de frequências espaciais altas representam variações bruscas na intensidade da imagem, tal como acontece nas bordas dos objetos na imagem. Componentes de frequências espaciais baixas representam variações suaves na imagem.

3.5.2 Imagens de Faixa Limitada - Transformada de Fourier Discreta - Teorema da Amostragem Bidimensional

Se as componentes de frequências de uma imagem encontram-se dentro de um intervalo finito, neste caso define-se a imagem de faixa limitada, ou seja,

$$\bar{I}(x, y) = \sum_{m=-M}^M \sum_{n=-N}^N C_{mn} \cdot \Phi_{mn}(x, y) \quad (3.22)$$

Mostra-se que é de faixa limitada ao seguinte intervalo de frequência,

$$|f_x| < \frac{N + 0.5}{X} = F_x \quad \text{ciclos/comprimento} \quad (3.23)$$

$$|f_y| < \frac{M + 0.5}{Y} = F_y \quad \text{ciclos/comprimento}$$

Nesta imagem o número de coeficientes não zeros C_{mn} é dado por $N_s = (2M + 1)(2N + 1)$. Para avaliar estes coeficientes, a integral (3.18) pode ser usada.

Métodos mais práticos de avaliação dos coeficientes C_{mn} de Fourier são baseados nas amostras espaciais da imagem de faixa limitada $\bar{I}(x, y)$ com N_s E.I. $\{(x_k, y_k), k = 1, \dots, N_s\}$. Para cada E.I. (x_k, y_k) , temos da equação (3.22),

$$\bar{I}(x, y) = \sum_{|m| \leq M} \sum_{|n| \leq N} C_{mn} \cdot \Phi_{m,n}(x, y) \quad (3.24)$$

Agora, se os E.I. são escolhidos tal que os N_s da equação (3.24) são linearmente independentes, logo é possível resolver para os N_s coeficientes desconhecidos C_{mn} sem avaliar a integral (3.18).

Um conjunto de tais pontos amostrados consiste num arranjo retangular de $N_x = 2N + 1$ colunas e $N_y = 2M + 1$ filas. Neste caso $N_s = N_x \cdot N_y$ e pode ser facilmente demonstrado para $|m|, |k| \leq M$ e $|n|, |l| \leq N$ que,

$$\sum_{i=0}^{N_y-1} \sum_{j=0}^{N_x-1} \Phi_{mn}(x_j, y_i) \cdot \Phi'_{kl}(x_j, y_i) = \frac{N_s}{XY} \cdot \delta_{mk} \cdot \delta_{nl} \quad (3.25)$$

Onde $x_j = \frac{X}{N_x} \cdot j$ e $y_i = \frac{Y}{N_y} \cdot i$. Aplicando este resultado na equação (3.24) define-se,

$$Cmn = \frac{XY}{N_s} \sum_{i=0}^{N_y-1} \sum_{j=0}^{N_x-1} \bar{I}(x_j, y_i) \cdot \Phi'_{kl}(x_j, y_i) \quad (3.26)$$

Equação que recebe o nome de Transformada Discreta de Fourier (DFT) para sequência de dados submetidos a amostragem.

Logo, amostrando $\bar{I}(x, y)$ usando um arranjo retangular de E.I. é suficiente para completamente especificar uma imagem de faixa limitada. Onde, $N_x = 2N + 1$ amostras deveriam ser tomadas horizontalmente e $N_y = 2M + 1$ verticalmente, as densidades de amostragem respectivas são,

$$D_x = \frac{2N + 1}{X} = 2 \cdot Fx \quad \text{ciclos/comprimento} \quad (3.27)$$

$$D_y = \frac{2M + 1}{Y} = 2 \cdot Fy \quad \text{ciclos/comprimento}$$

Comparando as equações (3.23) e (3.27) pode-se apreciar que cada densidade de amostragem é levemente maior ao dobro da correspondente frequência espacial mais alta contida na imagem. Isto na essência define o Teorema de amostragem de Nyquist, que estabelece que, se uma função é de faixa limitada a um intervalo de frequência W , pode-ser completamente especificada por amostras tomadas à taxa de amostragem de Nyquist de $2W$ amostras por unidade de tempo.

Para nossa imagem finita bidimensional, o Teorema de Amostragem fica,

$$\bar{I}(x, y) = \sum_{i=0}^{2M} \sum_{j=0}^{2N} \bar{I}(x_j, y_i) \cdot \chi(x - x_j, y - y_i) \quad (3.28)$$

onde,

$$\chi(u, v) = \frac{\sin[\pi(2N + 1)\frac{u}{X}] \sin[\pi(2M + 1)\frac{v}{Y}]}{N_s \sin(\pi\frac{u}{X}) \sin(\pi\frac{v}{Y})} \quad (3.29)$$

(sem considerar a simplificação de argumentos pequenos no denominador)

3.5.3 Espectro de Frequência de uma imagem amostrada

Supondo que temos uma imagem de faixa limitada $\bar{I}(x, y)$ que têm componentes de frequências C_{mn} apenas para $|m| \leq M$, $|n| \leq N$ e usando-se um arranjo de amostragem retangular para construir a imagem amostrada.

$$I_s(x, y) = \sum_{i=0}^{2M} \sum_{j=0}^{2N} \bar{I}(x_j, y_i) \cdot \delta(x - x_j) \cdot \delta(y - y_i) \quad (3.30)$$

Onde,

$$x_j = \frac{X}{2N + 1} \cdot j$$

e,

$$y_i = \frac{Y}{2M + 1} \cdot i$$

As componentes de frequência de Fourier de $I_s(x, y)$ são derivadas da equação (3.18).

O espectro de frequência de $I_s(x, y)$ é mostrado na figura (3.3).

Das equações (3.28) e (3.30) pode-se apreciar que a imagem de faixa limitada pode ser escrita em termos da imagem amostrada,

$$\bar{I}(x, y) = \int_0^X \int_0^Y I_s(w, z) \chi(x - w, y - z) dz dw \quad (3.31)$$

Posto que a função $\chi(u, v)$ (função “sinc”) têm um comportamento de filtro passa-baixa, ou seja, apenas retêm as componentes de frequência,

$$|fx| < Fx = \frac{N + 0.5}{X} \quad \text{ciclos/comprimento} \quad (3.32)$$

$$|fy| < Fy = \frac{M + 0.5}{Y} \quad \text{ciclos/comprimento}$$

Logo, nesta bandabase de frequência, o espectro de $\bar{I}(x, y)$ é exatamente o mesmo que apresenta a Imagem amostrada $I_s(x, y)$. Além, o espectro de frequência de $I_s(x, y)$ é o

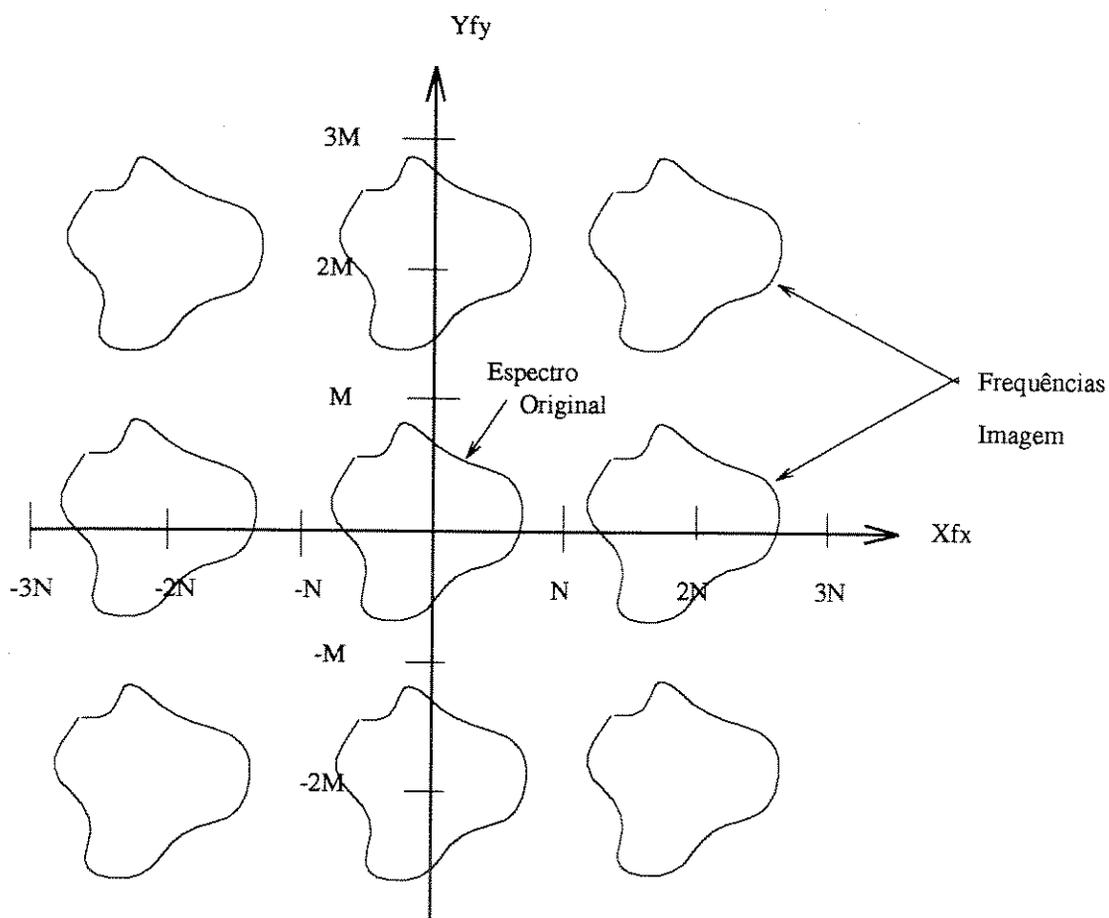


Figura 3.3: Espectro de frequência de Fourier da imagem amostrada $I_s(x, y)$. X_{fx} e Y_{fy} são respectivamente os ciclos horizontais e verticais. O intervalo de baixa frequência é a imagem $\bar{I}(x, y)$ original de banda limitada. Fora estão as frequências imagem.

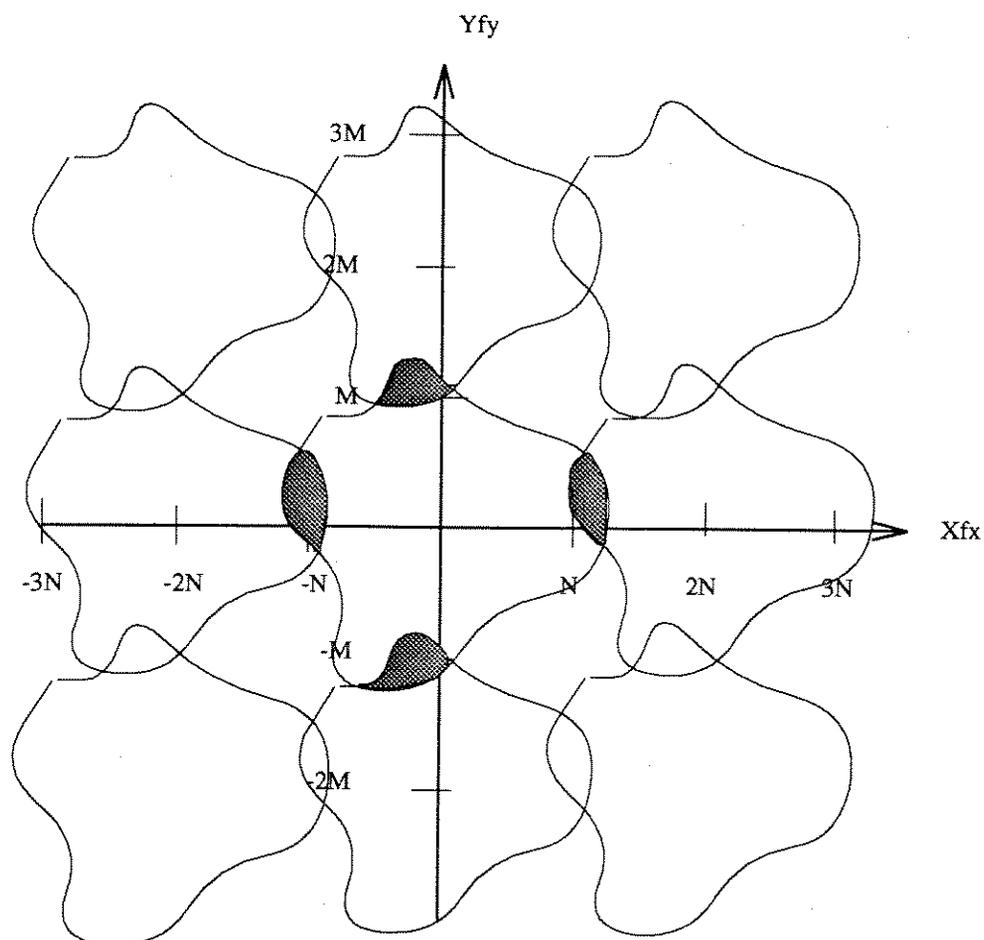


Figura 3.4: Espectro de Frequência da imagem amostrada $I_s(x, y)$ quando a imagem original $\bar{I}(x, y)$ contém componentes de frequência cujos valores estão sobre a metade da frequência da amostragem. Componentes na zona sombreada são afetadas pelas frequências imagem. Esta distorção é denominada “pre-aliasing” e é devido a insuficiência de pré - filtragem na imagem original $\bar{I}(x, y)$.

espectro original $I(x,y)$, periodicamente repetido como se apresenta na figura(3.3). As componentes de frequência fora desta bandabase são nominadas frequências “alias”.

Os defeitos apresentados na imagem previamente (seção 3.4), agora podem amostrar-se no espectro de frequência . Supondo que a imagem $\bar{I}(x, y)$ que vai ser submetida à amostragem, contém componentes de frequência de valor maior a metade da frequência de amostragem. Percebe-se na figura (3.4) que o espectro da imagem amostrada $I_s(x, y)$ possui áreas sobrepostas nas quais as componentes de frequências originais são corruptas pelas frequências “ alias”. A imagem original $\bar{I}(x, y)$ não pode ser recuperada a partir da imagem amostrada $I_s(x, y)$, apenas em casos especiais([9]).

Este tipo de distorção foi mencionado na seção 3.4 e é denominado “pre-aliasing” sendo causada por insuficiente pré-filtragem de $\bar{I}(x, y)$.

Agora, se a taxa de amostragem é o suficientemente alta para não ter “pre-aliasing” e o espectro apresenta-se como na figura(3.3), mas não se pode implementar o filtro ideal passa-baixa (função “sinc”) da equação(3.31), apenas pode-se empregar uma reconstrução sub-ótima,

$$\tilde{I}(X, Y) = \int_0^X \int_0^Y I_s(w, z) \cdot h_d(x - w, y - z) dz dw \quad (3.33)$$

Onde, a extensão de frequência do filtro $h_d(u, v)$ é apresentada pelo circulo segmentado da figura (3.5). Neste caso apareceram muitas componentes de frequências “alias” em $\tilde{I}(x, y)$ que não estavam no original $\bar{I}(x, y)$. Esta distorção é denominada “pós-aliasing”, é causada por inadequada pós-filtragem de $I_s(x, y)$. Tambem foi apresentada na seção 3.4.

Outros “patterns” de amostragem tambem podem ser usados como no caso de mosaico cruzado (veja referencia [9]).

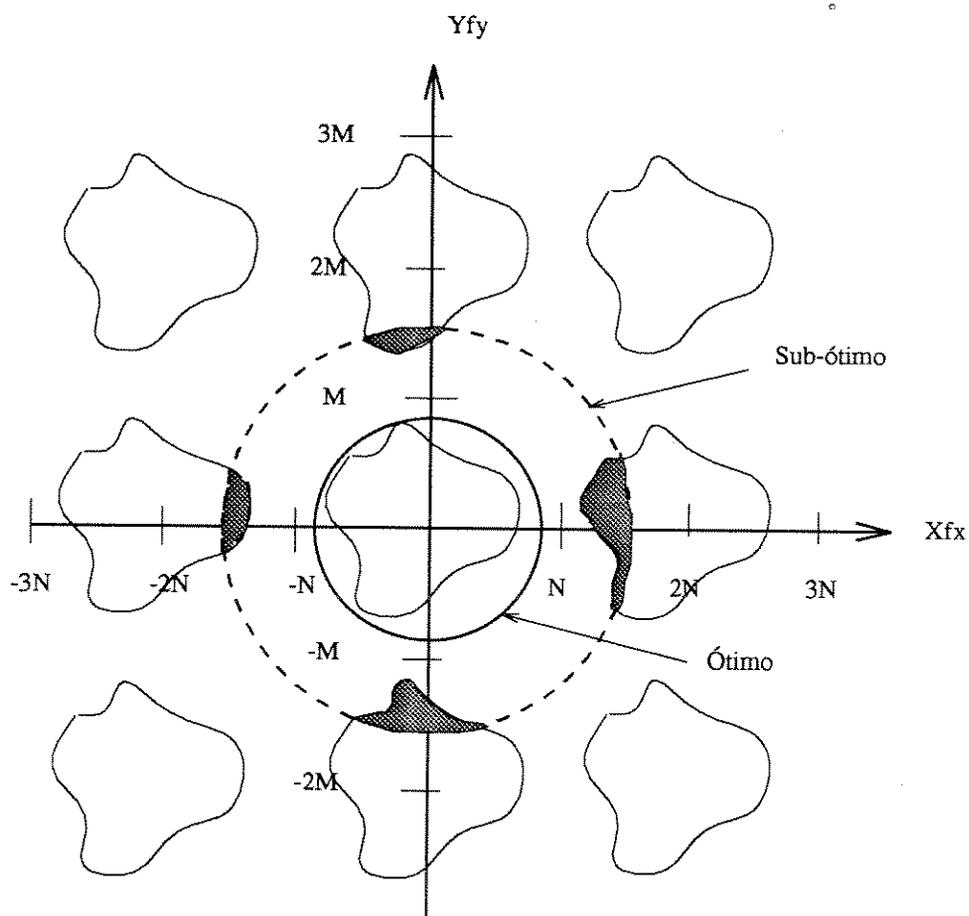


Figura 3.5: Pós-filtragem não ideal. A linha segmentada indica a extensão da frequência de um filtro sub-ótimo $h_d(u, v)$. A linha sólida indica o filtro ótimo. Componentes de frequências “alias” aparecem nas áreas sombreadas na imagem reconstruída. Esta distorção é denominada “post-aliasing”, e é devido a pós-filtragem inadequada de $I_s(x, y)$.

3.5.4 Espectro de Frequência para a Transformada Discreta Cosseno

Considera-se que a codificação de imagem usando transformada discreta de Fourier implica numa codificação mais eficiente aquelas componentes espectrais (frequências espaciais) onde o olho humano é mais sensível. Hall ([10], [11]) considerou esta situação e apresentou a Curva de Resposta Espacial Óptica que foi posteriormente medida por Mannos e Sakrison [14] para ponderar os coeficientes da Transformada antes de fazer a quantização e alocação de bit para obter deste modo uma eficiência de codificação melhorada.

No caso dos coeficientes da transformada DCT, a grande dificuldade para obter a característica prévia é que a distribuição espectral não é a que corresponde às componentes de frequência verdadeiras, como no caso da DFT. Neste caso, componentes espúrias de energia aparecem em localizações arbitrárias dentro do conjunto de coeficientes transformados. Clarke [15] mostrou que a maior parte da energia associada a uma determinada componente espacial reside em quase sua totalidade dentro dos três coeficientes adjacentes para o caso de blocos de coeficientes DCT de 16×16 , o que implica que deve-se usar diferentes funções de ponderação para estes coeficientes visando obter a eficiência da codificação reportada por Hall.

Por outro lado, o uso de transformadas ortogonais em compressão de dados de imagem é ligado a duas propriedades básicas das transformadas. Uma delas é a efetividade de decorrelação dos coeficientes transformados, e a outra é a discriminação de frequência no domínio da transformada. Com respeito à primeira propriedade a transformada visa decorrelacionar os dados dos blocos de imagem em forma exata, ou aproximada, diagonalizando as matrizes-covariância. Sabe-se que a transformada ótima que chega neste objetivo é a Transformada de Karhunen-Loeve (KLT) porque consegue diagonalizar completamente a matriz covariância. Porém é difícil, porque têm uma maior quantidade de passos de cálculos para determinar os autovalores da matriz covariância, com exceção dos casos especiais de transformadas. Uma delas, é a DCT que comporta-se como a transformada ótima KLT para modelos randômicos estacionários de Markov de primeira ordem que têm sido usados frequentemente como modelos naturais de cenas de imagens. Mais detalhes sobre este comportamento podem ser vistos na seção 3.13.

Com respeito à segunda propriedade que agora nos interessa, tem havido numerosos algoritmos para determinar a DCT usando variantes da DFT, mas a conversão direta tem sido difícil pelo fato de duplicar os dados de entrada ou fazer um re-ordenamento (“reshuffling”) dos dados de entrada sequencial. Isto faz com que a frequência natural espacial da entrada não seja facilmente discriminada no domínio da DCT e portanto essa análise não é atrativa do ponto de vista aplicativo. Por outro lado o que mais interessa no uso da DCT é visar a maior compressão sem afetar a qualidade da imagem e para procurar isso os processos de seleção e quantização são baseados na distribuição da energia no domínio da frequência espacial da DCT, não apenas pelo fato da compressão, mas também pela recuperação e melhoramento da imagem considerando nosso modelo de sistema visual humano (SVH). Mais para frente serão apresentados mais detalhes.

3.6 Introdução aos Conceitos Básicos de Codificação por Transformadas

O uso de transformadas unitárias e a justificação teórica correspondente é muito conhecida e tem sido bastante descrita na literatura. Devido ao surgimento de novas técnicas VLSI e processadores de vídeo a aplicação das transformadas em processamento de imagem digital mostra propriedades não descobertas, vantagens e também desvantagens.

A teoria de codificação por transformadas tal como é descrita nas referências [7], [8], [16] visa mostrar os motivos do uso das transformadas e qual é a transformada que tem a melhor eficiência.

Nesta introdução mostra-se um breve resumo das principais propriedades das transformadas e serão examinadas algumas propriedades no domínio das transformadas que são de importância neste trabalho.

3.7 Fatores que decidem a Escolha da Transformada

Os fatores que visam a determinação da transformada ortogonal com objetivo de uso em codificação de imagem são:

- A eficiência da codificação por transformada depende do ordenamento das variâncias dos coeficientes no domínio da transformada e a estratégia de alocação de bit [18].
- A eficiência da codificação por transformada também depende da escolha do tipo de transformada, da estratégia de quantização, do tamanho e forma do bloco de processamento [17].

Outros fatores visando a compressão de imagem e relacionados com a transformada ótima são:

- A Transformada Karhunen-Loève (KLT) é a transformada ótima pelo fato de sua propriedade de descorrelacionar completamente os dados de entrada (intra-quadro) [19].
- A eficiência da Transformada Cosseno Discreta (DCT) é comparável com a KLT e tem sido amplamente usada em processamento de imagem e sinais justamente pelo seu comportamento próximo ao ótimo.

Muitos outros fatores são conhecidos na literatura de trabalhos e livros relacionados [20] e [22] e conclui-se em [7] que, se uma transformada tem que ser escolhida para compressão em codificação de imagem em tempo real, é preferível adotar a Transformada Cosseno Discreta, mas considerando:

- Problemas de não casamento da transformada inversa.
- Problema de uso intensivo de processamento paralelo.
- Problema de complexidade de hardware

A DCT tem o melhor comportamento de compactação da energia para dados de entrada altamente correlacionados como o caso de imagem de vídeo [7], entretanto esse comportamento perde para dados descorrelacionados.

3.8 Propriedades do uso de Transformadas: a Transformada Ótima e Imagem Markoviana.

Codificação por Transformada explora a correlação e periodicidade mas deve-se ter cuidado e não confiar absolutamente nestes fatores para a ótima codificação por Transformada. As amostras a serem codificadas devem primeiro ser ordenadas dentro de blocos. Amostras dentro de um bloco não precisam ser continuadas no tempo como no caso da televisão, onde grupos de amostras são escolhidas de linhas adjacentes e quadros consecutivos. Cada bloco é linearmente transformado ao domínio da transformada, onde tem-se a propriedade de que o sinal fica concentrado em poucos coeficientes transformados comparando com o número de amostras do bloco original.

Na Codificação por Transformada Adaptativa (ATC), os coeficientes selecionados para a transmissão mudam de bloco para bloco, da mesma maneira como cada coeficiente faz na estratégia de quantização como se verá mais para frente.

Para a análise de processamento linear de imagem é adotada uma notação de vetor-espaco. O arranjo de entrada $\mathbf{f}(q,t)$ é representado como uma matriz $\mathbf{f}(q,t)$ ou como um vetor $\bar{\mathbf{f}}(q,t)$ obtido por varredura de fila ou coluna da matriz $\mathbf{f}(q,t)$. O índice t é a referência espacial. Para direcionar um EI no arranjo $\mathbf{f}(q)$ a notação $f(q_{x,y})$ é usada e para o domínio da transformada será $F(q_{u,v})$.

A transformada usada neste trabalho é de tipo real e unitária e é ortogonal se:

$$\mathbf{T}^{-1} = \mathbf{T}^t \quad (3.34)$$

A transformada ortogonal é separável, se,

$$\mathbf{T} = \mathbf{T}_c \otimes \mathbf{T}_r \quad (3.35)$$

onde,

\otimes = produto de Kronecker

\mathbf{Tc} e \mathbf{Tr} são operadores de colunas e filas de tamanho $N \times N$.

Como consequência a transformada bidimensional pode ser escrita como,

$$F(q, t) = \mathbf{Tc} \cdot f(q, t) \cdot \mathbf{Tr}^t \quad (3.36)$$

onde,

$F(q, t)$ = Matriz de saída de $N \times N$

$f(q, t)$ = Matriz de entrada de $N \times N$

\mathbf{Tc}, \mathbf{Tr} = Matriz de transformadas de colunas e filas de tamanho $N \times N$.

Posto que a transformada usada é simétrica e ortogonal as seguintes propriedades são válidas,

$$\mathbf{Tc} = \mathbf{Tr} = \mathbf{T} \quad (3.37)$$

$$\mathbf{T}^t \cdot \mathbf{T} = \mathbf{T} \cdot \mathbf{T}^t = \mathbf{I}$$

Logo, (3.36) fica,

$$F(q, t) = \mathbf{T} \cdot f(q, t) \cdot \mathbf{T}^t \quad (3.38)$$

A transformada bidimensional pode ser interpretada como a transformação de um arranjo $N \times N$ de EI $\mathbf{f}(q, t)$ num arranjo de $N \times N$ de coeficientes $F(q, t)$ na seguinte transformação direta:

$$F(q_{(u,v)}, t) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \mathbf{f}(q_{x,y}, t) \cdot \Phi_{u,v}(x, y) \quad (3.39)$$

E a transformada inversa,

$$\mathbf{f}(q_{x,y}, t) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(q_{u,v}, t) \cdot \Phi_{x,y}(u, v) \quad (3.40)$$

A transformada direta pode ser vista como um produto interno das funções bases $\Phi_{u,v}$ e o arranjo de dados de entrada $\mathbf{f}(q,t)$ que atuam como fator ponderador da soma das funções bases $\Phi_{u,v}$.

A transformada de Karhunen-Loève é ótima no sentido da decorrelação total dos dados de entrada na saída dos coeficientes. Para obter isso, primeiro deve-se calcular a matriz covariância do arranjo de entrada $\mathbf{f}(q,t)$ que é empilhada na forma de vetor $\bar{\mathbf{f}}(q,t)$ na equação seguinte,

$$[COV_{ff}] = E \left\{ [\bar{\mathbf{f}}(q,t) - E\{\bar{\mathbf{f}}(q,t)\}] \cdot [\bar{\mathbf{f}}(q,t) - E\{\bar{\mathbf{f}}(q,t)\}]^t \right\} \quad (3.41)$$

e,

$$[COV_{ff}] = E \left\{ \bar{\mathbf{f}}(q,t) \cdot \bar{\mathbf{f}}(q,t)^t \right\} - E \left\{ \bar{\mathbf{f}}(q,t) \right\} \cdot E \left\{ \bar{\mathbf{f}}(q,t)^t \right\}$$

Onde $E\{\cdot\}$ representa a Esperança e $\bar{\mathbf{f}}(q,t)$ o vetor empilhado. Usando a transformação unidimensional e substituindo em (3.41), a matriz Covariância congruente no domínio da transformada COV_{FF} fica,

$$[COV_{FF}] = [T] \cdot [COV_{ff}] \cdot [T]^t \quad (3.42)$$

Para que a transformada K-L seja ótima a matriz COV_{FF} a ser construída deve ser do tipo matriz diagonalizável. A matriz COV_{FF} pode ser representada por uma matriz diagonal se e somente se, a matriz COV_{ff} ou a transformada tem as funções bases formadas por autovetores. Quando $[T]$ consiste de autovetores a matriz COV_{FF} transforma-se em matriz diagonal Λ , cujos autovalores λ encontram-se em sua diagonal [23].

$$[COV_{FF}] = \lambda_{FF} \cdot [COV_{ff}] = [T_{KLT}] \cdot [COV_{ff}] \cdot [T_{KLT}]^t \quad (3.43)$$

Onde o tamanho das matrizes é $N^2 \times N^2$.

Embora a transformada K-L seja teoricamente a ótima, não é prática para implementação em tempo real. Para a determinação de um bloco da transformada K-L de $\mathbf{f}(q,t)$ os seguintes passos são necessários,

- Determinar a matriz autocovariância $[COV_{ff}]$.

- Diagonalizar a matriz previamente obtida e construir a matriz de transformação usando os auto-valores.
- Transmitir a matriz de transformação e transformar os dados considerados.

Imagem Markoviana

Matrizes da forma da equação(3.41), com elementos diagonais iguais são de muita importancia geral na teoria de processamento de sinais e são conhecidas como matrizes de tipo “Toeplitz”. Se a matriz no domínio espacial se normaliza com respeito à Variância, apresenta uma forma,

$$COV(x) = \sigma^2 \cdot \begin{pmatrix} 1 & \rho_1 & \rho_2 \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 \dots & \rho_{k-2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} \dots & 1 \end{pmatrix} \quad (3.44)$$

Onde,

ρ_k é o coeficiente de correlação no passo k, entre EI adjacentes.

Sabe-se que um deslocamento em sentido horizontal ou vertical nos EI adjacentes, não tem importancia na variância, já que ela mesma é um produto médio entre os elementos adjacentes [7]. Na teoria de propriedades estatísticas de imagem, existe um caso que é muito utilizado na representação de imagens e no qual existe uma relação funcional simple da correlação ρ_k entre os EI adjacentes e que permite posteriores simplificações, considerando que o coeficiente de correlação no passo k encontra-se relacionado ao coeficiente de passo k=1 por,

$$\rho_k = \rho^{|k|} \quad (3.45)$$

Onde,

$$\rho = \rho_1$$

Assim,

$$COR(x) = \begin{pmatrix} 1 & \rho & \rho^2 \dots & \rho^{k-1} \\ \rho & 1 & \rho^1 \dots & \rho^{k-2} \\ \cdot & \cdot & \cdot & \cdot \\ \rho^{k-1} & \rho^{k-2} & \rho^{k-3} \dots & 1 \end{pmatrix} \quad (3.46)$$

A matriz correlação (3.46) ou matriz covariância normalizada, é a matriz de um processo Markoviano de primeira ordem estacionário. Existe uma definição formal para esse conceito que é dada na referência, [30]. Apenas para a presente dissertação é possível definir como imagem Markoviana, aquela onde a influência histórica ou de vizinhança dos EI prévios numa linha ou numa fila ao elemento atual analisado esta apenas contido no elemento prévio. De fato este modelo se aproxima bastante na relação correlação/deslocamento na matriz de EI, para imagens bem comportadas e, ainda trabalha bem nos primeiros passos em imagens ativas [7].

O modelo Markoviano de primeira ordem é uma caracterização da correlação em uma dimensão definidos pelo conjunto de valores da primeira fila ou coluna da matriz (3.46). Isto significa que assumindo o coeficiente de correlação de passo $k=1$ com valor 0.95 (valor típico de imagens bem comportadas Markovianamente), os valores da primeira linha da matriz serão:

$$\rho_k = \rho^{|k|} = 1.000 \quad 0.950 \quad 0.903 \quad 0.857 \dots$$

$$k = 0 \quad 1 \quad 2 \quad 3 \dots$$

Também, e utilizado para o coeficiente de correlação comportado como modelo

Markoviano a seguinte equação,

$$\rho_k = \exp^{-\alpha \cdot |k|} \quad (3.47)$$

Onde, $\alpha = -\ln\rho$. No exemplo prévio, o valor $\rho=0.95$ corresponde a $\alpha = 0.0513$.

A imagem Markoviana de primeira ordem, é o modelo mais utilizado como representativo de imagens reais bem comportadas pelos investigadores com algumas limitações. Este será o modelo de referência nesta dissertação.

A KLT não será considerada nesta tese, pois sua implementação não é prática e ela não é a única transformada que explora estruturas redundantes. A KLT é mencionada apenas como referência com respeito a outras transformadas.

3.9 A Transformada Cosseno Discreta

A transformada bidimensional pode ser interpretada como um “mapping” de um arranjo de $N \times N$ E.I. $f(q,t)$ em um arranjo de $N \times N$ coeficientes $F(q,t)$. A transformada Cosseno Discreta (DCT) foi escolhida da teoria clássica e de várias simulações primárias [7] como a melhor transformada sub-ótima. Esta será a transformada usada neste trabalho.

Os argumentos conhecidos para o uso da DCT são:

- Comportamento é próximo ao ótimo da Transformada de Karhunen e Loève para uma imagem Markoviana [21].
- As funções bases de imagem adaptam-se melhor ao Sistema Visual Humano (SVH).

A Transformada Cosseno Discreta bi-dimensional é definida como:

$$F(q_{u,v}) = \frac{2}{N} C(u) C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(q_{x,y}) \cos\left(\frac{\pi u(2x+1)}{2N}\right) \cos\left(\frac{\pi v(2y+1)}{2N}\right) \quad (3.48)$$

Onde,

$u,v = 0,1,2,\dots,N-1$

A transformada Cosseno Discreta Inversa denominada IDCT, é dada por:

$$f(q_{x,y}) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)F(q_{u,v})\cos\left(\frac{\pi u(2x+1)}{2N}\right)\cos\left(\frac{\pi v(2y+1)}{2N}\right) \quad (3.49)$$

Onde,

$x,y = 0,1,2,\dots,N-1$

e,

x,y = coordenadas no domínio do espacial.

u,v = coordenadas no domínio da transformada.

Com,

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & \text{para } u,v=0 \\ 1 & \text{fora} \end{cases}$$

Estes algoritmos são implementados em linguagem Fortran nesta dissertação.

3.10 Implementação da DCT

A DCT pode ser implementada de várias maneiras, seja usando multiplicação por matrizes completas ou usando algoritmos rápidos com suas matrizes esparsas. De qualquer forma deve-se recorrer à precisão da entrada e da saída. O comprimento de palavra digital atual é de importância para o processo de codificação. Uma palavra muito pequena pode causar uma quantização e truncamento de maneira que os valores transformados não permitirão uma reconstrução com exatidão. Um comprimento de palavra digital que seja muito longa causará perda na eficiência do “codec”. Na literatura, encontram-se diferentes tipos de soluções com respeito a algoritmos rápidos; são dignos de mencionar Chen [24] 1977, Lee [25] 1985 e Vetterli/Nussbaumer [26] 1984 e outros mais recentes baseados na Transformada Generalizada de Chen (GCT) 1991 [27].

A base dos algoritmos rápidos é o tipo Cooley-Turkey. Este consiste de $\log N$ estados, por exemplo, para uma transformada de 8 pontos, são necessários $\log 8 = 3$ estados. Cada estado consiste de borboletas, que representam os elementos básicos para construir o algoritmo da matriz esparsa.

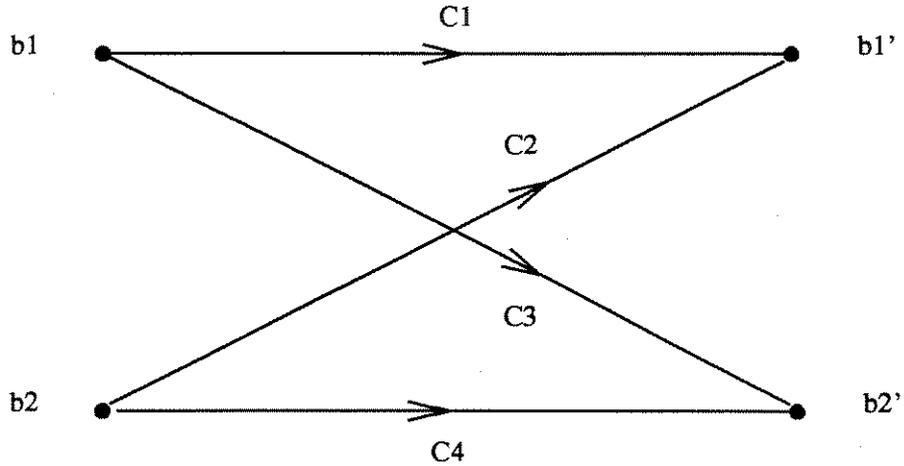


Figura 3.6: Exemplo de uma borboleta

Da figura (3.6) as seguintes equações podem ser formuladas:

$$\begin{cases} b'_1 = b_1 \cdot C_1 + b_2 \cdot C_2 \\ b'_2 = b_1 \cdot C_3 + b_2 \cdot C_4 \end{cases} \quad (3.50)$$

A linha inferior, ou seja $b_2 \rightarrow b'_2$ normalmente tem polaridade negativa. A matriz que pertence à borboleta e suas equações tem a forma.

$$[T] = \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix} \quad (3.51)$$

Para $N=2$ a matriz de transformação DCT é igual à matriz de transformação de Walsh Hadamard, $C_1 = C_2 = C_3 = C_4 = 1$. Para $N=4$ são necessários dois estados, duas borboletas como a figura (3.6) e uma borboleta com o dobro de comprimento da borboleta prévia. Esta aproximação é válida para solução com “deslocamento no lugar”. Na figura (3.7) duas soluções de 4 pontos são dadas.

A estrutura com “deslocamento no lugar” tem a estrutura de matriz esparsa dada por,

$$[S_N] = \begin{pmatrix} S_1 \frac{N}{2} & S_2 \frac{N}{2} \\ S_3 \frac{N}{2} & -S_4 \frac{N}{2} \end{pmatrix} \quad (3.52)$$

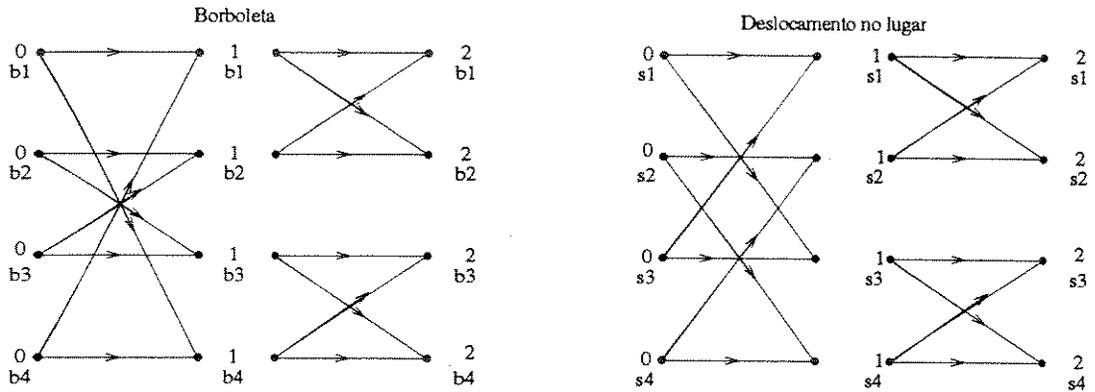


Figura 3.7: Duas soluções de 4 pontos

Onde $S_i \frac{N}{2}$ é a matriz identidade $[I_N]$ e $N \geq 2$.

O modelo borboleta tem como resultado uma estrutura de matriz dada por,

$$[S_N] = \begin{pmatrix} B_1 \frac{N}{2} & B_2 \frac{N}{2}^* \\ B_3 \frac{N}{2}^* & -B_4 \frac{N}{2} \end{pmatrix} \quad (3.53)$$

Onde $B_i \frac{N}{2}$ é a matriz identidade $[I_N]$ e $B_i \frac{N}{2}^*$ é a notação da matriz anti-diagonal com números iguais a um como entradas. As equações de ambas soluções serão,

Deslocamento no lugar	Borboleta
Primeira etapa	Primeira etapa
$s_1^1 = s_1^0 + s_3^0$	$b_1^1 = b_1^0 + b_4^0$
$s_2^1 = s_2^0 + s_4^0$	$b_2^1 = b_2^0 + b_3^0$
$s_3^1 = s_3^0 + s_1^0$	$b_3^1 = b_3^0 + b_2^0$
$s_4^1 = s_4^0 + s_2^0$	$b_4^1 = b_4^0 + b_1^0$
Segunda etapa	Segunda etapa
$s_1^2 = s_1^1 + s_2^1$	$b_1^2 = b_1^1 + b_2^1$
$s_2^2 = s_2^1 + s_1^1$	$b_2^2 = b_2^1 + b_1^1$

$$\begin{aligned}
 s_3^2 &= s_3^1 + s_4^1 & b_3^2 &= b_3^1 + b_4^1 \\
 s_4^2 &= s_4^1 + s_3^1 & b_4^2 &= b_4^1 + b_3^1
 \end{aligned}
 \tag{3.54}$$

Disto é possível concluir,

$$b_1^2 = s_1^2$$

$$b_2^2 = s_2^2$$

$$b_3^2 = s_3^2$$

$$b_4^2 = s_4^2$$

Estes exemplos são dados para explicar o problema do não casamento da IDCT. Este problema é causado pela implementação atual dos coeficientes, o truncamento do comprimento de palavra e a normalização dos resultados.

Para a implementação da DCT várias aproximações podem ser feitas. A multiplicação por matriz completa, denominada de “linha-base”, não precisa estrutura especial como no caso dos algoritmos rápidos, porém calcula a DCT por força bruta por meio da multiplicação da matriz previamente mencionada. Os componentes atuais desenvolvidos em VLSI podem realizar várias operações em um ciclo de relógio. Com a introdução de multiplicadores de matriz completa e mais rápidos, outras operações, como filtragem podem ser realizadas e a exatidão interna pode ser melhor entendida e controlada.

Na composição de algoritmos cada método tem sua própria exatidão, pois depende dos métodos de arredondamento usados. A figura (3.8) apresenta os comprimentos básicos das variáveis de entrada-saída na especificação do algoritmo da DCT.

3.11 Implementação da IDCT e Análise do Não Casamento

Foi mencionado que a causa do não casamento da IDCT, a qualidade do “codec” diminui dramaticamente. O método da implementação da IDCT deve ser especificado para assegurar a compatibilidade entre os “codecs”. O não casamento é um erro por causa da

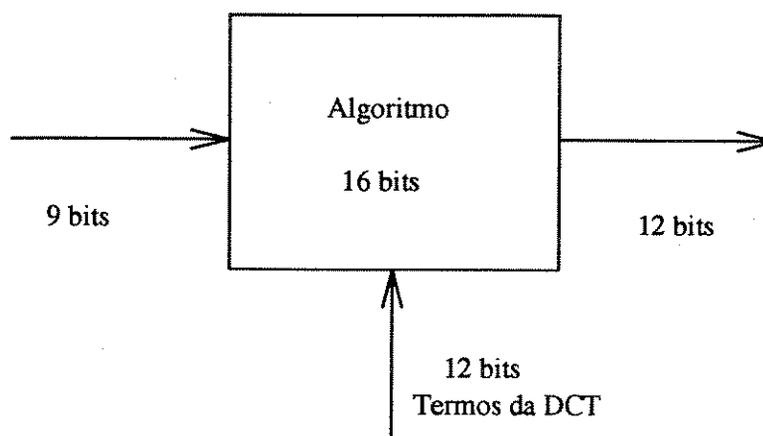


Figura 3.8: Especificação externa do algoritmo da DCT

aritmética de comprimento finito no cálculo da transformada ortonormal no codificador e no decodificador. Esta distorção aumenta em função do tempo. No lado do decodificador o erro sobe e a relação Sinal-Ruído da imagem decodificada degrada quando o número de quadros decodificados aumenta. Logo a implementação da IDCT precisa ser definida ou deve completar requerimentos pré-definidos e incorporar um mecanismo cíclico de “refresh”. O intervalo de “refresh” deverá ser menor que o tempo no qual a imagem é degradada pelo não casamento. O mecanismo de “refresh” não é apenas implementado para consertar o não casamento da IDCT, serve também para outros erros que serão apresentados mais para frente.

Para a especificação da IDCT dois parâmetros precisam ser especificados:

- O máximo não casamento permissível da IDCT.
- O máximo intervalo de “refresh”.

Para o cálculo do não casamento da IDCT é necessário ter em conta a configuração do laço básico do “codec”.

O comprimento da palavra interno é frequentemente truncado a certo número de bits (veja figura 3.8) e o processo de transformação é feito em várias fases de operações dependendo da constituição das borboletas, também o resultado pode ser escalado para ter uma determinada exatidão. A escolha do escalamento e outras implementações específicas

são entregues pelo fabricante do “chip” de tipo VLSI que efetua a transformada. Portanto o uso de implementação com diferentes algoritmos no codificador e decodificador causará a perda da eficiência do decodificador pelo não casamento da IDCT.

A explicação deste defeito pode ser feita no nível de implementação de “linha-base” do processo de transformação. Esta implementação consiste da multiplicação pela matriz completa dos coeficientes com exatidão de 16 bits. A transformada direta consiste na concatenação de 2 operadores tal como foi definido na equação (3.36) por exemplo, logo da multiplicação de matrizes do operador de fila \mathbf{Tr} , os valores são truncados a uma palavra de 16 bits, a saída do operador de coluna \mathbf{Tc} é arredondada a 12 bits e recortada aos maiores valores positivos e negativos. Para a operação inversa é recortada a 16 bits e a saída do operador de fila arredondada a 9 bits e recortada aos maiores valores positivos e negativos.

Na figura (3.9) apresenta o laço básico no codificador usando duas transformadas com as supostas fontes de erros. ΔDCT , ΔQ , Δ_{IDCT}^{cod} , Δ_{IDCT}^{dec} . É preciso notar que ΔDCT é o erro causado pela aritmética finita no cálculo da Transformada Direta Cosseno, ΔQ é o erro de arredondamento no processo de quantização, Δ_{IDCT}^{cod} é equivalente ao ΔDCT e Δ_{IDCT}^{dec} é o erro no decodificador. Frequentemente ΔDCT e Δ_{IDCT}^{cod} são do mesmo codificador e têm a mesma aritmética, mas isso não garante a mesma implementação no decodificador. A implementação depende do fabricante.

Se anotamos como $b(p,t)$ um bloco na posição p de um quadro $q(t)$ ao tempo t , o cálculo pode desenvolver-se na base de blocos, somando todos os resultados dos blocos por quadro. Logo o erro na base de quadros pode representar-se,

$$\Delta q(t) = \sum_{\forall b(p,t) \in q(t)} \Delta b(p,t) \quad (3.55)$$

A influência do número de quadros no cálculo é importante para a determinação do intervalo de “refresh”.

O bloco reconstruído $\hat{b}(p,t)$ no codificador da figura (3.9) pode-se expressar como,

$$\hat{b}(p,t) = bd(p,t) + \Delta_{DCT}(p,t) + \Delta Q(p,t) + \Delta_{IDCT}^{cod}(p,t) + \hat{b}(p,t-T) \quad (3.56)$$

Onde,

$$bd(p, t) = b(p, t) - \hat{b}(p, t - T) \quad (3.57)$$

Agora o bloco reconstruído $\check{b}(p, t)$ no decodificador é expressado por,

$$\check{b}(p, t) = bd(p, t) + \Delta_{DCT}(p, t) + \Delta Q(p, t) + \Delta_{IDCT}^{dec}(p, t) + \hat{b}(p, t - T) \quad (3.58)$$

O erro entre os dois sinais reconstruídos é,

$$\check{b}(p, t) - \hat{b}(p, t) = \check{b}(p, t - T) - \hat{b}(p, t - T) + \Delta_{IDCT}^{dec}(p, t) - \Delta_{IDCT}^{cod}(p, t) \quad (3.59)$$

O erro de não casamento entre os dois sinais decodificados da figura(3.9) será,

$$\Delta_{\text{n\~{a}ocas}}(p, t) = \check{b}(p, t) - \hat{b}(p, t) = \sum_{\forall b(p, t) \in q(t) | tr \leq t < to} \Delta_{IDCT}^{dec}(p, t) - \Delta_{IDCT}^{cod}(p, t) \quad (3.60)$$

Onde “tr” é o tempo do último “refresh”, “to” tempo ao qual será aplicado o “refresh”. O não casamento acumulado ao tempo atual é expressado como $\Delta_{\text{n\~{a}ocas}}(p, t)$. O intervalo entre tr e to é denominado intervalo de “refresh” que é determinado pela degradação tolerável.

O erro total por quadro segundo a equação(3.55),

$$\Delta_{\text{n\~{a}ocas}q}(t) = \sum_{\forall b(p, t) \in q(t)} \Delta_{\text{n\~{a}ocas}}(p, t) = \check{q}(t) - \hat{q}(t) \quad (3.61)$$

O erro entre o sinal original por quadro $q(p, t)$ e o sinal decodificado $\check{q}(p, t)$ pode expressar-se,

$$q(p, t) - \check{q}(t) = \Delta_{\text{n\~{a}ocas}q}(t) + \Delta_{DCT} \cdot q(t) + \Delta Q \cdot q(t) + \Delta_{IDCT}^{dec} \cdot q(t) \quad (3.62)$$

Se Δ_{DCT} , ΔQ , Δ_{IDCT}^{cod} , Δ_{IDCT}^{dec} não são correlados, o erro médio quadrático para as equações (3.61) e (3.62) será,

$$[\check{q}(t) - \hat{q}(t)]^2 = \overline{(\Delta_{IDCT}^{cod}(t))^2} + \overline{\Delta_{IDCT}^{dec}(t)^2} \cdot t \quad (3.63)$$

e,

$$(\check{q}(t) - q(t))^2 = \overline{\Delta Q(t)^2} + \overline{(\Delta_{IDCT}^{cod}(t))^2} + \overline{(\Delta_{IDCT}^{dec}(t))^2} \cdot t \quad (3.64)$$

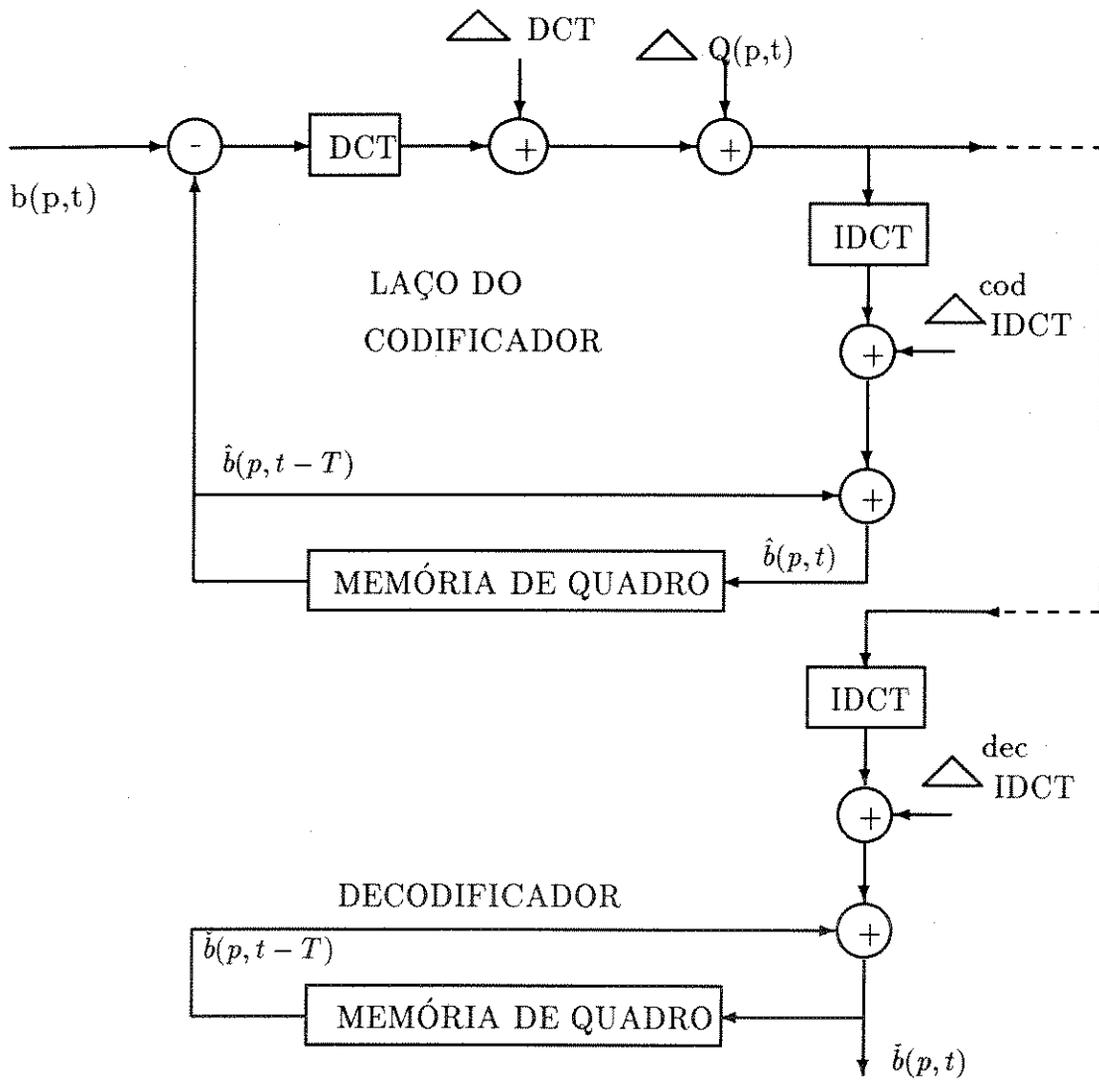


Figura 3.9: Análise do não casamento da IDCT para um esquema híbrido de codificação contendo duas transformadas no laço

Onde,

$$\Delta_{IDCT}^{cod}(t) = \Delta_{DCT}(t) \text{ no codificador.}$$

Se agora,

$\Delta_{IDCT}^{cod}(t)$ e $\Delta_{IDCT}^{dec}(t)$ são correlados ambas equações serão,

$$\overline{[\check{q}(t) - \hat{q}(t)]^2} = (\Delta_{IDCT}^{dec}(t) - \Delta_{IDCT}^{cod}(t))^2 \cdot t \quad (3.65)$$

e,

$$\overline{(\check{q}(t) - q(t))^2} = \overline{\Delta Q(t)^2} + (\Delta_{IDCT}^{dec}(t) - \Delta_{IDCT}^{cod}(t))^2 \cdot t \quad (3.66)$$

Destas equações conclui-se que o erro aumenta com o número de quadros decodificados representados pelo fator t .

Esta situação não acontece numa configuração onde o cálculo é desenvolvido exclusivamente no domínio da transformada. Tal configuração apresenta-se na figura (3.10), onde o laço básico de codificação não contém transformadas.

Considerando a DCT direta e a DCT inversa como uma espécie de pré e pós-processamento e usando a notação $B(p,t)$ para o bloco de entrada $b(p,t)$ transformado, pode ser demonstrado que não há influência da aritmética finita da transformada no processamento de codificação e decodificação.

O bloco $\hat{B}(p,t)$ reconstruído no codificador (veja figura 3.10) é expresso como,

$$\hat{B}(p,t) = B(p,t) - \hat{B}(p,t-T) + \Delta_{DCT}^{cod}(p,t) + \Delta Q(p,t) + \hat{B}(p,t-T) \quad (3.67)$$

Onde pode expressar-se,

$$B(p,t) - \hat{B}(p,t-T) = BD(p,t)$$

No decodificador,

$$\check{B}(p,t) = BD(p,t) + \Delta_{DCT}^{cod}(p,t) + \Delta Q(p,t) + \check{B}(p,t-T) \quad (3.68)$$

O erro entre $\hat{B}(p,t)$ e $\check{B}(p,t)$ é dado por,

$$\check{B}(p,t) - \hat{B}(p,t) = \check{B}(p,t-T) - \hat{B}(p,t-T) \quad (3.69)$$

O erro entre $b(p, t)$ e $\hat{b}(p, t)$ não pode ser dado diretamente devido ao cálculo que é desenvolvido exclusivamente no domínio da transformada. Portanto para calcular o erro devemos assumir a transformada inversa de $\hat{B}(p, t)$ usando a equação (3.69),

$$\hat{b}(p, t) - b(p, t) = \hat{b}(p, t - T) + \Delta_{IDCT}^{dec}(p, t) - \hat{b}(p, t - T) - \Delta_{IDCT}^{cod}(p, t) \quad (3.70)$$

O erro existe fora do laço básico de codificação e não afetará a eficiência de codificação.

Pode-se prevenir o não casamento da IDCT transmitindo o processo de arredondamento como informação lateral, mas para “codecs” de baixa taxa de bits isto não é eficiente devido ao aumento da quantidade de informação. A CCITT SG XV tem especificado o arredondamento e exatidão para a IDCT (apêndice g).

Experimentos flexíveis em “hardware” tem mostrado que o não casamento da IDCT acontece ainda que a IDCT tenha previamente considerado as especificações da CCITT.

Neste caso, o mecanismo do não casamento pode ser explicado para o caso simples de componentes diferentes de zero no domínio da frequência $F(u, v)$.

O valor reconstruído da IDCT $f(x, y)$ da equação (3.49), para no caso de $N=8$ pontos,

$$f(q_{x,y}) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v)F(q_{u,v})\cos\left(\frac{\pi u(2x+1)}{16}\right)\cos\left(\frac{\pi v(2y+1)}{16}\right) \quad (3.71)$$

re-ordenando,

$$f(q_{x,y}) = \sum_{u=0}^7 \sum_{v=0}^7 F(q_{u,v})\frac{1}{2}C(u)\frac{1}{2}C(v)\cos\left(\frac{\pi u(2x+1)}{16}\right)\cos\left(\frac{\pi v(2y+1)}{16}\right) \quad (3.72)$$

Onde a equação (3.72) pode ser representada segundo a equação (3.49) para $u, v = 0$ ou 4 , para cujo valor ambas partes da equação (3.72) conformam números racionais $\pm \frac{1}{2}$.

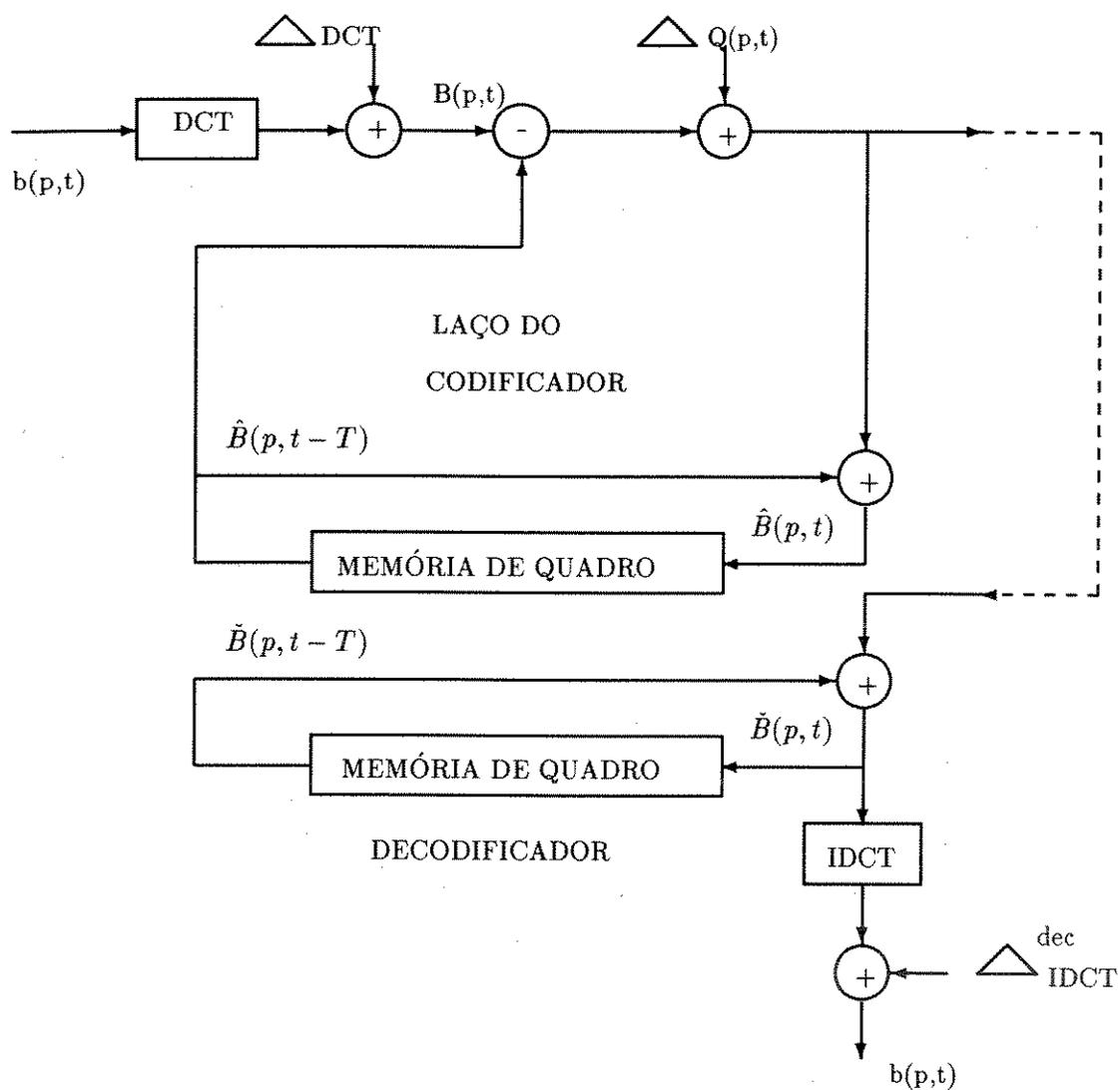


Figura 3.10: Análise do não casamento da IDCT para um esquema híbrido de codificação não contendo transformadas no laço

Se $F(q_{u,v})$ é igual a $(8m + 4)$, onde m é um inteiro, logo a saída da IDCT é $(m + \frac{1}{2})$ e por causa do fator " $\frac{1}{2}$ " a saída pode ser interpretada de diferente maneira no codificador e decodificador justificando o não casamento.

Neste caso, o não casamento é reduzido se os níveis de reconstrução do quantizador projetam-se de acordo com:

$$\begin{cases} r_i = g(2 \cdot a_i + 1) & a_i > 0 \\ r_i = -g(2 \cdot a_i - 1) & a_i < 0 \end{cases} \Bigg| g = \text{par} \quad (3.73)$$

$$\begin{cases} r_i = g(2 \cdot a_i + 1) - 1 & a_i > 0 \\ r_i = -g(2 \cdot a_i - 1) + 1 & a_i < 0 \end{cases} \Bigg| g = \text{impar} \quad (3.74)$$

Onde, g é um fator de escalamento que pode ser de 1 a 31 e considerando que o nível de representação " r_i " é zero para o nível de decisão " a_i " zero.

3.12 Introdução à Transformação de Imagem por Blocos

A configuração de codificação híbrida é construída em torno da DCT a qual representa os blocos de entrada como um conjunto de coeficientes os quais não estão diretamente relacionados com o sistema visual humano (SVH).

Os coeficientes não são decorrelacionados, ou seja, uma estrutura visual não pode ser representada apenas por um coeficiente. Além um quantizador uniforme, como será descrito mais para a frente, produz degradações típicas, portanto para uma ótima quantização é preciso conhecer a melhor relação entre os coeficientes e o SVH.

O problema dos modelos usados são a não estacionaridade do sinal, por isso a interpretação no domínio da transformada é limitada a algumas características da imagem. A eficiência da codificação pode ser otimizada se a informação estrutural primária pode ser reconhecida para adaptar-se à codificação.

3.13 Tamanho do Bloco

O principal objetivo do tamanho do bloco é obter uma redução do número de bits com o máximo de qualidade de imagem e com implementação simples quanto a complexidade do sistema. Neste caso a complexidade do sistema refere-se à quantidade de equipamento necessário no codificador e decodificador.

A escolha para o tamanho do bloco é função de,

- Compactação da Energia.
- Informação lateral a transmitir.
- Possibilidades de adaptação do sinal de entrada.
- Sensibilidade do erro.
- Complexidade da implementação.

Compactação da Energia

Uma das propriedades das transformadas é a compactação da energia. A energia em um bloco $f(p,t)$ pode ser uniformemente distribuída no tempo e ela pode ser transformada de tal maneira que apenas poucos coeficientes são necessários para a representação de toda a informação deste bloco. Das referências [18],[19],[28] conclui-se que quanto maior o tamanho do bloco maior é a compactação da energia e menor o número de blocos representativos, mas, quanto menor o tamanho do bloco, menores as perdas de informação de alta frequência e maior o número de blocos representativos. Quando o número de blocos representativos diminui, poucos bits são necessários de informação lateral. Logo a escolha do tamanho do bloco vai a depender de um compromisso entre a qualidade da imagem e a quantidade de bits de informação lateral por coeficiente. Detalhe desta escolha se verá no próximo capítulo.

O número de blocos representativos segundo o formato da SMPTE de 512x512 E.I., é representado na tabela(3.3).

Tamanho	4x4	8x8	16x16
Luminância(Y)			
Horizontal	128	64	32
Vertical	128	64	32
Total de Luminância	16384	4096	1024
Crominância(U,V)			
Horizontal	64	32	16
Vertical	64	32	16
Total de Crominância	4096	1024	256

Tabela 3.3: Comparação do número de blocos em um quadro com formato SMPTE

3.14 Características no Domínio da Transformada

A transformada utilizada usa a propriedade de separabilidade para o cálculo do bloco transformado em duas dimensões. As funções bases da DCT são dominadas por estruturas horizontais e verticais. Logo é mais fácil representar estes tipos de estruturas no domínio da transformada, do que representar estruturas oblíquas onde os coeficientes espalham-se em todo o bloco transformado.

Experimentos foram feitos nesta dissertação no sentido de detectar o comportamento dominante de orientação dos coeficientes dentro do bloco cujos resultados serão apresentados mais para frente (seção 4.4).

Estruturas periódicas particulares podem ser detectadas no domínio da transformada ao se examinar os valores dos coeficientes. É conhecido que as variâncias na matriz dos coeficientes transformados diminui com maior índice dentro do bloco.

Em codificadores de formas de ondas apenas explora-se as propriedades estatísticas, enquanto que em codificadores baseados no conhecimento experimentos são feitos para explorar a semântica na imagem. Estes últimos tipos usam técnicas de análise e síntese de imagem (Piramidal e Sub-Banda).

A DCT é sub-ótima no sentido que não produz a decorrelação completa das variáveis de entrada. Isto significa que os coeficientes não são independentes. Para coeficientes de estruturas específicas pode ser encontrada a característica para a reconstrução original.

Dos experimentos feitos, conclui-se que quanto maior o bloco menor o conhecimento que pode se deduzir no domínio da transformada. Em objetos inanimados com simetria oblíqua, a energia é concentrada na diagonal. Com este conhecimento pode se projetar uma codificação por zona do bloco, desprezando os coeficientes de alta frequência, mas mantendo intactos os coeficientes da diagonal. Contudo, para imagens naturais, esta estrutura não é tão óbvia.

Posto que as funções bases da DCT tem orientações nas direções horizontal e vertical, logo estruturas horizontais e verticais podem ser representadas com poucos coeficientes com respeito a estruturas de simetrias oblíquas. Portanto para conseguir uma espécie de adaptação é preciso previamente uma classificação do tipo de estrutura o que implica informação lateral necessária a transmitir para o receptor. De novo a eficiência de codificação determinará se a técnica é digna de implementar.

3.15 Quantização e Codificação dos Coeficientes da Transformada

3.15.1 Introdução

A técnica de quantização consiste na forma mais geral, em mapear N vetores desde uma fonte de informação análoga e converte-los numa coleção finita de M palavras digitais de acordo com uma lei específica (característica do quantizador) e transmiti-las através de um canal digital a um receptor, o qual decodifica e reproduz a fonte original.

O conjunto original de valores de entrada pode ser infinito (saída análoga duma fonte contínua), mas também pode ser finito como nesta dissertação que se trabalha com as fontes padrões da SMPTE(Society of Motion Picture and Television Engineer) digitalizadas.

A minimização da informação contida no número finito de palavras digitais produzidas pelo processo de quantização é denominado codificação.

Há vários problemas a considerar no processo de quantização:

1. Definição dum critério de distorção para obter os valores ótimos quantizados.
2. A exploração da correlação residual entre variáveis.
3. A seleção dos parâmetros de quantização de acordo com o critério apropriado ao SVH.
4. A maneira de controlar os dados produzidos pela variação dos parâmetros prévios.

3.15.2 Quantização, Distorção e Entropia

O problema fundamental da teoria da informação é minimizar a distorção “ δ ” para uma dada fonte de informação e sua posterior reprodução com fidelidade.

Conceito de **Entropia** (Distribuição de Primeiro Ordem)

Na teoria da informação, a noção fundamental de “surpresa ou incerteza” leva mais informação para eventos diferentes do que semelhantes. Assim, se um evento $x = X$ acontece com probabilidade $p(x)$, onde x é um EI arbitrário e X é um valor particular da luminância deste elemento, pode definir-se como “auto-informação de ocorrência” ou “conteúdo de informação” ([7], [29], [30]) a seguinte expressão :

$$i(x) = \log\left(\frac{1}{p(x)}\right) = -\log[p(x)] \quad (3.75)$$

Neste caso, o evento $x=X$ é a ocorrência de uma das 2^B saídas inteiras dum conversor A/D, e, a probabilidade de ocorrência de x , ($p(x)$) é o número de vezes que aparece o valor $x = X$, dividido pelo número total de ocorrência de todas as possíveis saídas do conversor. Esta interpretação de frequência relativa é a mais prática devido a sua simplicidade, ainda que existam outras na teoria de informação mais formais, porém complexas.

Dois valores da base do logaritmo da equação (3.75) são usados: o valor 2, onde $i(x)$ resulta

em “bits” como unidade de informação muito conveniente e prezada. O outro consiste em usar a base dos logaritmos naturais com grande aplicabilidade no desenvolvimento teórico da teoria da informação. No caso que as saídas são igualmente prováveis dentro duma possibilidade de saídas igual a 2^B , tem-se,

$$p(x) = \frac{1}{2^B}$$

logo,

$$i(x) = -\log_2[2^{-B}] = B \quad \text{bits} \quad (3.76)$$

Isto significa que uma palavra binária de B bits de comprimento tem B bits de “auto-informação”, ou, “conteúdo de informação”. Agora, o valor médio de “auto-informação” por EI nos diferentes níveis de luminância numa imagem, é determinado pela equação a seguir:

$$H = \sum_{x=1}^{2^B} p(x) \cdot i(x) = -\sum_{x=1}^{2^B} p(x) \cdot \log_2 p(x) \quad \text{bits} \quad (3.77)$$

Os níveis de variação da luminância também podem ir de $x=0$ ao 2^B-1 , deixando a equação anterior da seguinte forma:

$$H = \sum_{x=0}^{2^B-1} p(x) \cdot i(x) \quad \text{bits} \quad (3.78)$$

Observa-se que os valores de $i(x)$ e $H(x)$ dependem apenas das probabilidades de ocorrência dos níveis e não de seus valores. Logo, define-se como Entropia $H(x)$ ao valor médio de “conteúdo de informação” do arranjo, que representa a sequência de saídas dum conversor de imagem. Para o caso de saídas igualmente prováveis, o valor de $H(x)$ é máximo ([29]), a saber:

$$H = -\sum_{x=1}^{2^B} 2^{-B} \cdot \log_2[2^{-B}] = B \quad \text{bits} \quad (3.79)$$

Em cujo caso mostra o mínimo número médio de bits necessários por EI, para representar o arranjo sem perda de informação.

Todas as imagens produzidas naturalmente tem uma Entropia que parte desta situação uniforme e também podem representar-se, em média, com comprimentos menores que B bits por EI. Porém, o grau para o qual $H(x)$ é menor que B bits indicará a facilidade de codificar eficientemente uma imagem com respeito à média de reconstrução perfeita. Neste caso a codificação e reconstrução estão baseadas na base de EI individuais. No caso de aumento do valor de $H(x)$ com respeito a um valor original estaria indicando a existencia de uma fonte de erro adicional.

Devido à sua simplicidade, implementa-se nesta dissertação o algoritmo da equação (3.77) para a avaliação das imagens da SMPTE processadas, apesar de que existe a Distribuição de Segunda Ordem, mais complexa, que permite ser prognosticada a partir da Distribuição de Primeira Ordem com uma exatidão aceitável [7].

Existem duas estratégias para minimizar a distorção, a saber:

- Minimização da distorção δ tendo como limitante a Entropia H.
- Minimização da Entropia H tendo como limitante a distorção δ .

Para decidir as estratégias, a característica do quantizador é muito importante, ou seja, o mapeamento de N variáveis de entrada em M variáveis de saída.

A figura (3.11) mostra a aproximação básica.

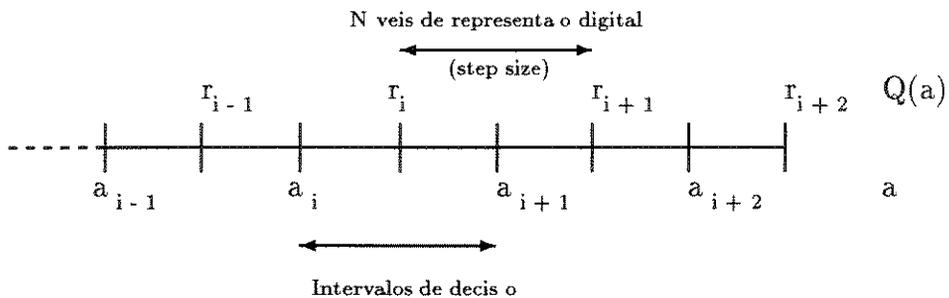


Figura 3.11: Quantização de Variáveis

Algoritmo do quantizador,

$$Q(a) = r_i, \quad \forall a \in (a_i, a_{i+1}) \quad | \quad i = 1, 2, \dots, M \quad (3.80)$$

No caso de um quantizador uniforme o número de bits representativos de acordo com o nível M é definido como,

$$R = \log_2 M \quad \text{bits/amostra} \quad (3.81)$$

Um quantizador uniforme é descrito pelos valores das saídas representativas e pelo seu intervalo uniforme, denominado “passo”. A posição dos valores representativos nos intervalos de decisão é o problema básico.

O quantizador não uniforme tem uma característica com níveis de decisão e valores representativos com diferente passo em função do índice do intervalo, ou seja, quanto maior o índice, maior o nível de decisão e representação.

3.15.3 Minimização da distorção δ tendo como limitante a Entropia H .

Este método, como o seguinte tem como objetivo minimizar o erro.

O método mais usado é o quantizador de Lloyd-Max [19]. O número total de bits usados para os M valores de saída é fixo. Este procedimento minimiza a distorção, sendo conhecida a função probabilidade $p(s)$ do sinal de entrada. A distorção é expressada como,

$$\delta^2 = \sum_{i=1}^M \int_{a_i}^{a_{i+1}} (s - r_i)^2 p(s) ds \quad (3.82)$$

Onde,

s = sinal de entrada

r_i = valor representativo de saída do quantizador.

O erro da equação (3.82) é equivalente ao erro no domínio dos EI. A minimização de (3.82) produz os níveis representativos digitais r_i e os intervalos de decisão a_i ,

$$r_i = \frac{\int_{a_i}^{a_{i+1}} (s) p(s) ds}{\int_{a_i}^{a_{i+1}} p(s) ds} \quad \left| \quad i = 1, 2, \dots, M \quad (3.83)$$

$$a_i = \frac{(r_{i-1} + r_i)}{2} \quad \left| \quad i = 2, 3, \dots, M \quad (3.84)$$

Os valores representativos digitais são as médias condicionais num intervalo de decisão ($d(x,y)$) dado e os limites dos intervalos encontram-se na metade dos valores representativos.

Um dos objetivos da quantização e da codificação entrópica é de encontrar a distribuição de bits ótima.

Para uma função distribuição de probabilidade (fdp), todas as probabilidades condicionais $p(y/x)$ devem ser consideradas. Seja M o conjunto de distribuições de bits, com o requerimento de que a informação mútua média I , por elemento não sobrepassa a taxa R ,

$$M = \{p(y/x) : I(x : y) \leq R\} \quad (3.85)$$

Logo, a distorção mínima deve ser encontrada dentro do conjunto M .

$$\delta(R) = \min_{p(y/x) \in M} \{E\{d(x, y)\}\} \quad (3.86)$$

A equação (3.86) minimiza a distorção, apenas se as variáveis sucessivas são estatisticamente independentes.

Considerando em conjunto a quantização e a distribuição de códigos de palavras, para o caso de um quantizador não uniforme, uma codificação entrópica adicional não produz uma grande melhora, já que a distribuição probabilística foi tomada em conta para os níveis de representação.

No caso de máxima saída entrópica (MSE) a saída do quantizador dará para cada intervalo uma ocorrência igual.(veja figura 3.12). Para um número dado de valores representativos os níveis de decisão devem ser construídos.

Tanto para o quantizador MSE, como do quantizador de Lloyd-Max, ambos tem um carácter não linear. Por exemplo no caso de um número fixo de valores representativos, a largura do intervalo chega a ser pequena para grandes valores de $p(x)$, e, aumenta para pequenos valores de $p(x)$.

O objetivo do quantizador de Lloyd-Max é minimizar a distorção e o objetivo do quantizador MSE é obter igual ocorrência. Isto significa que o ótimo é codificar os valores

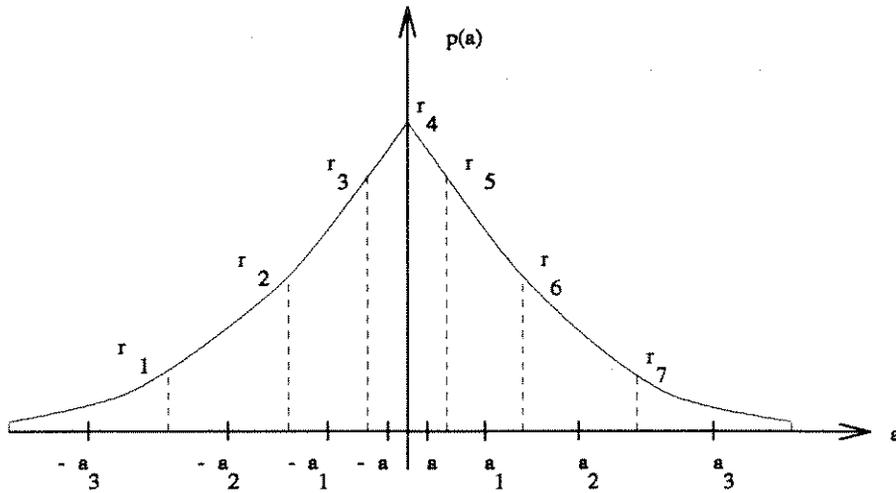


Figura 3.12: Função Densidade de Probabilidade do sinal de entrada nos intervalos de decisão com igual ocorrência

representativos com o mesmo número de bits. (Código de comprimento fixo, CCF). Na figura (3.13) mostra-se um exemplo da característica não linear de quantização.

Devido à não estacionariedade do sinal de entrada, a restrição na implementação de “hardware” e o controle da codificação, prefere-se a segunda estratégia, ou seja, minimizar a Entropia H tendo como limitante a distorção δ ([19],[18], [7], [5]).

3.15.4 Minimização da Entropia H tendo como limitante a distorção δ .

As mesmas equações do método prévio são usadas. O mapa de quantização do sinal de entrada no mais próximo nível de representação $M = 2^n$ é selecionado para minimizar o valor médio quadrático do ruído de quantização. O quantizador é caracterizado com o passo “ g ”, logo os intervalos de decisão e os níveis representativos são determinados com este passo e não mudam com o índice do intervalo como era o caso prévio. A característica do quantizador tem uma zona morta, ou limiar, e é do tipo de início de médio caminho (midtread), já que o do tipo de início de média elevação (midrise) aumentaria o ruído granular.

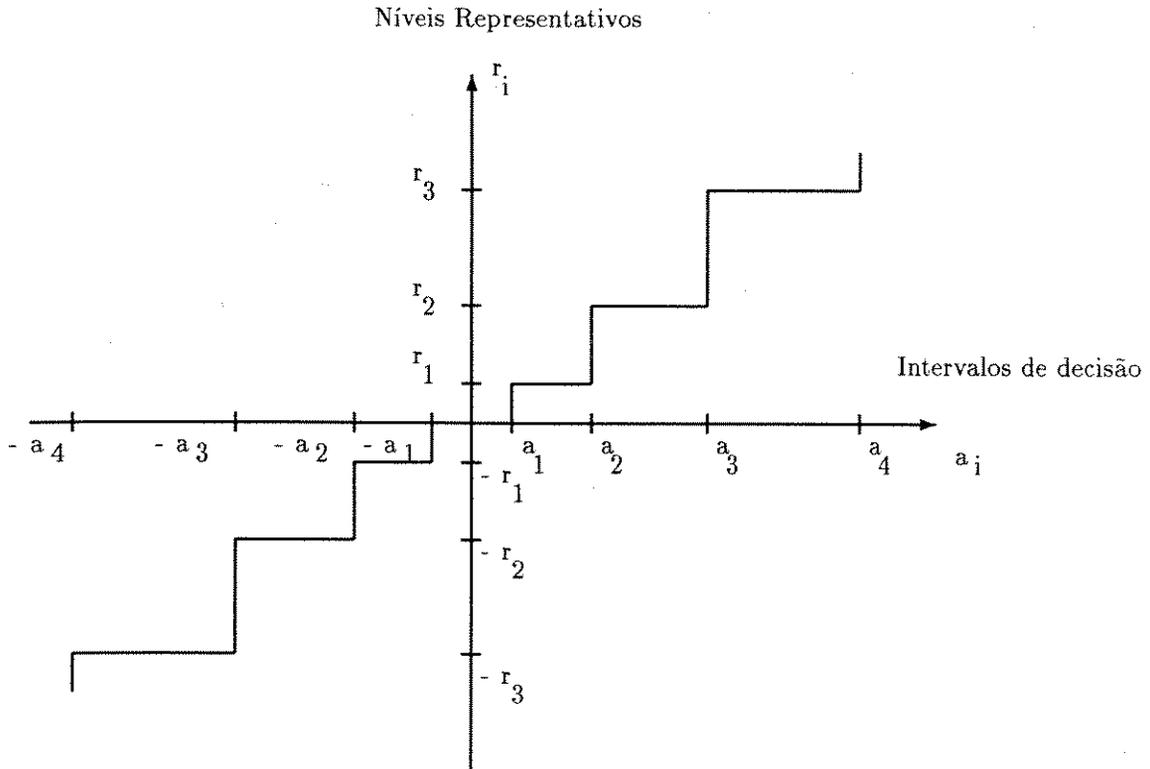


Figura 3.13: Característica não linear de quantização

Na figura (3.14) a_i mostra os níveis de decisão e r_i mostra os níveis representativos.

Os coeficientes são quantizados com uma característica de quantização uniforme. Como será mostrado mais adiante, o passo do quantizador é adaptado ao estado do “buffer”.

O limiar Th é derivado do passo “ g ”; Se $Th=g$, os níveis de decisão a_i e os valores representativos r_i são descritos como:

$$a_i = \begin{cases} Th + (i-1) \cdot g & | i \geq 1 \\ -Th + (i+1) \cdot g & | i \leq -1 \end{cases} \quad (3.87)$$

$$r_i = \begin{cases} \frac{a_{i+1} + a_i}{2} = a_i + \frac{g}{2} & | i \geq 1 \\ \frac{a_{i-1} + a_i}{2} = a_i - \frac{g}{2} & | i \leq -1 \end{cases} \quad (3.88)$$

Devido ao processo de quantização, os coeficientes são truncados, o que produz

o ruído de quantização cuja notação é $n(p,t)$. Este ruído depende do passo de quantização "g". O bloco reconstruído $\hat{B}(p,t)$ no lado do decodificador chega a ser:

$$\hat{B}(p,t) = \hat{B}(p,t-T) + \tilde{e}(p,t) + n(p,t) \quad (3.89)$$

Onde,

$\hat{B}(p,t-T)$ = Bloco transformado e reconstruído no quadro p, e no tempo $t - T$, no decodificador.

$\tilde{e}(p,t)$ = Erro de quantização na diferença de blocos no quadro p e no tempo t, no decodificador.

Uma forma de reduzir o ruído de quantização e implementar um filtro no laço do sistema híbrido, no lado esquerdo da memória de inter-quadro. Ver figura(5.1).

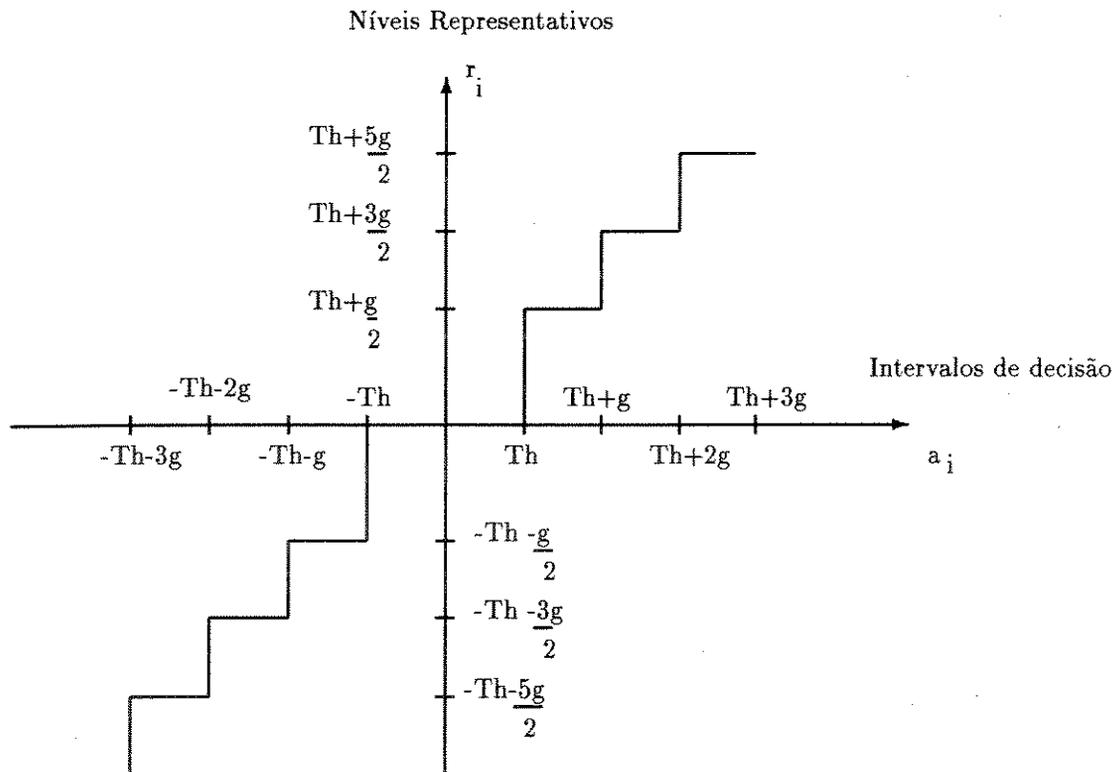


Figura 3.14: Característica de quantização uniforme.

3.16 Codificação dos valores Quantizados

Existem duas formas de selecionar os coeficientes quantizados, por codificação zonal e por codificação por limiar.

3.16.1 Codificação Zonal

Neste tipo de codificação o número de bits a ser assinalado vai depender da variância dos coeficientes. Ao aplicar codificação zonal apenas os coeficientes transformados numa certa zona geométrica são retidos, o resto é submetido a filtragem, zerando-os. Foram feitas várias provas sobre a geometria zonal, seção 4.4 (triangular, não isotrópica e isotrópica) e conclue-se que a forma mais simples é a triangular. Contudo os três tipos apresentam distorções para geometrias particulares que vão depender do conteúdo da imagem.

É importante adaptar o projeto do quantizador à distribuição probabilística dos coeficientes transformados. Existem diferentes suposições com respeito à distribuição probabilística dos coeficientes transformados, tais como distribuição Gaussiana [5] e distribuição Laplaciana [12]. Na referencia [13] (Reininger e Gibson) fizeram teste intensivos nos coeficientes transformados pela DCT num grande conjunto de imagens e concluíram que o modelo Laplaciano adaptava-se em melhor forma aos coeficientes não DC transformados.

O amostragem zonal é função da estrutura geométrica da distribuição da energia dos coeficientes. Esta estrutura depende do tipo de transformada usada, como também da estrutura da covariância da imagem tratada. Em forma geral tem-se considerado a estrutura de covariância de Markov como a mais representativa das imagens reais. Contudo, diferentes estruturas de covariâncias levam a diferentes zonas de amostragem. Por exemplo, uma imagem modelada com uma estrutura separável de Markov, apresenta no domínio da transformada DCT uma distribuição da energia dos coeficientes dada pelo modelo Laplaciano seguinte,

$$r(k, l) = \delta^2 \cdot \exp[-\alpha|k| - \beta|l|] \quad (3.90)$$

onde, α e β são constantes, k e l são as coordenadas dependentes.

3.16.2 Codificação por Limiar

A diferença entre a codificação zonal e a codificação por limiar é que a segunda é orientada ao coeficiente e é um procedimento para o qual todos os coeficientes transformados com uma magnitude, que exceda certo valor de limiar T_h , são quantizados e transmitidos. Neste caso o índice dos coeficientes precisa ser transmitido. O decodificador necessita conhecer qual coeficiente é recebido. Para reduzir a informação lateral para a transmissão dos índices, novas técnicas precisam ser implementadas.

As vantagens de usar codificação por limiar são duas: se reduz o comprimento médio por palavra usada para representar os valores dos coeficientes e pode-se usar em média menor que um bit para os valores de coeficiente zero.

3.16.3 Classes de Varredura

Para evitar a transmissão de muitos coeficientes de valor zero, técnicas de varredura eficiente dos coeficientes dos blocos transformados devem ser empregadas de acordo com a atividade do bloco. Esta eficiência é segundo o número de bits necessários de informação lateral.

Depois da Quantização deve-se transmitir os coeficientes de uma maneira eficiente. O maior potencial dos coeficientes é concentrado nos coeficientes de baixa ordem, especialmente para dados que tem uma alta correlação entre EI. A probabilidade que um coeficiente de alta ordem aconteça é baixa.

O bloco resultante depois da quantização consiste de coeficientes zeros e não zeros. Estes coeficientes são misturados no bloco resultante de acordo com o método de quantização usado. Uma maneira para cancelar a última cadeia de componentes zeros consecutivos é usar um método predefinido de transmitir os coeficientes em combinação com um código de palavra único de fim de bloco (Marcador "end of block", EOB).

Das varreduras: retangular, horizontal, vertical, diagonal e zig-zag, esta última é a classe de varredura mais promissora para transmitir coeficientes eficientemente pelo tempo de processamento. A vantagem do método retangular é que o EOB é desnecessário poupando assim, bits para as outras palavras, porém a informação lateral é grande. O

código EOB é incluído na árvore de codificação Huffman dos níveis quantizados. Uma medida para a eficiência da varredura é o número de zeros antes do último coeficiente não zero. A ocorrência deste coeficiente no bloco transformado pode ser adaptado com sua magnitude ao limiar T_h .

3.17 Distribuição de Bits e Codificação Entrópica

As N variáveis de entrada são quantizadas e mapeadas de acordo com a lei de quantização Q , em M valores de saída. O tamanho do passo é determinado pelo controle de codificação.

A característica de quantização adotada é uniforme, (veja figura 3.14). A distribuição de bit para a representação de valores é construída de acordo com o método Huffman. Este tipo de codificação minimiza a entropia, já que leva em conta a distribuição de frequência dos valores reconstruídos.

A codificação entrópica tem com objetivo melhorar a eficiência de codificação, embora tanto seu projeto bem como o projeto da memória elástica (“buffer”) são independentes do algoritmo a empregar e ambos são muito submetidos à evolução vertiginosa dos chips VLSI por não existir uma respectiva normalização. Além do mais, a própria seleção é baseada no gosto do projetista, visando conservar ou degradar a resolução espacial. Portanto, neste trabalho suas avaliações respectivas não são consideradas.

Capítulo 4

SISTEMA PROPOSTO, SIMULAÇÕES E AVALIAÇÕES

4.1 Introdução.

No capítulo anterior foram apresentados e analisados alguns conceitos básicos do sinal de vídeo e de Codificação por Transformadas.

Neste capítulo apresenta-se um método de compressão de imagem aplicado na codificação digital dos sinais de TV na forma de componentes Y , C_R , C_B , usando um algoritmo adaptativo de bloco variável à Transformada DCT em duas dimensões.

O algoritmo de codificação por Transformada Cosseno Discreta (DCT) convencional apresenta defeitos devido ao tratamento de todas as áreas segmentadas em blocos fixos não sobrepostos em forma indiscriminada. Este tipo de tratamento produz uma qualidade não uniforme para diferentes conteúdos de imagem não estacionários.

Tomam-se como referências as imagens padrões digitalizadas em componentes Y , C_R , C_B usando os processos de pré-filtragem e pós-filtragem da Tese de Doutorado : “Proposta de um Sistema de Codificação MCPD em 34 Mbit/s para Processamento de Sinais de TV em Componentes Y , C_R , C_B ” [1] (Martini), a conformação da imagem em componentes resume-se no uso das equações básicas apresentadas no capítulo 3 : (3.10), (3.11), (3.12) e (3.13) para pré e pós-filtragem e do proceso de digitalização e amostragem das cores primárias apresentada na figura(4.1) derivada da mesma Tese.

Além disso, a configuração completa baseada no algoritmo de bloco variável de compressão é desenvolvida com várias regras adotadas, diferentes das normalizadas para favorecer a apresentação das simulações do novo algoritmo e sem perder a objetividade dos resultados.

Propõe-se aqui uma estrutura de codificação eficiente e um algoritmo base diferente do sugerido pela CCIR e JPEG, visando: o avanço da tecnologia de componentes RISC (processamento paralelo) em processadores de vídeo (C-cube CL550-35), a compressão de imagem e a redução de taxa.

Foram realizados testes subjetivos e objetivos com a visualização das imagens padrão fornecidas pela SMPTE (Society of Motion Pictures and Television Engineers) através de pacotes de avaliação desenvolvidos nesta dissertação e complementados pelos sistemas de análises de imagens: KHOROS (The KHOROS Group. University of New Mexico. E.U.A.);

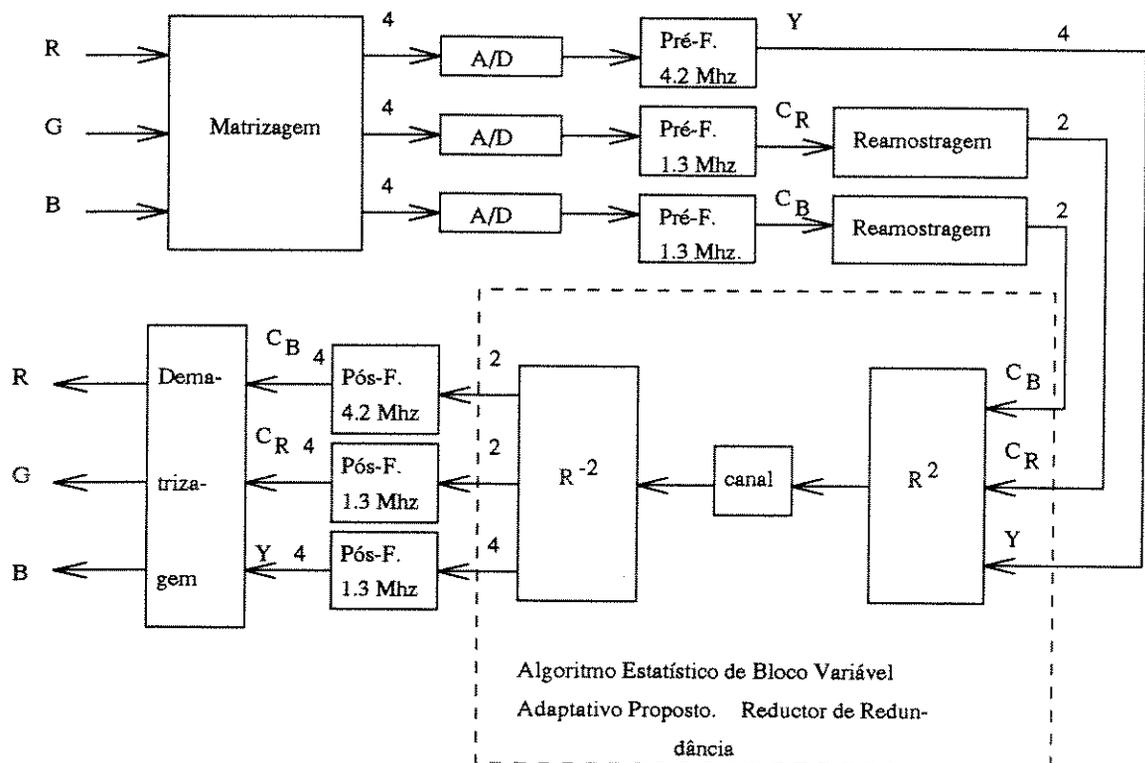


Figura 4.1: Esquema da utilização das sinais de imagem em componentes no algoritmo empregado nesta dissertação

ALV(Phill Everson, University of Bristol.U.K.) e XV (Version 2.21.John Bradley, University of Pennsylvania, 1992), presentes no ambiente público das "Workstations SUN".

Dentro das avaliações de carácter objetivo dos pacotes aqui desenvolvidos, salienta-se as medidas de Sinal/Ruído, Entropia Acumulada e características estatísticas de ocorrência de amostras (histogramas) para as imagens padrão e para sinais de teste determinísticos usados normalmente em avaliações de sinais de TV analógicos. Para avaliar o codificador. Utiliza-se o sinal de Luminância Y por ser mais crítico quanto a informação de vídeo (os sinais diferenças têm menos espectro de frequência ocupado, portanto são menos críticos).

A simulação do codificador ficou baseada no desenvolvimento de pacotes de "software" criados nesta dissertação para simular só as partes fundamentais do sistema, tais como: a Transformada Discreta Cosseno (DCT) em duas dimensões com segmentação de blocos fixo e variável, técnicas adaptativas de seleção de coeficientes e técnicas de quantização que permitirão recuperação da imagem com interferências imperceptíveis e com precisão de codificação de 2 a 3 bits/E.I.. As outras partes do esquema completo de codificação são mencionadas e algumas analisadas lateralmente mas não avaliadas, porque outros autores têm trabalhado exaustivamente tais partes (DPCM, codificação entrópica, memórias elásticas, detecção e correção de erros, multiplexador, etc..) [1], [2].

Neste trabalho são fornecidos resultados de intensos recursos computacionais empregando "workstations SUN" da UNIX de tipo Sparc-2, com programas utilitários tais como: formatador de texto LATEX, graficador GNUPLOT e desenhador XFIG, trabalhando com duas áreas perfazendo em conjunto uma área ocupada de 100 Mbytes comprimidos, associada ao uso da fita magnética e minidiscos de 3 1/2 polegadas de alta e baixa densidade.

A maioria dos programas de simulação foram criados usando linguagem Fortran 77, programação em Postscript e alguns em linguagem C. O tratamento dos programas criados e empregados não são objetos de análise deste trabalho porque constituem um desenvolvimento longo e paralelo cujo objetivo é fornecer os resultados para análise.

4.2 Generalidades do Sistema Empregado

4.2.1 Introdução

A configuração proposta é conhecida como codificador por transformada DCT-2D de Bloco Variável Adaptativo. Adaptativo porque o sistema escolhe o tamanho do bloco de acordo com as ferramentas estatísticas de primeira e segunda ordem.

A estrutura genérica é mostrada na figura (4.2). Nela apresenta-se a conformação dum quadro SMPTE para visar o algoritmo estatístico de bloco variável adaptativo do modelo empregado nesta Tese.

Esta estrutura é baseada nos seguintes programas computacionais desenvolvidos neste trabalho e cujos resultados atingem um tempo superior de avaliação de 300 hrs. de workstations compartilhadas no tempo com outros usuários:

- Matriz de Imagem de formato SMPTE 512x512. Programa “bidefr.for”. Transforma a imagem binária da SMPTE em decimal.
- Segmentação em Blocos. Programa “arruord.for”, produz um ordenamento do arranjo de 512x512 em blocos de 16x16, 8x8, 4x4 e tira estatísticas completas por blocos. Veja figura (4.3)
- Bloco Variável Adaptativo. Programa “adaptativo.for”, produz a seleção do tipo de bloco de acordo com a Média e o Desvio Padrão.
- Série a Paralelo e Transformada Direta. Programa “co2d.for”, produz a multiplicação dos blocos de 16x16, 8x8 e 4x4 para tirar os coeficientes transformados em blocos de 16x16, 8x8 e 4x4.
- Compressão e Quantização. Programa “compress.for”, produz a filtragem de coeficientes em forma adaptativa por zona triangular e por limiar para blocos de 16x16, 8x8 e 4x4 e a respectiva quantização.
- Quantização Inversa. Programa “dequant.for”, produz a quantização inversa adaptativa do sinal recebido.
- Transformada Inversa e Paralelo a Série. Programa “dec.for”, produz a decodificação do sinal em blocos decimais de 16x16, 8x8 e 4x4.

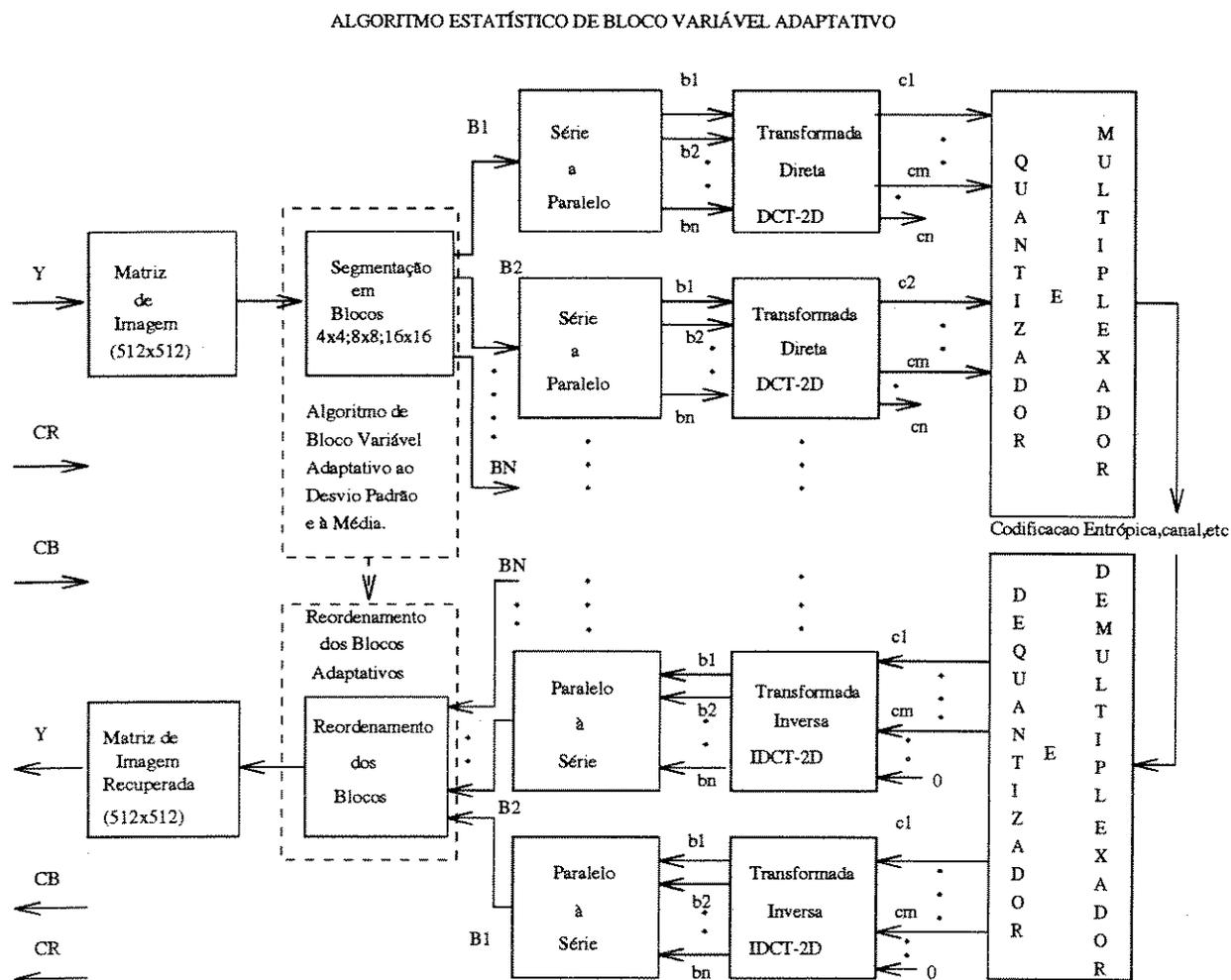


Figura 4.2: Algoritmo Estatístico de Bloco Variável Adaptativo Empregado.

- Reordenamento de Blocos. Programa “recup.for”, produz o ordenamento final do arranjo de 512x512 em inteiros.
- Matriz de Imagem recuperada. Programa “debifr.for”, produz a imagem recuperada no formato digital da SMPTE.

Parte explicativa da aplicação dos programas encontra-se no Apêndice (d).

Ainda que os dados duma imagem real possam não coincidir com o modelo de imagem Markoviano de primeira ordem [7], resultados experimentais tem demonstrado que a DCT (com exceção da KLT) (seção 3.7) supera outras transformações ortogonais

BLOCO No. : 8 16	BLOCO No. 34 47
MÉDIA: 124.508 VARIÂNCIA 419.953 DES. ST. 20.4928	MÉDIA: 91.7812 VARIÂNCIA: 86.7959 DES.ST.: 9.31643
ENERGIA TOTAL DO BLOCO: 15922.1	ENERGIA TOTAL DO BLOCO: 8510.59
116. 116. 115. 114. 111. 111. 106. 101. 97. 97. 96. 93. 88. 95. 97. 97.	118. 110. 104. 101. 105. 101. 100. 94.
114. 110. 106. 105. 101. 99. 95. 94. 94. 95. 96. 97. 97. 102. 105. 105.	115. 110. 106. 95. 97. 95. 95. 92.
101. 96. 95. 87. 92. 91. 92. 97. 96. 97. 102. 106. 107. 111. 113. 114.	105. 96. 91. 92. 87. 88. 88. 93.
91. 95. 95. 95. 96. 101. 100. 100. 103. 105. 107. 111. 114. 111. 117. 120.	113. 101. 93. 87. 86. 85. 89. 87.
92. 96. 100. 102. 105. 105. 108. 109. 109. 111. 113. 115. 116. 117. 119. 116.	103. 95. 87. 86. 88. 87. 83. 83.
100. 102. 103. 111. 110. 111. 115. 112. 108. 117. 117. 115. 121. 121. 122. 121.	95. 94. 90. 81. 86. 86. 82. 86.
104. 107. 110. 109. 110. 112. 113. 116. 114. 116. 114. 119. 122. 120. 122. 125.	89. 90. 87. 81. 82. 83. 80. 80.
110. 108. 110. 113. 111. 118. 114. 116. 123. 118. 119. 125. 125. 126. 129. 132.	91. 87. 81. 84. 80. 82. 86. 80.
110. 111. 111. 108. 115. 117. 121. 126. 127. 132. 135. 133. 137. 140. 142. 146.	
111. 111. 114. 115. 118. 119. 127. 139. 139. 139. 145. 145. 146. 151. 152. 154.	BLOCO No. : 65 104
114. 120. 119. 120. 127. 127. 137. 143. 147. 148. 149. 152. 151. 153. 152. 158.	MÉDIA: 139.625 VARIÂNCIA : 14.1094 DES.ST.: 3.75624
125. 126. 129. 132. 131. 135. 139. 147. 149. 153. 154. 155. 155. 156. 154. 156.	ENERGIA TOTAL DO BLOCO : 19509.2
136. 136. 136. 137. 138. 136. 143. 148. 151. 150. 154. 157. 154. 155. 155. 154.	147. 140. 140. 135.
138. 138. 135. 139. 142. 141. 143. 147. 152. 154. 154. 156. 155. 155. 157. 158.	143. 143. 137. 137.
133. 138. 141. 140. 139. 140. 142. 144. 151. 154. 152. 155. 155. 157. 156. 155.	144. 138. 136. 135.
140. 141. 141. 142. 141. 143. 143. 147. 149. 154. 153. 155. 157. 154. 157. 160.	142. 144. 139. 134.

Figura 4.3: Estatística por bloco de 16x16, 8x8, 4x4.

[7], [6], [20]. Devido a sua alta eficiência e à implementação dos requisitos de velocidade computacional em VLSI e processadores de vídeo (C-cube Cl550-35), atualmente a DCT no contexto da codificação por transformada tem chegado a ser a de maior uso normalizado.

Tomando como referência básica nesta dissertação o codificador convencional por transformada DCT da figura (4.4). Pode-se dizer neste sistema básico, que a imagem de entrada é dividida em pequenos blocos adjacentes não superpostos. Cada bloco dos dados de imagem é transformado usando a DCT que produz na saída uma forma mais compacta dos dados (seção 3.16) facilitando o processo de compressão e filtragem que consiste em selecionar os coeficientes da DCT de acordo com um limiar mínimo de seleção em zona, (figura 4.46) e posterior quantização (capítulo 3, seção 13). Os coeficientes logo são lidos com varredura em zig-zag [20], [22], tomando os valores dos coeficientes não zeros e o número de coeficientes sequencialmente zeros para ser entropicamente codificados.

4.2.2 Defeitos do Sistema Convencional

Em forma geral, um tamanho de segmentação de bloco maior, causa uma maior eficiência de compressão (capítulo 3, seção 16), assumindo que os dados da imagem mantém

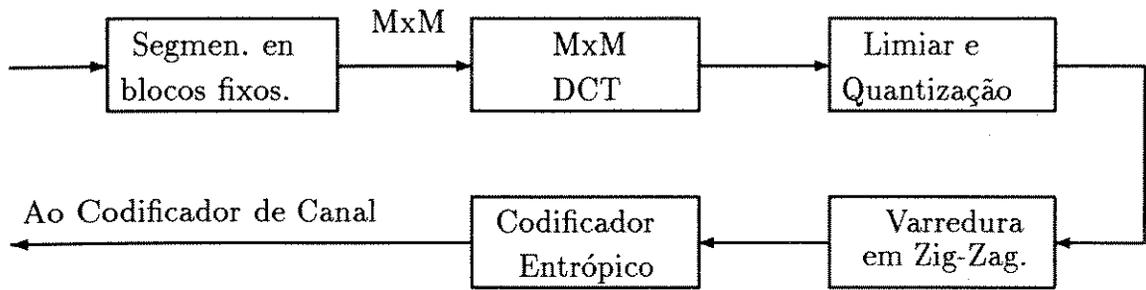


Figura 4.4: Codificador convencional

a estacionariedade espacial. Contudo, esta suposição não se mantém no caso da maioria das imagens típicas, onde os dados dos blocos apresentam não estacionariedade se o tamanho do bloco for grande, (figura 4.3). Através de simulações, utilizando o programa “arruord.for”, mostra-se que os valores da Variância e do Desvio Padrão diminuem à medida que o tamanho do bloco se reduz; contudo essa dependência é relativa em forma particular às transições abruptas do conteúdo da imagem.

Verificou-se através de simulações que a qualidade da imagem sofre degradação para blocos não-estacionários. Por outro lado, um tamanho de bloco pequeno resulta usualmente em baixa capacidade de compressão, porém produz uma alta qualidade de imagem, já que a suposição de estacionariedade mantém-se para quase todos os blocos e, se tiver alguns efeitos serão menos determinantes. Este resultado foi observado avaliando um sistema convencional de blocos fixos não superpostos, para segmentações de 16x16, 8x8 e 4x4, e são apresentados na tabela (4.4)(pag. 87) medidas de Relação Sinal/Ruído e os seus correspondentes tamanhos relativos de segmentação são mostrados para as imagens padrões SMPTE01Y e SMPTE15Y nas figuras (4.5) e (4.6) respectivamente.

Encontrar um compromisso que permita aproveitar as vantagens da capacidade de compressão de blocos grandes e a melhor qualidade da imagem para blocos pequenos constituem a motivação principal e a contribuição desta Tese. A investigação de um algoritmo de blocos variáveis, onde os tamanhos dos blocos são segmentados de acordo com o conteúdo da imagem previamente analisado, permitem uma compressão adaptativa empregando transformada DCT-2D que fornece uma taxa média próxima a 2 bits/amostra.

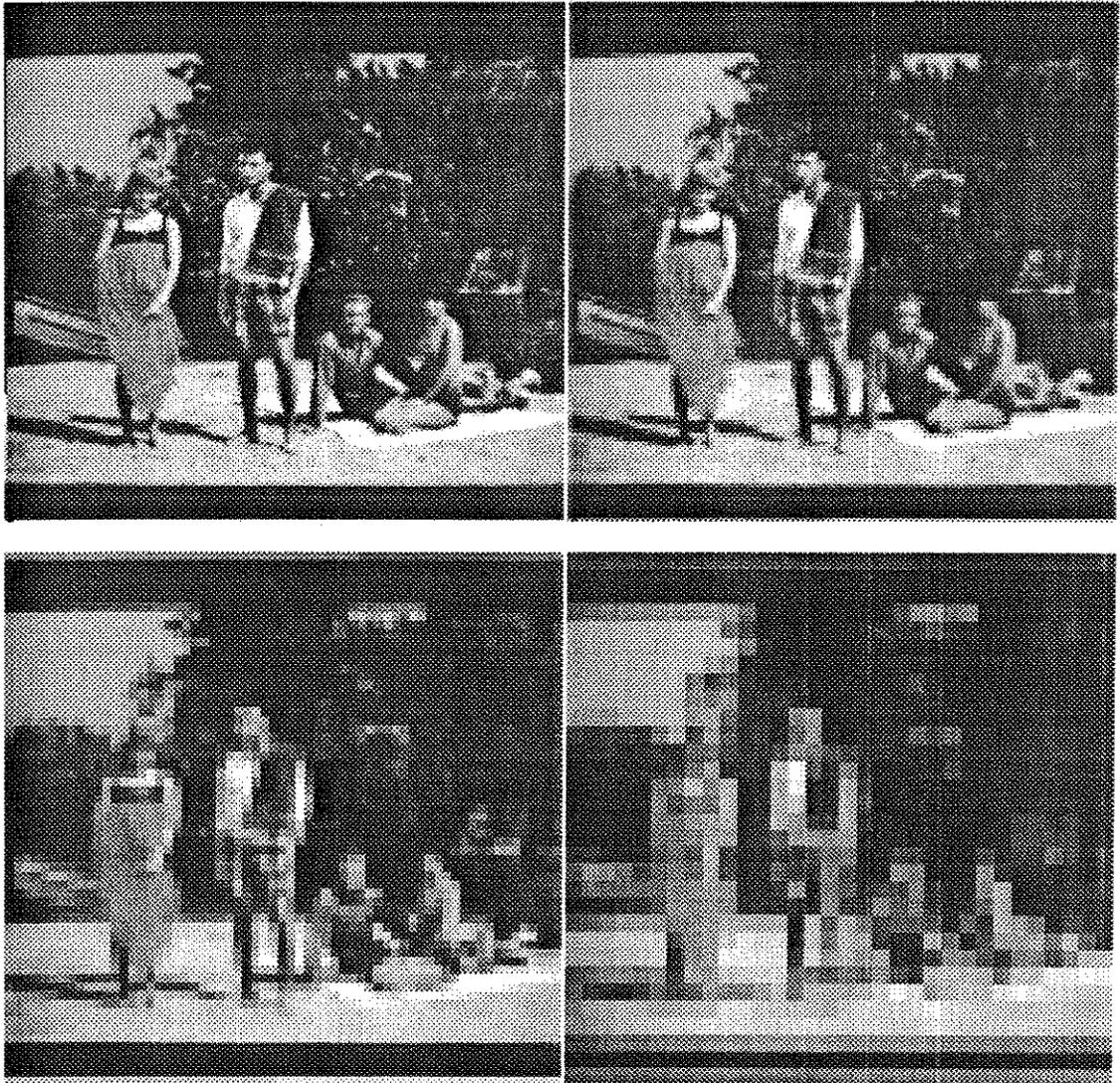


Figura 4.5: Segmentação de tamanho de blocos 1x1, 4x4, 8x8 y 16x16 para a imagem padrão SMPTE01Y.

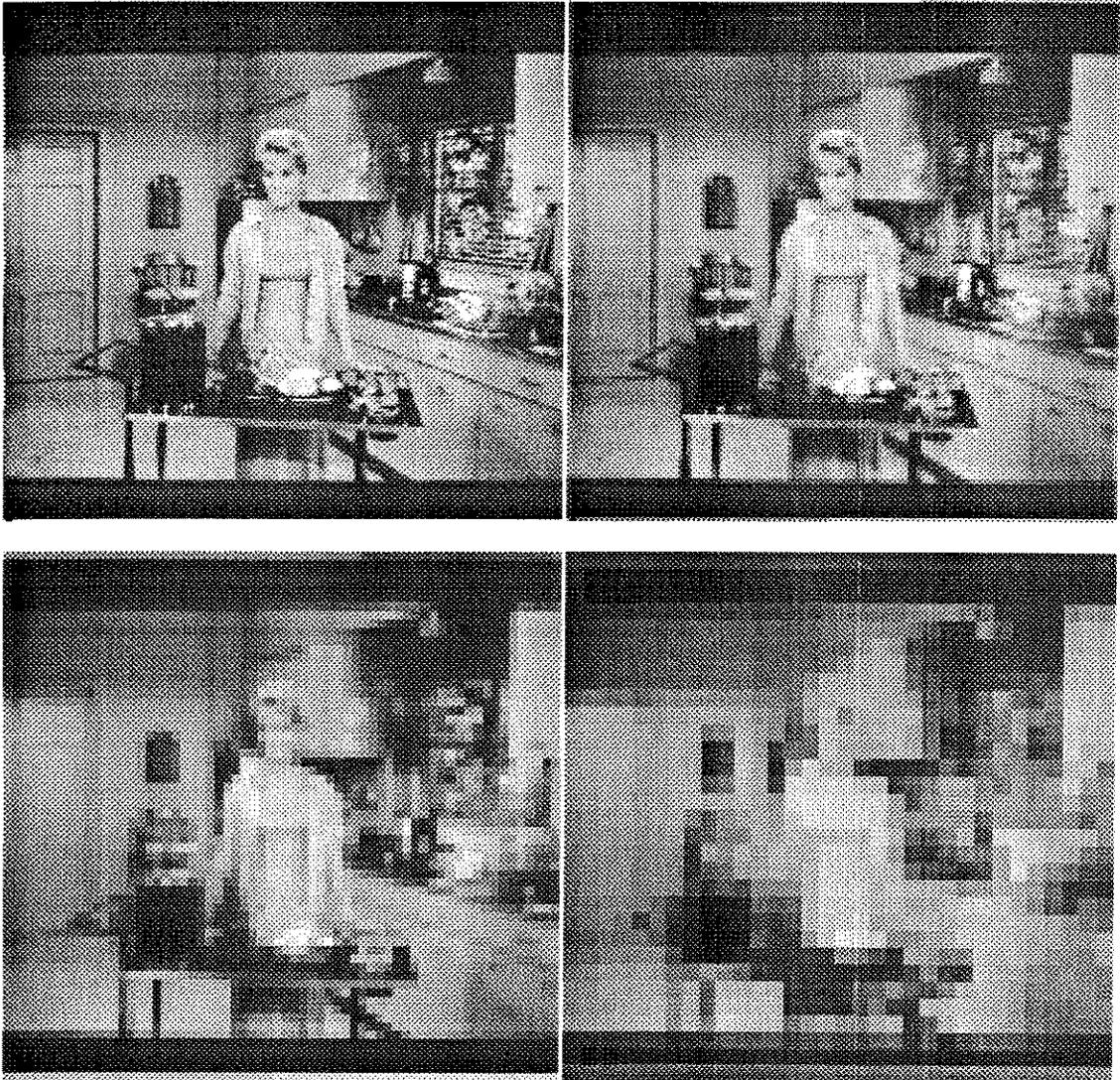


Figura 4.6: Segmentação de tamanho de blocos 1x1, 4x4, 8x8 y 16x16 para a imagem padrão SMPTE15Y.

4.3 Imagens Padrão e Ferramentas de Avaliação.

4.3.1 Introdução - Imagens Digitais Padrão

Em geral a avaliação do desempenho de um sistema ou equipamento de comunicações é realizada através do uso de aparelhos de teste que medem os parâmetros que caracterizam tal sistema ou equipamento. Naturalmente, no caso de desenvolvimento de projetos, tais medidas objetivas são imprescindíveis, porém é também possível se obter uma avaliação do desempenho através de simulações do sistema. As simulações geralmente realizadas com a ajuda de um computador permitem testar uma ampla gama de variações de parâmetros com diferentes condições aplicadas ao equipamento físico real a fim de se prever o comportamento do mesmo nessas situações. Assim a vantagem das simulações em relação à implementação física real é evidente quando é grande o número de opções que se tem para se construir o equipamento empregado. A confiabilidade nos resultados obtidos através de simulações é limitada pela fidelidade do modelo usado para representar o equipamento real. Uma vez que comumente é impossível modelar exatamente a realidade, as simulações têm em geral um caráter de apoio ao projeto, sendo que o desempenho de fato deve ser obtido no equipamento fisicamente implementado.

As imagens padrão primárias da SMPTE são originalmente derivadas de uma forma de diapositivo que é atravessada por uma luz branca para ser lida por uma espécie de dispositivo de ponto flutuante (flying spot scanner) e posteriormente o sinal analógico é amostrado em cerca de 10 Mhz e digitalizado em 8 bits paralelos de forma a se obterem 512 amostras por linha e cujo mosaico fica alinhado na vertical, logo é armazenado em um quadro de 512x512 amostras de 8 bits. Um fato relevante a se destacar é que para as normas de TV que consideram como padrões as frequências de amostragem $3f_{sc}$ e $(8/3)f_{sc}$ as amostras colhidas em 10 Mhz são utilizáveis diretamente devido à proximidade das frequências (nesta dissertação assume-se essa condição), porém para $4f_{sc}$ é necessária a criação de amostras intermediárias já que a diferença entre a frequência de amostragem $4f_{sc}$ e a frequência usada de 10 Mhz é da ordem de 50%. A criação dessas amostras intermediárias é feita por interpolação digital usando-se filtros digitais tipo FIR ("Finite Impulsive Response"). As referências [1] e [32] contém grande detalhe de exemplos para as diferentes frequências de amostragem normalizadas.

4.3.2 Imagens de Teste

Das quinze imagens-padrão digitalizadas da SMPTE, escolheram-se para simulações as imagens com algumas características mais diferenciadas entre si. Dessa forma, usaram-se as imagens em componentes: smpte01y e smpte15y. Além disso, usaram-se outras imagens ; Lena e Babuíno. Essas imagens de teste são apresentadas na figura (4.7) e suas características subjetivas são apresentadas na tabela(4.1). Suas características objetivas estatísticas são calculadas por programas desenvolvidos em fortran 77, e eles resumem-se na tabela(4.2).

IMAG. TESTE	CARACTERÍSTICAS
SMPTE01Y	Muito detalhe, fundo rico, luz diurna, ambiente exterior, bastante iluminação. Comportamento Markoviano*.
SMPTE15Y	Muito detalhe, fundo rico, luz diurna, ambiente interior, bastante iluminação, ênfase na pessoa. Comportamento Markoviano*
LENA	Quantidade razoável de detalhes, luz diurna, ambiente interior, fundo pobre, cabeça e ombro, bastante iluminação, ênfase na pessoa. Comportamento Markoviano*.
BABUÍNO	Muito detalhe localizado, luz diurna, ênfase na cabeça do animal, bastante iluminação. Corportamento não Markoviano*.

Tabela 4.1: Características subjetivas das quatro imagens de teste. * Ver secção 3.8

IMAGEM DE TESTE	Média	Mediana	Moda	Desv.Pad.	Máx.EI	Mín.EI
SMPTE01Y	135.231	103	255	61.2205	255	53
SMPTE15Y	141.579	140	255	56.9956	255	45
LENA	133.929	141	157	42.2286	255	0
BABUÍNO	165.843	168	162	42.3236	210	11

Tabela 4.2: Características estatísticas das quatro imagens de teste

Observa-se que a imagem que apresenta a Média mais baixa é a Lena e a que apresenta o maior Desvio Padrão é a imagem da praia (smpte01y) para todos os 512x512 EI analisados do quadro.

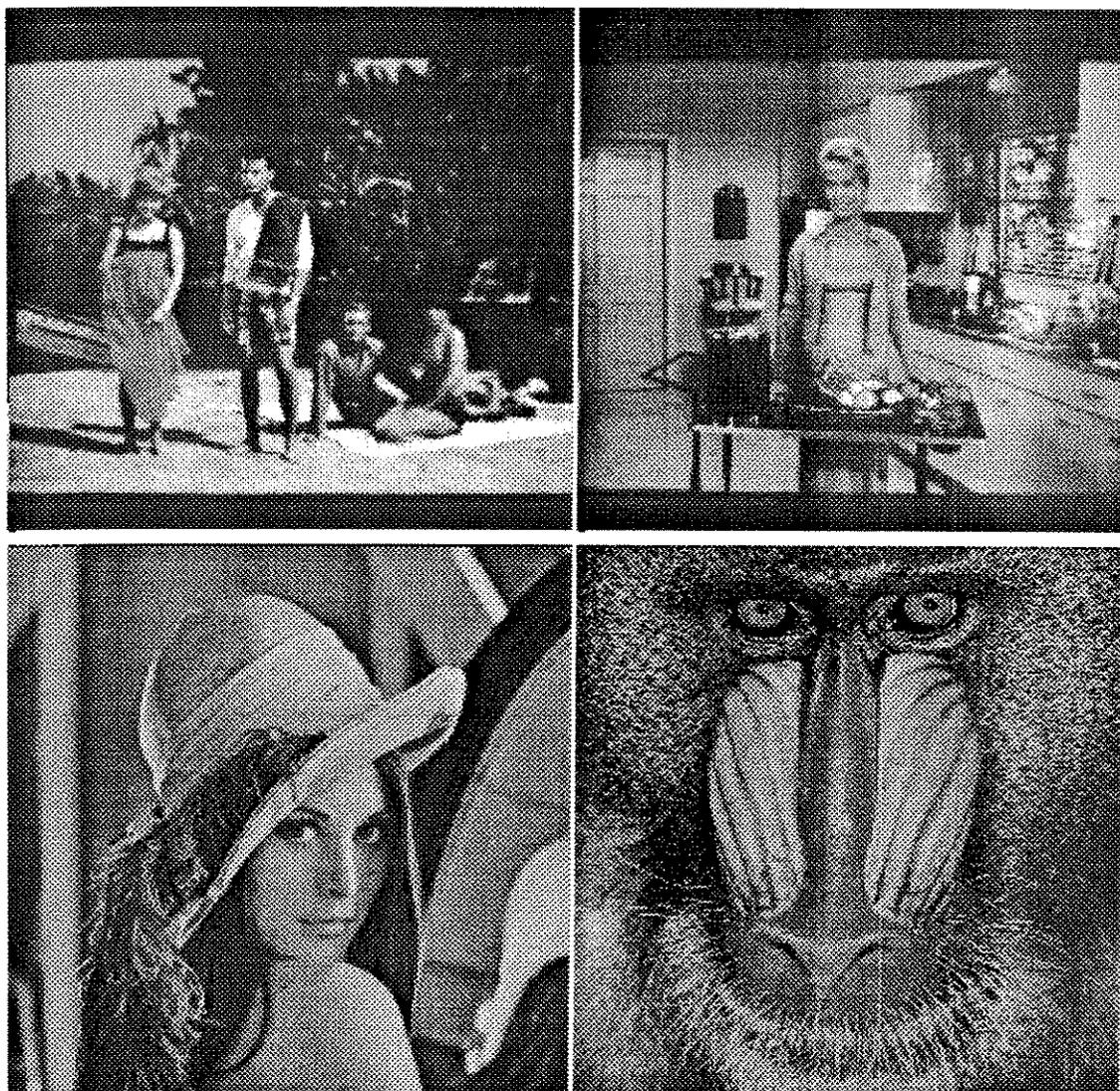


Figura 4.7: Imagens de Teste usadas nesta dissertação: smpte01y, smpte15y, Lena e Babuíno

Como imagem de teste mais objetiva utiliza-se o multiburst ou sinal de teste CCIR II (SST 10) . Os sinais de teste CCIR são determinísticos embora sejam de difícil ocorrência em casos práticos, permitem acompanhar o comportamento das diversas partes do sistema em regiões do sinal com características específicas, já que a sua forma é a priori perfeitamente conhecida. Eles foram padronizados pelo “CCIR” e são gerados tanto por software como por meio eletrônico, veja referência [2].

O sinal de teste CCIR II (SST 10) é um sinal que apresenta uma grande dificuldade à redução de taxa de bits devido a sua complexidade. Seu desenho apresenta-se na figura (4.8) e sua representação na tela na figura (4.9).

Possíveis medidas que podem ser feitas com esta imagem de teste:

- Resposta de Frequência do sistema às diferentes salvos de frequências: 0.5Mhz, 1.0Mhz, 2.0Mhz, 3.0Mhz, 3.58Mhz e 4.2Mhz. Esta medida é importante pois permite verificar o desempenho em termos de faixa de frequência do sistema.
- Ganho Diferencial de Amplitude de Crominância.
- Ganho Diferencial de Fase.
- Intermodulação Crominância-Luminância.

Nessa dissertação apenas utiliza-se para observar a resposta em frequência do sistema.

4.3.3 Pacotes de software desenvolvidos para avaliação do sistema empregado.

São pacotes que permitem avaliar a degradação da imagem processada levando em consideração três fatores diferentes : Relação Sinal-Ruído, Entropia e Histogramas de diferentes parâmetros.

- Média e Desvio Padrão. Programa “ arruor1.for”, permite obter os dados da média e Desvio Padrão, parâmetros que são de controle para o programa “ adaptativo.for” e para determinar seus respectivos histogramas.

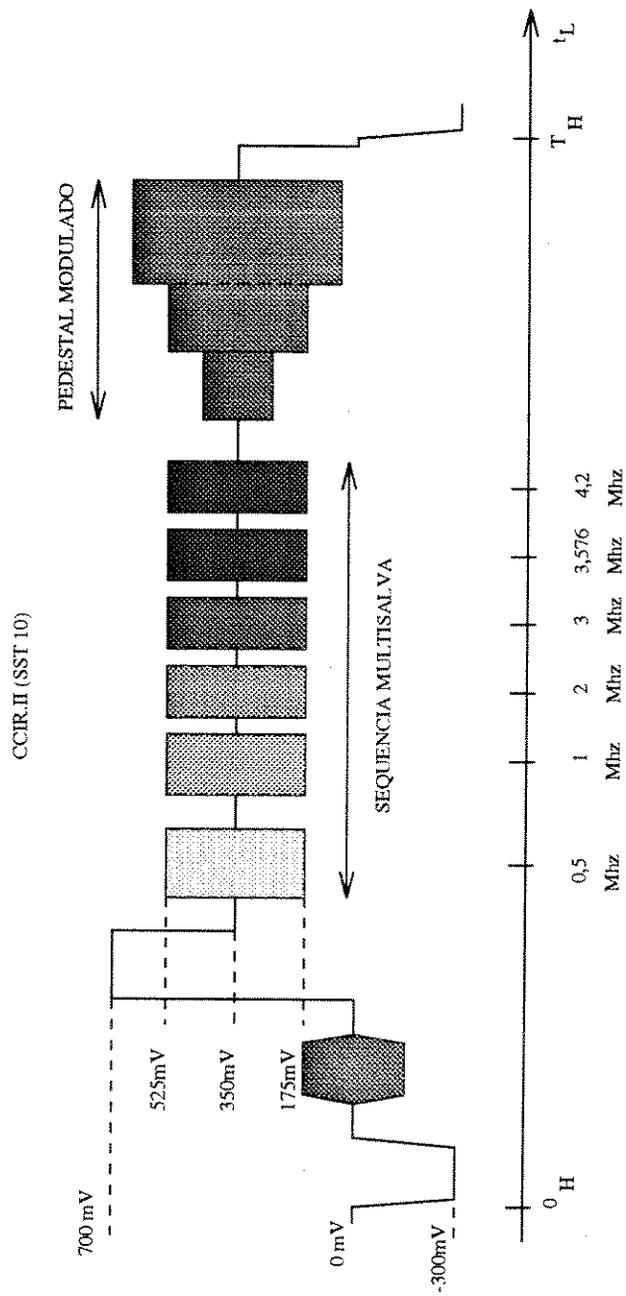


Figura 4.8: Sinal de Teste CCIR II ou Multiburst

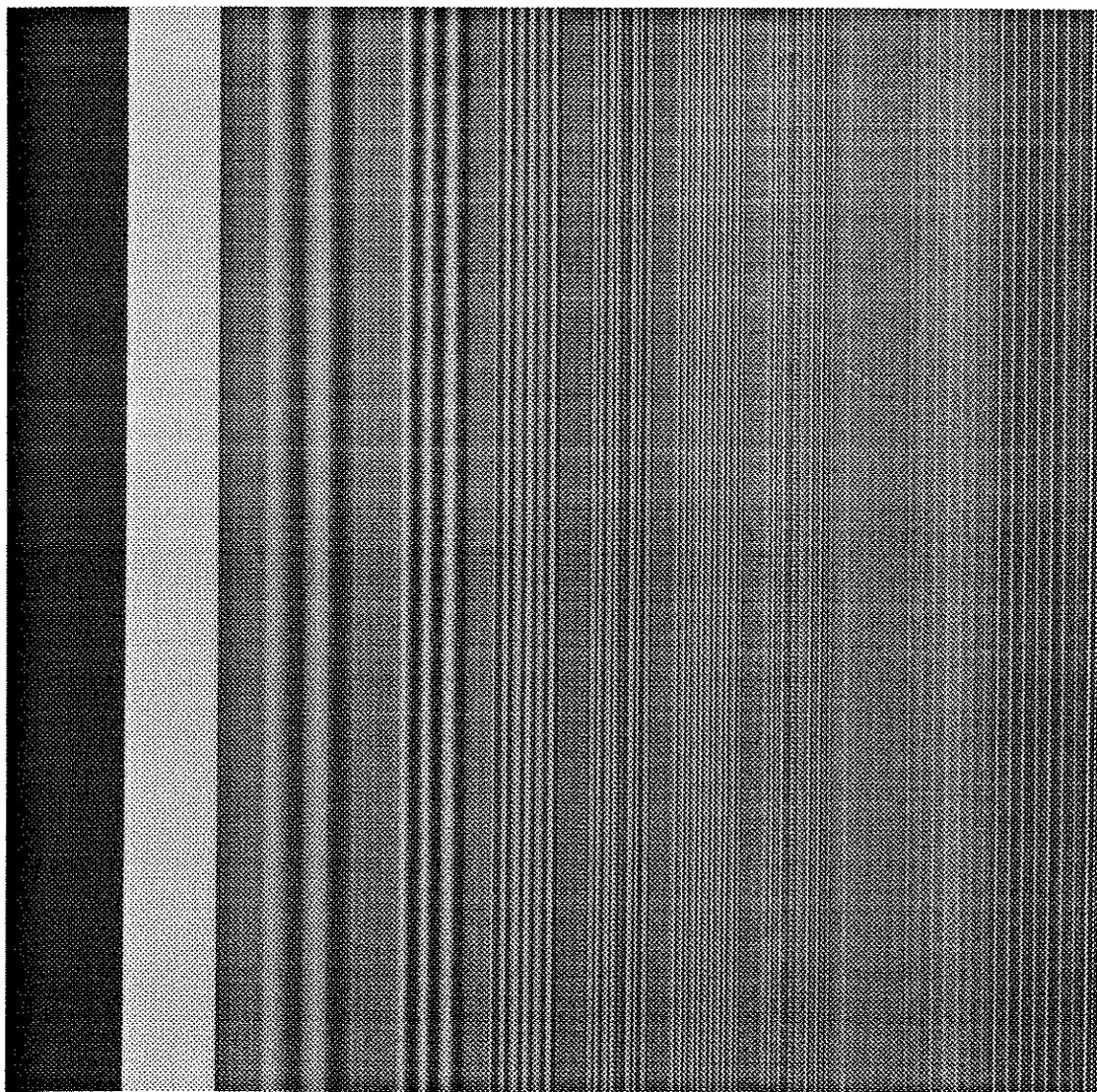


Figura 4.9: Sinal de Teste CCIR II, na tela

- Relação Sinal/Ruído. Programa “snr1.for”, permite comparar o sinal original com o sinal recuperado, tirar a diferença entre os respectivos E.I. dos sinais e defini-la como Ruído, obtendo a relação Sinal/Ruído do sinal recuperado.
- Entropia ou Distribuição de Primeira Ordem. Programa “entrop.for” permite determinar a degradação de uma imagem processada, usando a relação (3.78).
- Histograma dos E.I.. Programa “hist.for”, permite determinar a frequência dos distintos níveis do sinal de Luminância recuperado. Sua comparação com o histograma do sinal original, permite observar a degradação da imagem processada.
- Histograma do Desvio Padrão e da Média “hstestdesv.for” e “hstmed.for” permitem determinar o intervalo apropriado dos parâmetros estatísticos de controle dos blocos variáveis e de filtragem dos coeficientes por zona da transformada.
- Programas combinados, “entrhist.for” e “enthstsnr.for”, permitem avaliar os conceitos de entropia-histogramas e entropia-histogramas e relação Sinal/Ruído em forma combinada.

Parte explicativa da aplicação dos programas encontram-se no Apêndice (d).

4.4 Avaliações para o Sistema Convencional de Bloco Fixo.

4.4.1 Introdução

Nestas primeiras avaliações, depois de implementada a codificação direta e inversa da transformada DCT-2D através dos programas: “bidefr.for”, “arruord.for”, “co2d.for”, “dec2d.for”, “recup.for” e “debifr.for”, procurou-se estudar as características dos coeficientes transformados com o objetivo de determinar a estrutura geométrica mais apropriada para selecionar coeficientes no domínio da transformada. Desta maneira se produz uma filtragem de coeficientes que minimize a degradação da imagem recuperada e como consequência produza um aumento na eficiência de compressão. Esses programas também

foram usados para testar o comportamento das diferentes segmentações de blocos (figura 4.10).

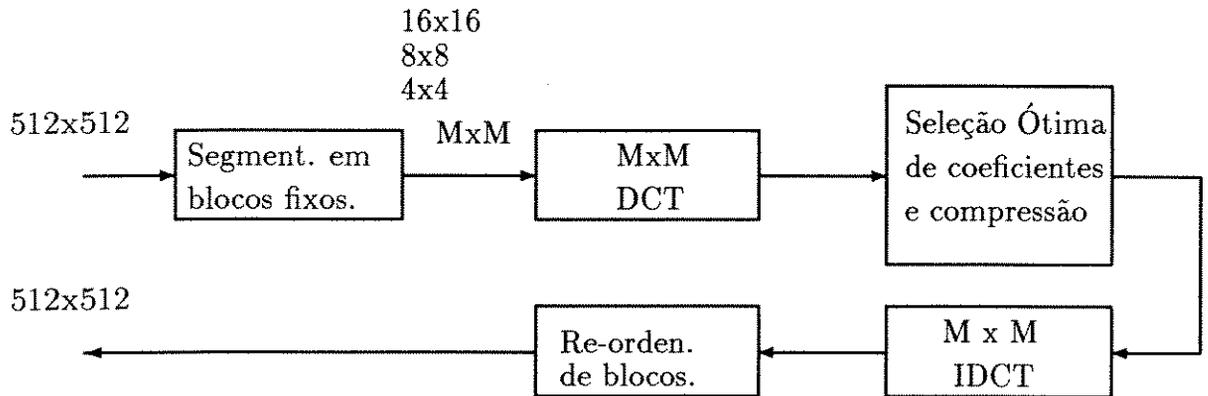


Figura 4.10: Avaliação e seleção ótima de coeficientes.

4.4.2 Características no domínio da transformada

A transformada DCT, é de tipo sub-ótima no sentido de não decorrelacionar totalmente os dados de entrada. Isto significa que os coeficientes não são independentes, ou seja, é possível encontrar uma estrutura específica que possa ser mais importante para a reconstrução da imagem.

Foram feitos quatro tipos de teste com a seleção aproximada de 25 % dos coeficientes por bloco e filtrando a zero o resto. Aproveitando a propriedade de compactação da energia que tem a DCT, já que o 25 % dos coeficientes representa quase o 95 % da energia total dos E.I. de entrada por bloco, [7]. As estruturas geométricas de seleção de coeficientes são construídas em cada bloco, baseando-se nos seguintes programas:

- Programa “cmprcir.for”. Produz uma seleção isotrópica de coeficientes a partir de onde existe a maior concentração da energia.
- Programa “cmprlip.for”. Produz uma seleção elíptica não isotrópica de coeficientes, a partir da maior concentração da energia pelas bordas do bloco.

- Programa “cmprbd.for”. Produz uma seleção pelas bordas do bloco de coeficientes a partir da maior concentração da energia.
- Programa “cmprtrg.for”. Produz uma seleção triangular de coeficientes a partir da maior concentração da energia por bloco.

Nas figuras (4.11) e (4.12) mostram-se as diversas estruturas dos coeficientes selecionados por bloco testadas, para um processamento de bloco de 16x16.

O comportamento dos coeficientes no domínio da frequência vai depender do tipo da geometria da imagem ou estrutura dos dados de entrada, e também do tamanho do Bloco de analisado.

Devido às funções bases da Transformada Cosseno terem estruturas geométricas nas direções horizontais e verticais (figura 4.13), onde observa-se a simetria dos valores das funções bases para a transformada DCT 4x4, consegue-se melhor compactação de energia neste tipo de estruturas no domínio da transformada do que as estruturas de dados de entrada com geometrias diagonais, como pode-se observar na figura (4.14) a.-, b.- e c.- Teoricamente, tem-se já analisado (seção 3.15) o modelo Laplaciano da distribuição da energia dos coeficientes transformados da DCT, segundo os dados de entrada de uma imagem que segue um comportamento Markoviano representante da maioria das imagens reais. Pode-se portanto usar este modelo como referência para determinar a estrutura geométrica mais adequada para selecionar os coeficientes.

O comportamento dos coeficientes vai depender do tamanho do bloco; quanto maior o tamanho do bloco, mais confuso é o comportamento devido à falta de estacionariedade. Para o caso de figuras geométricas simples e blocos pequenos é possível projetar uma codificação controlada pelo comportamento de compactação da energia dos coeficientes, por exemplo: para uma figura retangular que sofre uma pequena inclinação, (figura 4.14 c.-), a energia que ficava compactada nos coeficientes das bordas do bloco transformado começa a redistribuir-se na diagonal principal que nasce nos coeficientes mais baixos, que em princípio fornecem uma “pista” para projetar um sistema adaptativo em função da inclinação da figura retangular. Contudo, para imagens naturais comprovou-se que esta concentração da energia não segue um padrão tão bem comportado [4].

Esta primeira avaliação determina a necessidade de um estudo intenso do com-

BLOCO No.: 12 9

2575.25	139.59	-155.78	139.00	-34.71	-25.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-192.67	94.36	-9.84	-73.07	-7.33	36.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
64.13	117.41	-58.16	-155.56	12.85	-14.42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-83.46	-71.48	29.81	61.05	2.51	18.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
58.73	-0.92	20.23	20.08	-0.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
21.71	31.47	21.02	10.45	-8.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

1. FILTRAGEM ISOTRÓPICO DE 33 COEFICIENTES EFETIVOS NO BLOCO

BLOCO No. 19 2

1794.00	-356.48	14.02	-80.32	123.99	-38.78	5.59	-6.28	5.88	-6.74	5.88	-6.74	-3.36	5.55	-4.91	1.17
196.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
26.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-1.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-4.84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.62	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-4.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-1.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

2. FILTRAGEM FORA DAS BORDAS DO BLOCO DE 31 COEFICIENTES

Figura 4.11: Estruturas de Coeficientes Transformados em blocos de 16x16 (a).

BLOCO No. 7 29

1488.00	25.79	-22.01	-8.10	-1.28	-8.89	-3.65	-4.45	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-138.65	-38.41	28.68	13.65	-1.41	11.23	7.75	3.72	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-24.66	48.93	19.87	5.81	5.86	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
61.54	31.04	-36.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-7.97	-28.66	13.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11.72	6.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-7.12	5.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3.02	2.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

3. FILTRAGEM NÃO ISOTRÓPICO NO BLOCO DE 35 COEFICIENTES

BLOCO No. 19 2

1794.00	-356.48	14.02	-80.32	123.99	-38.78	5.59	-6.28	5.87	0.00	0.00	0.00	0.00	0.00	0.00	0.00
196.88	108.21	12.11	77.32	95.23	-25.43	3.18	3.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
26.45	31.06	8.23	-12.67	60.34	12.92	5.71	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-1.76	10.74	17.32	25.62	17.63	6.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-4.84	-21.13	8.04	-3.21	4.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.56	10.65	7.91	1.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.45	-0.83	4.89	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.08	3.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

4. FILTRAGEM TRIANGULAR DE 45 COEFICIENTES

Figura 4.12: Estruturas de Coeficientes Transformados em blocos de 16x16 (b).

0.500000	0.500000	0.500000	0.500000
0.653281	0.270598	-0.270598	-0.653282
0.500000	-0.500000	-0.500000	0.500000
0.270598	-0.653281	0.653282	-0.270598

Figura 4.13: Matriz de Funções Bases da DCT 4x4

portamento estatístico dos blocos para poder classificá-los segundo a Média e o Desvio Padrão. Pode-se perceber (figura 4.14) que as mudanças abruptas provocadas pelas bordas nas figuras geométricas simples são rapidamente detectadas pelo Desvio Padrão do bloco. Portanto, para ter um controle da orientação das imagens, é preciso uma classificação dos blocos em função dos parâmetros de controle estatístico de primeira e segunda ordem anterior ao processo de transformação. Essa classificação pode controlar o tamanho da transformada. Não obstante, também é preciso encontrar um compromisso entre a informação paralela adicional desta classificação para o receptor e a eficiência da codificação do sistema.

4.4.3 Resultados

Das quatro estruturas testadas (figuras 4.11 e 4.12), observa-se na tabela (4.3) que o pior comportamento na recuperação da imagem foi para a filtragem fora das bordas (figura 4.11 2.-). Esse tipo de estrutura é mais válido para imagens de figuras retangulares simples do que para o caso de imagens naturais ou reais.

A tabela (4.3) mostra o comportamento da Relação Sinal/Ruído do sinal recuperado para a imagem *smpte15y* dos quatro casos testados.

Nos três casos restantes, a diferença em qualidade (SNR) não é tão grande. No caso da seleção por estrutura triangular, a qualidade é ligeiramente maior porque tem considerado perto de dez coeficientes a mais. Em todo caso, na implementação de um sistema adaptativo, o caso triangular é o que oferece o controle mais simples. Mas, nos três casos tem-se um compromisso na distribuição dos bits (secção 3.17) e a maior eficiência da distribuição resulta da opção de medir a atividade por bloco. Finalmente, pode-se dizer que

MÉDIA : 255 DES. ST.: 0.0		
255 255 255 255		1020 0000 0000 0000
255 255 255 255	→	0000 0000 0000 0000
255 255 255 255		0000 0000 0000 0000
255 255 255 255		0000 0000 0000 0000
a.- DOMÍNIO DO TEMPO	→	DOMÍNIO DA FREQUÊNCIA
CARACTERÍSTICA:		CARACTERÍSTICA:
Bloco Estacionário.		Grande compactação da energia apenas num coeficiente.
MÉDIA: 204.50 DES. ST.:87.4686		
255 255 255 255		0818.00 0263.92 -202.00 0109.32
255 255 255 255	→	0000.00 0000.00 0000.00 0000.00
255 255 255 255		0000.00 0000.00 0000.00 0000.00
053 053 053 053		0000.00 0000.00 0000.00 0000.00
b.- DOMÍNIO DO TEMPO	→	DOMÍNIO DA FREQUÊNCIA
CARACTERÍSTICA:		CARACTERÍSTICA :
Borda horizontal abrupta		Compactação apenas na borda superior horizontal.
MÉDIA: 130.00 DES. ST.:96.8246		
255 255 255 055		0520.00 0223.05 0000.00 0015.85
255 255 055 055	→	0223.04 0100.00 -092.39 0000.00
255 055 055 055		0000.00 -092.39 0100.00 0038.27
055 055 055 055		0015.85 0000.00 0.038.27 0100.00
c.- DOMÍNIO DO TEMPO	→	DOMÍNIO DA FREQUÊNCIA
CARACTERÍSTICA:		CARACTERÍSTICA:
Borda diagonal abrupta		Coefficientes significativos perto da diagonal principal do bloco
MÉDIA: 154.00 DES. ST : 101.0000		
255 255 053 053		0616.00 0.00 0.00 0.00
255 255 053 053	→	0373.25 0.00 0.00 0.00
255 255 053 053		0000.00 0.00 0.00 0.00
255 255 053 053		-154.60 0.00 0.00 0.00
d.- DOMÍNIO DO TEMPO	→	DOMÍNIO DA FREQUÊNCIA
CARACTERÍSTICA:		CARACTERÍSTICA:
Borda vertical abrupta		Compactação apenas na borda lateral esquerda

Figura 4.14: Estruturas de Geometrias Simples.

ESTRUTURA DE FILTRAGEM DE BLOCOS	SNR
ISOTRÓPICA	26 db
FORA DAS BORDAS	15 db
NÃO ISOTRÓPICA	25.5 db
SELEC. TRIANGULAR	28 db

Tabela 4.3: Relação Sinal/Ruído da Imagem smpte15y recuperada utilizando apenas 25 % dos coeficientes transformados

a complexidade da forma da imagem tem menos prioridade que a informação estatística que vai controlar a otimização do sistema adaptativo.

O comportamento da segmentação por bloco fixo no algoritmo de compressão convencional, quanto à qualidade na recuperação da informação, fica refletido na seguinte tabela (4.4), onde tem-se selecionado para as segmentações de 4x4, 8x8 e 16x16 perto de 40 % dos coeficientes transformados em estrutura triangular e o resto zerado. Utiliza-se como imagem de teste a smpte15y e smpte01y.

	4x4	8x8	16x16
No. DE COEF./BL.	6	28	91
SNR SMPTE15Y	34.1 db.	32.3 db.	31.2 db.
SNR SMPTE01Y	33.7 db.	31.3 db.	29.8 db

Tabela 4.4: Relação Sinal/Ruído para as segmentações de bloco fixo 4x4, 8x8 e 16x16 com quantização uniforme e em torno de 40% de coeficientes selecionados.

Pode-se observar pela tabela, que o comportamento quanto à qualidade é superior a uma segmentação menor devido ao bloco ficar mais estacionário e isso permite uma maior decorrelação entre os coeficientes transformados e uma compactação que permite absorver os coeficientes de alta frequência e produzir uma filtragem e compressão mais

eficiente sem causar muita degradação na imagem.

No algoritmo de compressão convencional de blocos fixos da DCT, o limiar e o passo do quantizador também são fixos para todos os blocos. Ainda que, a suposição de estacionariedade dos dados seja estável para todos os blocos, a característica ou natureza de cada bloco pode ser diferente. Por exemplo, o Sistema Visual Humano (S.V.H.) é mais tolerável à degradação nas áreas de muito detalhe, e mais sensível nas áreas simples e suaves. Além disso, a codificação por transformadas explora a redundância subjetiva nas imagens, a diferencia do sistema CMPD onde uma amostra submetida pela quantização afeta relativamente a uns poucos EI na reconstrução da imagem, a compressão e quantização de um coeficiente transformado afeta a todos os EI no bloco recuperado, este espalhamento do erro de quantização faz que o projeto do quantizador para os coeficientes da Transformada seja diferente do sistema CMPD. Este efeito apresenta-se de modo similar ao prévio do S.V.H., por exemplo, em muitos casos no domínio espacial parte de um bloco pode ser um área suave de baixo detalhe e parte pode ser um área cheia de variação transitória estável, que denominaremos daqui para frente como "área de textura de ruído estável". Nesta situação um erro de quantização num coeficiente eventualmente causará uma forma de distorção que chega a ser visível na área suave de baixo detalhe e devido ao mascaramento espacial esta distorção não será visível na área de textura de ruído estável [8]. Ainda que um pequeno limiar e passo do quantizador garantam uma alta qualidade de imagem reconstruída, a taxa de bits será muito alta para aquelas áreas de muito detalhe, onde a degradação não é perceptível pelo S.V.H. e pelo mascaramento espacial na reconstrução dos coeficientes transformados. Em consequência, uma taxa de bits menor produziria a mesma qualidade de imagem. Como resultado, é desejável usar a possibilidade do limiar e tamanho do passo do quantizador variável para discriminar as áreas de muitos detalhes das que são suaves tendo como objetivo de reduzir a taxa de bits sem causar degradação visual, sendo esse o objetivo de usar as tabelas de quantização psico-visuais (secção 4.5.5).

4.4.4 Determinação dos Intervalos Adequados dos Parâmetros Estatísticos de Controle para as Imagens de Teste

4.4.4.1 Introdução

Mediante os seguintes programas, é possível obter informação gráfica do comportamento da Média, do Desvio Padrão, da Variância e da Energia Total por bloco para as Imagens de Teste, eles são respectivamente: medhist.for, despahis.for, varhist.for e ethist.for.

Foram feitas intensas avaliações no comportamento da Média e do Desvio Padrão nos diferentes blocos para determinar os possíveis limites de controle adaptativo adequado para a segmentação de blocos e a seleção dos coeficientes transformados para os casos de processamento: 16x16, 8x8 e 4x4. Levou-se em consideração os valores dos parâmetros estatísticos que mais se repetem para cada tamanho de bloco, através da equação (4.1).

$$n = f(VE) \implies P(VE) = \frac{n}{NTD} \quad (4.1)$$

n = frequência de cada nível

VE = Valor estatístico da Média ou Desvio Padrão.

NTD= N_o total de dados

A análise dos dados é feita, por interpretações subjetivas das gráficas dos histogramas, devido à grande quantidade de dados que contém uma imagem.

4.4.4.2 Histogramas dos Parâmetros de Controle Estatístico para a Imagem smpte01y usando processamento em blocos de 16x16

Na figura (4.15) mostra-se o histograma do Desvio Padrão nos blocos de 16x16 que constituem a imagem smpte01y.

Pode-se observar que os valores de Desvio Padrão situados entre 2 e 16 apresentam frequência de ocorrência superior a 38 chegando a atingir valores acima de 70 enquanto que valores de Desvio Padrão acima 16 não ultrapassam a frequência de ocorrência de 20. Isto mostra uma certa estacionariedade nos valores baixos de Desvio Padrão. As mudanças abruptas da imagem apresentam-se nos valores de Desvio Padrão de 73 a 90, porém com menor frequência ocorrência.

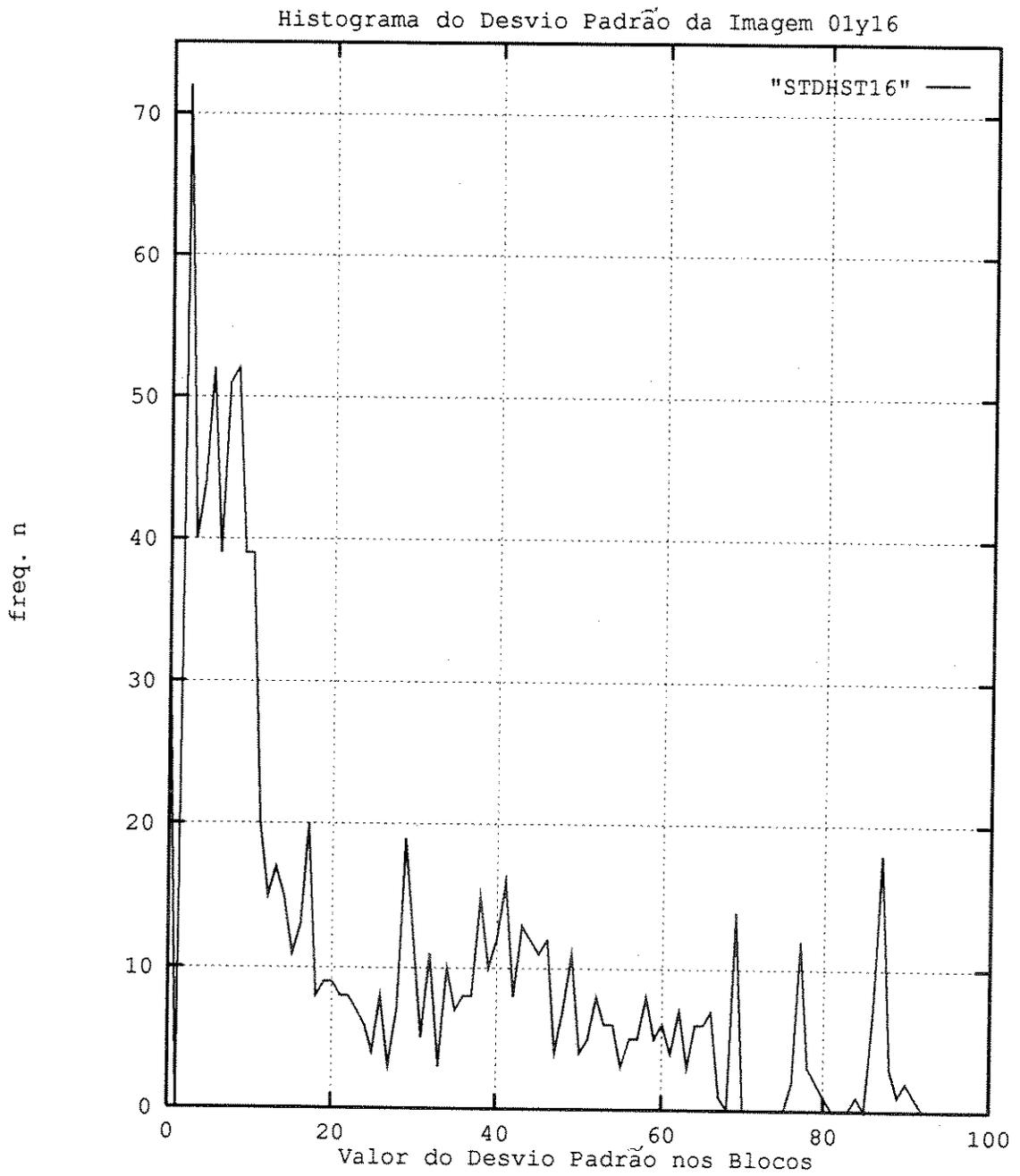


Figura 4.15: Comportamento do Desvio Padrão na segmentação de 16x16 na imagem smp-te01y.

Na figura (4.16) mostra-se o histograma da Média nos blocos de 16x16 para a imagem *smpte01y*.

Pode-se observar que os valores da Média situados entre 70 e 98 apresentam frequência de ocorrência superior a 5 chegando a atingir valores acima de 27 enquanto que valores da Média acima de 98 não ultrapassam a frequência de ocorrência de 18. É preciso notar que o valor 255 de EI é o mais frequente pois nas imagens da SMPTE é usado no ajuste do início e do final da imagem.

Apenas para este caso e como uma forma de acrescentar maiores informações sobre este tipo de segmentação, fez-se o histograma da Variância e da Energia Total por Bloco apresentados nas figuras (4.17) e (4.18).

Observa-se que o comportamento destas duas últimas figuras coincidem com as anteriores, demonstrando que a estacionariedade da imagem *smpte01y* encontra-se dentro do comportamento do modelo de Markov de primeira ordem [7].

4.4.4.3 Histogramas dos Parâmetros de Controle Estatístico para a Imagem *smpte15y* para processamento em blocos de 16x16

Na figura (4.19) mostra-se o histograma do Desvio Padrão nos blocos de 16x16 que constituem a imagem *smpte15y*.

Pode-se observar que os valores de Desvio Padrão situados entre 2 e 14 apresentam frequência de ocorrência superior a 30 e entre 3 e 8 sua frequência é superior a 40, chegando a atingir valores acima de 60 enquanto que valores de Desvio Padrão acima de 16 não ultrapassam a frequência de ocorrência de 25. Isto, também mostra uma certa estacionariedade nos valores baixos de Desvio Padrão. As mudanças abruptas da imagem apresentam-se nos valores de Desvio Padrão de 75 a 96, porém com menor frequência de ocorrência.

Na figura (4.20) mostra-se o histograma da Média nos blocos de 16x16 para a imagem *smpte15y*.

Pode-se observar que os valores da Média situados entre 55 e 75 apresentam frequência de ocorrência superior a 5 chegando a atingir o valor de 15, enquanto que valores

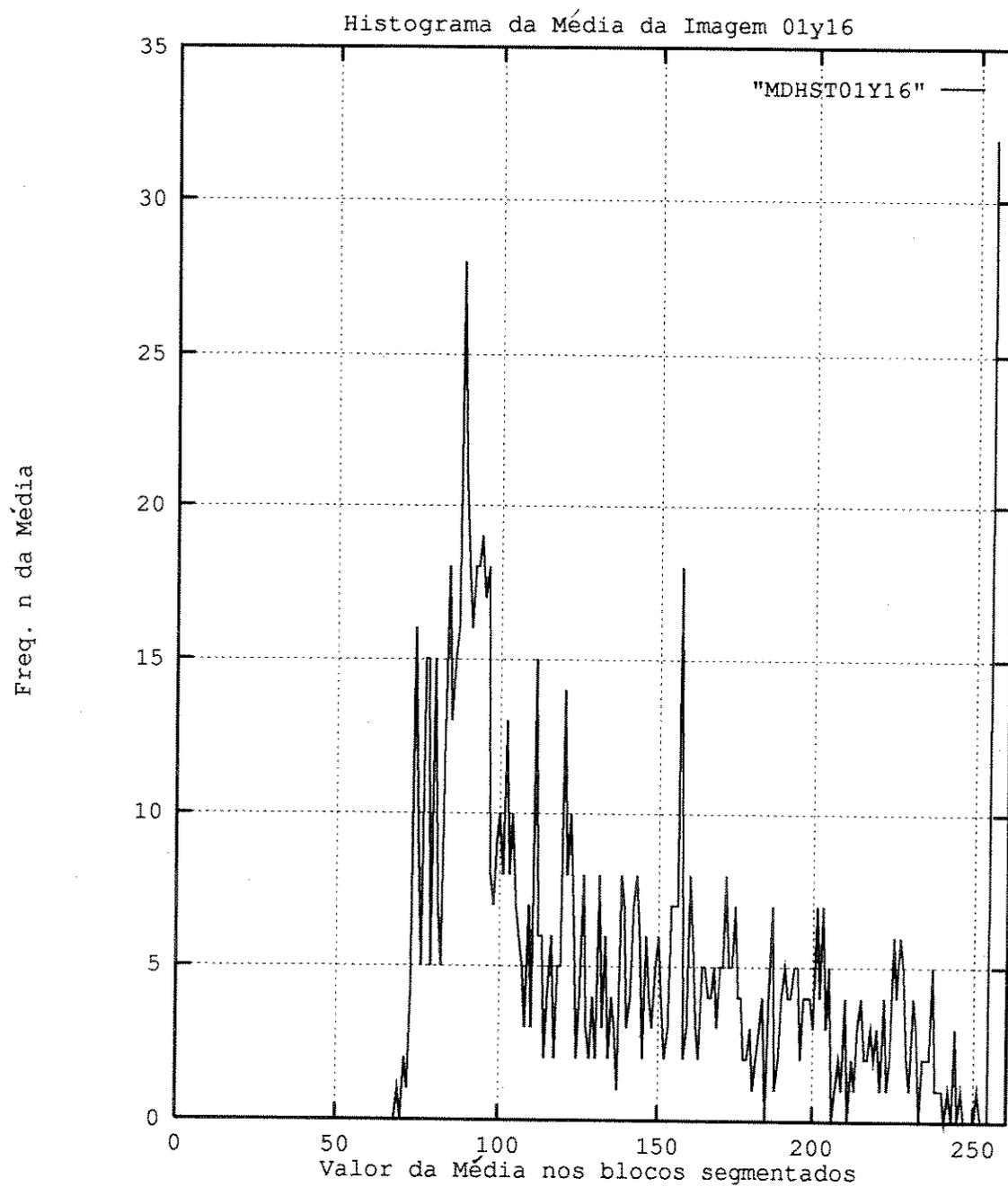


Figura 4.16: Comportamento da Média na segmentação de 16x16 na imagem smpte01y.

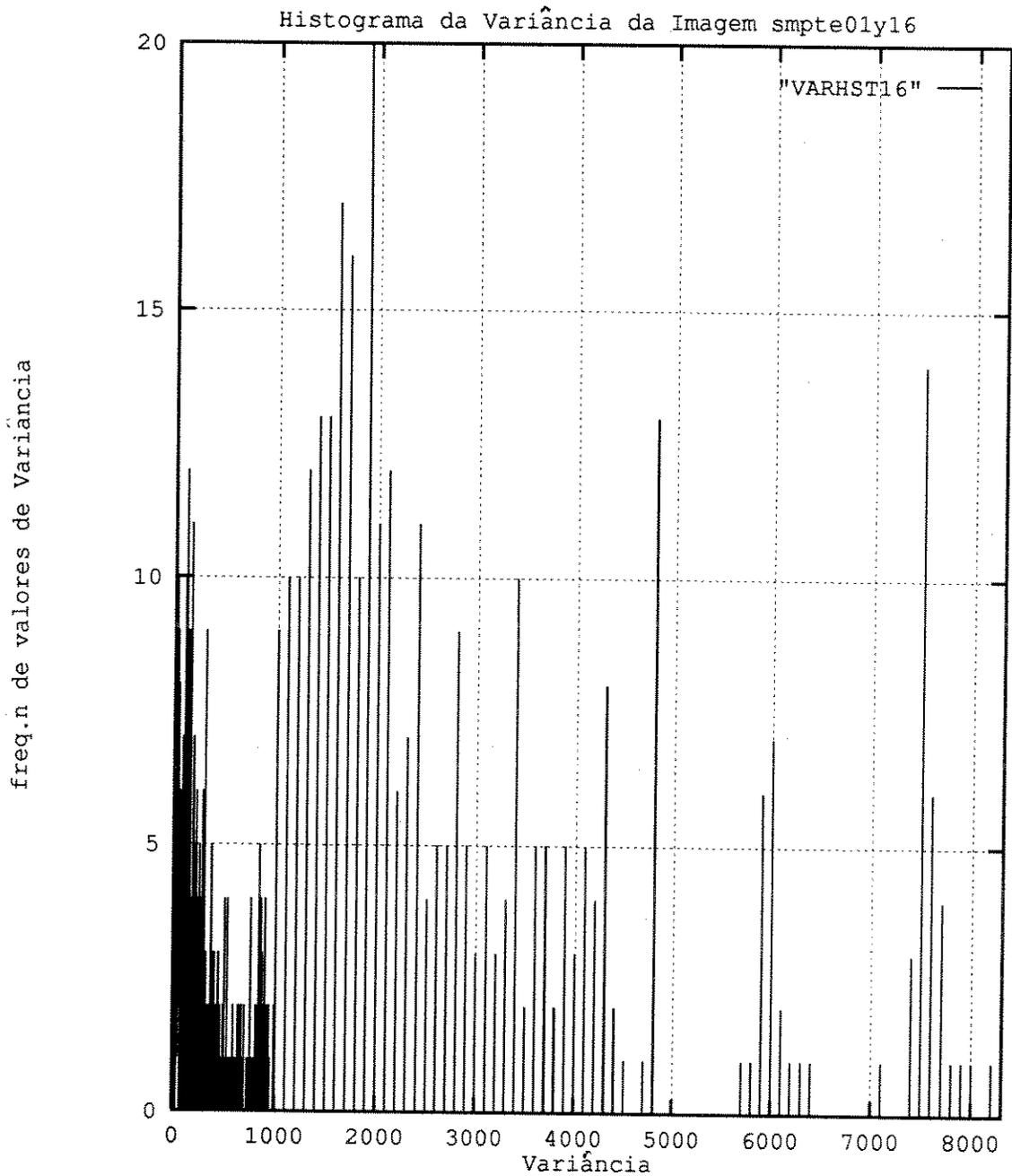


Figura 4.17: Comportamento da Variância na segmentação de 16x16 na imagem smpte01y.

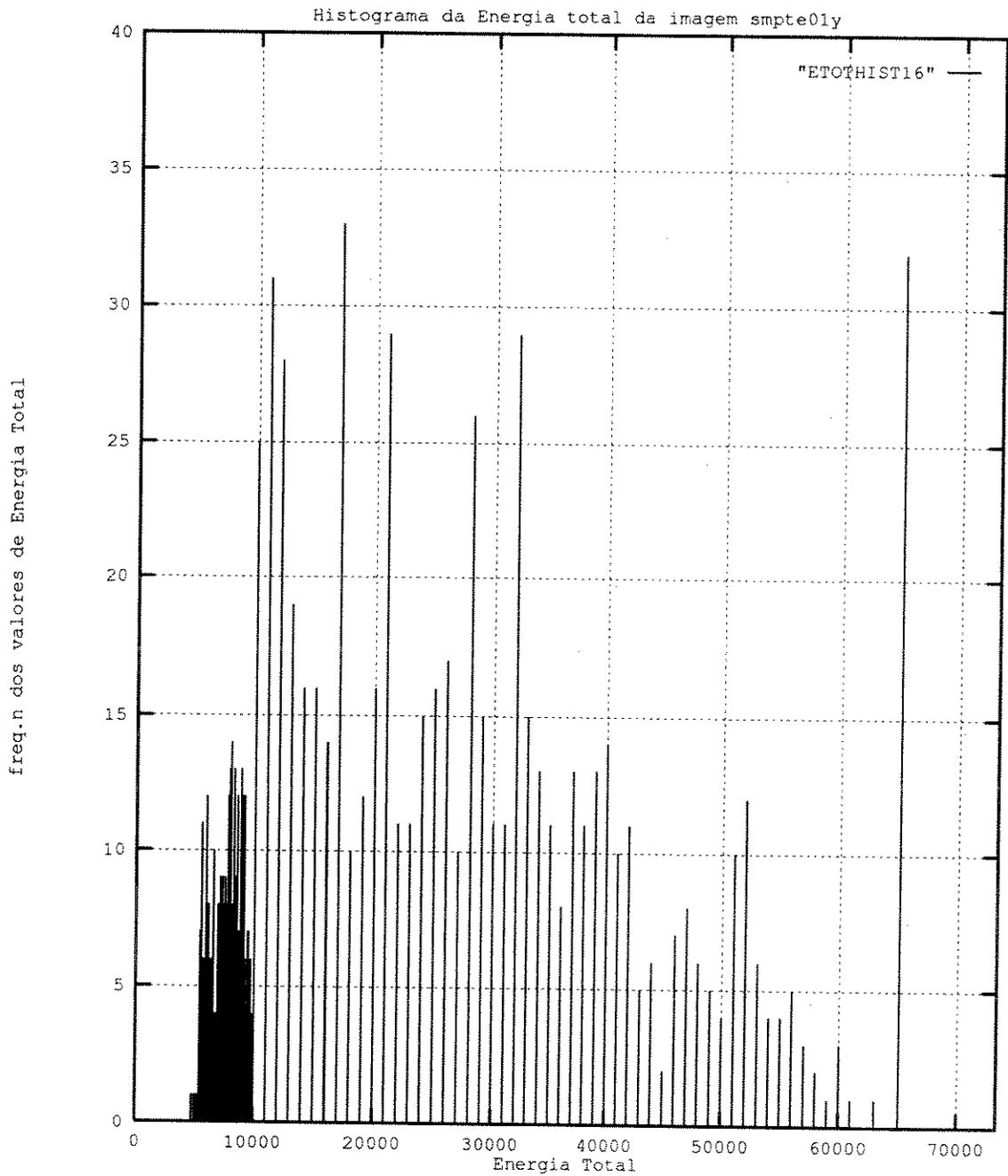


Figura 4.18: Comportamento da Energia Total na segmentação de 16x16 na imagem smpte01y.

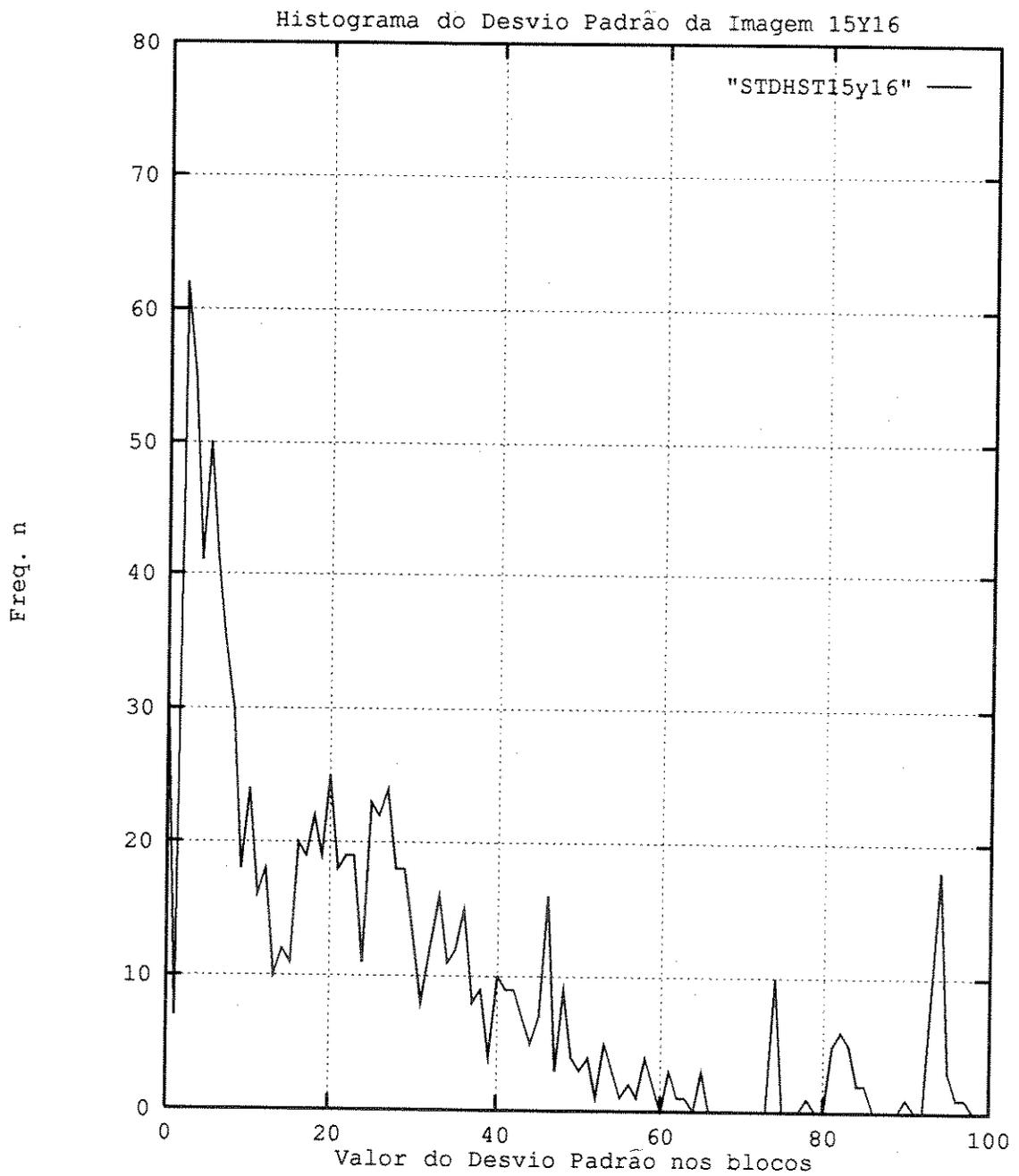


Figura 4.19: Comportamento do Desvio Padrão na segmentação de 16x16 na imagem smp-te15y.

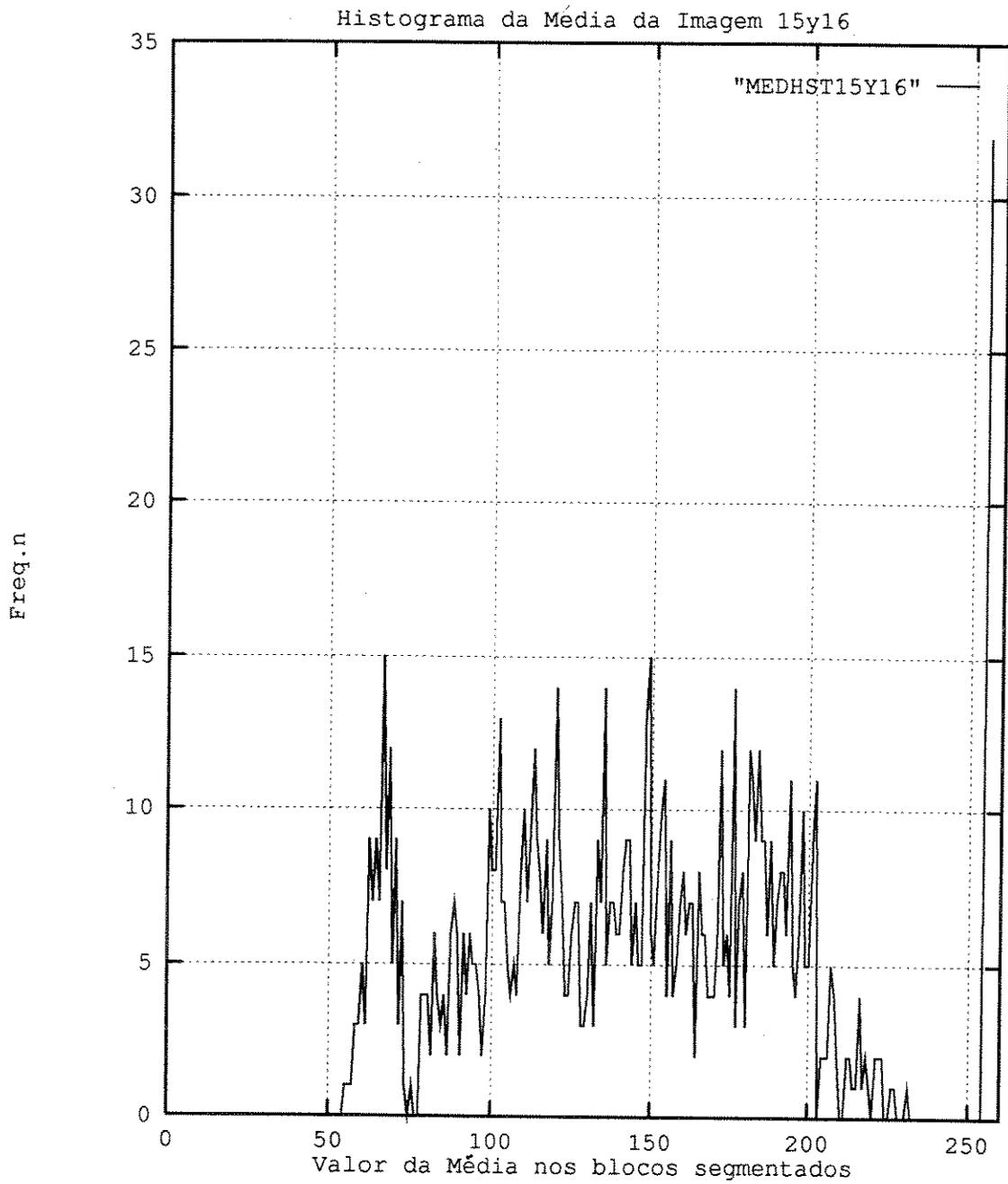


Figura 4.20: Comportamento da Média na segmentação de 16x16 na imagem smptel5y.

da Média acima de 78 não ultrapassam a frequência de ocorrência de 15. É preciso notar que o valor 255 de EI é o mais frequente pois nas imagens da SMPTE é usado no ajuste do início e do final da imagem.

4.4.4.4 Histogramas dos Parâmetros de Controle Estatístico para a Imagem Lena para processamento em blocos de 16x16

Na figura (4.21) mostra-se o histograma do Desvio Padrão nos blocos de 16x16 que constituem a imagem Lena.

Também, pode-se observar que os valores de Desvio Padrão situados entre 2 e 16 apresentam frequência de ocorrência superior a 32 e entre 3 e 8 sua frequência é superior a 50, chegando a atingir valores acima de 88 enquanto que valores de Desvio Padrão acima 16 não ultrapassam a frequência de ocorrência de 25. Isto mostra uma maior estacionariedade nos valores baixos de Desvio Padrão do que os casos prévios. As mudanças abruptas da imagem apresentam-se nos valores de Desvio Padrão de 73 a 76, porém com muito menor frequência ocorrência do que os casos prévios.

Na figura (4.22) mostra-se o histograma da Média nos blocos de 16x16 para a imagem Lena.

Pode-se observar que uma grande quantidade de valores da Média situados no intervalo entre 55 e 175 apresentam frequência de ocorrência superior a 5 chegando a atingir o valor de 20, enquanto que valores da Média acima de 178 não ultrapassam a frequência de ocorrência de 8.

4.4.4.5 Histogramas dos Parâmetros de Controle Estatístico para a Imagem Babuíno para processamento em blocos de 16x16

Na figura (4.23) mostra-se o histograma do Desvio Padrão nos blocos de 16x16 que constituem a imagem Babuíno.

Pode-se observar que os valores de Desvio Padrão situados entre 9 e 43 apresentam frequência de ocorrência em torno de 20 e a maioria superior a 22, chegando a atingir valores acima de 46 enquanto que, valores de Desvio Padrão acima 44 não ultrapassam a

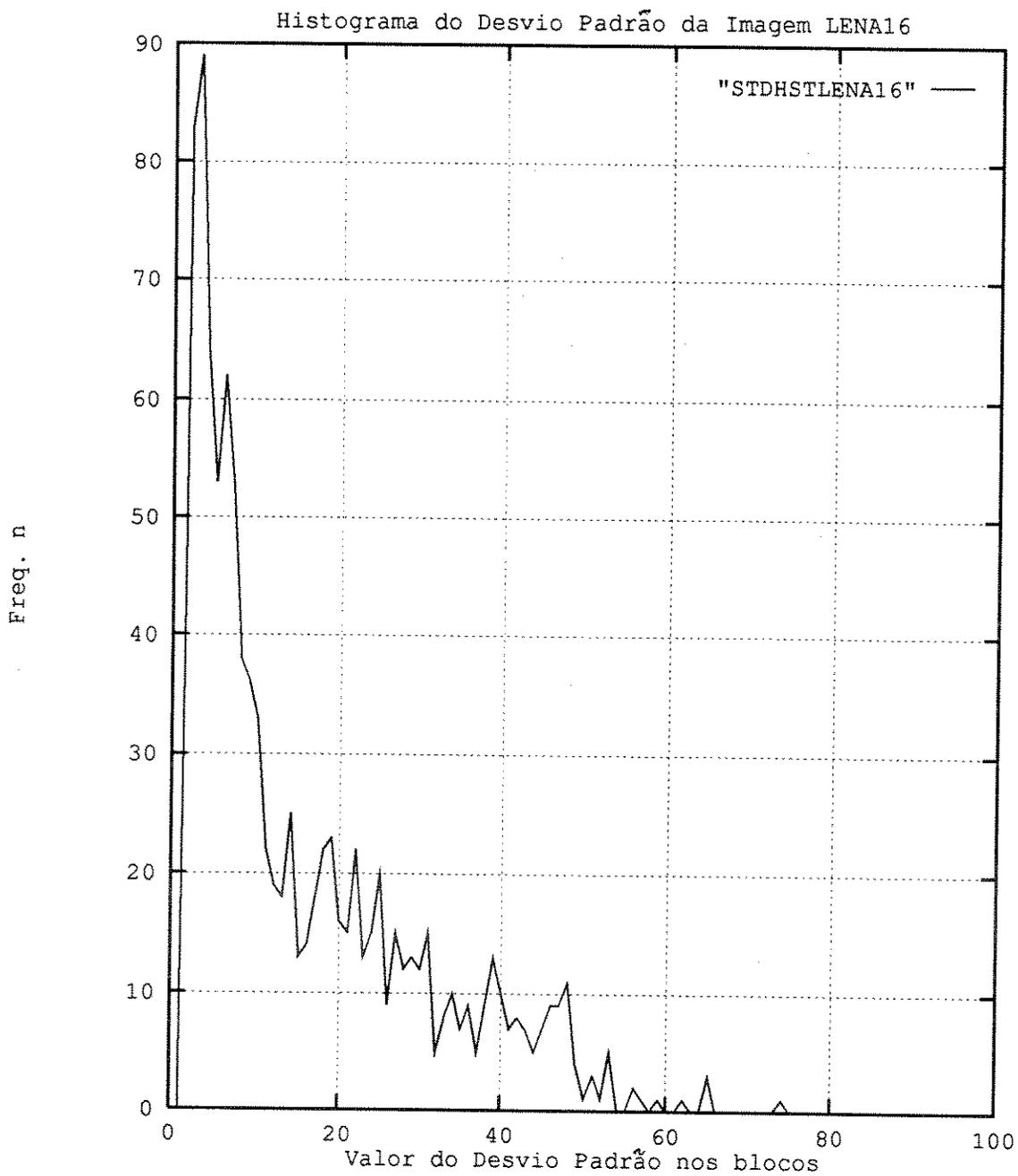


Figura 4.21: Comportamento do Desvio Padrão na segmentação de 16x16 na imagem Lena.

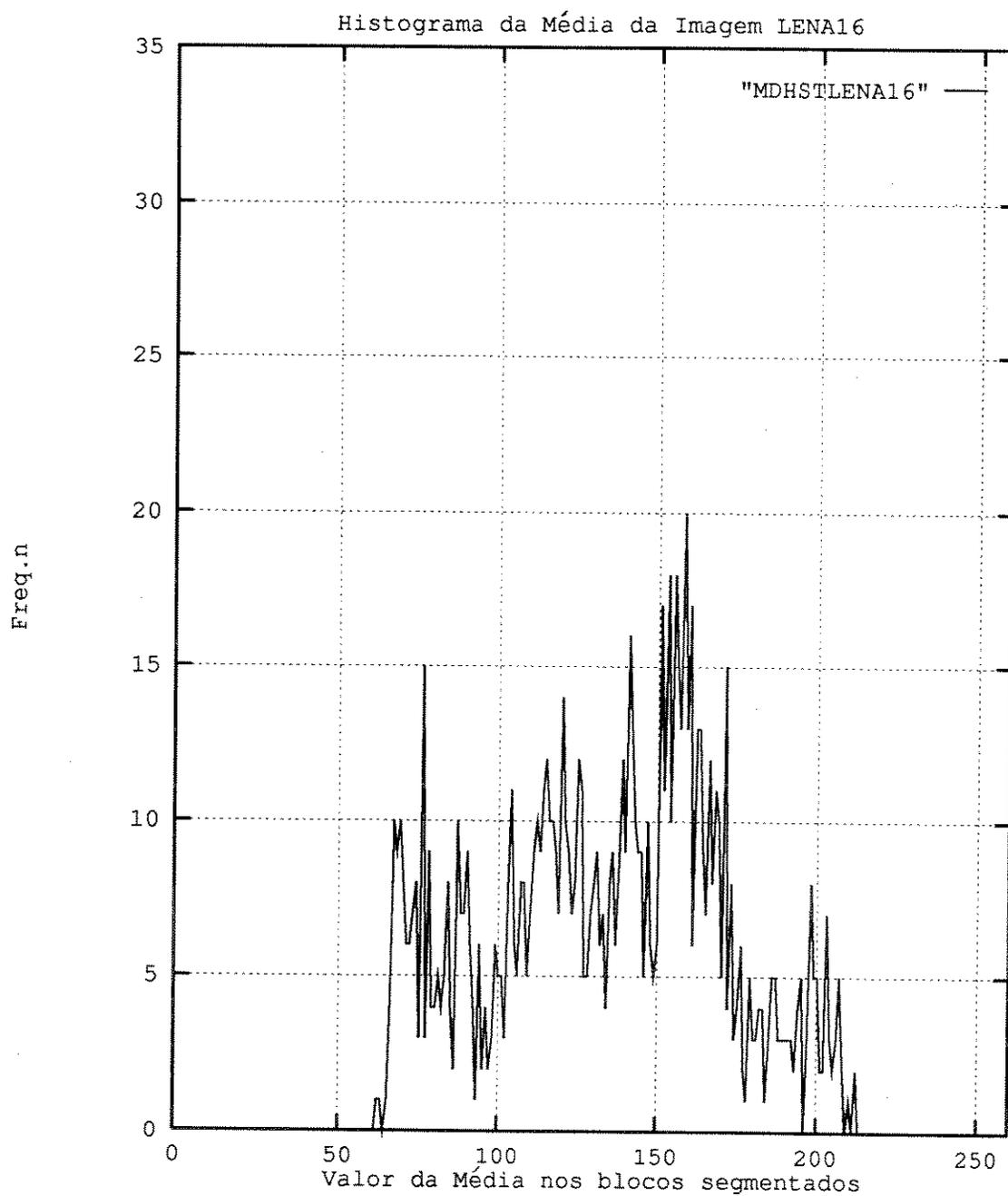


Figura 4.22: Comportamento da Média na segmentação de 16x16 na imagem Lena.

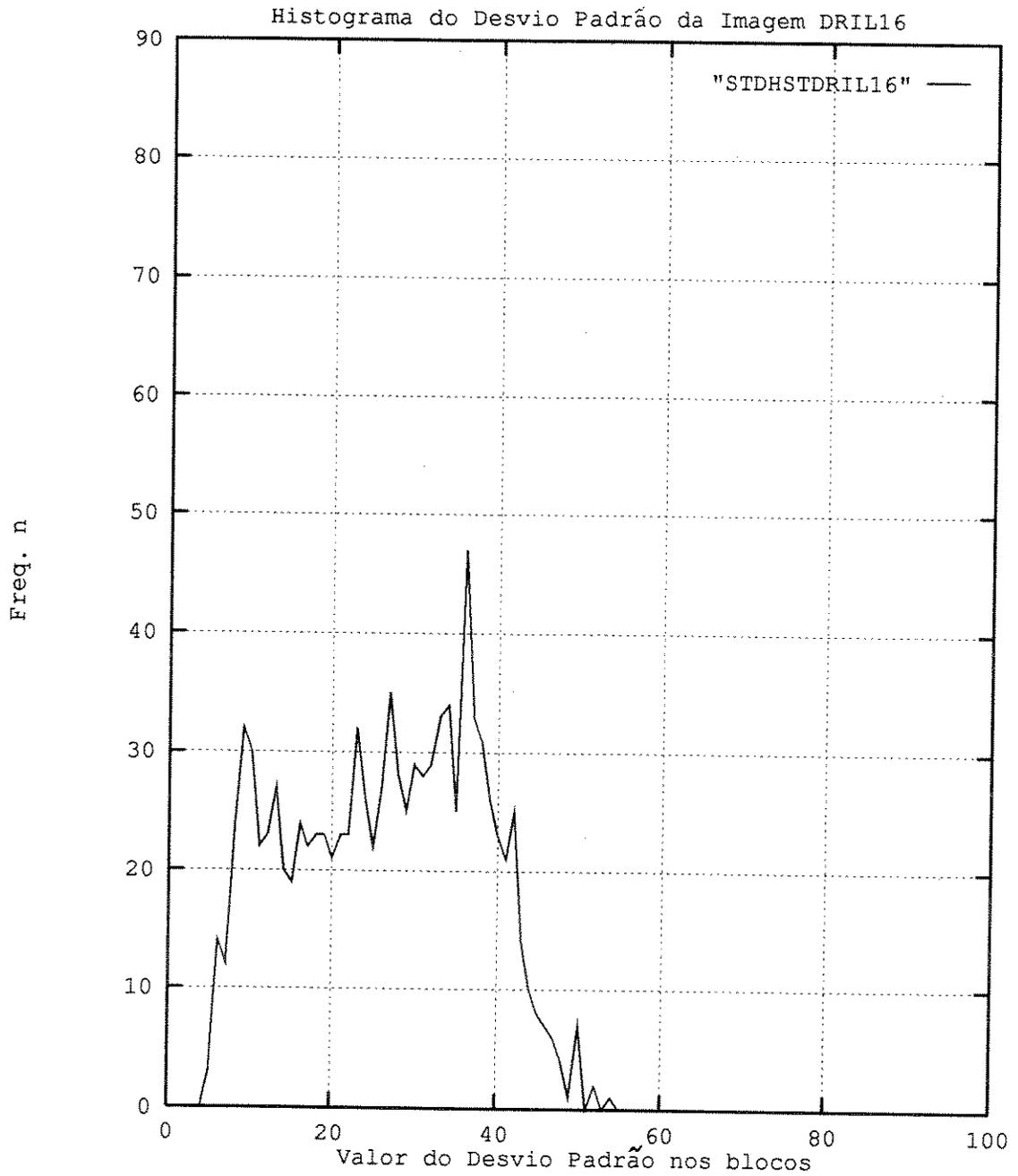


Figura 4.23: Comportamento do Desvio Padrão na segmentação de 16x16 na imagem Babuíno.

frequência de ocorrência de 8. Isto mostra uma zona de muito detalhe estável nos valores mencionados de Desvio Padrão. As mudanças abruptas da imagem são imperceptíveis.

Pode-se dizer, que esta imagem tem um comportamento de mudanças de níveis abruptos de pouca amplitude entre picos e com muito detalhe (devido à frequência manter uma simetria em relação aos diferentes valores sequenciais da Média de Luminância nos blocos, figura (4.24)), pouco perceptível ao Sistema Visual Humano (SVH), representando uma quase estacionariedade carente de mudanças marcadas com respeito ao fundo da imagem. Portanto para este tipo de imagem, a melhor segmentação pensando na eficiência da transformada e da codificação seria usar na maior parte dela blocos de 16x16.

Na figura (4.24) mostra-se o histograma da Média nos blocos de 16x16 para a imagem Babuíno. Pode-se observar que uma quantidade regular de valores da Média situados no intervalo entre 86 e 175 apresentam frequência de ocorrência superior a 10 chegando a atingir valores acima de 20, enquanto que valores da Média baixos e acima de 178 não ultrapassam a frequência de ocorrência de 8.

4.4.4.6 Conclusões sobre o Comportamento do Histograma do Desvio Padrão para o processamento de 16x16

Na figura (4.25) mostra-se a comparação dos histogramas do Desvio Padrão nas diferentes imagens de teste como resultado da segmentação de 16x16.

Nesta figura observa-se que os valores maiores de frequência de ocorrência estão relacionados aos menores valores do Desvio Padrão situados no intervalo entre 2 e 15 não chegando a atingir a frequência de repetição de valor igual a 90, onde os valores da imagem Lena ficam em torno deste valor quando comparados às outras três imagens de teste. O fato de ter os valores menores do Desvio Padrão nos blocos com um maior valor da frequência de ocorrência está indicando uma mudança de níveis mais suave e um comportamento quase estacionário nos mesmos. O Babuíno foi a exceção à esta característica.

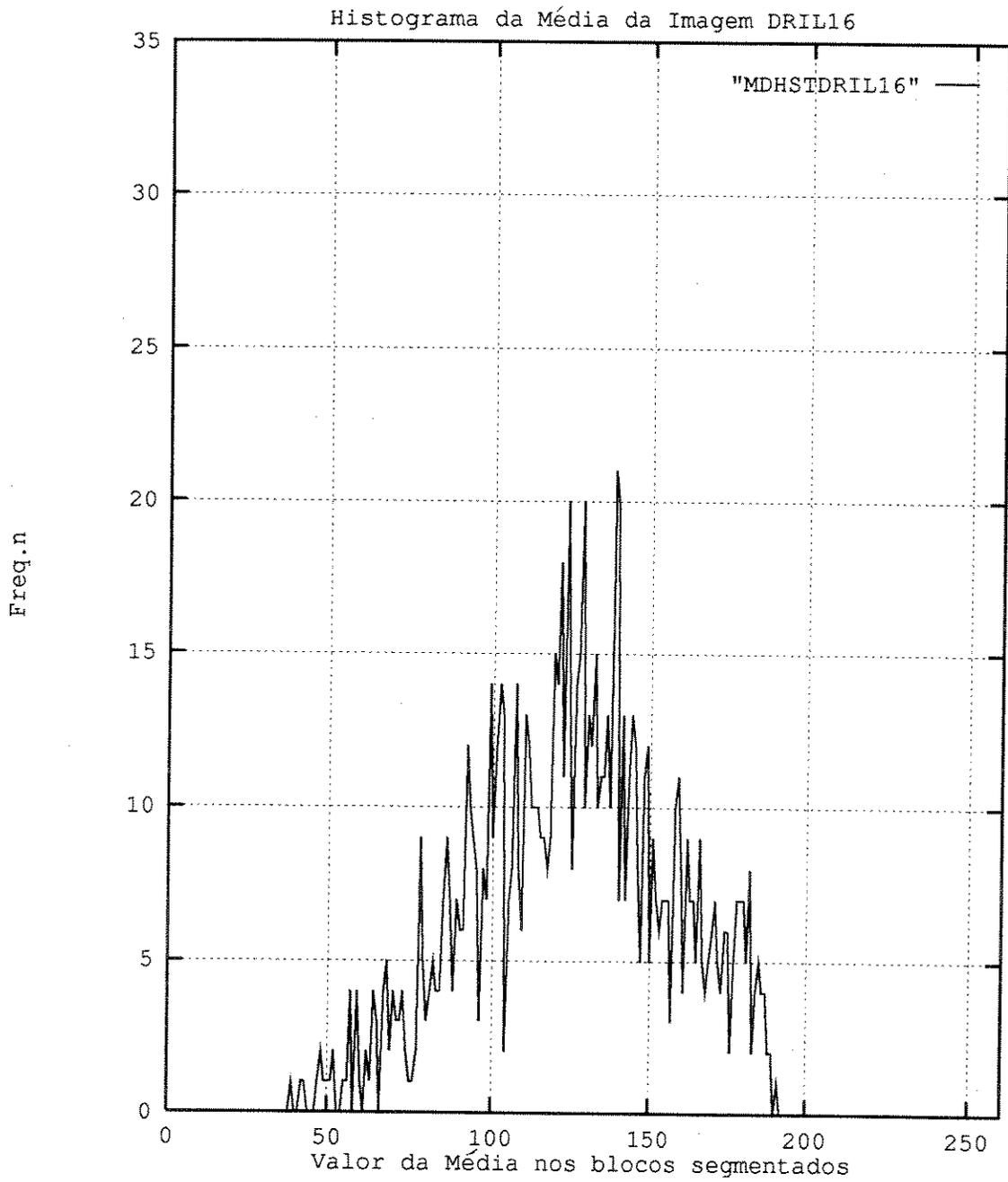


Figura 4.24: Comportamento da Média na segmentação de 16x16 na imagem Babuíno.

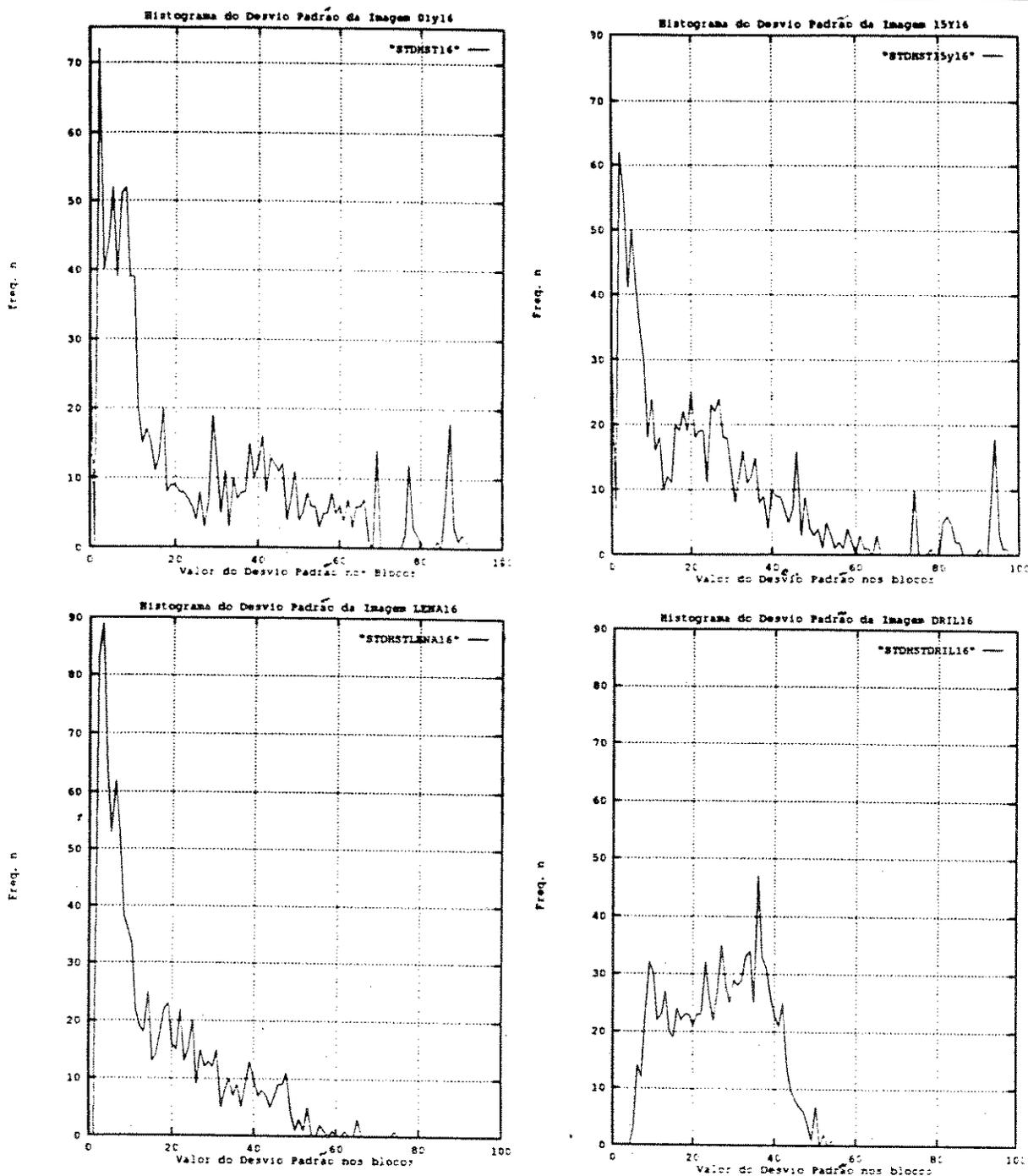


Figura 4.25: Comportamento do Desvio Padrão na segmentação de 16x16 nas diferentes imagens de teste.

4.4.4.7 Conclusões sobre o Comportamento do Histograma da Média para o processamento de 16x16

Na figura (4.26) mostra-se a comparação dos histogramas da Média nas diferentes imagens de teste como resultado da segmentação de 16x16.

Nesta figura pode-se ver que os valores da Média tem um comportamento quase simétrico no intervalo entre 50 e 225 para as imagens Lena e Babuíno, porém para as imagens da SMPTE tem marcado com maior frequência da Média os intervalos de baixos valores entre 53 e 98 para a imagem smpte01y e no caso da imagem smpte15y os maiores valores de frequência da Média sem atingir o valor de 30 estão no intervalo de valores baixos entre 60 e 70 depois, ainda que ambas apresentem uma grande diferença de luminosidade média, pois uma apresenta a claridade do dia numa praia (smpte01y) enquanto que a outra possui a luminosidade do interior de uma cozinha, ambas apresentam um comportamento quase simétrico nos intervalos dos valores da média entre 102 e 160 e entre 100 e 200, respectivamente.

4.4.4.8 Histogramas dos Parâmetros de Controle Estatístico para a Imagem smpte01y usando processamento em blocos de 8x8

Neste tipo de menor segmentação é de esperar uma maior estacionariedade por bloco, o que significa um maior poder de compactação da energia nos coeficientes da transformada e uma maior quantidade de blocos analisados com uma maior frequência dos parâmetros estatísticos (ver seção 3.16).

Na figura (4.27) mostra-se o histograma do Desvio Padrão nos blocos de 8x8 que constituem a imagem smpte01y.

Pode-se observar que os valores de Desvio Padrão situados entre 3 e 10 apresentam frequência de ocorrência superior a 100 chegando a atingir valores acima de 700 enquanto que valores de Desvio Padrão acima de 10 não ultrapassam a frequência de ocorrência de 70. Isto mostra uma certa estacionariedade nos valores baixos de Desvio Padrão. As mudanças abruptas da imagem apresentam-se nos valores de Desvio Padrão de 75 a 95, porém com menor frequência de ocorrência.

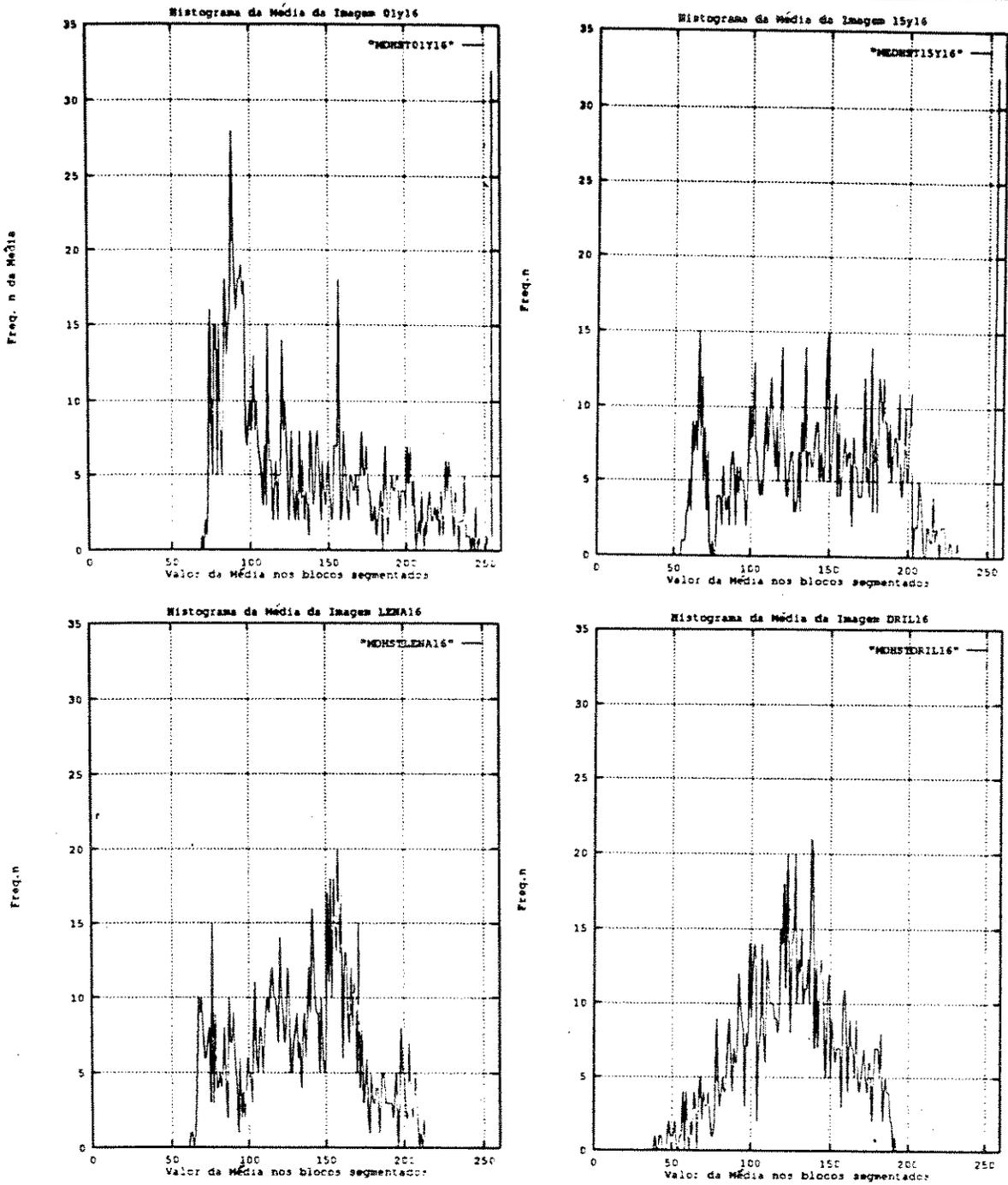


Figura 4.26: Comportamento da Média na segmentação de 16x16 nas diferentes imagens de teste.

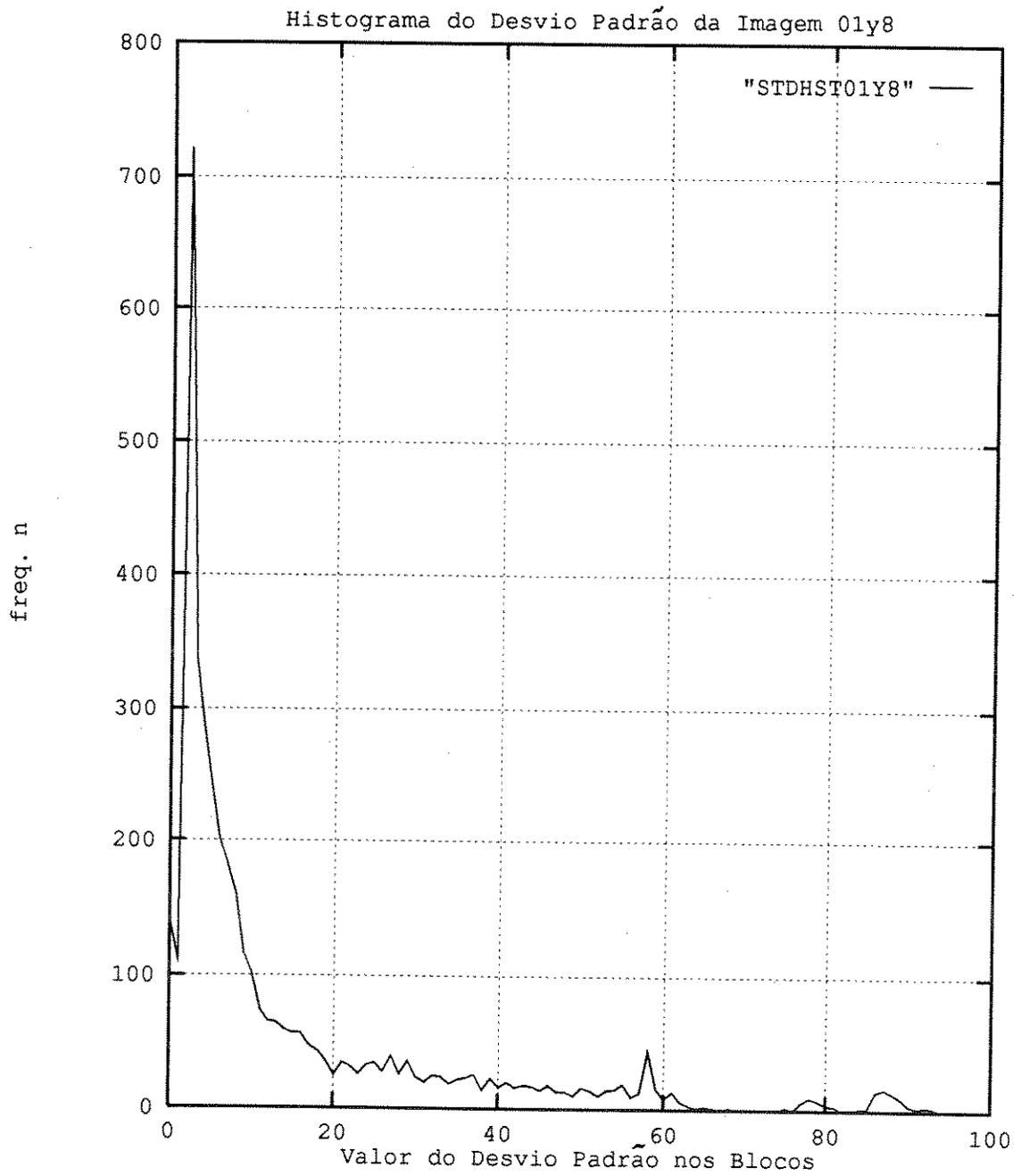


Figura 4.27: Comportamento do Desvio Padrão na segmentação de 8x8 na imagem smp-te01y.

Na figura (4.28) mostra-se o histograma da Média nos blocos de 8x8 para a imagem *smpte01y*.

Pode-se observar que os valores da Média situados entre 70 e 100 apresentam frequência de ocorrência superior a 40 chegando a atingir valores acima de 80, enquanto que valores da Média acima de 100 não ultrapassam a frequência de ocorrência de 43. É preciso notar que o valor 255 de EI é o mais frequente pois nas imagens da SMPTE é usado no ajuste do início e do final da imagem.

4.4.4.9 Histogramas dos Parâmetros de Controle Estatístico para a Imagem *smpte15y* usando processamento em blocos de 8x8

Na figura (4.29) mostra-se o histograma do Desvio Padrão nos blocos de 8x8 que constituem a imagem *smpte15y*.

Pode-se observar que os valores de Desvio Padrão situados entre 3 e 10 apresentam frequência de ocorrência superior a 100 chegando a atingir valores acima de 750 enquanto que valores de Desvio Padrão acima 16 não ultrapassam a frequência de ocorrência de 80. Isto mostra uma certa estacionariedade nos valores baixos de Desvio Padrão. As mudanças abruptas da imagem apresentam-se nos valores de Desvio Padrão de 82 a 98, porém com menor frequência de ocorrência.

Na figura (4.30) mostra-se o histograma da Média nos blocos de 8x8 para a imagem *smpte15y*.

Pode-se observar que os valores da Média situados entre 60 e 76 apresentam frequência de repetição superior a 20 e entre 65 e 75 apresentam frequência de ocorrência superior a 40 chegando a atingir valores acima de 115 enquanto que valores da Média acima de 75 não ultrapassam a frequência de ocorrência de 40. É preciso notar que o valor 255 de EI é o mais frequente pois nas imagens da SMPTE é usado no ajuste do início e do final da imagem.

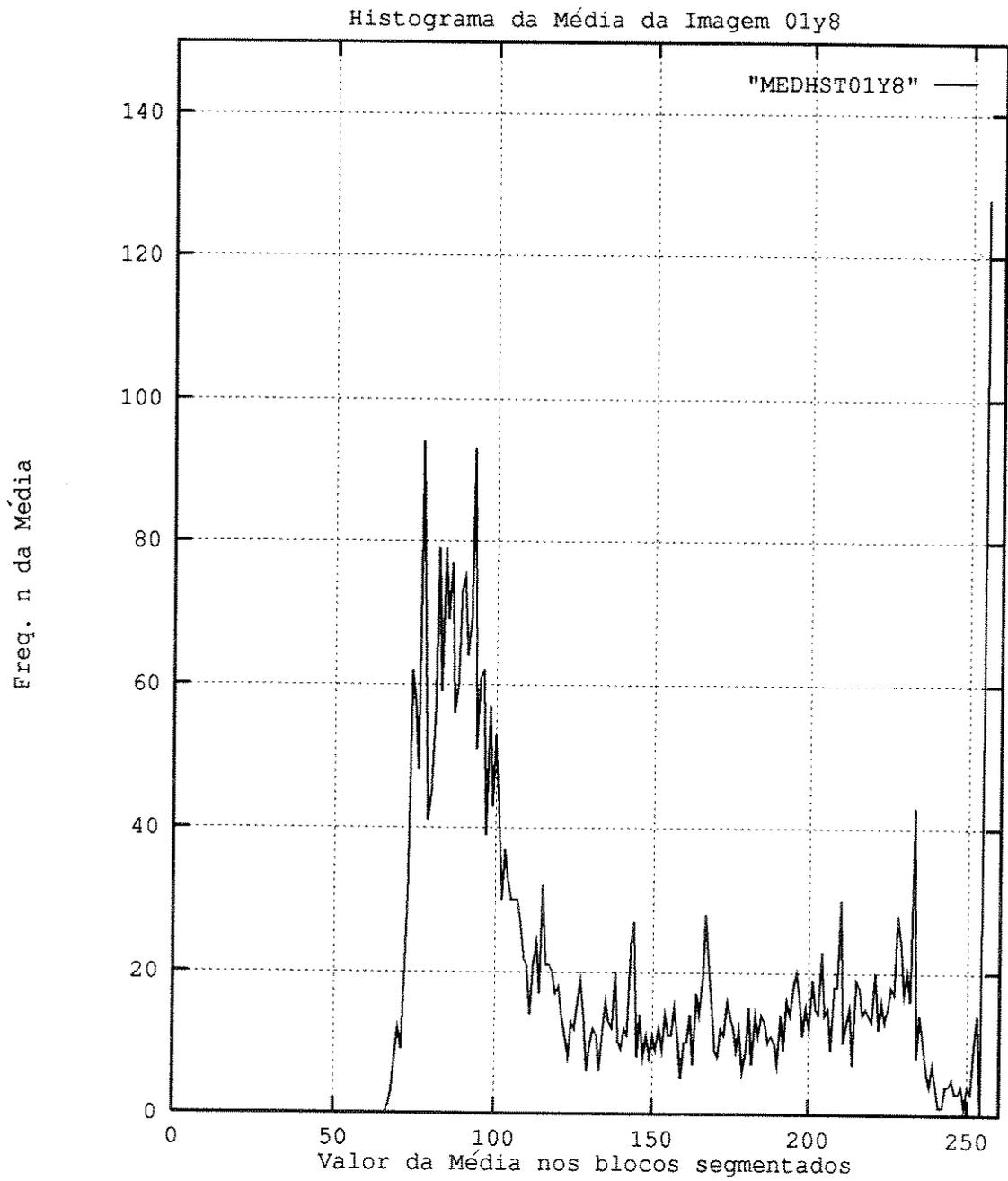


Figura 4.28: Comportamento da Média na segmentação de 8x8 na imagem smpte01y.

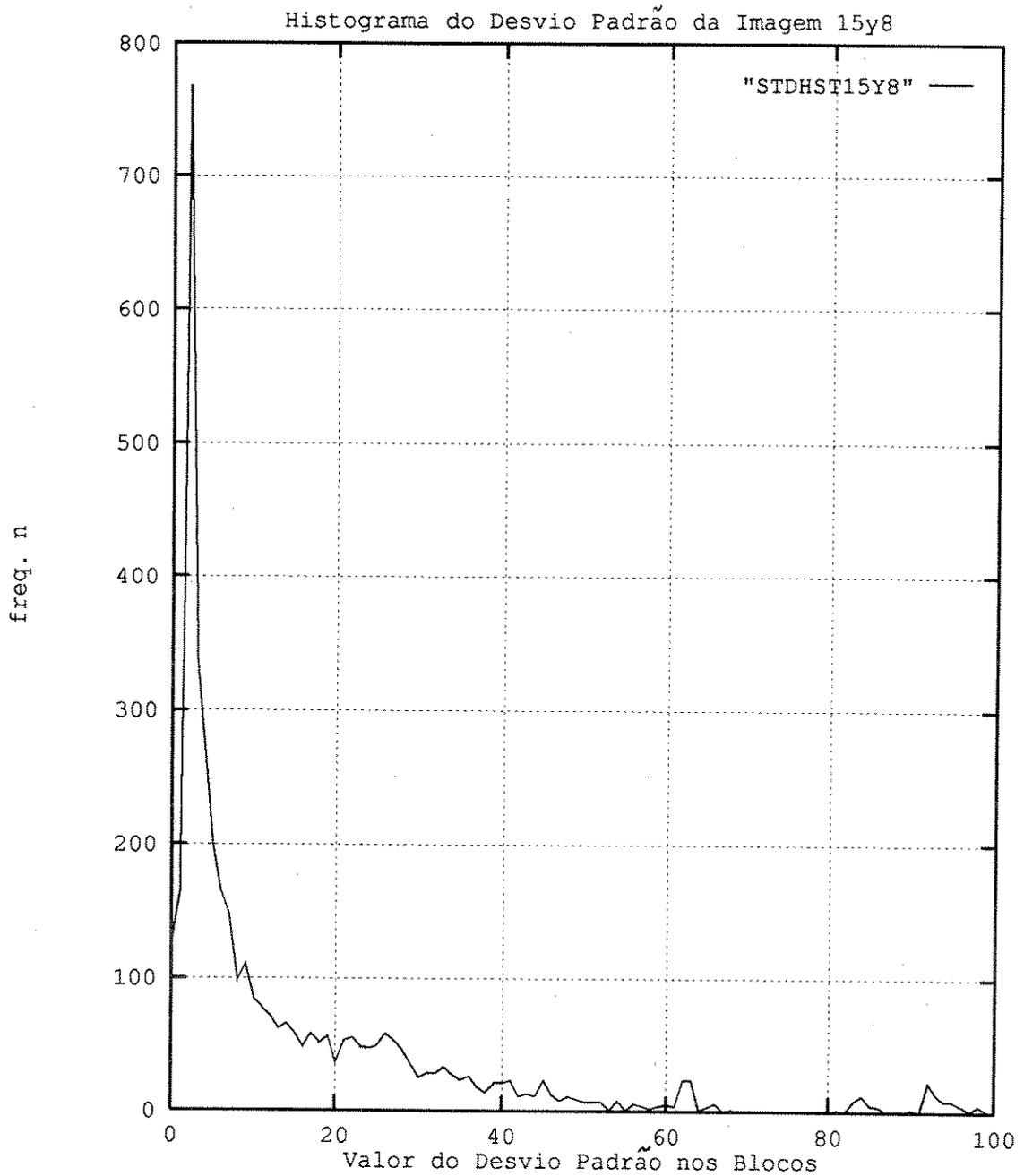


Figura 4.29: Comportamento do Desvio Padrão na segmentação de 8x8 na imagem smp-te15y.

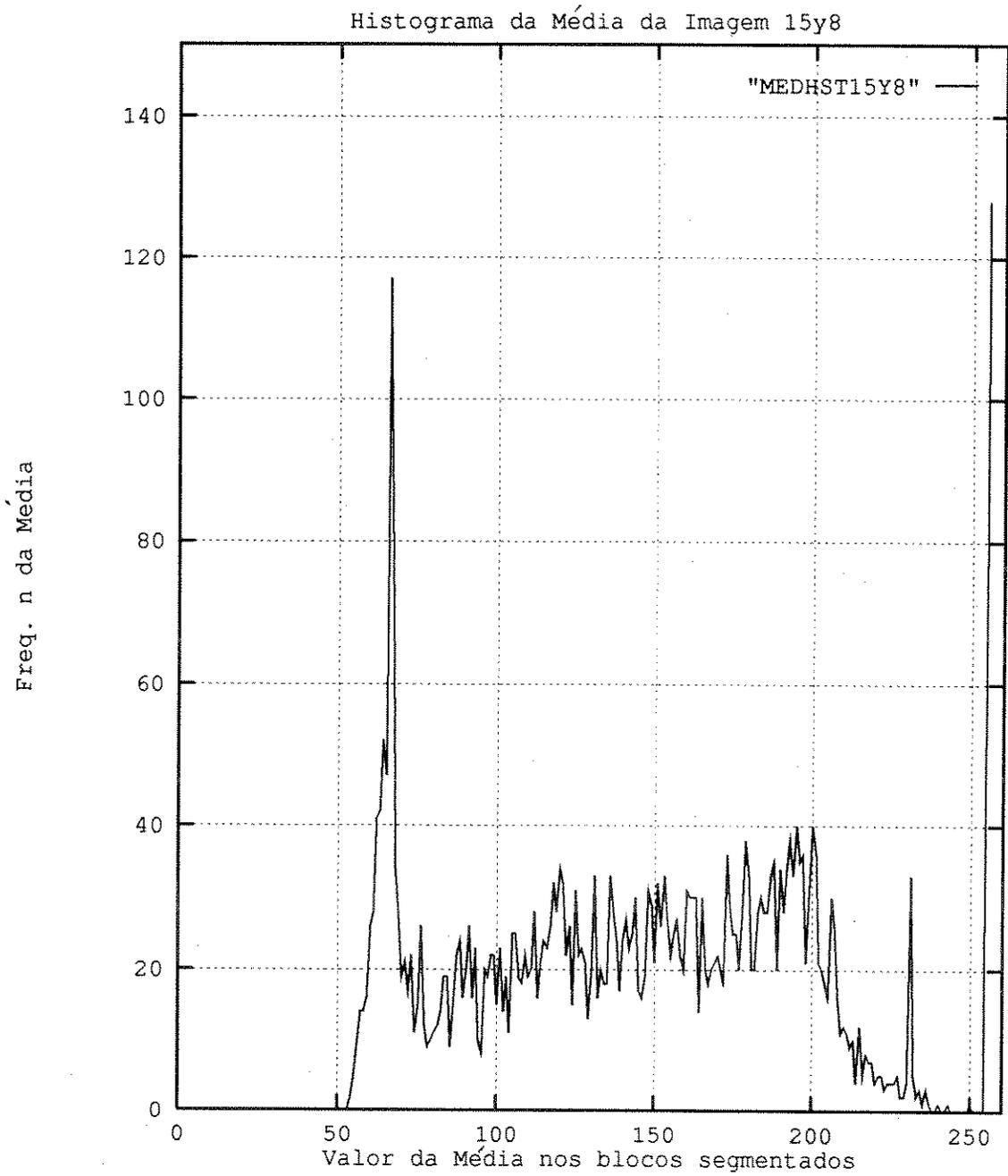


Figura 4.30: Comportamento da Média na segmentação de 8x8 na imagem smpte15y.

4.4.4.10 Histogramas dos Parâmetros de Controle Estatístico para a Imagem Lena usando processamento em blocos de 8x8

Na figura (4.31) mostra-se o histograma do Desvio Padrão nos blocos de 8x8 que constituem a imagem Lena.

Pode-se observar que os valores de Desvio Padrão situados entre 2 e 10 apresentam frequência de ocorrência superior a 100 chegando a atingir valores acima de 860 enquanto que valores de Desvio Padrão acima 18 não ultrapassam a frequência de ocorrência de 60. Isto mostra uma certa estacionariedade nos valores baixos de Desvio Padrão. Aparentemente as mudanças abruptas da imagem são mais graduais e apresentam-se nos valores de Desvio Padrão de 60 a 72, porém com menor frequência de ocorrência.

Na figura (4.32) mostra-se o histograma da Média nos blocos de 8x8 para a imagem Lena.

Pode-se observar que os valores da Média situados entre 60 e 76 apresentam frequência de repetição superior a 20 e entre 118 e 160 apresentam frequência de ocorrência superior a 17 chegando a atingir valores acima de 50 enquanto que valores da Média acima de 208 não ultrapassam a frequência de ocorrência de 10. Em forma global os valores da Média apresentam um comportamento mais estável entre 60 e 208 do que os casos prévios.

4.4.4.11 Histogramas dos Parâmetros de Controle Estatístico para a Imagem Babuíno usando processamento em blocos de 8x8

Na figura (4.33) mostra-se o histograma do Desvio Padrão nos blocos de 8x8 que constituem a imagem Babuíno.

Pode-se observar que os valores de Desvio Padrão situados entre 5 e 42 apresentam frequência de ocorrência superior a 75 chegando a atingir valores acima de 150 enquanto que valores de Desvio Padrão acima 42 não ultrapassam a frequência de ocorrência de 40. Isto mostra um grande espalhamento nos valores baixos de Desvio Padrão em relação ao observado nas outras imagens de teste. O espalhamento prévio estaria indicando a existência de variações transitórias estáveis de distintos níveis de Luminância. Aparentemente não tem mudanças abruptas.

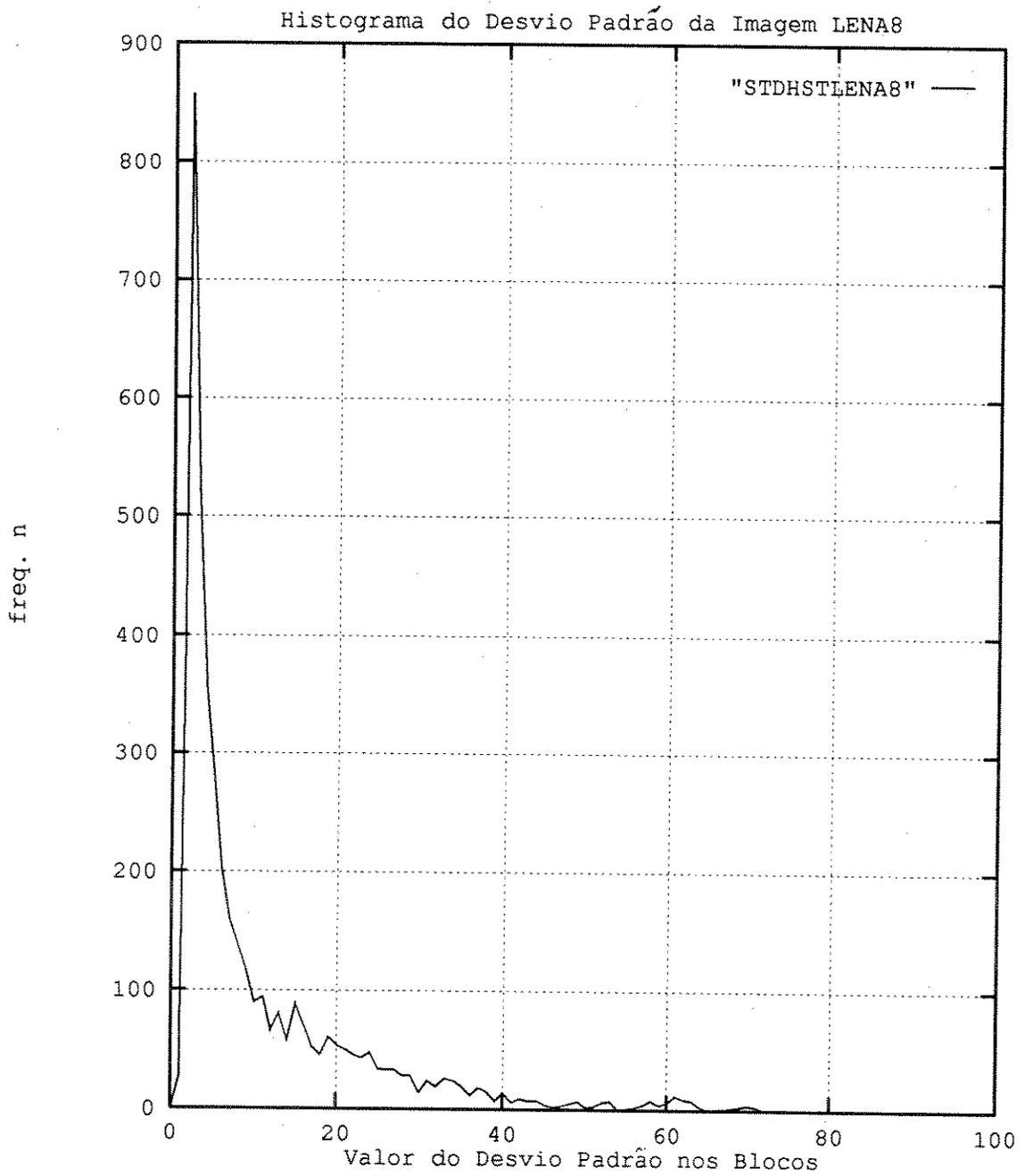
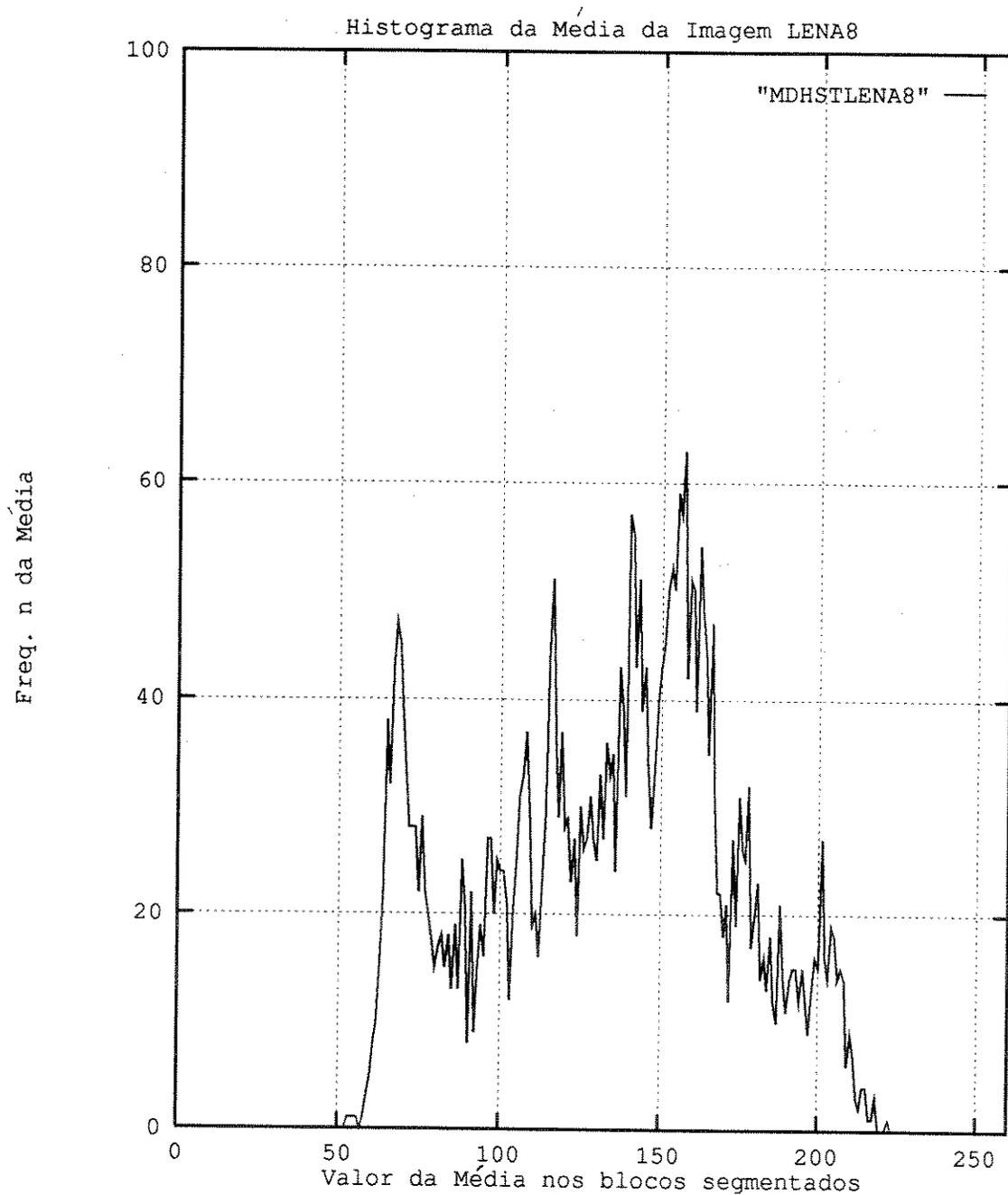


Figura 4.31: Comportamento do Desvio Padrão na segmentação de 8x8 na imagem Lena.

Figura 4.32: Comportamento da Média na segmentação de 8×8 na imagem Lena.

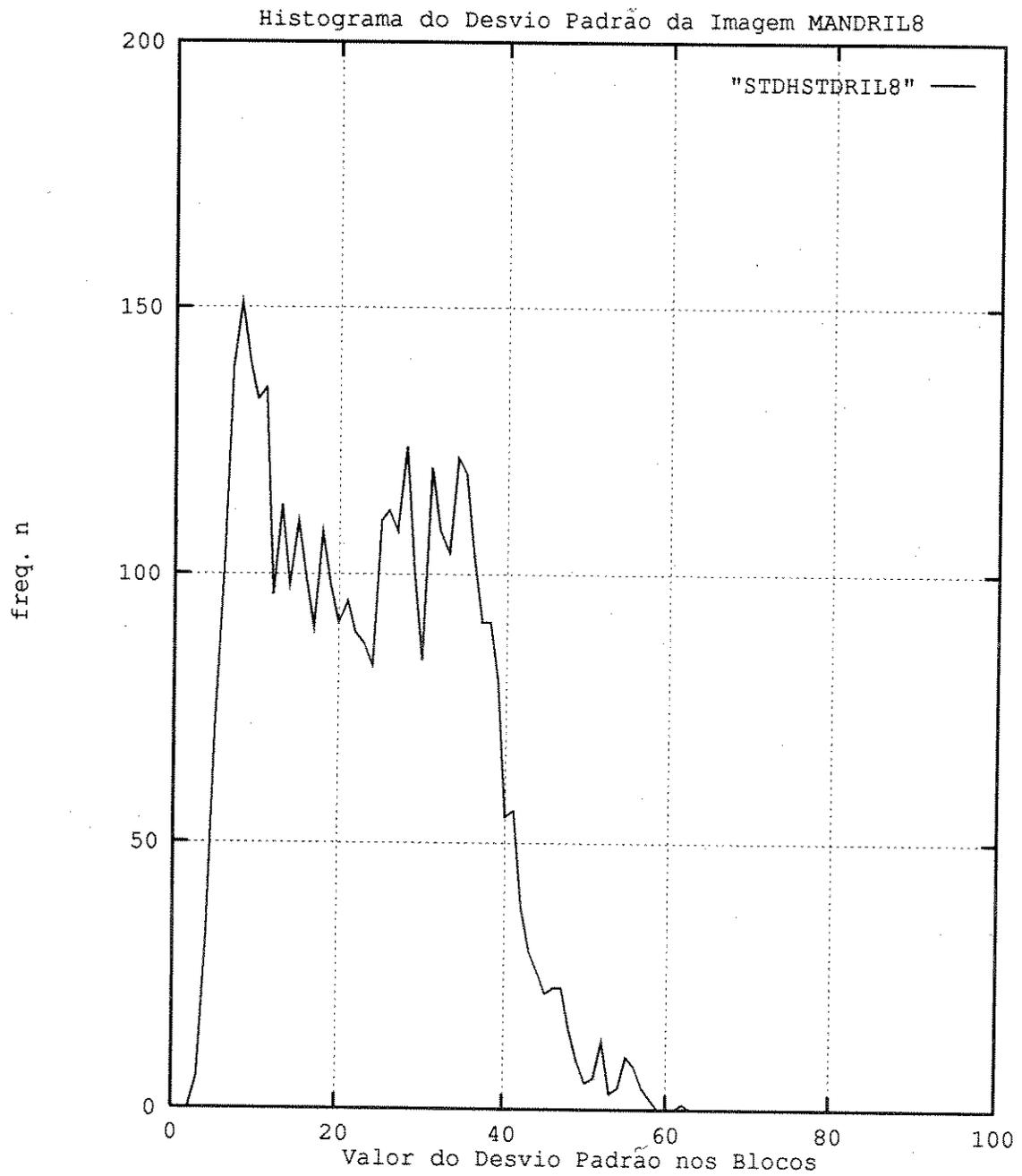


Figura 4.33: Comportamento do Desvio Padrão na segmentação de 8x8 na imagem Babuíno.

Na figura (4.34) mostra-se o histograma da Média nos blocos de 8x8 para a imagem Babuíno. Pode-se observar que os valores da Média situados entre 80 e 162 e, também entre 175 e 178 apresentam frequência de repetição superior a 20, diminuindo gradualmente fora destes intervalos. Observa-se a existência da variação transitória estável nos diferentes níveis de luminância da Média.

4.4.4.12 Conclusões sobre o Comportamento do Histograma do Desvio Padrão para o processamento de 8x8

Na figura (4.35) mostra-se o comportamento do Desvio Padrão nas diferentes imagens de teste como resultado da segmentação de 8x8.

Nesta figura observa-se que os valores maiores de frequência de ocorrência estão relacionados aos menores valores do Desvio Padrão situados no intervalo entre 2 e 15 não chegando a atingir a frequência de repetição de valor igual a 900, onde os valores da imagem Lena ficam em torno deste valor quando comparados às outras três imagens de teste. O fato de terem os valores menores do Desvio Padrão um valor maior da frequência ocorrência está indicando uma mudança de níveis mais suave e um comportamento quase estacionário nos mesmos. O Babuíno foi a exceção à esta característica.

4.4.4.13 Conclusões sobre o Comportamento da Média para o processamento de 8x8

Na figura (4.36) mostra-se o comportamento da Média nas diferentes imagens de teste como resultado da segmentação de 8x8.

Nesta figura pode-se ver que os valores da Média tem um comportamento quase simétrico no intervalo entre 25 e 225 para as imagens Lena e Babuíno, porém para as imagens da SMPTE tem marcado com maior frequência da Média os intervalos de baixos valores entre 70 e 102 para a imagem smpte01y e no caso da imagem smpte15y os maiores valores de frequência da Média sem atingir o valor de 160 estão no intervalo de valores baixos da Média entre 60 e 70 depois, ainda que ambas apresentem uma grande diferença de luminosidade média, (como foi comprovado na tabela(4.2) do comportamento estatístico das 4 imagens de teste escolhidas) já que uma é de claridade do dia (smpte01y) e a outra

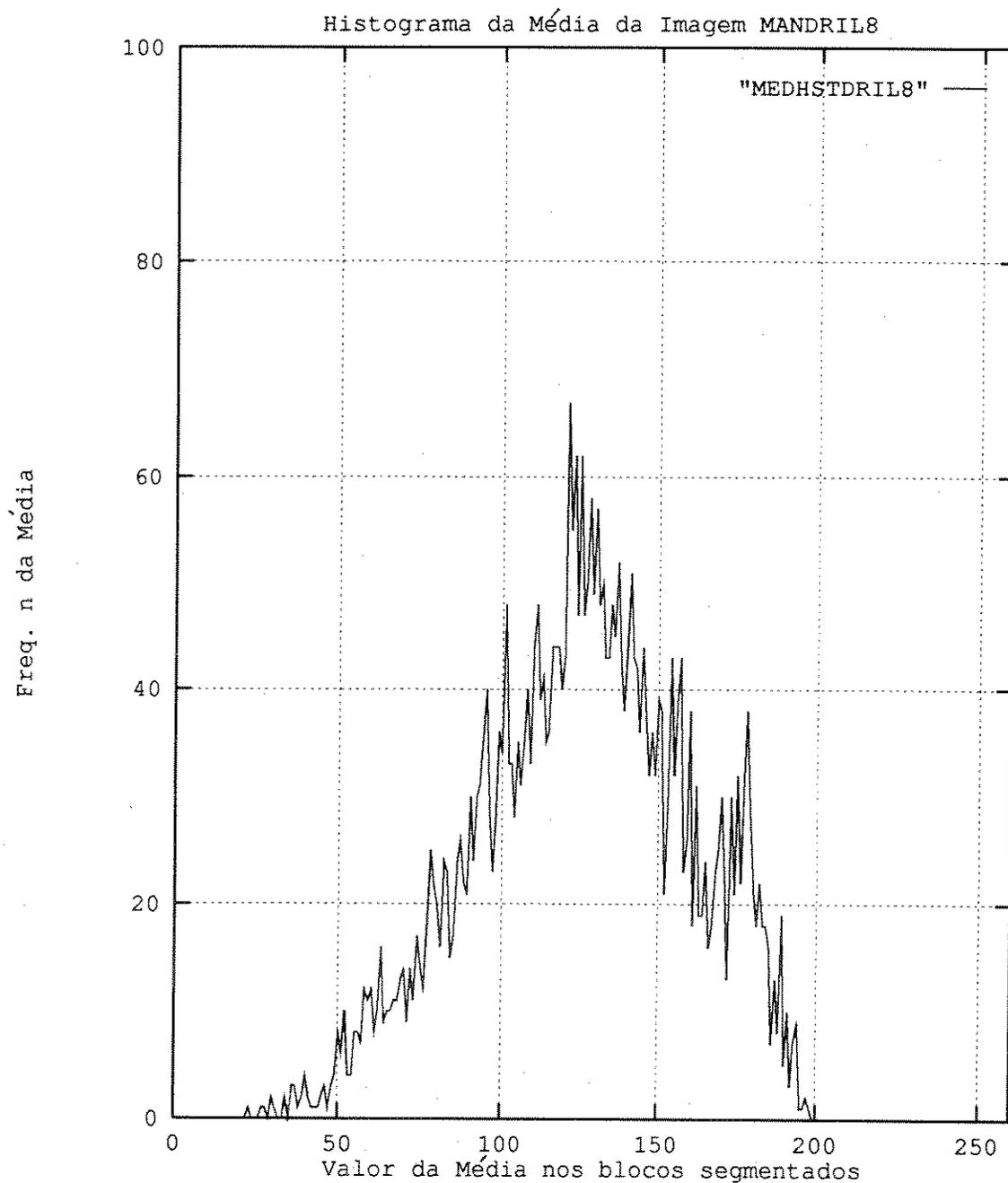


Figura 4.34: Comportamento da Média na segmentação de 8x8 na imagem Babuíno.

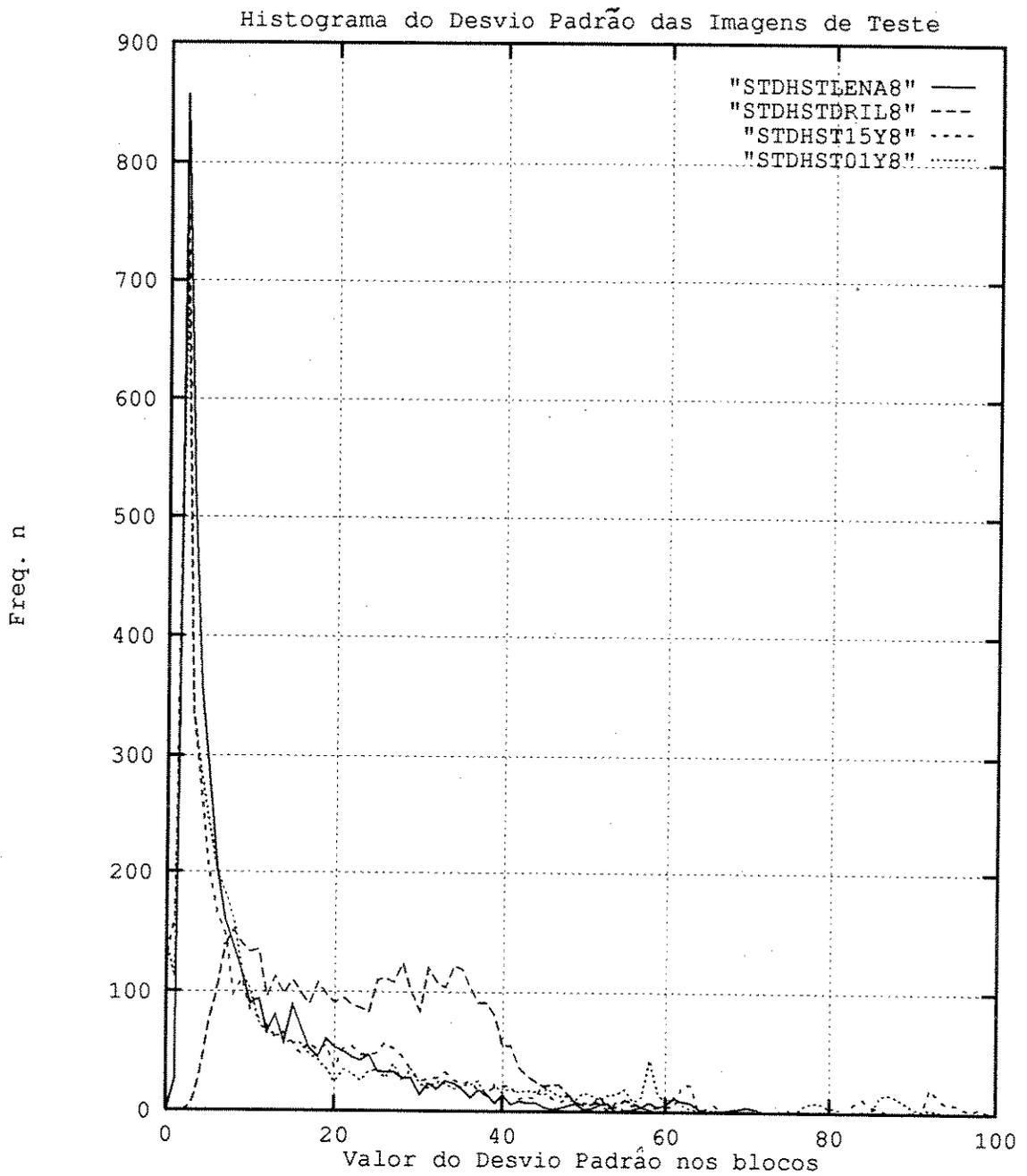


Figura 4.35: Comportamento do Desvio Padrão na segmentação de 8x8 nas diferentes imagens de teste.

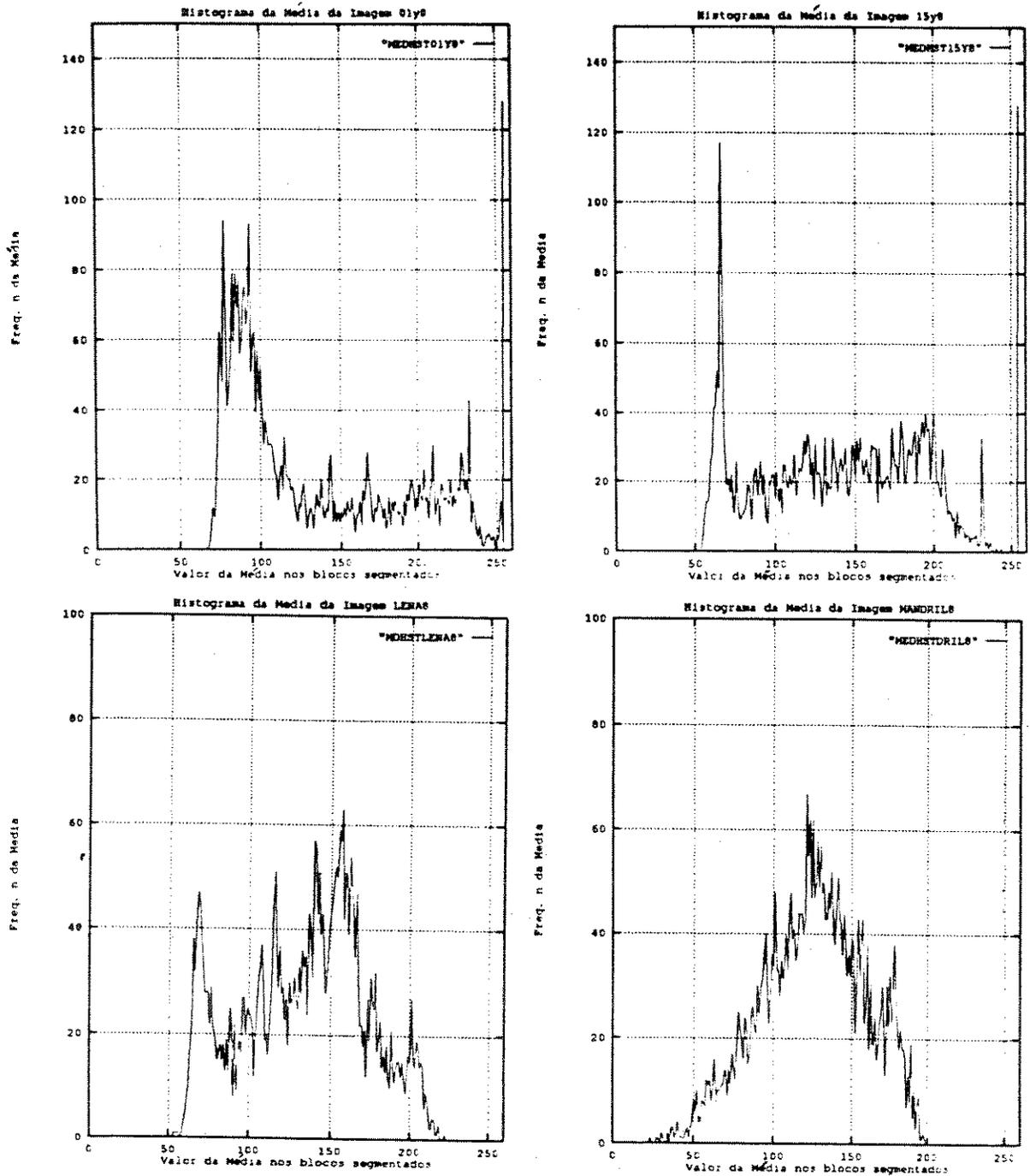


Figura 4.36: Comportamento da Média na segmentação de 8x8 nas diferentes imagens de teste.

é de claridade do interior de uma cozinha e ambas apresentam um comportamento estável em frequência nos intervalos de valores da Média entre 110 e 240 para a imagem *smpte01y* e nos intervalos de 70 e 210 de valores da Média na imagem *smpte15y*.

Observa-se dos parágrafos prévios, que o comportamento dos parâmetros estatísticos já estão seguindo uma forma mais definida, tanto para a Média como para o Desvio Padrão. Portanto, apenas serão vistas as conclusões do comportamento nas diferentes imagens de teste para a segmentação de bloco para 4x4.

4.4.4.14 Conclusões sobre o Comportamento do Desvio Padrão para o processamento de 4x4

Na figura (4.37) mostra-se o comportamento do Desvio Padrão nas diferentes imagens de teste como resultado da segmentação de 4x4.

Nesta figura tem-se que os maiores valores de frequência encontram-se relacionados aos baixos valores do Desvio Padrão entre 2 e 10 sem atingir a frequência de repetição de 4750, sendo os valores da imagem *Lena* os que ficam mais perto deste valor quando comparados aos das três imagens de teste. Pode-se observar da tabela (3.3) que este valor de frequência é apenas maior que 25 % da quantidade de blocos analisados (16.384) e o fato de ter valores menores para o Desvio Padrão nos blocos com um valor maior de frequência, está indicando uma variação dos níveis de forma mais suave e de comportamento quase estacionário dos mesmos pela redução de tamanho do bloco.

4.4.4.15 Conclusões sobre o Comportamento da Média para o processamento de 4x4

Na figura (4.38) mostra-se o comportamento da Média nas diferentes imagens de teste como resultado da segmentação de 4x4.

Nesta figura pode-se observar que a Média tem um comportamento quase simétrico no intervalo de valores entre 20 e 225 para as imagens *Lena* e *Babuíno*. Para as imagens da *SMPTE*, os intervalos de valores baixos ocorrem com maior frequência. Para a imagem *smpte01y* estes valores variam entre 70 e 102. No caso da imagem *smpte15y* estes

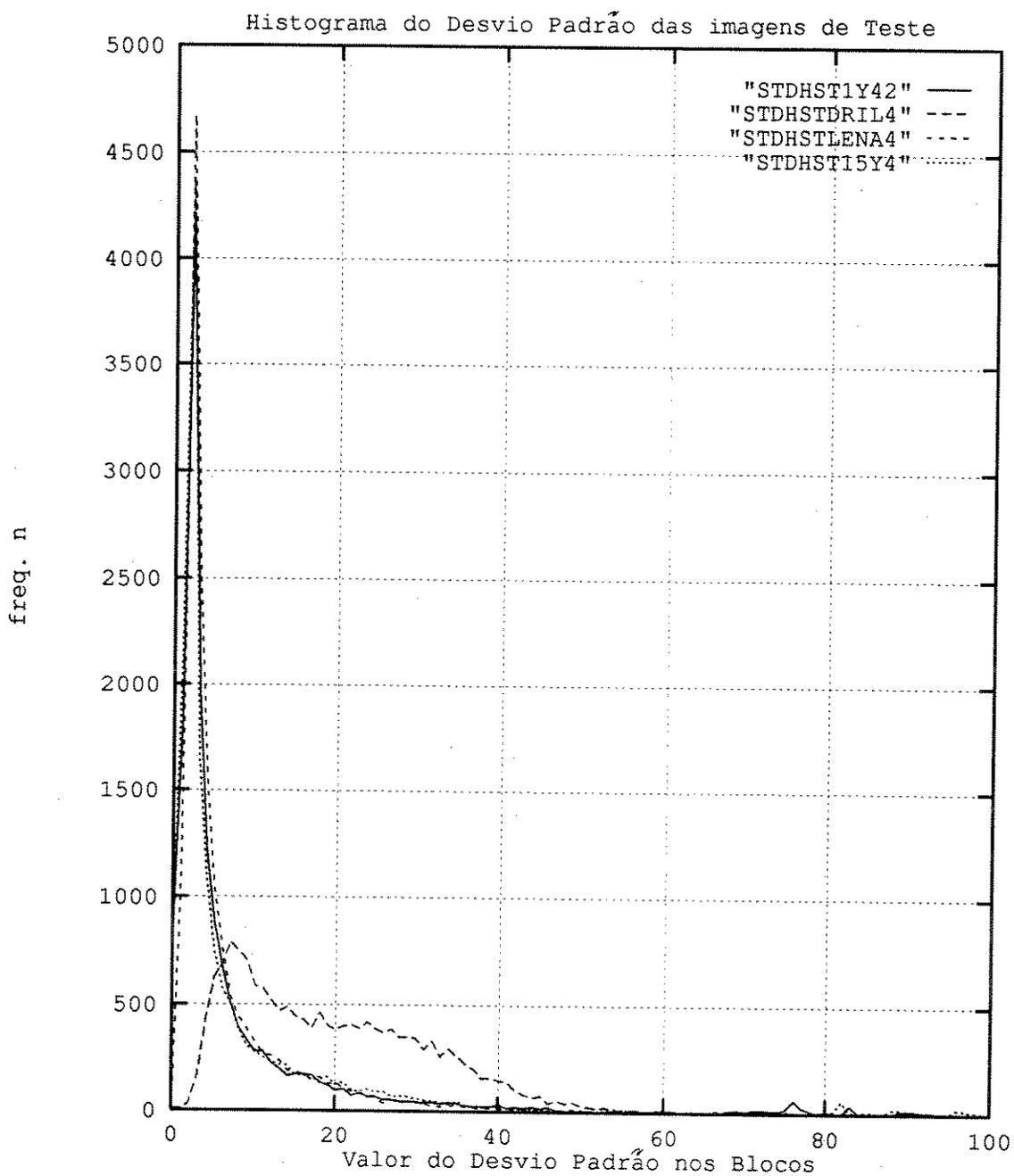


Figura 4.37: Comportamento do Desvio Padrão na segmentação de 4x4 nas diferentes imagens de teste.

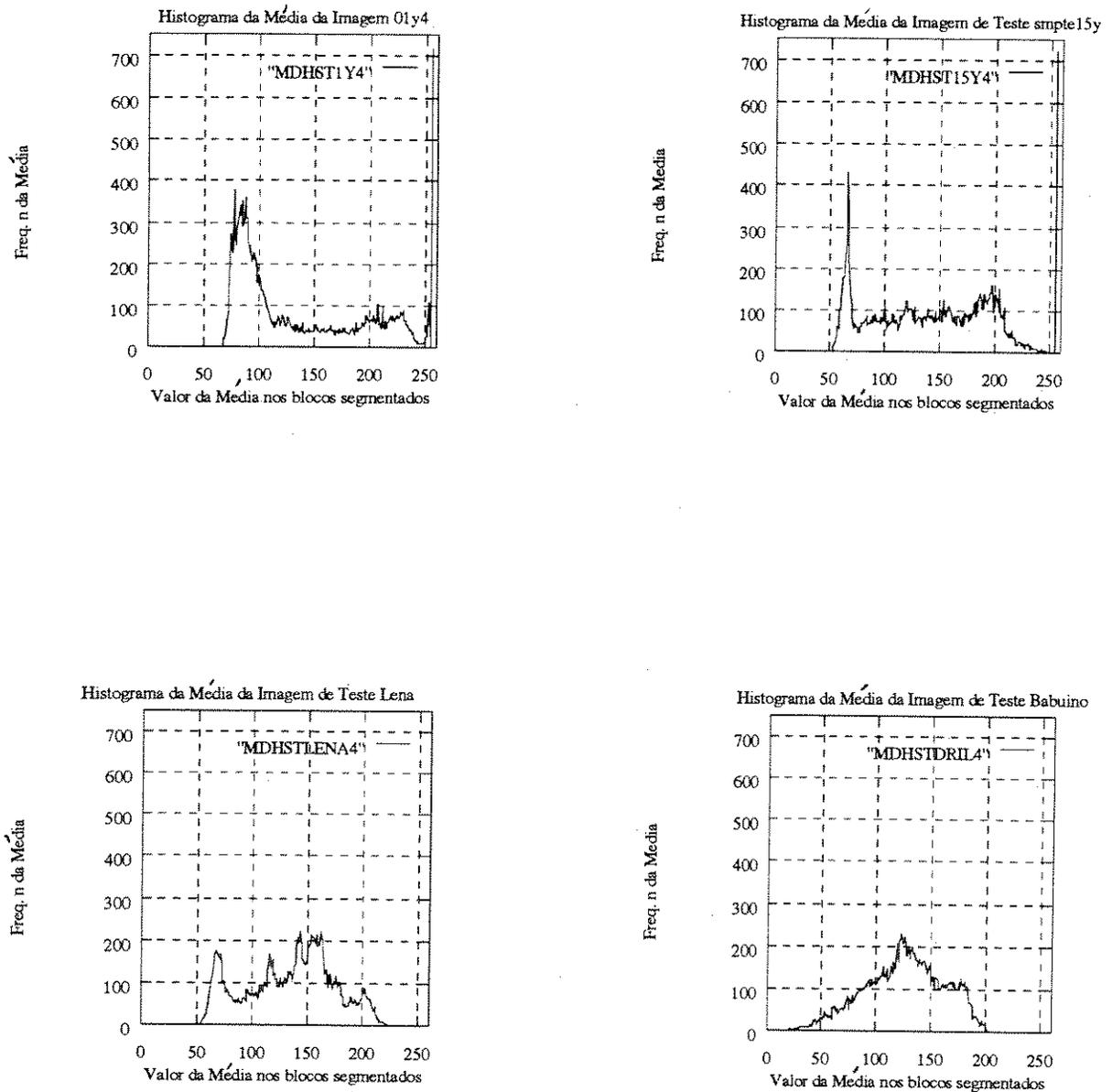


Figura 4.38: Comportamento da Média na segmentação de 4x4 nas diferentes imagens de teste.

valores variam entre 60 e 70 com frequência de ocorrência menor que 450. Estas imagens apresentam uma diferença de luminosidade média (como foi comprovado na tabela(4.2) do comportamento estatístico das 4 imagens de teste escolhidas), já que uma é ao ar livre e de dia (smpte01y) e a outra é no interior de uma cozinha e ambas apresentam um comportamento estável em frequência nos intervalos entre 110 e 240 para a imagem smpte01y e nos intervalos entre 70 e 210 de valores da Média na imagem smpte15y.

4.4.5 Resumo dos Parâmetros Estatísticos de Controle para as Imagens de Teste.

O sistema empregado prevê a utilização de segmentação de bloco variável. Esta segmentação é baseada na comparação dos parâmetros estatísticos do Desvio Padrão e da Média com valores empiricamente determinados pela análise comparativa dos gráficos de histogramas feitos previamente para as quatro imagens de teste, na segmentação do tamanho do bloco de 16x16 para 8x8 e finalmente para 4x4, ou usando o algoritmo no sentido inverso, juntar 4 sub-blocos de 4x4 e formar um sub-bloco de 8x8 e depois juntar 4 sub-blocos de 8x8 e conformar um bloco máximo de 16x16. O primeiro modo é o implementado nesta dissertação por ser o mais comum.

A tabela (4.5) apresenta os valores empíricos adotados como controle de segmentação por Desvio Padrão, que permitem discriminar as mudanças abruptas da imagem, das variações suaves no fundo. O comportamento para a segmentação de 4x4, apresenta os mesmos intervalos de controle que o dois casos prévios.

	Segmen. 16				Segmen. 8			
	x		16		x		8	
Valor	mín.s.	máx.s.	mín.ab.	máx.ab.	mín.s.	máx.s.	mín.ab.	máx.ab.
Des.Pd.	2.0	70	75	95	2.0	70	75	95

Tabela 4.5: Resumo dos Valores limites de Intervalos Adequados do Desvio Padrão usado para o Controle da segmentação das Imagens de Teste.

Onde,

mín.s.=mínimo suave; máx.s.=máximo suave; mín.ab.=mínimo abrupto; máx.ab.=máximo abrupto.

A tabela resumo (4.6) apresenta os limites de variação da Média e como referência empírica à diferença entre os blocos adjacentes igual a 10, que vai fazer o controle de segmentação do primeiro nível ($k=1$) do passo de 16×16 para 8×8 ou a juntura de quatro blocos de 8×8 formando um de 16×16 . A outra referência empírica limite é para a diferença entre os blocos igual a 20, que define o controle de segmentação de segundo nível ($k=2$) do passo de 8×8 para 4×4 ou a junção de quatro blocos de 4×4 formando um de 8×8 .

Valor da	Segmen. 16 x 16			Segmen. 8 x 8		
	mín.	máx.	Dif.($k=1$)	mín.	máx.	Dif.($k=2$)
Média	33	252	10	25	253	20

Tabela 4.6: Resumo dos Valores limites de Intervalos Adequados da Média usado para o Controle da segmentação das Imagens de Teste.

Os valores apresentados nestas tabelas servem para projetar na forma adequada o controle do algoritmo empregado.

4.5 Algoritmo Adaptativo Empregado

4.5.1 Introdução

Vários algoritmos DCT com quantização adaptativa ao conteúdo da imagem têm sido pesquisados [20] e [22] mas, considerando tamanho de bloco fixo, em [22], um esquema de algoritmo DCT-2D com seleção triangular adaptativa de K -ésima ordem tem sido empregado, onde o limiar e o passo do quantizador é escolhido adaptativamente baseado no Desvio Padrão para a seleção do coeficiente k em cada bloco fixo. Deste modo obtém-se melhores resultados.

Visto que o limiar adaptativo e o quantizador adaptativo fornecem como resultado uma qualidade de imagem superior com menor taxa de bits, neste trabalho tem-se como objetivo adicionar essas características no algoritmo de bloco variável usando a DCT.

4.5.2 Diagrama em blocos adaptativo simplificado

O diagrama em blocos é apresentado na figura (4.39). Como na codificação por transformada convencional, a imagem de entrada é segmentada em blocos adjacentes não sobrepostos de tamanhos adequados $M \times M$, onde o valor típico de M pode ser 32×32 ou 16×16 . Neste trabalho utiliza-se 16×16 , logo cada bloco é segmentado em sub-blocos de tamanho variável de acordo com os parâmetro estatísticos mencionados previamente.

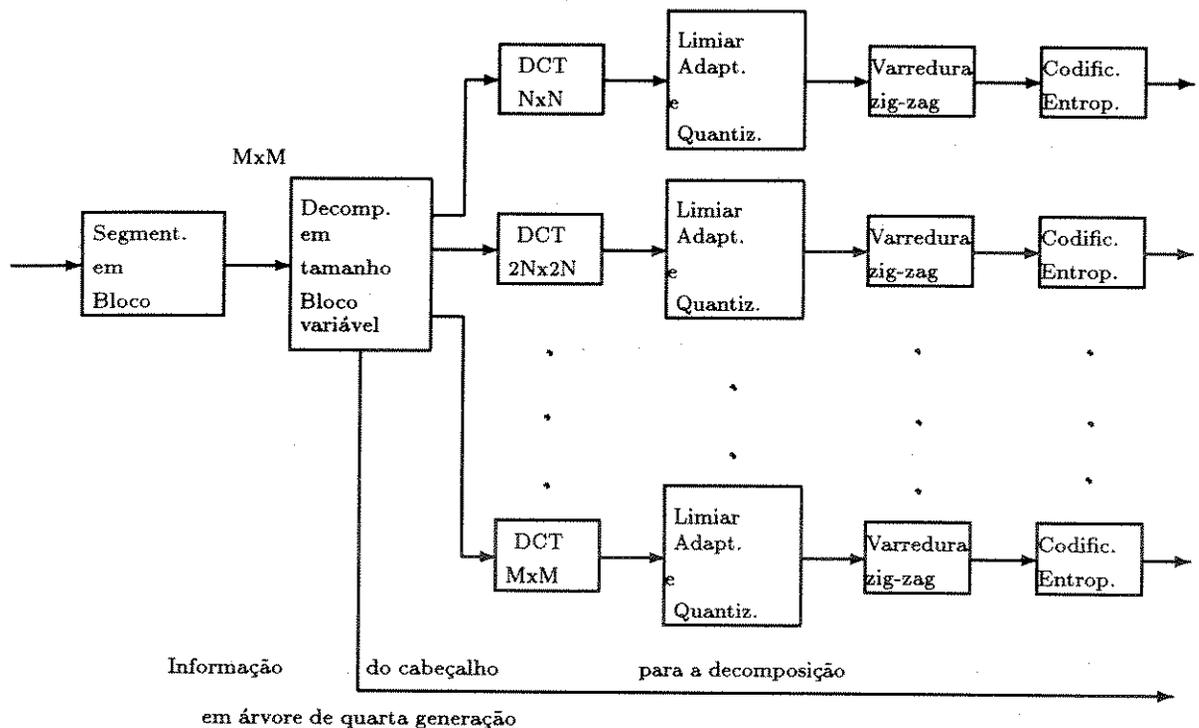


Figura 4.39: Codificação por Transformada adaptativa com segmentação de tamanho de bloco variável.

O objetivo de segmentar o bloco em blocos menores, é organizar os dados da

imagem numa forma mais estacionária, de tal modo, que pode-se obter na saída da DCT uma maior decorrelação que permite uma melhor compressão da imagem.

Se permite-se que o menor tamanho de bloco seja $N \times N$, a relação entre o maior e o menor bloco será dado por (4.2),

$$M = N \cdot 2^{L-1} | L = 1, 2, \dots \quad (4.2)$$

Onde, L = nível de profundidade da segmentação (seção 4.5.3).

Neste caso como $L=3$, logo cada bloco de $M \times M$ é primeiro decomposto em quatro sub-blocos de $4^{L-2} \cdot N \times N$. Para cada destes $4^{L-2} \cdot N \times N$ sub-blocos adjacentes, aplicando o critério de separação ou união baseado nos parâmetros estatísticos que são apresentado mais adiante, permitirá voltar a testar se estes quatro sub-blocos podem ser separados. Ou seja, se o teste é positivo, estes quatro sub-blocos são separados em outros 4 sub-blocos de $4^{L-3} \cdot N \times N$, chegando ao tamanho mínimo $N \times N$. Um exemplo é apresentado na figura(4.40).



Figura 4.40: Segmentação de tamanho de bloco por junção ou união de quatro sub-blocos adjacentes.

Não é fácil encontrar um critério de segmentação de bloco perfeito. Como observou-se previamente, a degradação da codificação por DCT é mais severa em blocos que contém bordas de imagem(mudanças abruptas), os quais não cumprem com a estacionariedade espacial quando o tamanho do bloco é grande. Portanto um critério de segmentação de bloco desejado teria que classificar as bordas da imagem com uma categoria que permitisse usar tamanho do bloco pequeno. Também, para tirar vantagem do fato que o S.V.H. não é sensível à degradação nas áreas de grande detalhe estável ou de textura ruidosa, logo este critério deveria classificar nessas áreas na categoria de utilização de tamanho de bloco

maior, para uma melhor compressão. Na referência [33], um critério simples envolvendo a Média dos quatro sub-blocos adjacentes é considerado. Ainda que a qualidade da imagem resultante é muito boa, o critério pode não ser adequado para segmentar as áreas de alto detalhe estável ou de textura ruidosa em blocos de tamanho maior e como consequência a taxa de bits seria desnecessariamente alta se a imagem possuísse uma boa quantidade para esse tipo de área.

Propõe-se uma estratégia de segmentação que supere esta situação. Portanto, além da Média (Estatística de primeira ordem) usada na referência [33], utiliza-se também, a estatística de segunda ordem, tal como a Variância ou Desvio Padrão de quatro sub-blocos adjacentes. Seja, $m_k(i)$ e $\sigma_k(i)$, a Média e o Desvio Padrão respectivamente, onde $i=1,2,3,4$ representam os sub-blocos adjacentes e k é o índice do nível de segmentação que determina o tamanho do sub-bloco em uso, k em geral encontra-se num intervalo de 1 a $L-1$, neste caso, se $L=3$, logo k toma valores de 1 a 2.

O critério para separar ou reagrupar quatro sub-blocos adjacentes usa as seguintes decisões baseadas em testes de hipóteses implementadas no programa " adaptivo.for":

Testes de hipóteses :

1. Se, $a_k < \sigma_k(i) < b_k$ para todo i , não se separam.
2. E se, $|m_k(i) - m_k(j)| < t_k$ para todo $i \neq j$, não se separam.
3. Qualquer outro caso, separam-se.

Onde,

a_k , b_k y t_k são limiares pré-determinados.

O primeiro teste (1.), é usado principalmente para discriminar entre outras as áreas de textura de ruído estável (secção 4.4.3), onde o S.V.H. não é sensível à degradação e um tamanho de bloco grande deve ser usado para obter uma melhor compressão sem degradar a qualidade visual. Verificou-se para as 4 imagens de teste empregadas que o Desvio Padrão das áreas de textura de ruído estável é limitado por valores moderados para imagens típicas, enquanto que resultam valores altos em blocos que possuem bordas de alto contraste, cuja ondulação produzirá efeitos visíveis na codificação por DCT [17]. Logo é desejável não separar em quatro sub-blocos adjacentes, se todos os valores do Desvio Padrão

são limitados por valores pré-determinados de a_k y b_k . Neste caso de acordo com a tabela (4.5) obtida a partir das imagens de teste os parâmetros limites para o Desvio Padrão são: $a_k = 2.0$ para o fator de união de bloco $k=1,2$ e $b_k = 70$ também para $k=1,2$. No caso da tabela (4.6), os valores de parâmetros limites para a Média são: $t_1 = 10$ e $t_2 = 20$ indicando os limites do valor da diferença entre Médias consecutivas, $k=1$, e representando o fator de união de 4 sub-blocos adjacentes com tamanho 16×16 e com possibilidade (teste (1.), positivo) de separar-se a 8×8 , $k=2$, representa o fator de união de 4 sub-blocos adjacentes de tamanho 8×8 e com possibilidade (teste (1.), positivo) de separar-se a 4×4 . Estes parâmetros foram empiricamente escolhidos para obter resultados de segmentação adequado para uma classe de imagens com uma grande variação de características. A distribuição de blocos resultante para cada tamanho da segmentação é apresentada na tabela (4.7)(pág. 146).

Nos resultados apresentados mais adiante, demonstra-se que a inclusão do teste relacionado ao Desvio Padrão dá como resultado um tamanho de segmentação de bloco melhor para as imagens de teste. Testes (2.) e (3.) são usados para discriminar as bordas de alto contraste das áreas simples e suaves. É preciso notar, que a escolha primária do limiar da diferença entre blocos t_k foi 8, valor que não deu certo para frisar o uso de blocos de 16×16 no caso do Babuíno, para conseguir o efeito mudo-se para 10. Logo, os limiares de decisão (a_k , b_k e t_k) podem ser utilizados para segmentar tamanhos de blocos diferentes.

Em resumo, devido à natureza não estacionária dos dados para blocos de grande tamanho, a DCT frequentemente introduz degradação visível, sobretudo se o bloco contém bordas de alto contraste. O esquema da descomposição por tamanho de bloco variável classifica aquelas áreas possuindo bordas para serem tratadas com tamanho de bloco pequeno, e em consequência resulta uma estrutura mais estacionária de dados que ao aplicar a DCT e fazer um filtragem dos coeficientes na compressão, produz uma qualidade subjetivamente melhor para a imagem. Devido ao S.V.H. não ser sensível à degradação daquelas áreas de alto detalhe ou de textura de ruído estável(secção 4.4.3), este esquema vai classificar esta situação na categoria de blocos de tamanho maior possível, neste caso 16×16 , para recolher a vantagem da DCT de possuir a compressão mais eficiente para estes blocos.

4.5.3 Representação por Regiões Estruturadas

A classificação destes blocos de acordo com a estacionariedade levou a pesquisar diferentes métodos, dos quais escolheu-se o processamento paralelo de regiões usando árvores de quarto grau ou quarta geração do tipo linear, referencias [34] e [35] onde apresenta-se como é possível acelerar a computação de regiões com propriedades geométricas a um fator “p” usando computação paralela com “p” processadores (CREW-PRAM)⁽¹⁾.

Uma região da árvore de quarto grau é uma estrutura hierárquica de dados compactos para transmissão comprimida ou armazenamento de imagens digitais. Ela segmenta blocos de $2^m \times 2^m$ não estacionários em quatro quadrantes iguais, onde precisa-se transmitir com antecedência um cabeçalho de informação por cada bloco $M \times M$ a segmentar para facilitar o processo de decodificação no terminal receptor.

Um exemplo extremo fornecidos por vários pesquisadores [34], é que a partir de um bloco não homogêneo é possível decompor em subquadrantes menores até chegar a uma unidade de blocos homogêneos. Neste caso aplica-se apenas a base destes conceitos para procurar a máxima estacionariedade e não a homogeneidade, visto que busca-se explorar a vantagem da compressão da DCT, não sendo objetivo desta Tese procurar um processamento paralelo intenso que comprometa esta compressão de imagem, e que necessite de maior armazenamento em memória.

Para uma imagem binária, no caso extremo, as definições são as seguintes: em um bloco homogêneo que contenha apenas regiões escuras ou então somente regiões claras, cada bloco pode ser representado como um nó em uma árvore de quarto grau e os quatro quadrantes correspondem aos quatro nós “filhos” ou de primeira geração. Logo, a raiz (nível 0) representa o bloco inicial de segmentação $2^m \times 2^m$ ($m=4$, neste caso) e qualquer outro nó ao nível ou profundidade k , representa a região quadrada $2^{m-k} \times 2^{m-k}$. Nós terminais (ou “folhas”) são pretos ou brancos (neste caso, estacionários). Nós interiores, se denominam cinza. Um exemplo desta segmentação por regiões e seu correspondente “Quadtree” com a identificação de nós é mostrada na figura (4.41). Nesta representação objetivando poupar memória apenas os nós terminais homogêneos pretos são armazenados,

¹CREW-PRAM: máquina que tem “p” processadores com acesso comum à mesma memória global. Qualquer número dos processadores podem ler simultaneamente da mesma memória em um tempo constante; também qualquer processador pode escrever numa constante de tempo mínima, embora escrita na mesma posição não seja permitida.[34]

denominando-se representação “Quadtree” sem ponteiro [36],[37] ou “Árvore de quarto grau linear”.

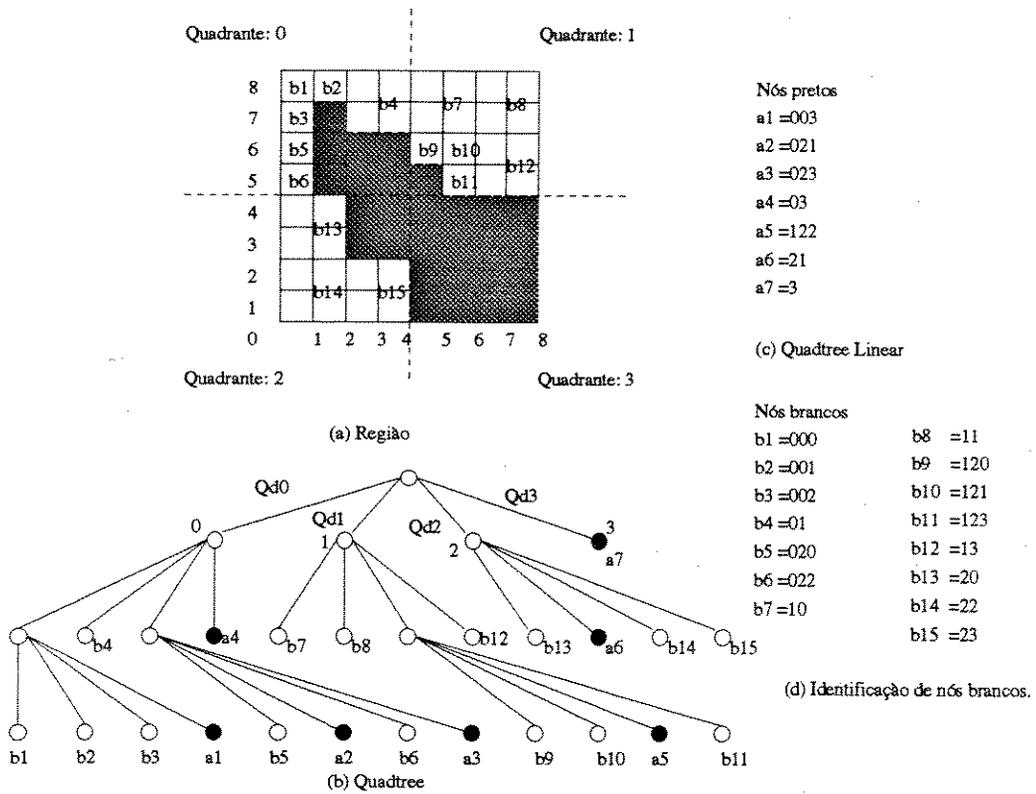


Figura 4.41: Imagem segmentada por Regiões e sua estrutura em árvore de quarto grau.

Esta mesma idéia é utilizada para produzir uma grande compactação para o caso de modelamento de sólidos, CAD-CAM, processamento de imagens, robótica e computação gráfica, onde utiliza-se como base um “Oct-tree linear” (Árvore de octavo grau linear). A figura(4.42) mostra um objeto conformado apenas pelos nós terminais homogêneos pretos e sua correspondente representação.

4.5.4 Indexação codificada de Árvore de Quarto Grau.

Neste caso, com objetivo de conseguir uma identificação eficiente do “Quadtree” similar ao caso linear mencionado previamente, quando o bloco atingir a estacionariedade

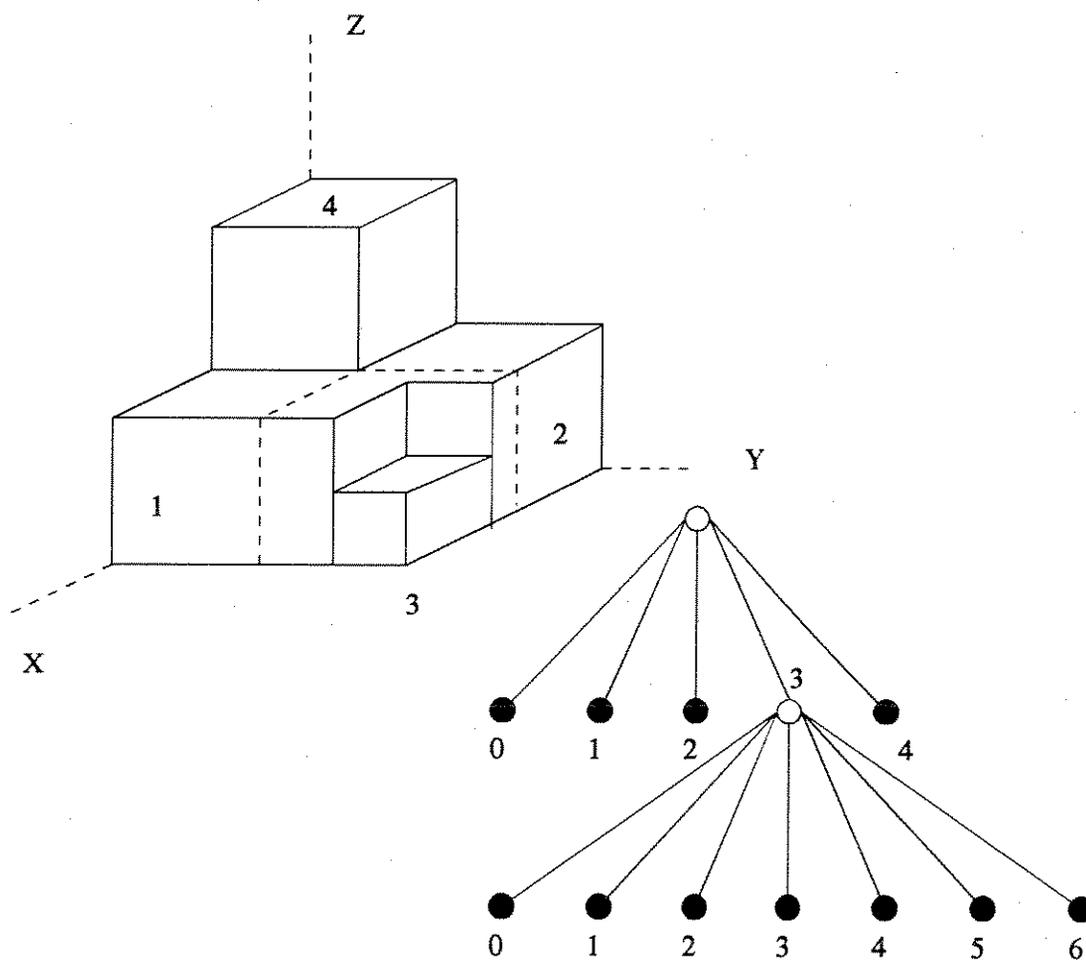
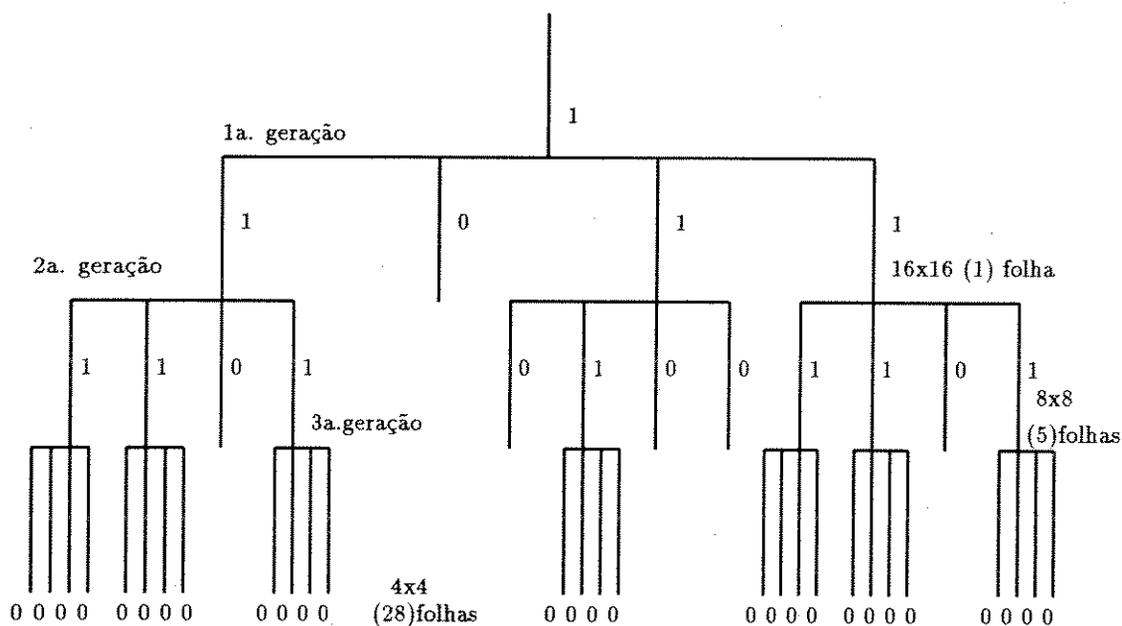


Figura 4.42: Oct-tree de nós terminais escuros e sua representação

adequada para processamento por transformada DCT, este receberá o código “zero”, para que desta maneira na codificação de zeros por comprimento variável apenas se contabilizem estes zeros, melhorando a eficiência da compressão na transmissão.

Um exemplo da codificação eficiente usada para uma árvore de quarto grau é apresentado na figura (4.43). Este corresponde à segmentação pelo tamanho de bloco da figura (4.40), na qual representa teoricamente como seria tratada uma borda de imagem pela segmentação classificada pelos parâmetros estatísticos. Tais parâmetros, são necessários transmitir quando o tamanho de bloco do nível zero é grande (como neste caso, 16x16).



CODIGO DE ÁRVORE DE 4o. GRAU : 1 1011 1101 0100 1101

Figura 4.43: Árvore de quarto grau correspondente à figura (4.39)

4.5.5 Seleção de Coeficientes e Quantização por Blocos Segmentados

Depois que os blocos são segmentados (veja, figura(4.39)), cada sub-bloco é transformado pela DCT com o tamanho apropriado determinado pelo "Quadtree" de grau 4. A DCT usada é de transformação simétrica par [7], [19], cuja implementação segue as equações definidas na seção 3.9 e o intervalo de variação dos coeficientes transformados vai depender do tamanho do bloco, o que significa a necessidade de normalização e de definição de uma tabela de Quantização de acordo com o tamanho do bloco.

Usando a classificação dos blocos com respeito à estacionariedade, obtém-se uma determinada velocidade de variação dos valores dos E.I.. No caso dos blocos maiores (16x16) a energia ficará concentrada no intervalo de frequências baixas em uns poucos coeficientes transformados pela DCT-2D, fazendo uma compressão eficiente para dados altamente correlatos devido à estacionariedade dos mesmos.

Conceitualmente, a DCT-1D equivale à Transformada de Fourier tomando apenas a parte real da equação (cosseno).

No caso da DCT-2D, esta pode ser obtida ao efetuar a DCT-1D nas colunas e depois efetuar a DCT-1D nas linhas devido a que esse tipo de transformada cumpre as condições de ser ortogonal e separável (secção 3.8). Os coeficientes transformados são ordenados de tal forma que o valor médio (o coeficiente DC) encontra-se na parte superior esquerda dos blocos e os coeficientes de altas frequências vão subindo de frequência à medida que aumenta a distancia do coeficiente DC. Frequências altas com traços verticais são representadas pelas filas de localização maior e frequências altas de traços horizontais são representadas pelas colunas superiores, no caso de pontos horizontais simétricos e harmônicos às funções bases de imagem da transformada, como no caso do sinal Multisalva, a energia dos dados de entrada é compactada apenas nos coeficientes transformados da primeira coluna do bloco (figura 4.44), fazendo muito crítica sua recuperação original depois do quantizador no caso de usar codificação por transformadas(secção 4.4.3).

A seleção dos coeficientes transformados segue uma lei zonal adaptativa do K-ésimo coeficiente ao valor do Desvio Padrão do bloco e a magnitude do coeficiente vai depender da distribuição da energia em sua localização dentro do bloco. A distribuição

668.00	0.00	0.00	0.00
-22.30	0.00	0.00	0.00
-6.00	0.00	0.00	0.00
-1.58	0.00	0.00	0.00

Figura 4.44: Localização das faixas verticais de frequência do sinal Multisalva (pontos horizontais simétricos e harmônicos às funções bases de imagem da transformada), no domínio da transformada DCT-2D

de energia dos coeficientes segue um modelo Laplaciano se a imagem de entrada tem um comportamento Markoviano de primeira ordem [7],[13] válido para a grande maioria das imagens reais (secção 3.14, 3.15). Sua representação tridimensional é mostrada na figura (4.45), seguindo o modelo da equação 3.88, e sua representação zonal bidimensional compara-se com a zona de seleção triangular adaptativa ao Desvio Padrão, na figura (4.46), onde pode perceber-se que a zona triangular é mais conservativa do que a zona Laplaciana.

Modelo Laplaciano fornecido por 3.88,

$$r(u, v) = \exp[-\alpha \cdot |u| - \beta \cdot |v|]$$

onde, $\alpha = \beta = -\ln 0.95$

O parâmetro K-ésimo de seleção adaptativa de coeficientes [22] é aplicado aos blocos de tamanho variável e representa o número de coeficientes a ser retidos por bloco. Neste caso, os números de coeficientes são 105, 36 e 16 para 16x16, 8x8 e 4x4 respectivamente.

Depois da seleção adequada dos coeficientes, vem o processo de quantização em valores que dependem do tamanho do bloco e da tabela de quantização respectiva, (figura 4.47). A tabela de quantização tem por objetivo reduzir a magnitude dos coeficientes por normalização e aumentar o número de coeficientes com valor zero (para melhorar a eficiência de compressão no codificador entrópico).

Quantizadores uniformes são escolhidos (veja figura 3.13) para os diferentes tamanhos de blocos. O tamanho do passo é variado de acordo à localização dos coeficientes

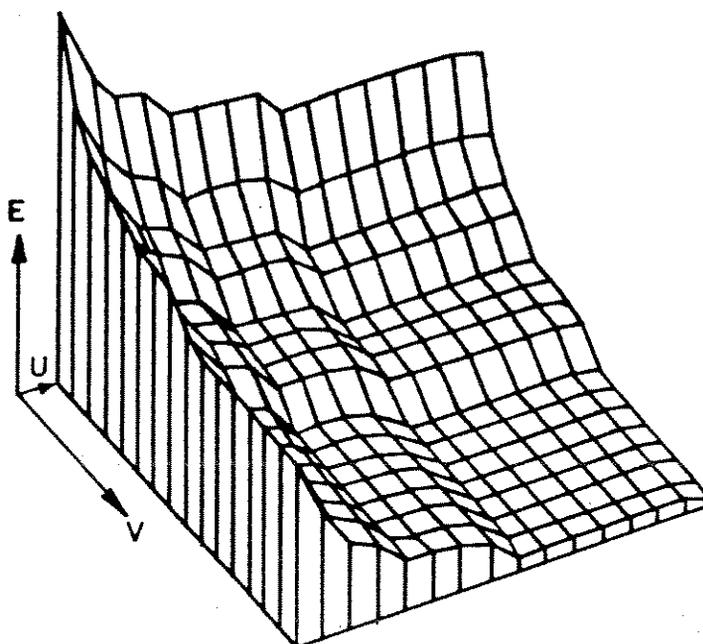


Figura 4.45: Distribuição da energia dos coeficientes usando a equação 3.88, variando os valores das coordenadas U e V .

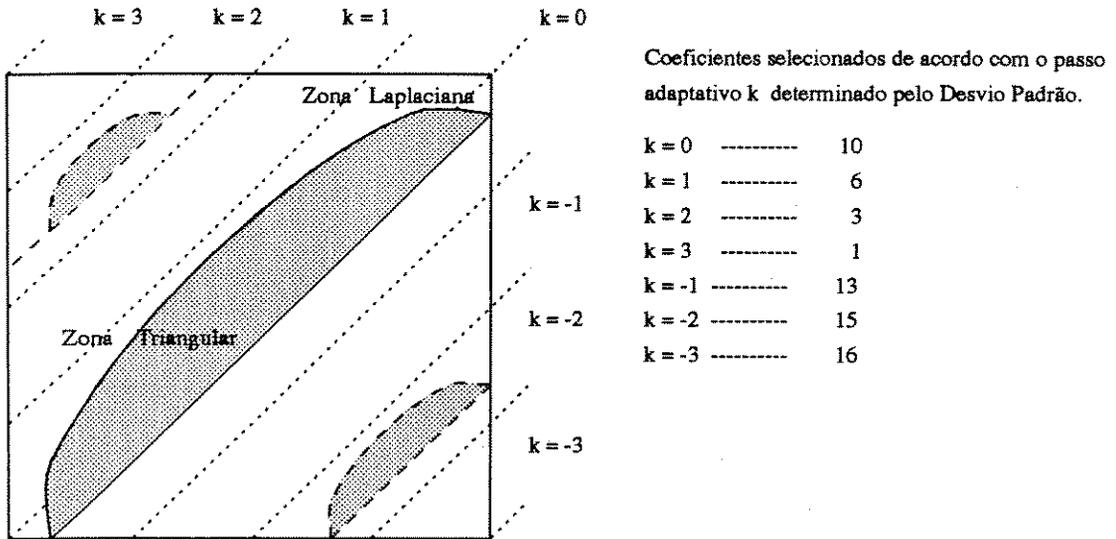


Figura 4.46: Controle de seleção adaptativa ao Desvio Padrão aplicado à amostragem por zona de coeficientes. Mostra-se a diferença entre a seleção triangular e a seleção por zona Laplaciana para um bloco de 4x4. Neste caso usa-se k=-3.

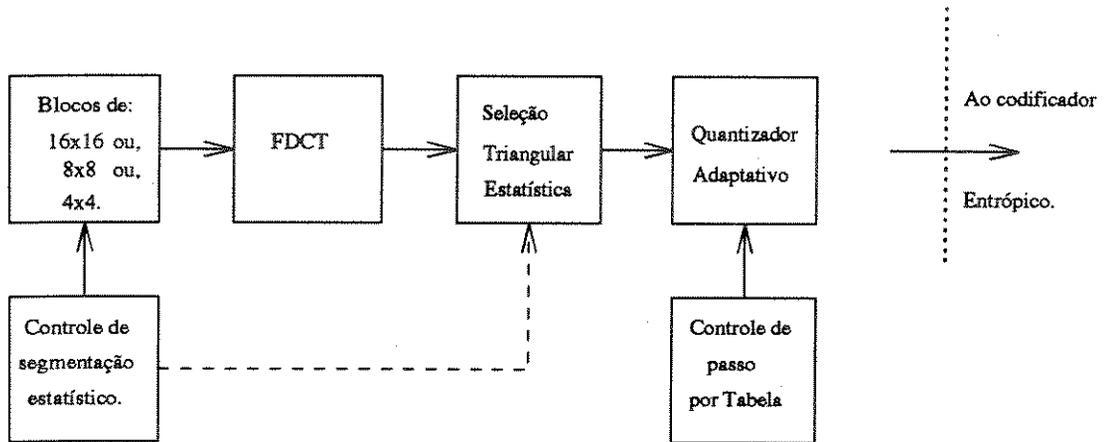


Figura 4.47: Processo de quantização empregando Tabela.

dentro do bloco e para cada coeficiente $F(q_{v,u})$ (da equação 3.48) é possível obter um tamanho de passo a partir do correspondente elemento Q_{vu} da tabela de quantização especificada na transmissão pelo parâmetro de quadro Tq_i (apêndice e) no cabeçalho de informação paralela adicional. Este parâmetro se simulou nas avaliações dependente da localização dos coeficientes. Logo o quantizador é definido pela equação (4.3), onde Fr representa um fator de arredondamento ao inteiro mais próximo.

$$Sq_{vu} = Fr \cdot \frac{F(q_{vu})}{Q_{vu}} \quad (4.3)$$

Sq_{vu} é o coeficiente que tem sido selecionado por zona triangular adaptativa e que finalmente é normalizado pelo tamanho de passo do quantizador. No decodificador esta normalização é removida pela equação (4.4) que define a quantização inversa.

$$R_{vu} = Sq_{vu} \cdot Q_{vu} \quad (4.4)$$

A figura (4.48) mostra as tabelas de Quantização para os 3 tipos de blocos segmentados. Estas tabelas são projetadas em função de fatores de ponderação psicovisuais aplicado ao S.V.H. para a componente de Luminância. No caso da segmentação 8x8, usa-se a tabela determinada pela norma 601 da CCIR para imagem [38], para as outras segmentações (4x4 e 16x16) as tabelas são determinadas por interpolação linear da tabela determinada pela norma.

Em um processador de vídeo estas tabelas de quantização são carregadas pelo processador hóspede ("host processor") no começo da operação de compressão ou descompressão para fornecer o valor individual de quantização para cada coeficiente transformado pelo tamanho do bloco respectivo (256, 64 ou 16), que em nosso caso encontram-se incorporado nos programas `compress.for` e `dequant.for`. Na saída para a distribuição de bits por bloco, o coeficiente DC de qualquer tamanho de bloco começa com o maior número de bits significativos (7 bits) logo começa a decrescer a quantidade usada de acordo ao K-ésimo coeficiente colhido na varredura em zig-zag.

Um resumo do processamento a partir da seleção de blocos por árvores de quarto grau, envolvendo os coeficientes transformados, sua seleção adaptativa fixa baseada no K-ésimo coeficiente de acordo com o Desvio Padrão e sua quantização terminal frisando

TABELAS DE QUANTIZAÇÃO PARA OS DIFERENTES BLOCOS.

16 x 16: Componente Y.															
32	27	22	21	20	26	32	40	48	60	80	91	102	112	122	132
28	28	28	26	24	30	35	43	50	74	98	105	111	114	117	120
24	24	24	26	28	33	38	46	53	85	117	119	120	116	111	107
26	26	25	27	30	37	44	55	66	89	112	121	129	121	112	104
28	27	26	29	32	40	48	64	80	97	114	126	138	125	112	99
28	29	30	34	38	46	53	72	91	118	144	147	149	134	118	103
28	31	34	39	44	51	58	80	102	138	174	167	160	142	124	106
32	37	42	51	59	73	87	103	119	156	196	190	183	161	139	117
36	40	44	59	74	95	116	126	136	177	218	212	206	180	154	128
42	49	56	75	93	108	122	135	148	181	213	215	216	193	169	146
48	59	70	91	111	120	128	145	162	185	208	217	226	205	184	163
73	86	99	117	134	143	151	168	184	205	225	230	234	214	193	173
98	113	128	142	156	165	174	190	206	224	242	241	240	221	202	183
122	139	156	165	173	179	155	201	217	219	221	222	223	212	200	189
144	164	184	187	190	193	196	210	224	220	216	211	206	202	198	194
167	190	212	210	207	207	207	220	233	206	179	184	189	193	196	199

8 x 8: Componente Y								4 x 4: Componente Y			
16	11	10	16	24	40	51	61	8	5	12	26
12	12	14	19	26	58	60	55	7	8	20	35
14	13	16	24	40	57	69	56	9	19	34	57
14	17	22	29	51	87	80	62	25	39	54	60
18	22	37	58	68	109	103	77				
24	35	55	64	81	104	113	92				
49	64	78	87	103	121	120	101				
72	92	95	98	112	100	103	99				

Figura 4.48: Tabelas de quantização baseadas em funções de ponderação projetadas para o Sistema Visual Humano.(Psychovisual Weighting Functions).

as avaliações finais que requerem um intenso uso de processamento computacional, é apresentada na figura (4.49). Por facilidade gráfica utiliza-se o tamanho de bloco de 4x4.

4.5.6 Plano de distribuição de bits e sequencia de varredura em zig-zag.

O processo de alocação de bits depende do critério de erro usado e das funções de densidade de probabilidade dos coeficientes transformados. Isto envolve a alocação de mais bits aos coeficientes escalares com variâncias maiores (espectro de baixa frequência) e poucos bits aos escalares com variâncias menores (espectro de alta frequência dentro do bloco). O algoritmo de referencia neste problema foi desenvolvido por Huang-Schultheiss para distribuição Gaussiana e posteriormente modificado por Kamangar-Rao para distribuição Laplaciana e codificação híbrida, referencias [7], [22], empregando o critério de "MMSE" (mínimo erro quadrático medido).

ALOCAÇÃO DE BITS

$$B_i = \frac{B}{N} + \frac{1}{2} \log_2 \left[\frac{\sigma_i^2}{\left(\prod_{i=1}^N \sigma_i^2 \right)^{\frac{1}{N}}} \right] \quad (4.5)$$

donde, $1 \leq i \leq N$

B_i = N_0 de bits por coeficiente.

B = N_0 total de bits requeridos para os coeficientes do bloco.

N = N_0 total de coeficientes significativos por bloco.

σ^2 = Variância do i -ésimo coeficiente escalar.

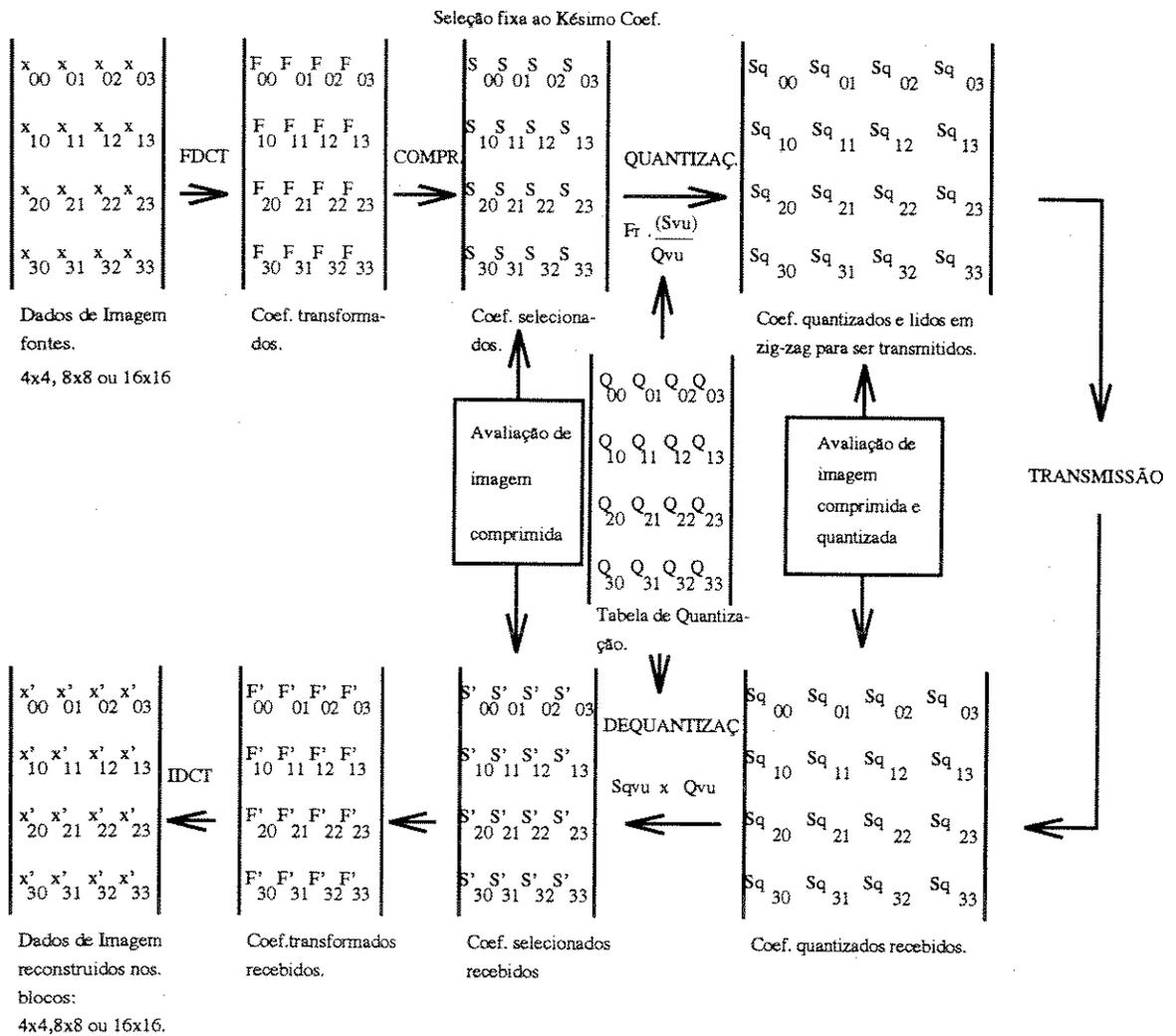


Figura 4.49: Resumo do processamento a nível de blocos segmentados e suas avaliações finais que requerem um intenso uso de facilidades computacionais considerando e não considerando o teste (1.).

Os coeficientes nos tamanhos de blocos respectivos são lidos por uma varredura na forma zig-zag (secção 3.16.3) formando um esquema que reordena os coeficientes em uma sequência unidimensional dependente das Variâncias (Método de Parsons e Tescher) [7]. Esta metodologia supõe uma diminuição contínua da variância dos coeficientes e apenas modificada pela magnitude quadrática do último coeficiente quantizado e codificado dentro do tamanho do bloco respectivo. Além disso, vai existir um ordenamento com respeito às frequências crescentes dentro do bloco respectivo (esquerda à direita e de acima para baixo).

Em resumo, os coeficientes dos diferentes tamanhos de bloco são recolhidos pela varredura em zig-zag, figura (4.50), mostrada por simplicidade para o bloco de 4x4. Eles vão ter sua quantização e alocação de bits feitas por estimação das variâncias e MMSE segundo esta equação (4.5) e usando três mapas de alocação de bits para os tamanhos de blocos respectivos, representados na figura (4.51), com quantização fina (7 bits) para os coeficientes transformados mais significativos de baixa frequência que representam a parte substancial da imagem, e com quantização grossa para os coeficientes menos significativos de cada bloco. Desta maneira consegue-se a precisão de codificação apresentada na tabela das avaliações finais (4.8) que determina uma codificação eficiente para a excelente qualidade de imagem obtida se comparada com a tabela de compressão e qualidade de imagem usada pela JPEG tabela (4.10). Além disso, este ordenamento ajuda a facilitar a codificação entrópica que vem posteriormente para acrescentar o poder de compressão como um sistema e que vai depender dos requerimentos do organismo a cargo de implementar este algoritmo empregado como parte de um futuro sistema para compressão e transmissão de imagem. De fato esta codificação entrópica pode se realizar de duas maneiras: como codificação Huffman ou como codificação Aritmética (como no caso apresentado a nível de proposições na norma JPEG). Em todo caso, a codificação entrópica é considerada como um processo de dois passos. O primeiro passo é converter a sequência em zig-zag dos coeficientes quantizados (nos quais existirão usualmente poucos coeficientes não zeros comparados com a existência dos coeficientes zeros), em uma sequência unidimensional intermediária de símbolos. O segundo passo é transformar os símbolos a um fluxo de dados nos quais não têm limites externos identificáveis mediante a alocação de um código de comprimento variável dependente da probabilidade de ocorrência dos símbolos distribuídos. Quanto menor a ocorrência, maior o comprimento da palavra digital, e quanto maior a ocorrência menor o comprimento da palavra. Desta maneira consegue-se ainda uma maior eficiência de codificação.

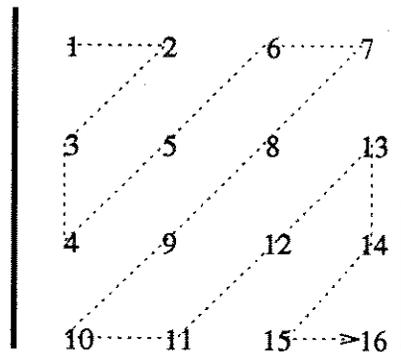


Figura 4.50: Esquema de varredura em zig-zag para um bloco de 4x4.(Método de Parsons-Tescher)

4.5.7 Avaliações Finais do Algoritmo Empregado na Recuperação das imagens de Teste

4.5.7.1 Introdução

As avaliações feitas foram de dois tipos: objetivas-determinísticas e subjetivas.

Para as objetivas-determinísticas utilizam-se as seguintes ferramentas:

ENTROPIA DE PRIMEIRA ORDEM ACUMULATIVA[7].

$$H(x) = \log_2 Q + \frac{1}{Q} \sum_{i=1}^L \frac{Q_i}{L} \log_2 \frac{L}{Q_i} \quad (4.6)$$

Q = N₀ de elementos (512 × 512)

L = N₀ de níveis (256)

Comentário

A equação (4.6) já foi apresentada na secção 3.15.2 e representa uma contribuição válida para avaliar imagens naturais ou reais bem comportadas e cujas características seguem de perto o comportamento de um processo Markoviano de primeiro ordem. No caso

7	6	5	5	4	4	4	3	3	3	2	2	2	1
6	5	5	4	4	4	3	3	3	2	2	1	1	
5	5	4	4	4	3	3	3	2	2	1	1		
5	4	4	4	3	3	3	2	2	1	1			
4	4	4	3	3	3	2	2	1	1				
4	4	3	3	3	2	2	1	1					
4	3	3	3	2	2	1	1						
3	3	3	2	2	1	1							
3	3	2	2	1	1								
3	2	2	1	1									
2	2	1	1										
2	1	1											
2	1												
1													

$16 \times 16 = 105$ Coeficientes = 282 bits

7	6	5	4	3	3	2	2
6	5	4	3	3	2	1	
5	4	3	3	2	1		
4	3	3	2	1			
3	3	2	1				
3	2	1					
2	1						
2							

7	6	5	5
6	5	4	4
5	4	3	2
5	4	2	1

$4 \times 4 = 16$ Coeficientes = 68 bits

$8 \times 8 = 36$ Coeficientes = 107 bits

Figura 4.51: Plano de alocação de bits para os diferentes tamanhos de blocos.

de imagens com zonas de textura ruidosa estável (secção 4.4.3), na recuperação apresentam-se com um erro cumulativo que do ponto de vista subjetivo não é perceptível pelo S.V.H. devido a mascaramento espacial que sofrem as variações rápidas transitórias, sobretudo em níveis de luminância altos e em transições coloridas [8].

No caso de imagens bem comportadas, uma entropia recuperada inferior à entropia original significa claramente, entre outras coisas, perda em altas frequências com degradação nos detalhes finos, e uma entropia recuperada maior pode significar um deslocamento da energia do sinal para frequências mais altas, resultando em desagradável distorção de côr e principalmente de luminância. Em geral uma entropia recuperada maior resulta em espalhamento do sinal com a conseqüente degradação nas altas frequência devido ao processo de pós-filtragem na recuperação do sinal[1].

RELAÇÃO SINAL RUÍDO[1].

$$SNR = 10 \log \frac{\bar{S}^2}{\bar{N}^2} \quad (4.7)$$

Onde,

$$\bar{S}^2 = \frac{1}{N_L} \sum_{i=1}^{N_L} \sum_{j=1}^{N_a} \frac{S_{i,j}^2}{N_a}$$

$$\bar{N}^2 = \frac{1}{N_L} \sum_{i=1}^{N_L} \sum_{j=1}^{N_a} \frac{(\hat{S}_{i,j} - S_{i,j})^2}{N_a}$$

\bar{S}^2 = Sinal quadrática média

\bar{N}^2 = Ruído quadrático médio

$(\hat{S}_{i,j} - S_{i,j})$ = Ruído

N_L = N^o de linhas

N_a = N^o de amostras por linhas

$S_{i,j}$ = i-ésima amostra de referência da j-ésima linha.

$\hat{S}_{i,j}$ = i-ésima amostra de recuperada da j-ésima linha.

Comentário.

A medida da relação Sinal-Ruído média não é suficiente para garantir um bom desempenho do sistema, pois o sinal pode ser recuperado com baixa relação Sinal-Ruído e apresentar degradação visual imperceptível para o S.V.H. ou alta Sinal-Ruído e apresentar degradação de faixa que se caracteriza entre outras coisas, por interferência espectral e principalmente por perdas na qualidade do sinal recuperado em detalhes finos da imagem. Contudo, para imagens bem comportadas, é uma ferramenta que tem sido usada por muitos pesquisadores para avaliar imagens reais. Porém em zonas de textura ruidosa estável é afetada do mesmo modo que a Entropia Acumulada para as variações rápidas transitórias.

HISTOGRAMA

$$n = f(L) \implies P(L) = \frac{n}{Q} \quad (4.8)$$

n = frequência de cada nível

L = Valor de Luminância dos E.I.

Q = quantidade de amostras

AVALIAÇÃO SUBJETIVA

A avaliação subjetiva depende da comparação da imagem original com a recuperada na tela da estação de trabalho SUN usando um "display" de 8 bits para as imagens preto e branco e uma paleta de 24 bits para as imagens coloridas. Procurou-se observar apenas as diferenças entre as imagens originais e recuperadas devido à falta de resolução da tela, classificando-as "in-situ" pelo autor numa tabela de 5 graus [8].

Comentário.

A classificação da visualização subjetiva da imagem fica muito difícil, pois necessita de uma quantidade mínima de 20 observadores comuns para determinar a Qualificação de Parecer Médio ("Mean Opinion Score"), ou de observadores altamente especializados e reconhecidos em avaliar imagens. Além disso, também é necessário uma quantidade de equipamento de alto custo inacessíveis, para poder diagnosticar com uma certa confiabilidade qualquer problema relacionado com a perda de resolução e interferência espectral (contudo, no caso de especialistas, eles observam com maior sensibilidade as imperfeições de imagens e podem não ser representativos de observadores comuns[8]). Devido à falta de equipamento adequado e com a pouca disponibilidade de tempo de uso, se fez uma tabela

de classificação de 5 graus da imagem recuperada comparada com a original na mesma tela da estação colorida de trabalho SUN. Deve ser ressaltado que a caracterização resultante será subjetiva do observador e naturalmente não pretende ser única.

E finalmente utiliza-se a avaliação subjetiva da comparação do sinal Mutisalva recuperado com a original, tomando as restrições mencionadas para este tipo do sinal na secção 4.5.5.

Além das avaliações mencionadas, comparam-se as tabelas resultantes da utilização do algoritmo empregado, quanto ao efeito da segmentação por árvore de quarto grau usando o controle por parâmetro estatístico com o resultado na precisão da codificação e com consequência na eficiência de codificação.

4.5.7.2 Figuras e Tabelas resultantes do Algoritmo Empregado

As seguintes figuras apresentam a comparação da Distribuição de blocos com o controle estatístico, num caso considerando o teste 1. do Desvio Padrão e em outro caso apenas considerando o teste da Média, ou seja sem o Desvio Padrão para as quatro imagens de teste. Através delas observa-se facilmente por comparação a quantidade de blocos distribuídos por tamanho com o controle positivo do Desvio Padrão, que leva ao uso de blocos maiores, e em consequência a uma melhor eficiência na codificação.

Das figuras mencionadas e tabelas que se seguem pode-se observar que a segmentação em blocos de tamanho variável com o primeiro teste de decisão referente ao Desvio Padrão mostra que não difere muito para imagens compostas de fundo e transições suaves no caso de decisão positiva ou negativa, como por exemplo o caso de Lena, ver figura (4.54) e tabela (4.7).

Não obstante, no caso de uma imagem de muito transitórios, como no Babuíno, onde sem a condição do Desvio Padrão a área do pêlo do animal é equivalente a uma textura de ruído estável (secção 4.4.3) e é processada em blocos de 4x4 (para $t_kx = 8$), em contraste do caso de considerar o critério do teste do Desvio Padrão e a Média, onde essa mesma área é processado em blocos de 16x16. Ainda que a qualidade da imagem é muito alta como resultado de processar a imagem por codificação DCT-2D sem considerar o teste estatístico, também isto fornece como resultado uma precisão de codificação alta,

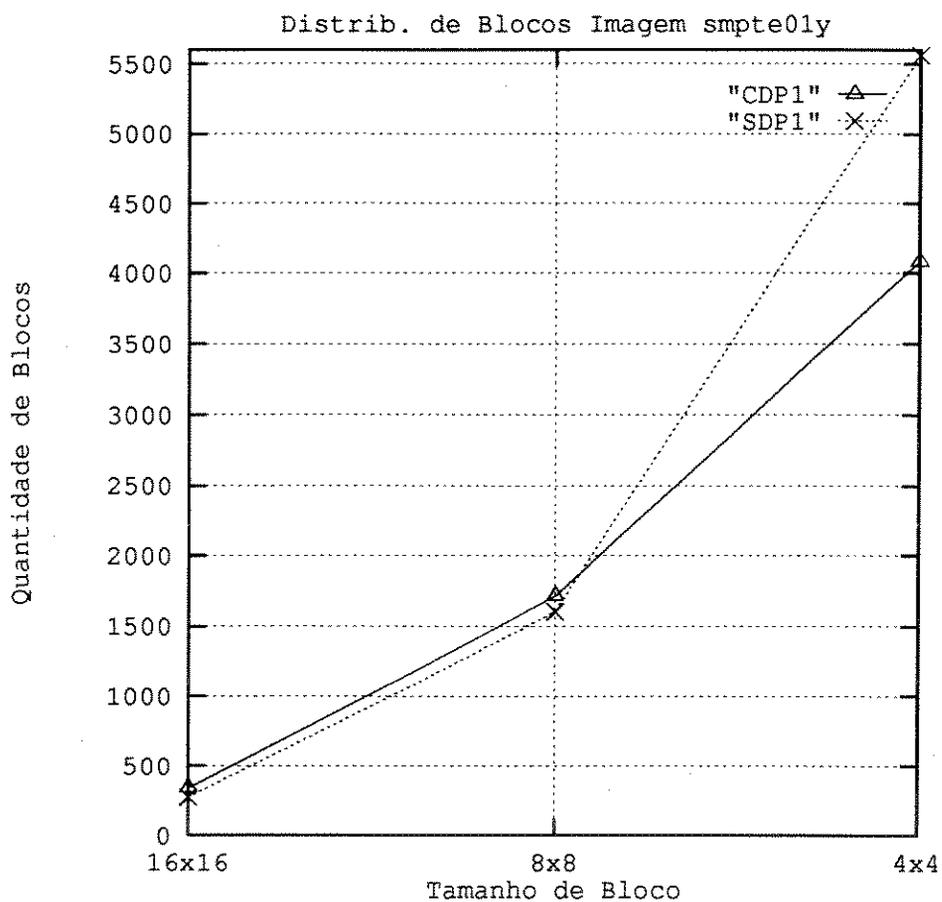


Figura 4.52: Distribuição de Blocos por Controle Estatístico para imagem smpte01y.

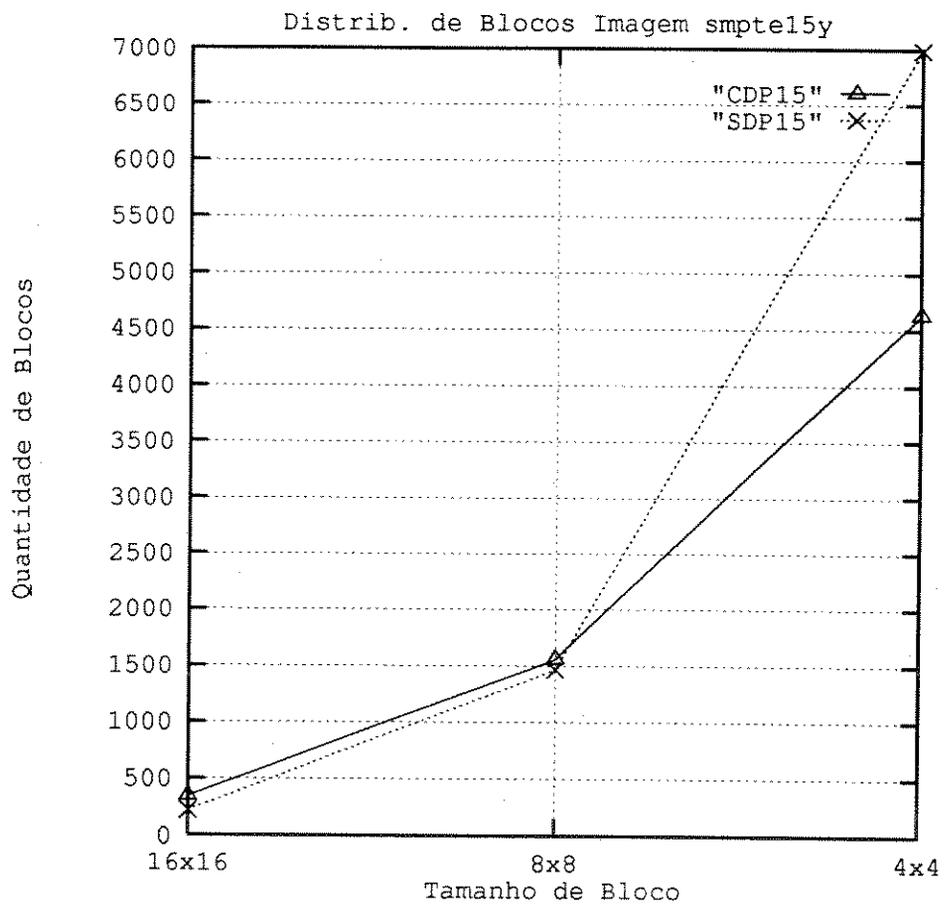


Figura 4.53: Distribuição de Blocos por Controle Estatístico para imagem smpte15y.

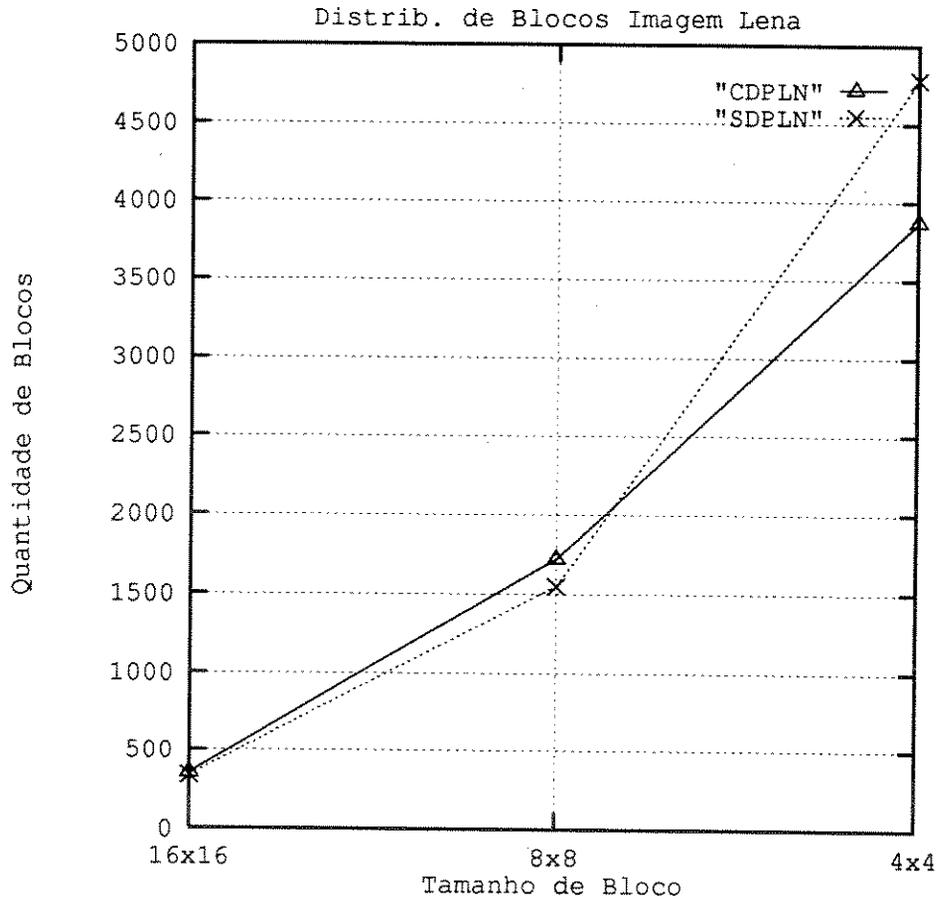


Figura 4.54: Distribuição de Blocos por Controle Estatístico para imagem Lena.

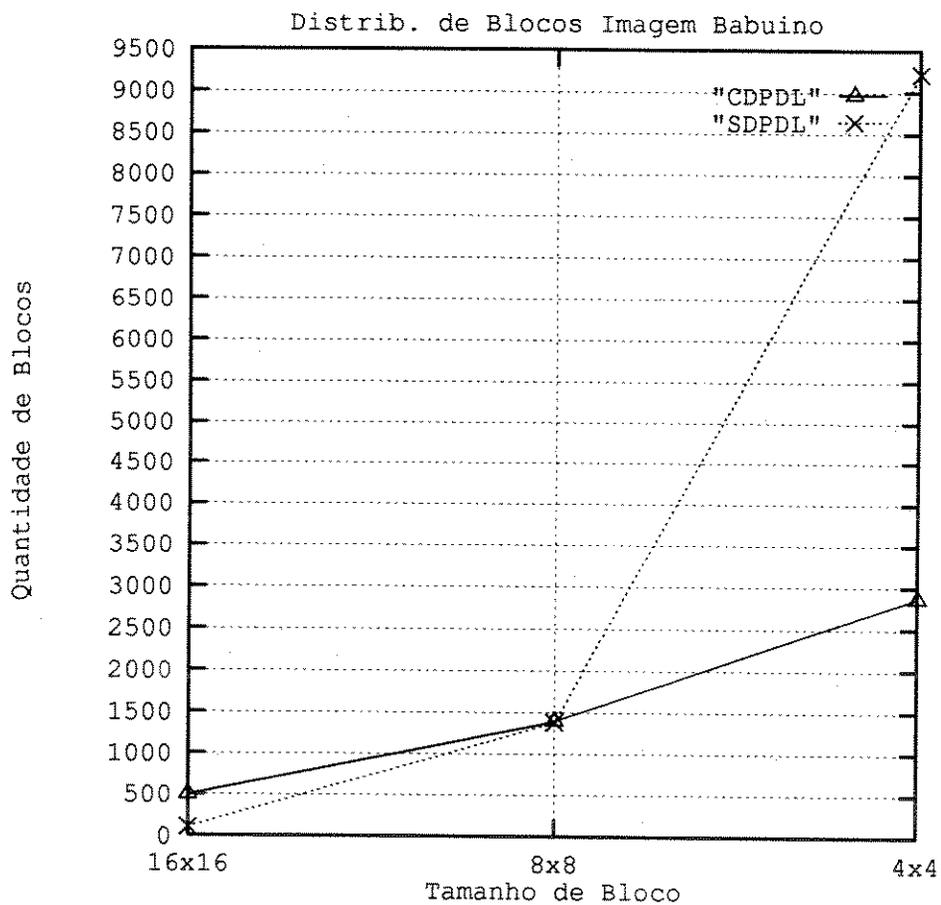


Figura 4.55: Distribuição de Blocos por Controle Estatístico para imagem Babuíno.

como mostra a tabela (4.8). Como se mencionou previamente, nessas áreas de alta variação estável o S.V.H. não é sensível à degradação, logo um tamanho de bloco o maior possível seria o mais desejado. Dessa maneira uma precisão de codificação menor pode ser obtida sem introduzir degradações apreciáveis no uso deste algoritmo empregado e a consequente eficiência de codificação final.

Teste	smpte 01Y		smpte 15Y		Lena		Babuíno	
	CDP	SDP	CDP	SDP	CDP	SDP	CDP	SDP
16x16	340	275	345	221	352	338	496	103
8x8	1715	1605	1558	1464	1720	1548	1395	1379
4x4	4084	5564	4632	6992	3872	4784	2868	9220

Tabela 4.7: Distribuição de blocos por imagem de teste. Onde, CDP = com Desvio Padrão e SDP= sem Desvio Padrão.

A precisão da codificação correspondente às imagens de teste está na tabela (4.8). As imagens reconstruídas e obtidas dos mesmos parâmetros excluindo o teste (1.) do Desvio Padrão na segmentação do bloco, são visualmente as mesmas consideradas no teste (1.). A correspondente precisão de codificação considerando o teste (1.) e vice-versa, está na tabela prévia para comparação. É preciso notar que imagens de fundos de variação suave como no caso de Lena, a precisão de codificação não tem grande variação. Esta situação tem grande diferença na precisão da codificação da imagem de teste Babuíno, onde resulta um número bem menor de blocos 4x4.

	smpte01y	smpte15y	Lena	Babuíno
CDP	2.1486	2.2335	2.1078	1.8650
SDP	2.4226	2.6821	2.2619	3.1062

Tabela 4.8: Precisão de codificação. (Unidade: bits/E.I.)

Comparando a imagem visual reconstruída (4.56) com ou sem o teste (1.) não apresentou diferença, o que leva a concluir que o acréscimo da precisão de codificação por não usar o teste (1.) não é necessário para as quatro imagens escolhidas.



Figura 4.56: Comparação da Imagem Lena, esquerda, apenas com o teste de segmentação usando a Média; Direita, usando o teste 1. do Desvio Padrão empregado nesta dissertação.

A tabela (4.9) mostra o acréscimo da informação de segmentação por árvore que é muito pequeno, pois é multiplicado por um fator de 10^{-2} . Ela também mostra a pequena diferença de considerar o teste (1.), onde o fato de utilizar tamanhos de blocos maiores ajuda a reduzir o número de bits de informação lateral do cabeçalho, importante na eficiência da codificação e na recuperação da informação no receptor.

	smpte01y	smpte15y	Lena	Babuíno
CDP	2.3400	2.4900	2.2700	1.8100
SDP	2.8400	3.3100	2.5500	4.0900

Tabela 4.9: Cabeçalho da árvore incluído na precisão de codificação. (Unidade: $\times 10^{-2}$ bits/E.I.)

Pode-se observar que os resultados da tabela (4.8) apresentam um valor de precisão de codificação perto ou superior do valor de 2 bit/EI, o que significa do ponto de vista subjetivo de qualidade de acordo com a tabela (4.10) da JPEG [38], que este resultado não é perceptível entre a imagem original e a recuperada e é suficiente para a maioria das aplicações usando imagens paradas.

COMPRESSÃO(bits/EI)	QUALIDADE
0.25-0.5	De moderada a boa, suficiente para algumas aplicações.
0.5-0.75	De boa a muito boa, suficiente para muitas aplicações.
0.75-1.5	Excelente, suficiente para a grande parte das aplicações.
1.5-2.0	Não perceptível da original, suficiente para a maioria das aplicações.

Tabela 4.10: Características subjetivas da qualidade em função da compressão usada pela JPEG.

4.5.7.3 Resultados do Ponto de vista de Histogramas, Entropias Acumuladas e as correspondentes Relações Sinal-Ruído.

De acordo com a figura (4.49), serão comparadas objetivamente e deterministicamente os sinais recuperados das imagens de teste prévio ao processamento pelas tabelas de quantização, os sinais finais recuperados considerando as tabelas de quantização com o teste da Média e finalmente os sinais recuperados usando o teste 1(Desvio Padrão) deste algoritmo de compressão estatístico com as imagens de teste originais.

Na figura (4.57) compara-se o histograma do sinal original e recuperado da imagem *smpte01y*, apenas submetida a compressão com seleção adaptativa dos coeficientes com passo fixo de $k = -3$ (16 coeficientes) para blocos de 4×4 , $k = 0$ (36 coeficientes) para blocos de 8×8 e $k = 3$ (105 coeficientes) em função do Desvio Padrão que também segmenta os blocos por árvore de quarto grau da imagem *smpte01y*. Também, mostra-se neste caso a correspondente Relação Sinal-Ruído.

Da figura (4.57) pode-se observar que a seleção adaptativa dos coeficientes para a compressão causa uma mínima diferença de valores de magnitude dos EI da imagem e uma relação Sinal-Ruído bastante alta.

Agora, a Entropia Acumulada é comparada na figura (4.58), onde não dá para perceber um mínimo acréscimo da informação de imagem.

Na continuação apresentam-se estes mesmos gráficos de histogramas e entropias acumuladas para as imagens *smpte15y*, *Lena* e *Babuino* nas figuras (4.59), (4.60), (4.61), (4.62), (4.63) e (4.64) considerando apenas a seleção adaptativa de coeficientes por parâmetros estatísticos.

Como resultado destas figuras pode-se concluir que a seleção adaptativa por parâmetros estatísticos de coeficientes transformados por zonas opera de forma correta para as imagens teste que seguem de perto um comportamento Markoviano de primeira ordem (secção 3.8). Isto comprovou-se na comparação das ferramentas de avaliações: Entropia, Histograma e Relação Sinal-Ruído (SNR), nas imagens recuperadas *smpte01y*, *smpte15y* e *Lena* com respeito às originais onde não é perceptível uma grande diferença. Não obstante no caso de imagens cujo comportamento difere da característica Markoviana, o resultado determinístico da Entropia e do Histograma acusam variações na comparação com a ima-

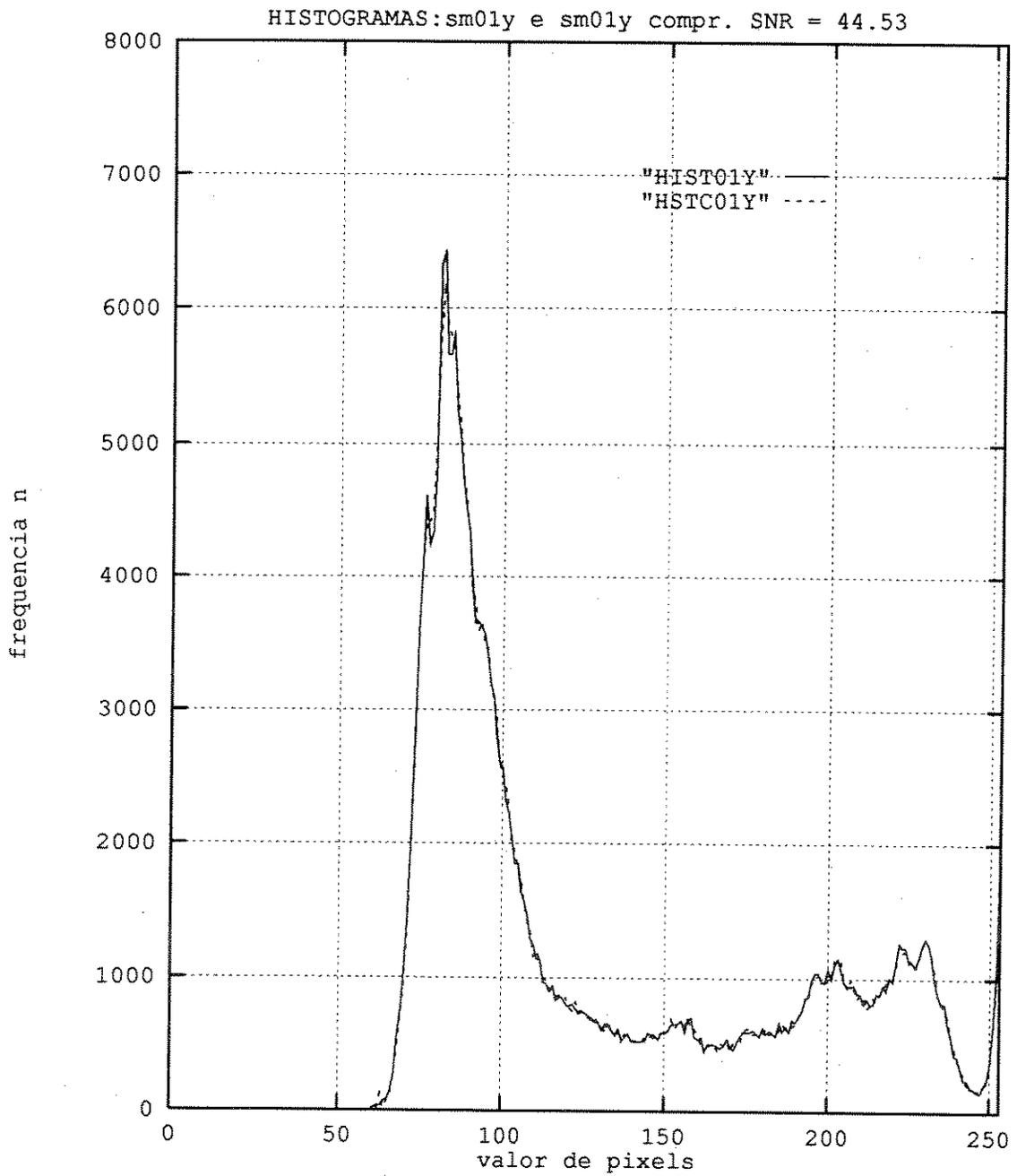


Figura 4.57: Histograma da Imagem recuperada smpt01y, apenas comprimida pelo algoritmo empregado e comparada com a original.

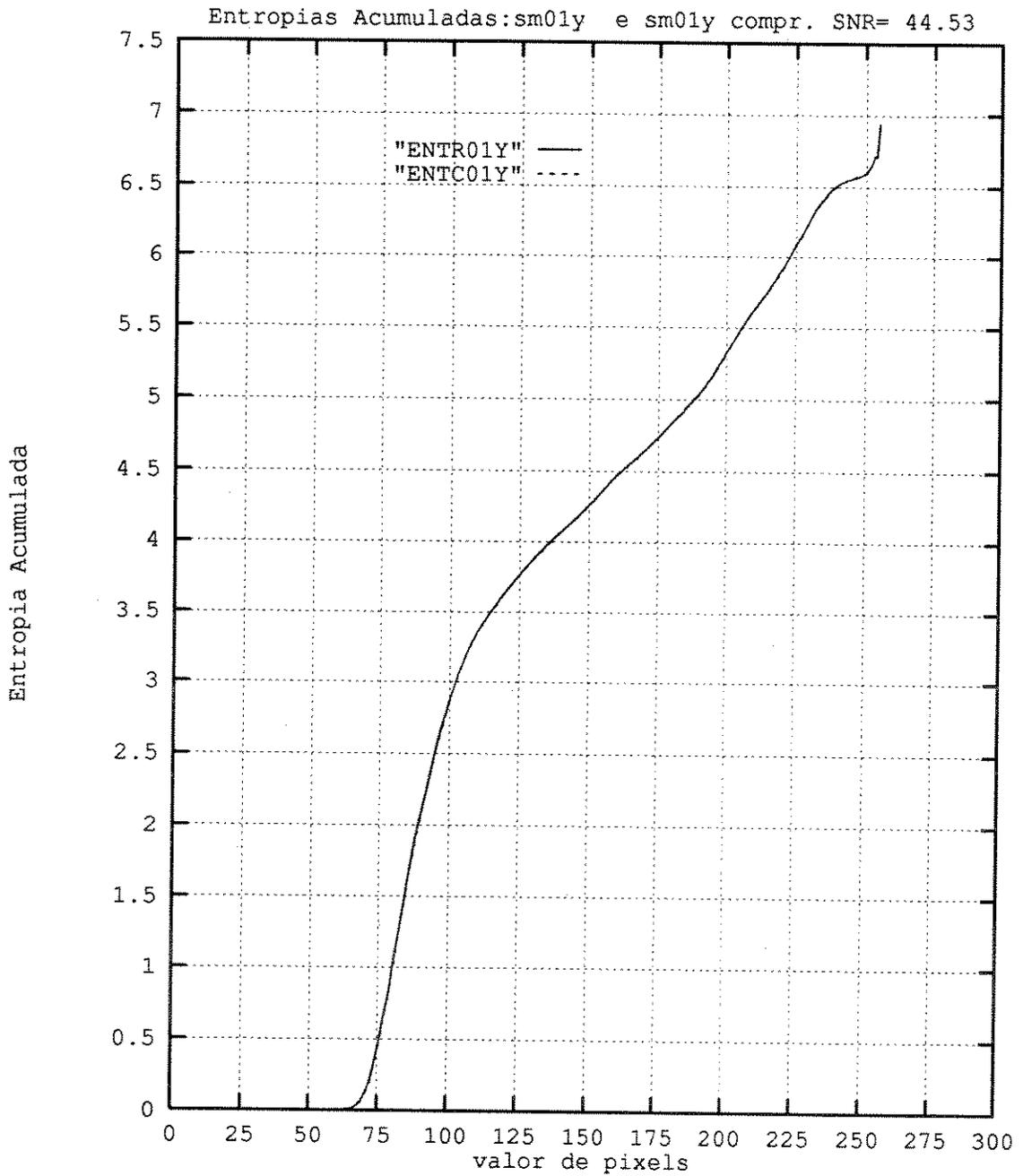


Figura 4.58: Entropia Acumulada da Imagem recuperada smpt01y, apenas comprimida adaptativamente com os parâmetros estatísticos do algoritmo empregado e comparada com a original.

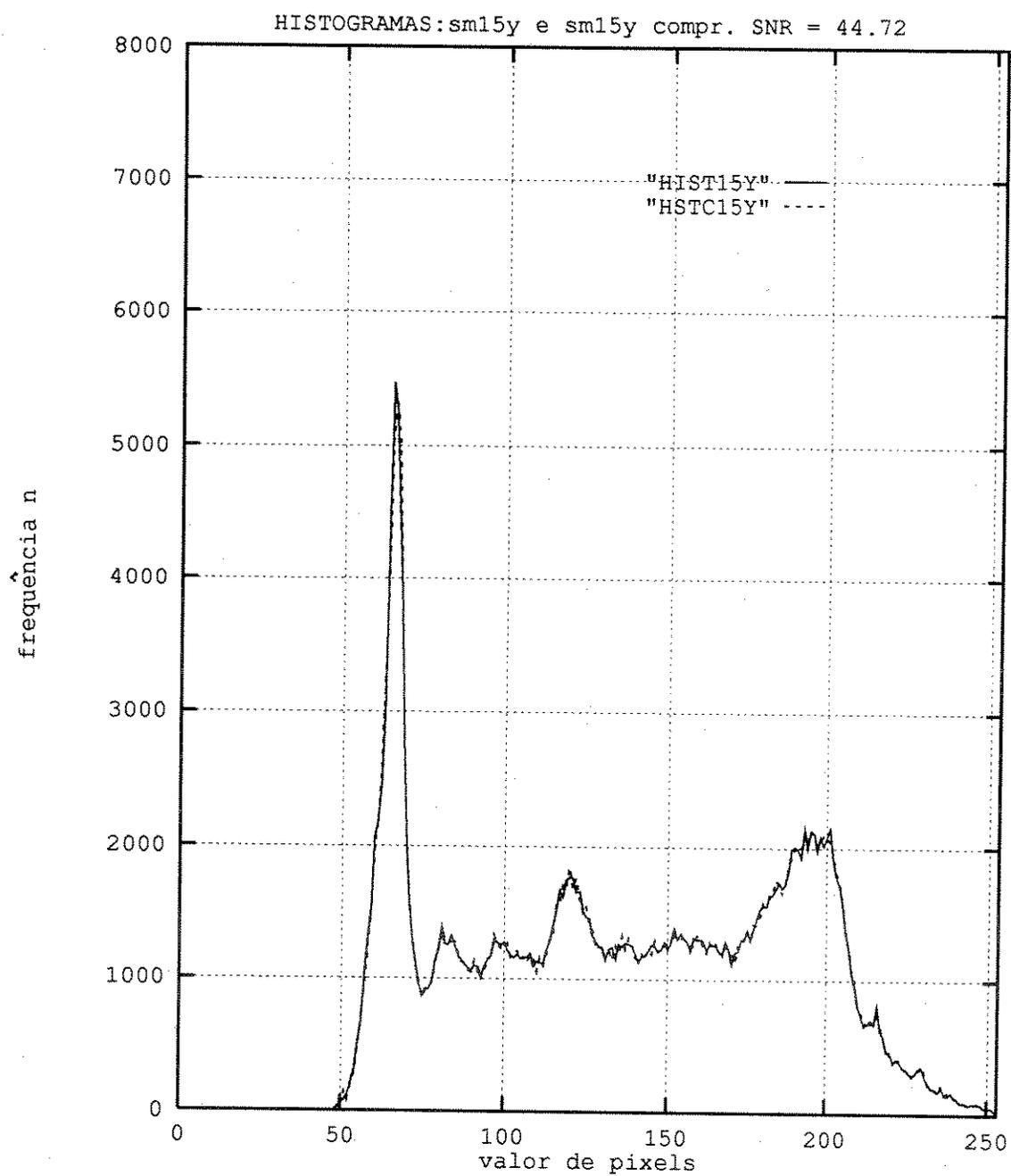


Figura 4.59: Histograma da Imagem recuperada smpt15y, apenas comprimida pelo algoritmo empregado e comparada com a original.

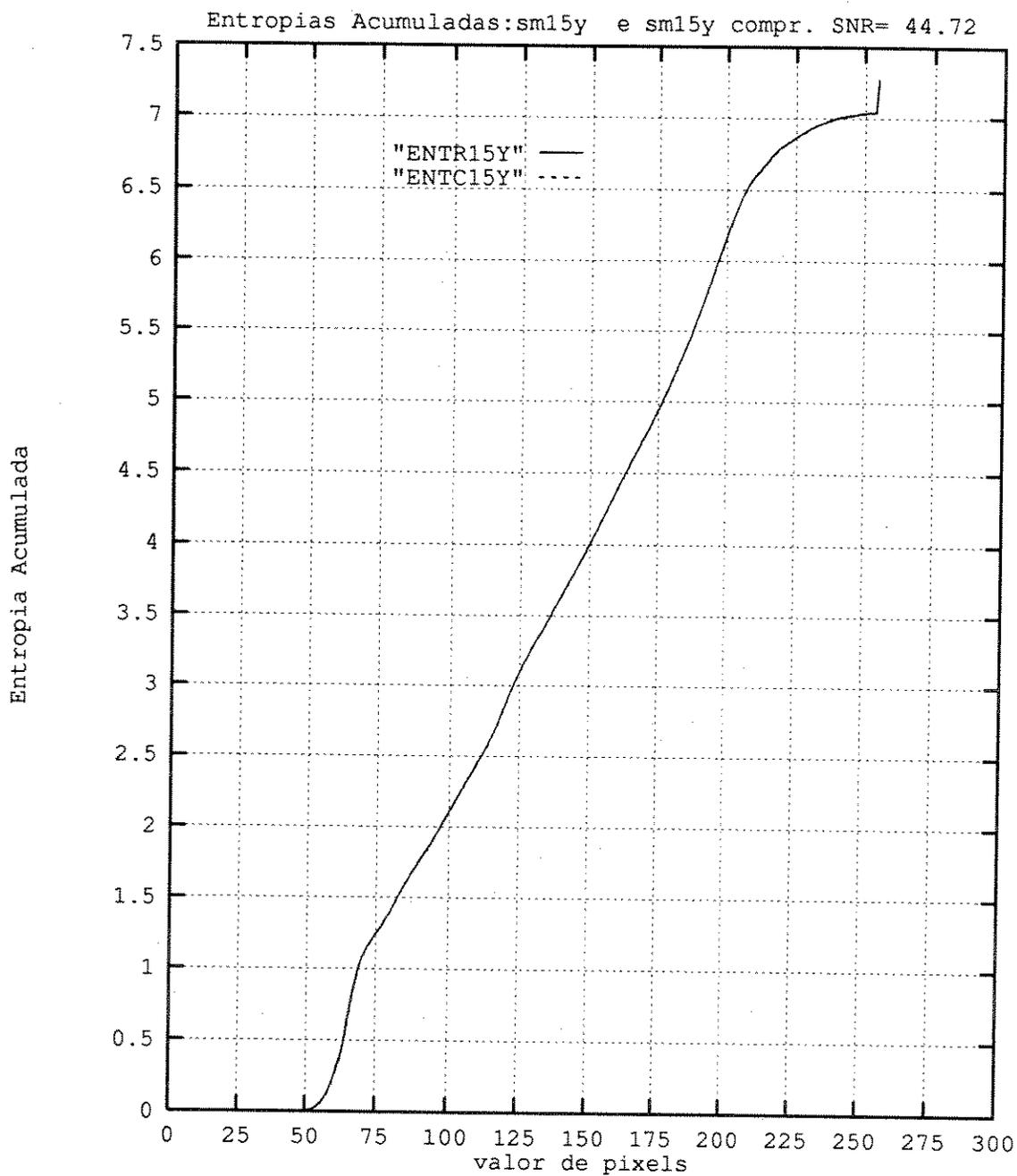


Figura 4.60: Entropia Acumulada da Imagem recuperada smpt15y, apenas comprimida adaptativamente com os parâmetros estatísticos do algoritmo empregado e comparada com a original.

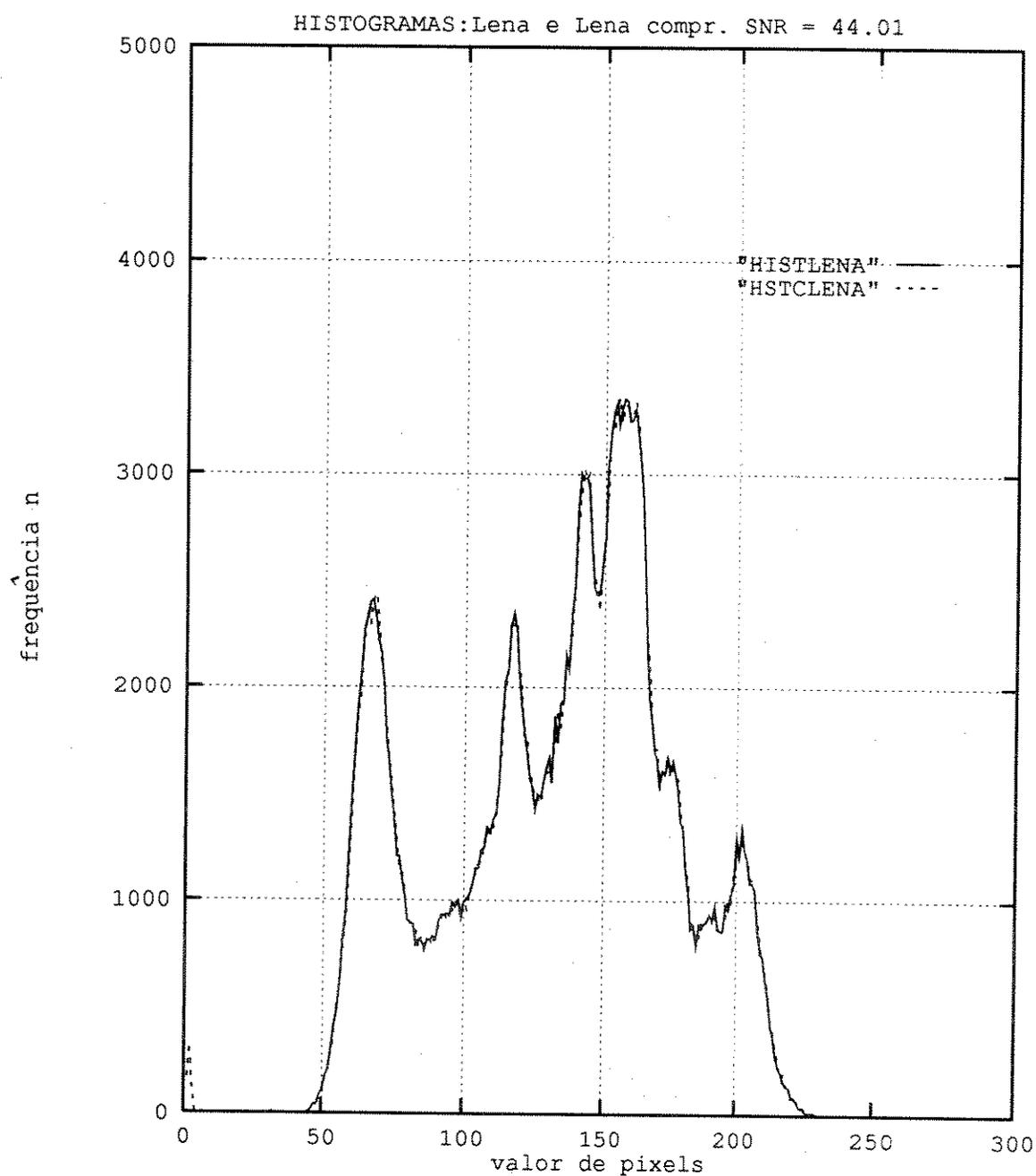


Figura 4.61: Histograma da Imagem recuperada Lena, apenas comprimida pelo algoritmo empregado e comparada com a original.

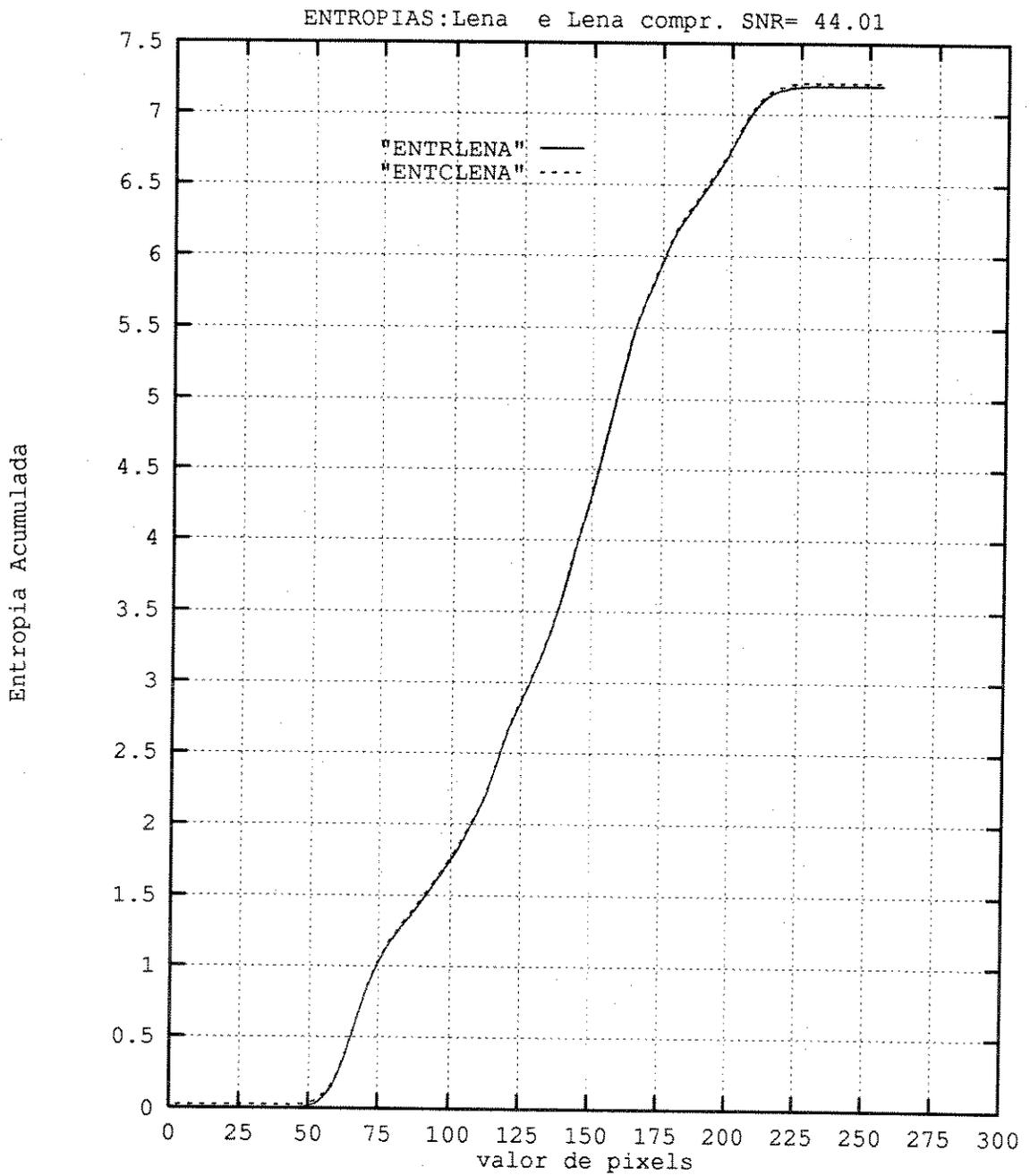


Figura 4.62: Entropia Acumulada da Imagem recuperada Lena, apenas comprimida adaptativamente com os parâmetros estatísticos do algoritmo empregado e comparada com a original.

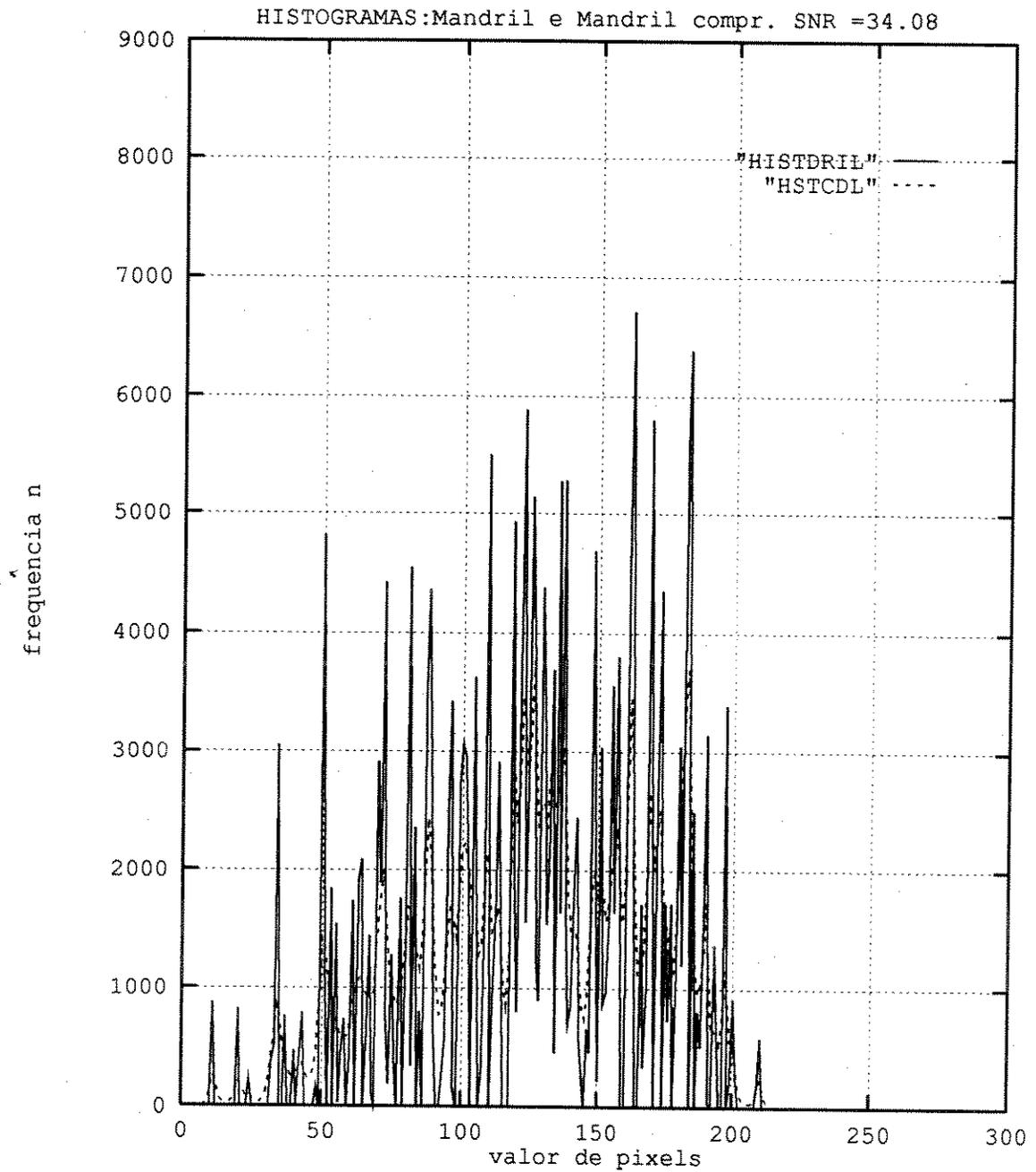


Figura 4.63: Histograma da Imagem recuperada Babuíno, apenas comprimida pelo algoritmo empregado e comparada com a original.

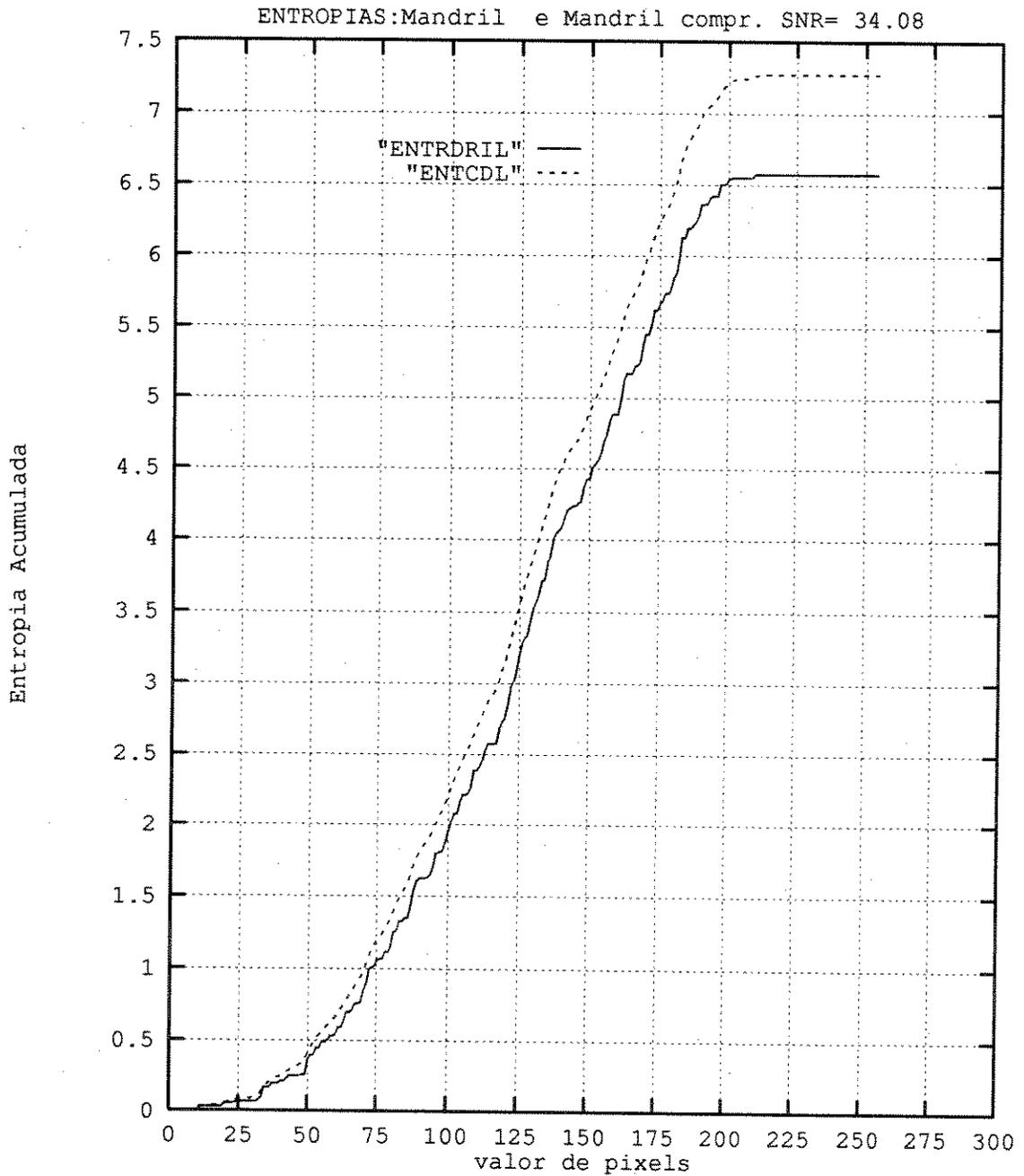


Figura 4.64: Entropia Acumulada da Imagem recuperada Babuíno, apenas comprimida adaptativamente com os parâmetros estatísticos do algoritmo empregado e comparada com a original.

gem original, tal como se aprecia no caso da imagem Babuíno na figura (4.63) cujo conteúdo de imagem é de textura ruidosa estável que não segue o comportamento histórico de um fenômeno de Markov. Contudo o conteúdo mencionado previamente não é de grande importância na visualização da imagem, pois o S.V.H. permite degradação determinística neste tipo de detalhe estável que não é apreciável subjetivamente na recuperação da imagem.

Na continuação se apresentam as mesmas avaliações prévias na recuperação da imagem, depois do quantizador (veja figura (4.49)) usando os parâmetros estatísticos empregado nesta dissertação, e considerando no mesmo gráfico de Histograma e Entropia acumulada a comparação dos dois casos estatísticos: apenas usando o teste (2.) da Média e usando os fatores de decisão completo, onde encontra-se incluído o teste (1.) do Desvio Padrão.

Na figura(4.65) mostra-se que na comparação das entropias acumuladas da imagem *smpte01y* não se percebe grande diferença entre a imagem original e as imagens submetidas a processamento. Não obstante a SNR acusa as diferenças de valores dos E.I. (veja equação(4.7)) caindo apenas de um valor de 44,53 db antes da quantização para um valor de 43,70 db com o teste da Média, e logo caindo para 43,19 db usando o teste empregado do Desvio Padrão. Este último decréscimo é compensado com a menor precisão de codificação, passando de 2,4226 bits/EI para 2.1486 bits/EI.

Na continuação apresentam-se as comparações de entropias acumuladas nas figuras (4.66), (4.67), (4.68) para as restantes imagens de teste, tendo todas um comportamento similar à anterior com exceção da imagem Babuíno, que por não seguir um comportamento Markoviano estatístico apresenta degradação do ponto de vista determinístico. Não obstante, na forma subjetiva como se apreciará nas imagens resultantes, figura (4.74), não se percebe diferença alguma (secção 4.5.7.1) na tela da estação de trabalho.

Na continuação apresentam-se a comparação dos Histogramas nas figuras (4.69), (4.70), (4.71) e (4.72), das imagens recuperadas afetados pelo controle estatístico e comparadas com a original de cada imagem de teste. Nesta característica pode-se observar o comportamento diferente da imagem Babuíno na figura (4.72) comparada com o resto das imagens de teste que possuem um comportamento Markoviano. Neste caso a Relação Sinal-Ruído acusa uma grande degradação depois do quantizador, pois cai de 34.08 db para 26.83 db (secção 4.5.7.1) com o algoritmo empregado do Desvio Padrão devido a que o projeto

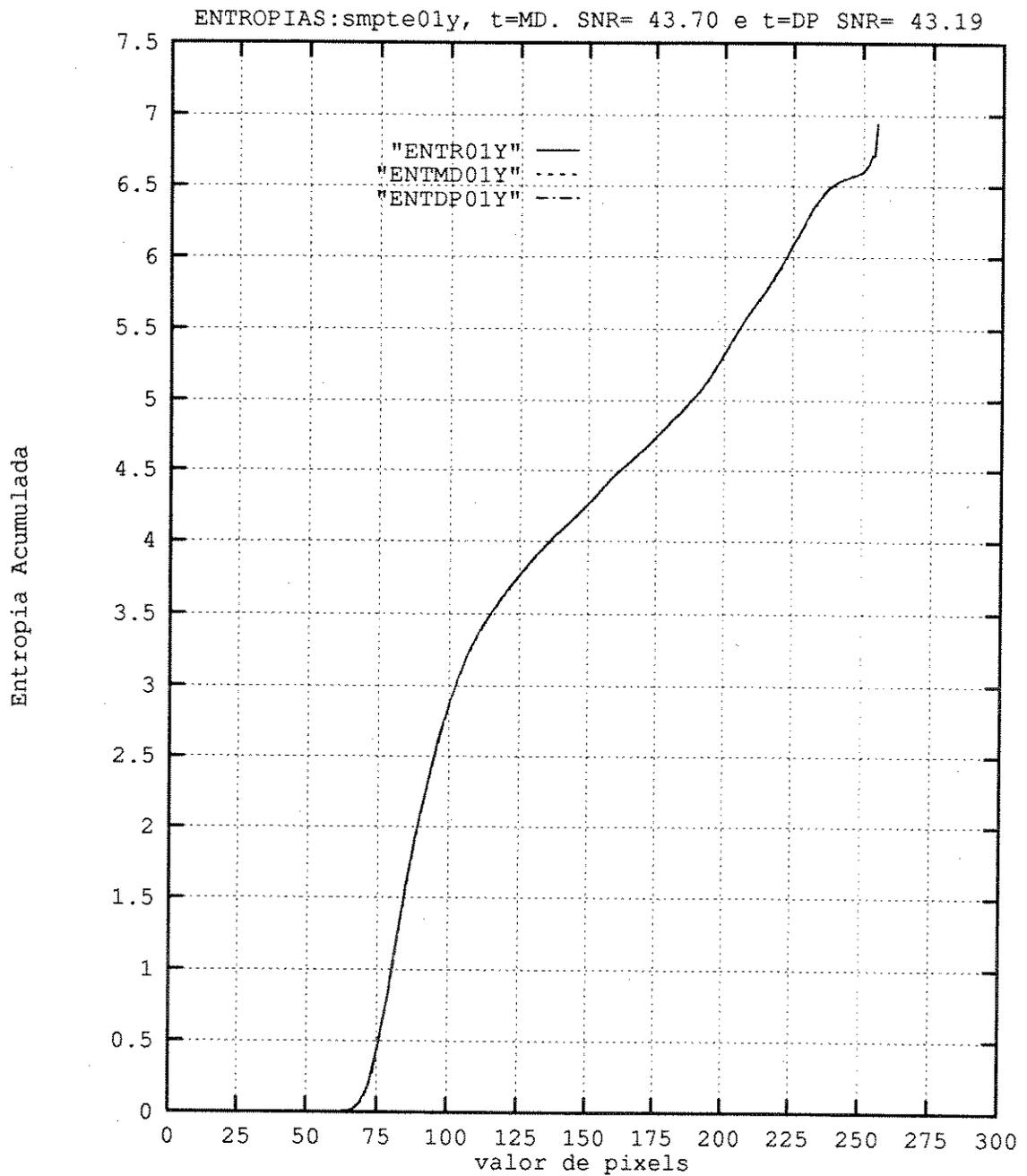


Figura 4.65: Comparação das Entropias acumuladas para a imagem recuperada smpte01y, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.

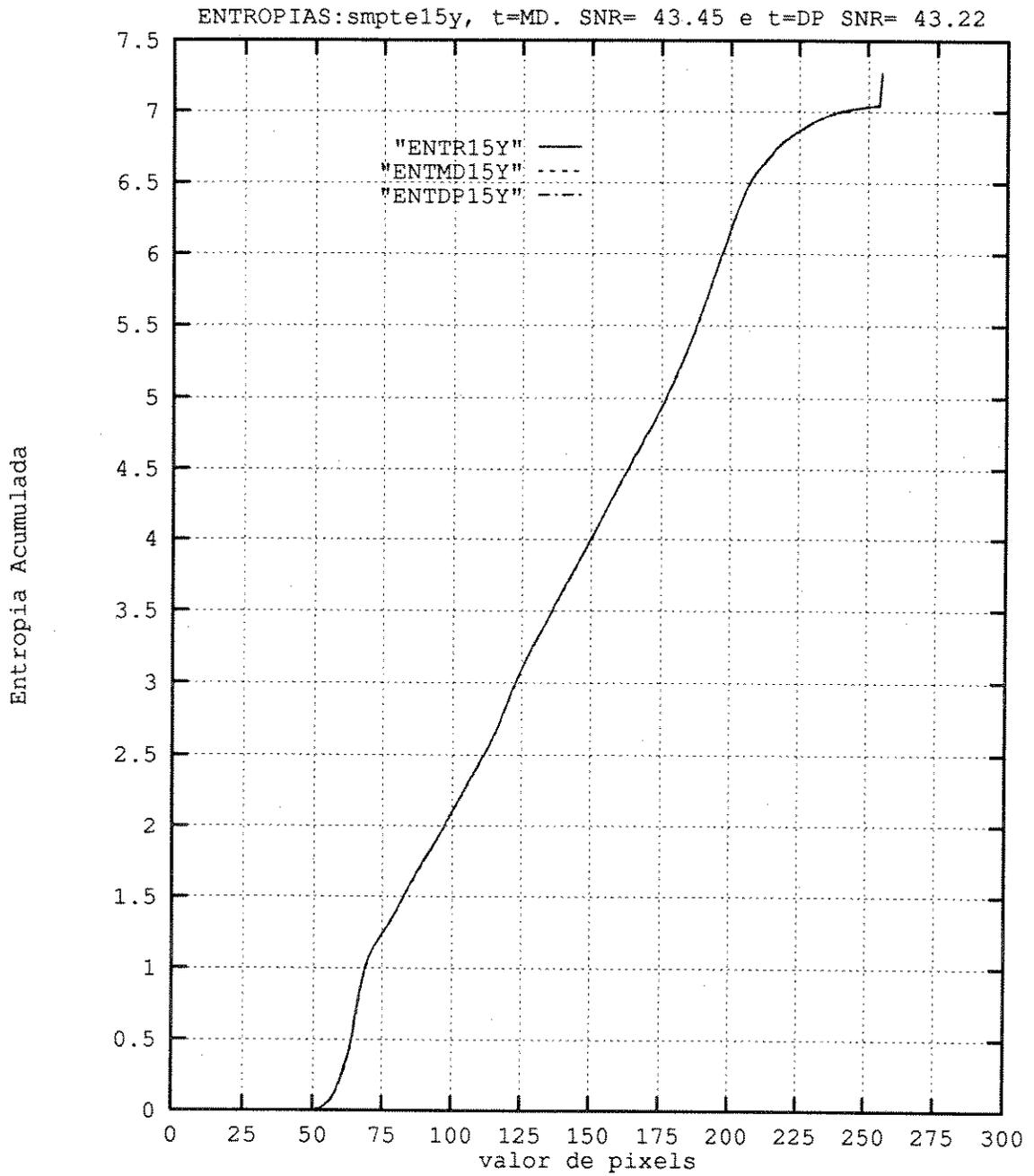


Figura 4.66: Comparação das Entropias acumuladas para a imagem recuperada smpte15y, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.

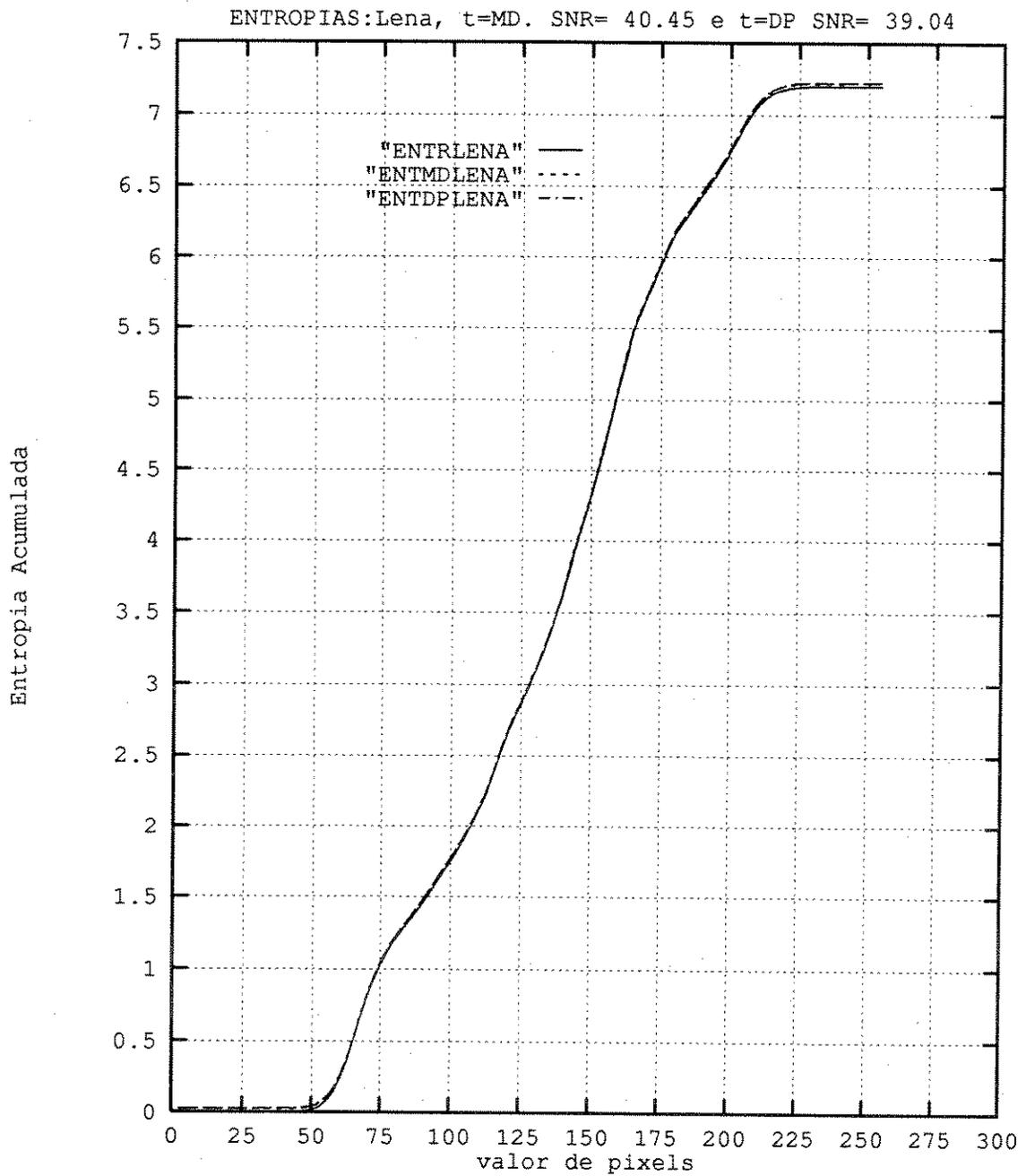


Figura 4.67: Comparação das Entropias acumuladas para a imagem recuperada Lena, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.

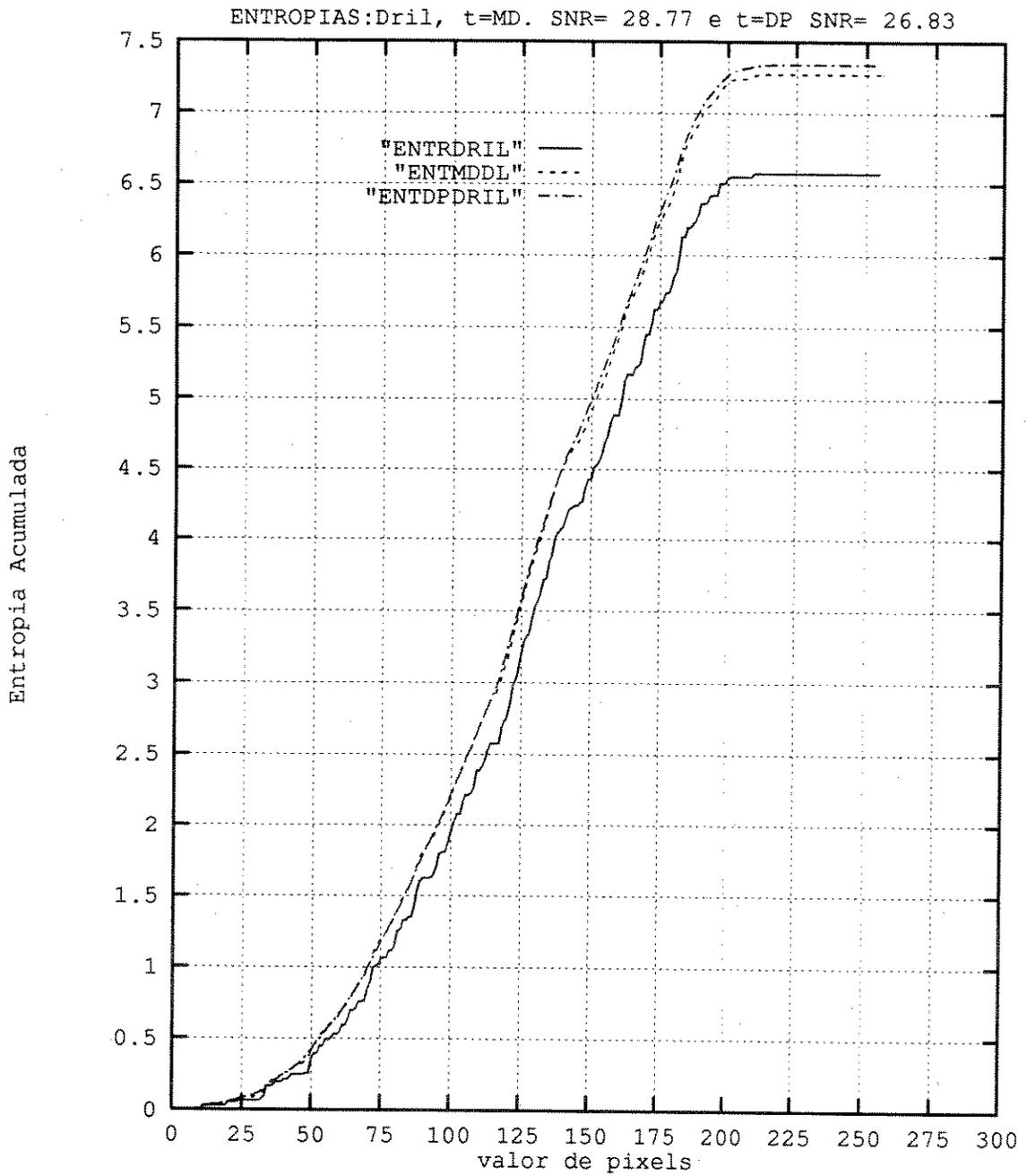


Figura 4.68: Comparação das Entropias acumuladas para a imagem recuperada Babuíno, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.

do quantizador é baseado em pautas psico-visuais da maioria das imagens reais [38]. Não obstante, do ponto de vista subjetivo do S.V.H. não é possível perceber grande diferença, (veja figura (4.74)), na escala de 5 grau da tabela (4.13) a degradação é imperceptível.

4.5.7.4 Resultados do ponto de vista Estatístico e No de Tonalidades Resultantes

Na tabela (4.11), apresentam-se as características estatísticas das quatro imagens de teste, antes e depois do quantizador segundo figura (4.49), submetidas ao controle da Média e do Desvio Padrão. Percebe-se que tanto na Média como no Desvio Padrão, as diferenças não são tão notórias.

IMAGEM DE TESTE	Média	Mediana	Moda	Desv.Pad.	Máx.EI	Mín.EI
SMPTE01Y	135.231	103	255	61.2205	255	53
smc01y	135.348	103	255	61.2026	255	54
smmd01y	135.461	103	255	61.1838	255	45
smdp01y	135.348	103	255	61.1990	255	51
SMPTE15Y	141.579	140	255	56.9956	255	45
smc15y	141.696	140	255	56.9756	255	43
smmd15y	141.811	140	255	56.9986	255	35
smdp15y	141.695	140	255	56.9715	255	45
LENA	133.929	141	157	42.2286	255	0
dcclena	134.193	141	155	42.2107	255	0
dcmlena	134.075	141	157	42.1950	255	0
dcplena	134.198	141	155	42.1884	255	0
BABUÍNO	165.843	168	162	42.3236	210	11
dccdrl	166.091	168	166	42.2487	223	4
dcmddril	166.002	168	166	42.0539	253	0
dcdpdril	166.096	168	166	41.8971	255	0

Tabela 4.11: Características estatísticas das quatro imagens de teste recuperadas antes e depois do quantizador, submetidas ao controle dos parâmetros estatísticos, Média e Desvio Padrão

A tabela (4.12), mostra a comparação do comportamento da Relação Sinal-

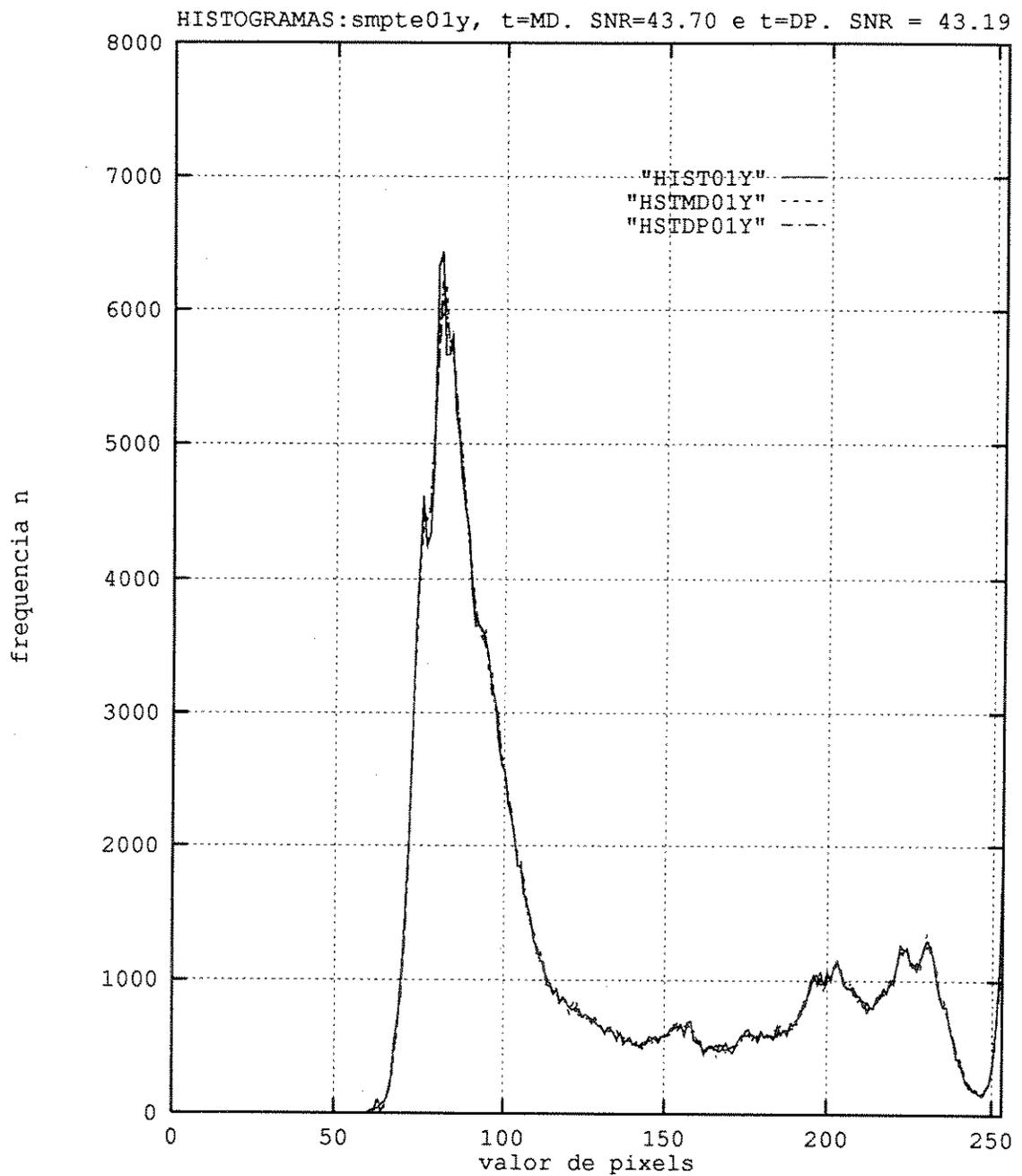


Figura 4.69: Comparação dos Histogramas para a imagem recuperada smpte01, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.

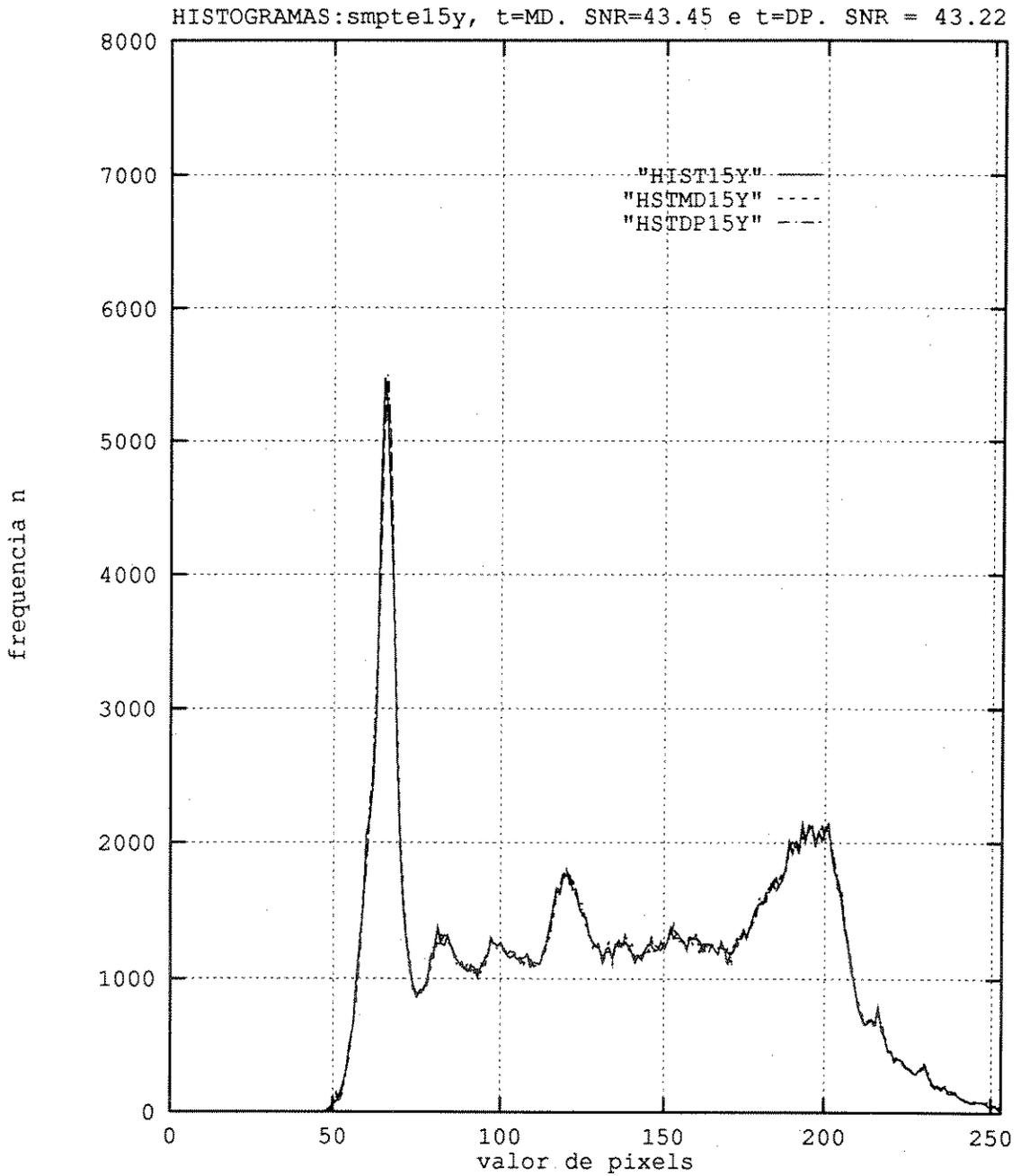


Figura 4.70: Comparação dos Histogramas para a imagem recuperada smpte15, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.

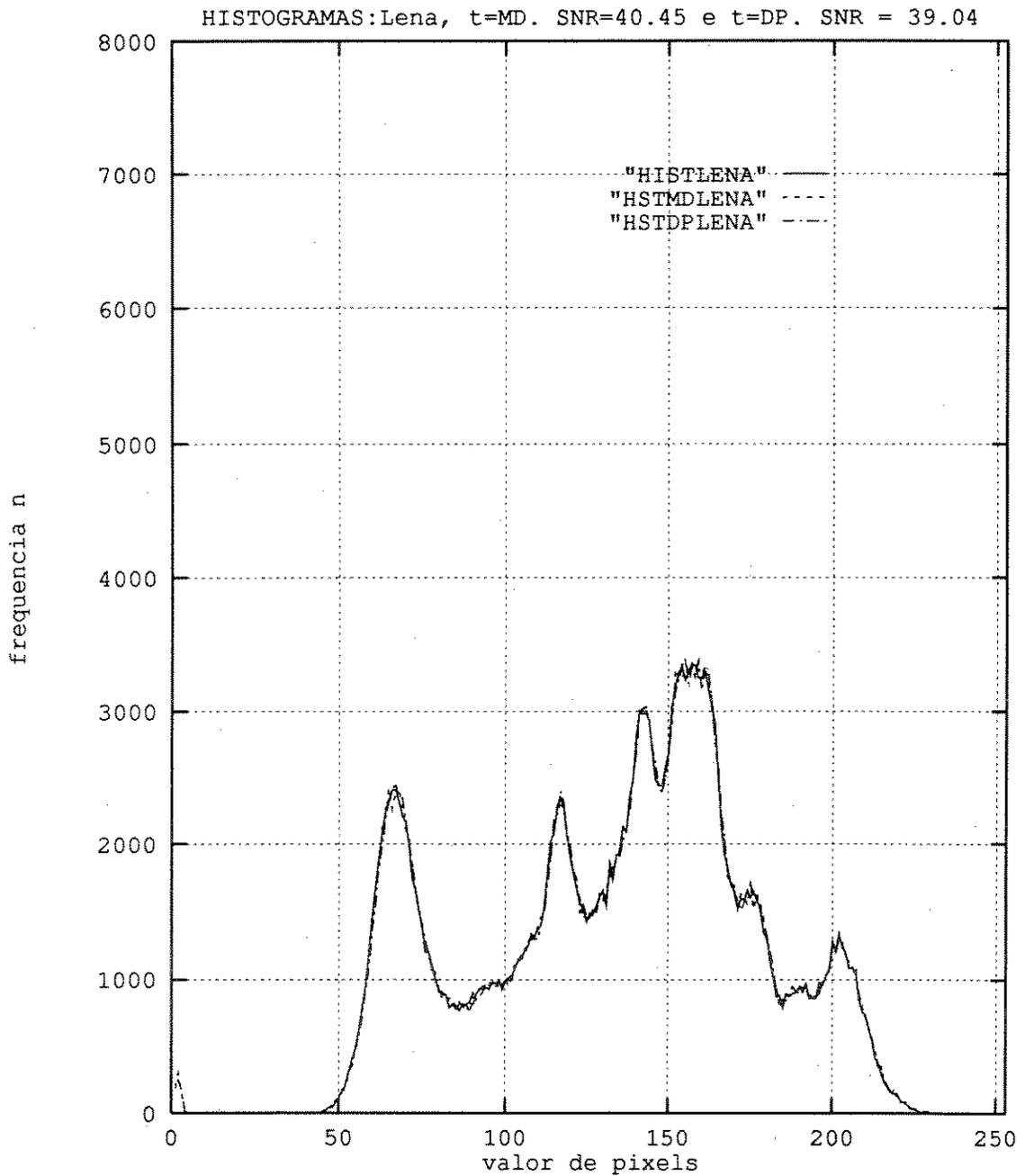


Figura 4.71: Comparação dos Histogramas para a imagem recuperada Lena, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.

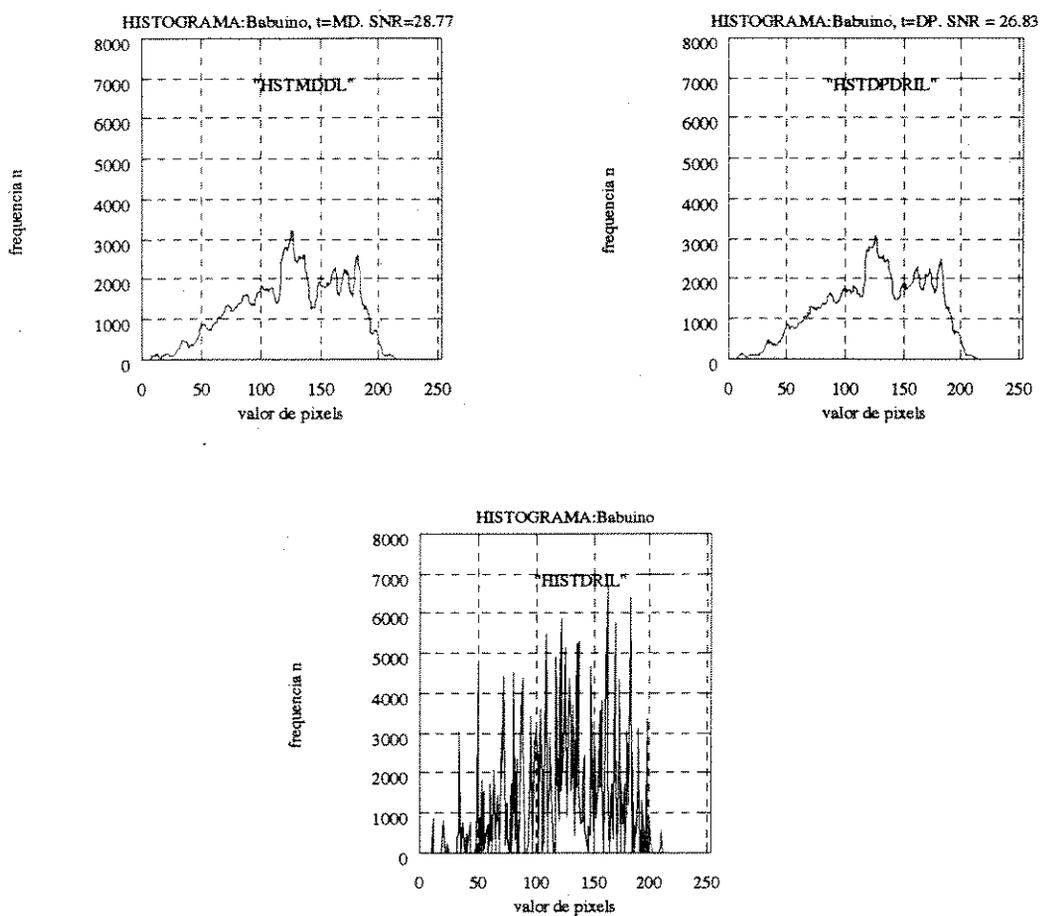


Figura 4.72: Comparação dos Histogramas para a imagem recuperada Babuíno, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.

Ruido e do No de tonalidades, que cujo acréscimo é similar ao caso da entropia. Observa-se, que na recuperação das imagens de teste com referência à Relação Sinal-Ruído , não existe uma grande variação, com exceção da imagem Babuíno, que por não ter um comportamento Markoviano, sua Relação Sinal-Ruído é bastante reduzida. Usando o No de tons a imagem smpte01y é a que melhor se adapta a este tipo de controle estatístico.

IMAGENS	TONALIDADES	SNR (DB)
SMPTE01y	199	REF.
smc01y	200	44.53
smmd01y	200	43.70
smdp01y	200	43.19
SMPTE15y	209	REF.
smc15y	211	44.72
smmd15y	213	43.45
smdp15y	215	43.22
LENA	216	REF.
dcclena	222	44.01
dcmdlena	229	40.45
dcdplena	227	39.04
BABUÍNO	134	REF.
dcdril	218	34.08
dcmdrill	227	28.77
dcdpdril	230	26.83

Tabela 4.12: Quantidade de tonalidades reproduzida na tela da estação de trabalho de acordo com a quantização de tons de cinza permitida e uma comparação do comportamento da Relação Sinal-Ruído para a imagens de teste recuperadas.

4.5.7.5 Resultados do Ponto de vista Subjetivo

Nas figuras (4.73) e (4.74) apresentam-se as imagens de teste recuperadas usando o algoritmo empregado nesta dissertação controlado pelos parâmetros estatísticos e são comparadas subjetivamente com a original de acordo com a Tabela de 5 graus[2],[8], considerando as restrições da secção 4.5.7.1. para esse tipo de tabela subjetiva.

Na figura (4.73) apresenta-se a comparação das imagens de teste da SMPTE, onde se apresenta a cena da praia, smpte01y, e a cena da cozinha, smpte15y.

Comprovou-se que não existe uma grande diferença e que a qualidade quando é observada na tela da estação de trabalho SUN, é de boa a excelente dentro da resolução que permite a tela da estação colorida SUN, e a degradação imperceptível segundo a tabela (4.13)[8],[2], com as considerações prévias.

Na figura (4.74) apresenta-se a comparação das imagens de teste Lena e Babuíno, recuperada e original, Lena tem um comportamento idêntico às imagens da SMPTE. No caso do Babuíno, onde a relação Sinal - Ruído apresenta degradação notória, embora no momento da visualização da figura desta imagem não é possível perceber grande diferença, o que significa que subjetivamente para a qualidade segundo a Tabela (4.13), encontra-se no nível 4 a 5 e para a degradação no nível 5, [2],[8] com a mesma consideração prévia.

ESCALA DE 5 GRAUS.	
QUALIDADE	DEGRADAÇÃO
5.-Excelente	5.- Imperceptível
4.- Boa	4.- Perceptível mas não perturbadora.
3.- Regular	3.- Levemente perturbadora.
2.- Ruim	2.- Perturbadora.
1.- Péssima	1.- Muito Perturbadora.

Tabela 4.13: Escala de 5 graus para medida subjetiva de qualidade ou degradação da imagem.

Na figura (4.75) apresenta-se o sinal Multisalva recuperada, e consegue-se apreciar uma mínima degradação no intervalo das altas frequências ao comparar com a original da página seguinte figura (4.76), isto é devido às restrições impostas neste tipo de sinal na codificação por transformadas (secção 4.5.5).

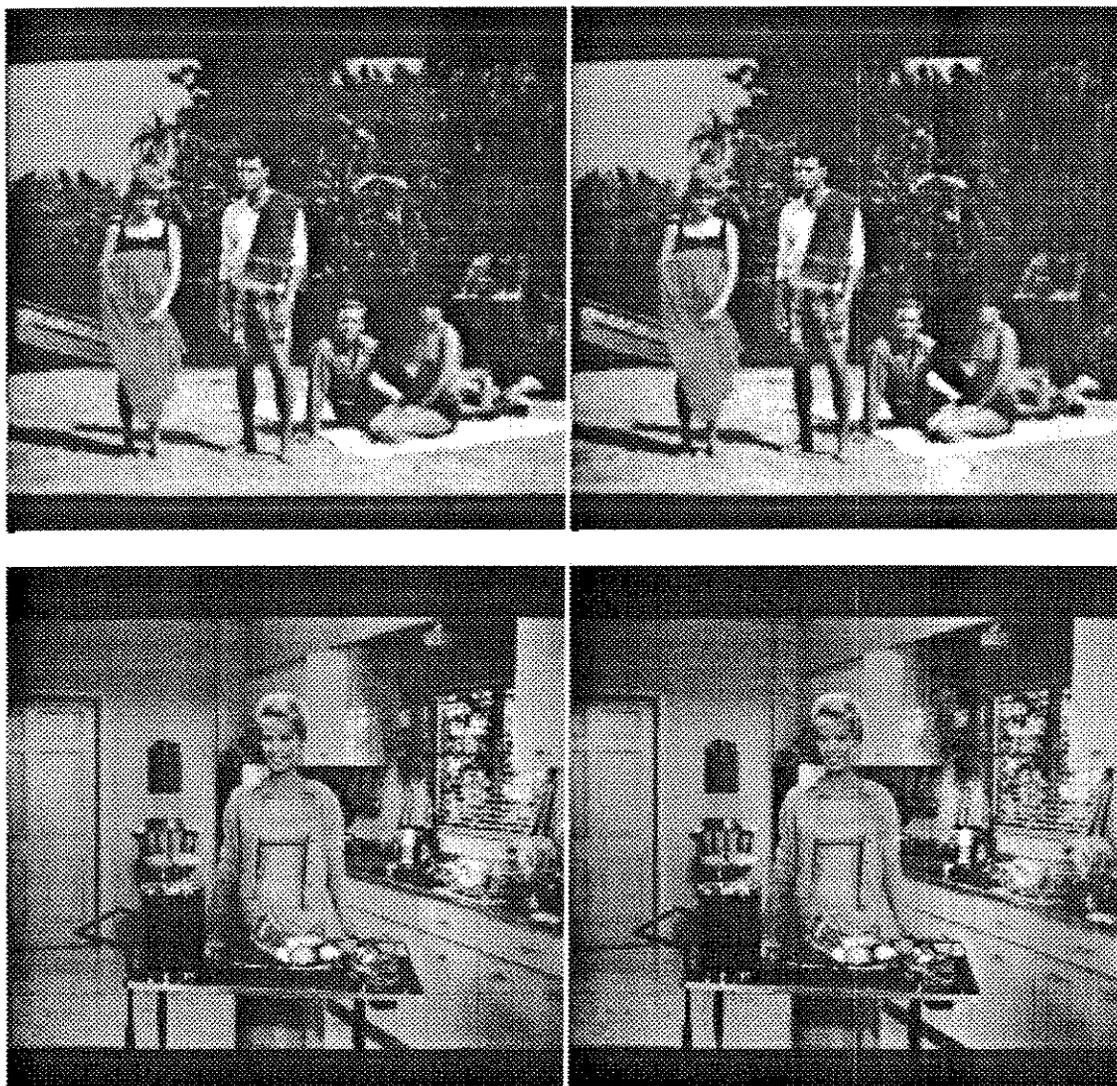


Figura 4.73: Comparação das Imagens de Teste originais (esquerda) com as recuperadas (direita), parte superior: smpte01y e parte inferior: smpte15y.

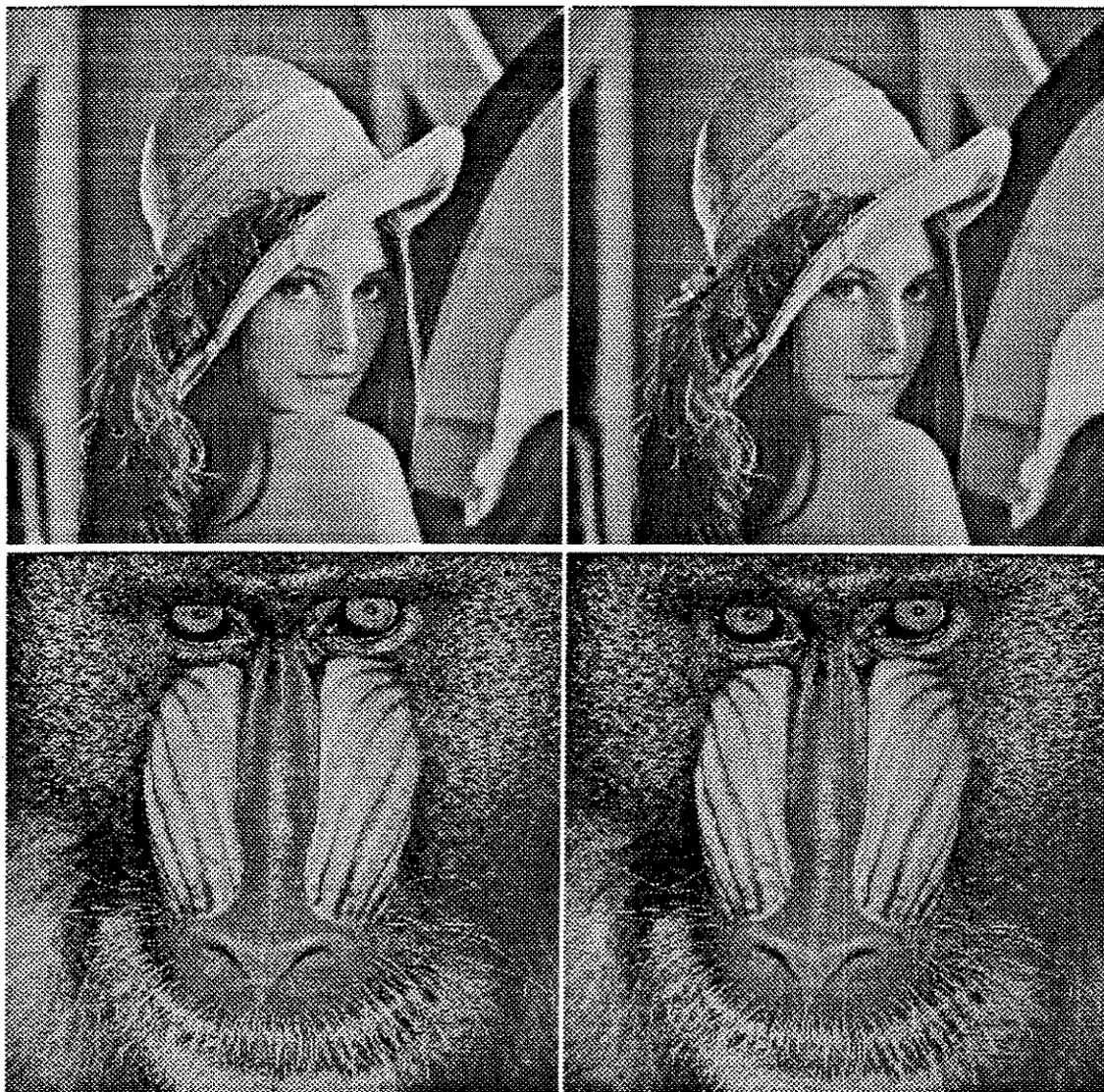


Figura 4.74: Comparação das Imagens de Teste originais (esquerda) com as recuperadas (direita), parte superior: Lena e parte inferior: Babuíno.

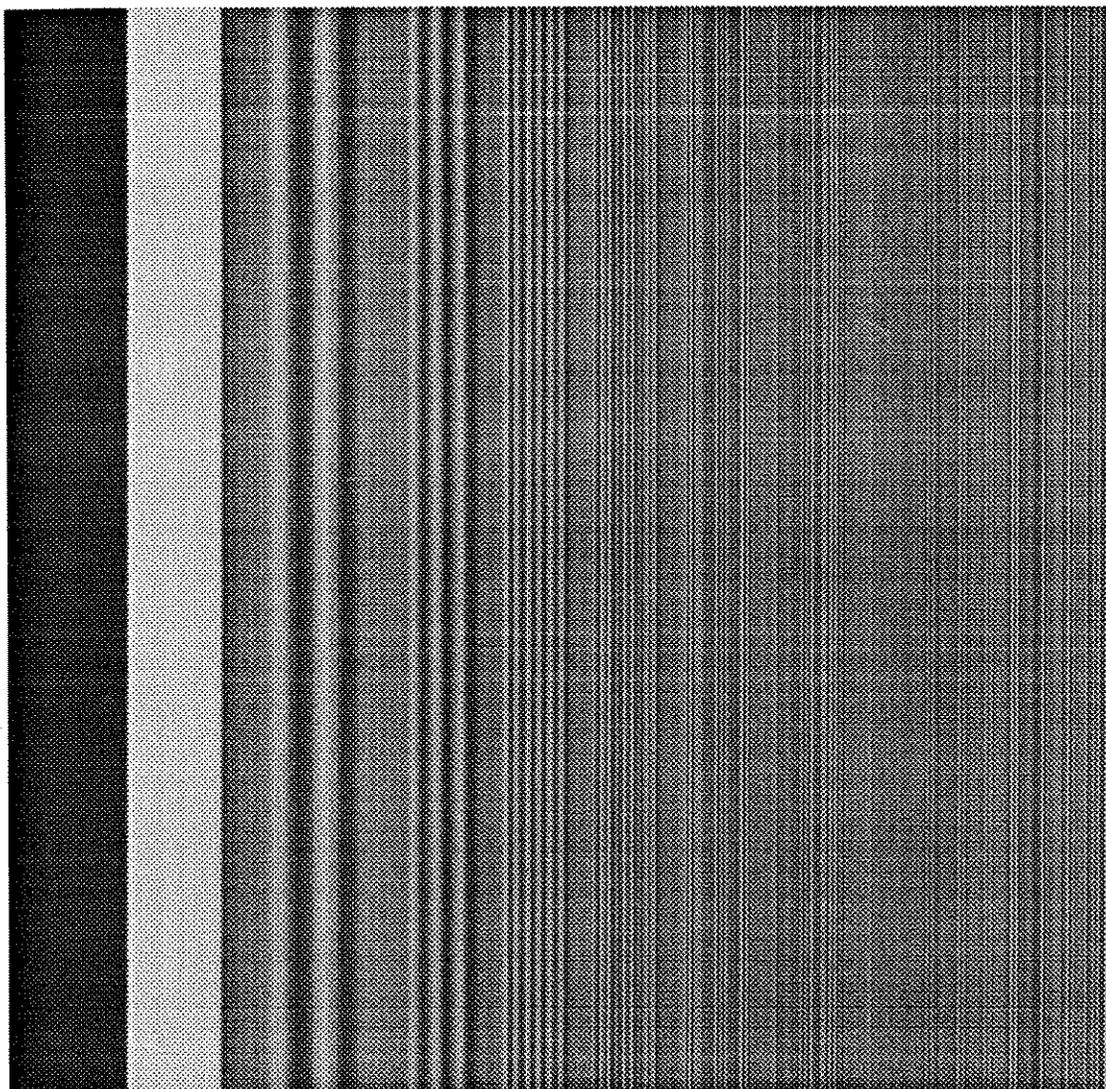


Figura 4.75: Sinal Multisalva recuperada, com mínima degradação no intervalo das altas frequências.

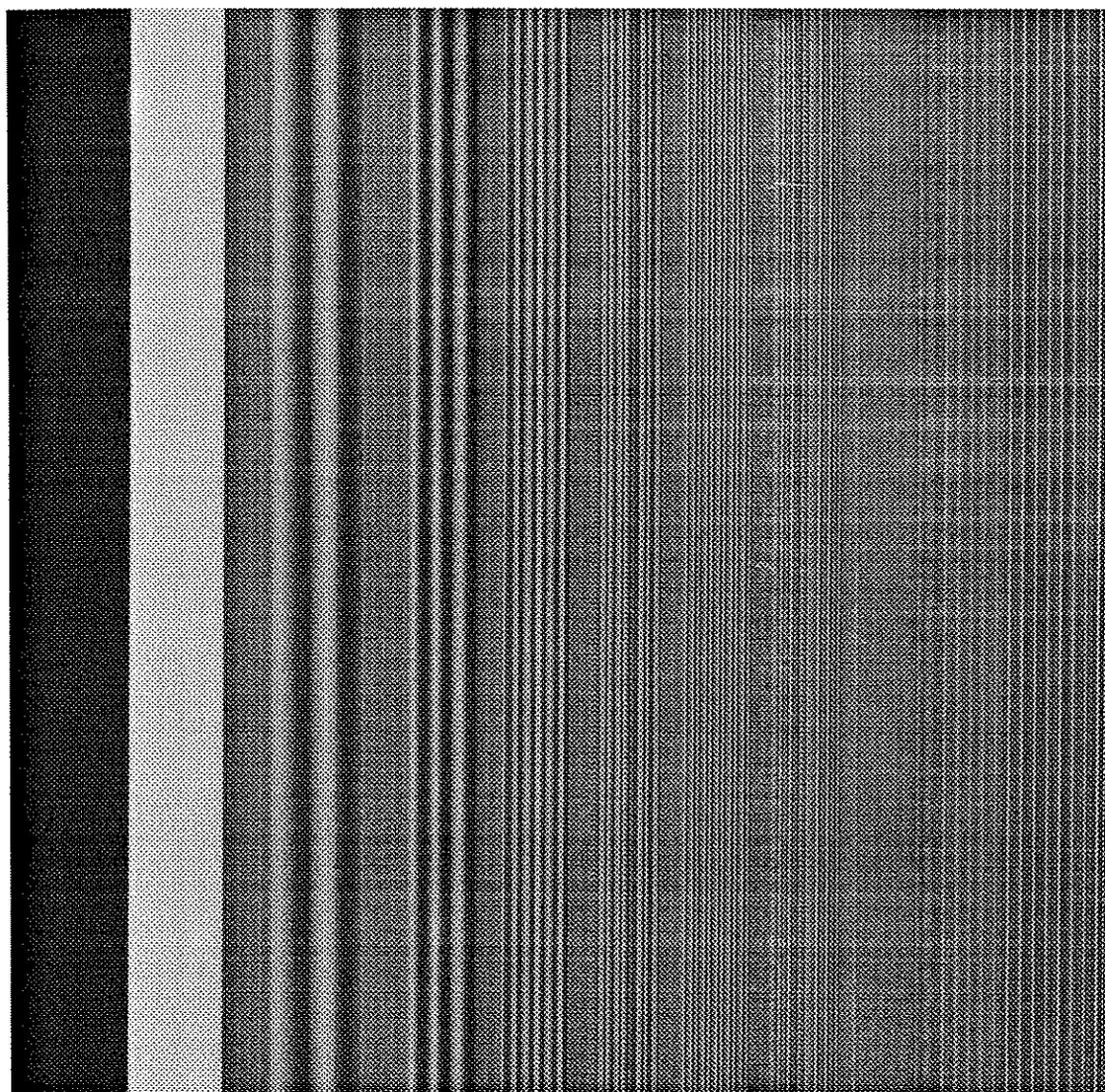


Figura 4.76: Sinal de Teste CCIR II, na tela

Capítulo 5

CONCLUSÕES

Nesse trabalho foram investigados os aspectos referentes à compressão estatística de imagens aplicadas na codificação digital dos sinais na forma de componentes: Y , C_R e C_B , de uso mais universal do que na forma composta, usando um algoritmo adaptativo de bloco variável aplicado à Transformada Cosseno Discreta (DCT) em duas dimensões.

Visou-se conquistar um melhor desempenho comparado ao método convencional empregando codificação por DCT convencional, que trata indistintamente todas as áreas segmentadas em blocos fixos não sobrepostos, produzindo uma qualidade não uniforme para diferentes conteúdos de imagem, devido à não estacionariedade dos dados nos blocos fixos.

Para a redução das redundâncias do sinal de vídeo, utilizou-se a transformada DCT em duas dimensões adaptativa na seleção de coeficientes, de acordo com os blocos e os parâmetros estatísticos de primeira e segunda ordem. Usou-se a estratégia de bloco variável para melhorar a precisão de codificação e, em consequência, a eficiência de codificação. O algoritmo de segmentação e a escolha do tamanho apropriado foi feito de acordo com testes da Média e Desvio Padrão de cada bloco, no início com 16×16 e, se os testes são positivos, na decisão separam-se para passar a blocos de 8×8 e assim até chegar a 4×4 . Posto que um tamanho de bloco maior causa uma melhor eficiência de compressão quando existe suficiente estacionariedade espacial dos dados, procurou-se uma segmentação que ressalte este objetivo e utilizou-se o método de processamento paralelo de regiões usando árvore de quarto grau ou quarta geração do tipo linear ("linear quad-tree") que frisa a homogeneidade na segmentação em aplicações no modelamento de sólidos, robótica, CAD-CAM, etc.. [34], [35]. Mas, como neste caso o objetivo foi a estacionariedade dos dados, os parâmetros estatísticos da Média e do Desvio Padrão foram os encarregados de discriminar esta característica nos respectivos blocos. A grande vantagem em usar a estrutura de árvore de quarto grau para segmentar uma imagem em blocos variáveis é que o tamanho e a localização (índice) de cada sub-bloco estão pré-determinados com um pequeno cabeçalho que informa se os sub-blocos vão separar-se ou vice-versa.

Foram realizados testes subjetivos e objetivos com a visualização das imagens padrão fornecidas pela SMPTE (Society of Motion Pictures and Television Engineers) através de pacotes de avaliação desenvolvidos nesta dissertação e complementados pelos sistemas de análises de imagens: KHOROS (The KHOROS Group. University of New Mexico. E.U.A.); ALV(Phill Everson, University of Bristol.U.K.) e XV (Version 2.21.John Bradley, University

of Pennsylvania, 1992), presentes no ambiente público das estações de trabalho SUN.

Dentro das avaliações de caráter objetivo dos pacotes aqui desenvolvidos, salienta-se as medidas de relação Sinal/Ruído média, Entropia Acumulada e características estatísticas de ocorrência de amostras (histogramas) para as imagens padrão e para o sinal de teste determinístico, usado também nas avaliações de sinais de TV analógicos. Para avaliar o comportamento do codificador, utilizou-se o sinal de Luminância Y por ser mais crítico quanto à informação de vídeo (os sinais diferenças de cor têm menos espectro de frequência ocupado, portanto são menos críticos).

A maioria dos programas de simulação foram criados usando linguagem Fortran 77, programação em Postscript e alguns em linguagem C. O tratamento dos programas criados e empregados não foram objetos de análise deste trabalho porque constituem um desenvolvimento longo e paralelo, cujo objetivo é fornecer os resultados para análise.

A simulação do codificador foi baseada no desenvolvimento de pacotes de “software” criados nesta dissertação para simular só as partes fundamentais do sistema, tais como: a Transformada Discreta Cosseno (DCT) em duas dimensões com segmentação de blocos fixo e variável, técnicas adaptativas de seleção de coeficientes e técnicas de quantização que permitiram a recuperação integral da imagem livre de interferências perceptíveis inerentes de um processamento com perdas na compressão e com precisão de codificação do ordem de 2 bits/E.I.. As outras partes do esquema completo de codificação foram mencionadas e algumas analisadas lateralmente, mas não foram avaliadas, porque outros autores têm trabalhado exaustivamente em tais partes(DPCM, codificação entrópica, memórias elásticas, detecção e correção de erros, multiplexador, etc..)

Frisa-se que os objetivos de se ter usado a segmentação ou a partição da matriz 512x512 de dados de imagem nesta dissertação foram dois: primeiro, a organização dos dados de imagem numa forma mais estacionária, para conseguir na saída da DCT uma maior decorrelação e uma compressão mais eficiente; segundo, este tipo de segmentação pertence à nova classe de algoritmos matriciais particionados para adaptar-se mais facilmente na implementação futura de calculadores rápidos de matrizes em “VLSI” com grande aplicação em processamento de imagem e aplicações científicas e em tempo real, usando modelos de computação paralela, onde é possível obter um acréscimo de velocidade usando “p” processadores de sub-matrizes (“chips” aritméticos) independentes [34], [39].

Logo, propõe-se aqui uma estrutura de codificação eficiente e um algoritmo base diferente do sugerido pela CCIR e JPEG, visando: o avanço da tecnologia de componentes RISC (processamento paralelo) em processadores de vídeo (C-cube CL550-35), a compressão de imagem e a redução de taxa.

Intensas avaliações foram feitas para estabelecer a tabela empírica dos parâmetros de controle estatísticos das quatro imagens de teste, Média e Desvio Padrão. Se, além disso, são consideradas as avaliações finais de Entropia Acumulada, Histogramas, Nível de Sinal-Ruído, tabelas de Tonalidades e tabelas Estatísticas, a utilização dos diferentes dados armazenados em forma comprimida ocupou um espaço de memória superior a 200 Mbytes, o que levou a analisar os dados apenas utilizando resultados de curvas gráficas para projetar o algoritmo por árvore de quarta geração e deduzir as conclusões finais.

Das avaliações finais previamente mencionadas e apresentadas no capítulo quatro, conclui-se que os resultados obtidos determinam uma comprovação das afirmações feitas na introdução do capítulo um e capítulo quatro e neste ao início.

Foi demonstrado a grande inconveniência do uso da codificação convencional por bloco fixo que, ao não considerar o comportamento não estacionário das imagens reais, produz perdas de degradação da imagem pelo fato de não usar na segmentação pequenos blocos apropriados para as mudanças abruptas de bordas e com seleção de coeficientes adaptativos nessas mudanças. Por outro lado, o fato de não poder usar bloco maior para uma situação estacionária da imagem, causa a perda da utilização da grande vantagem da transformada DCT com respeito a sua propriedade de compactação de energia que incide diretamente na eficiência de compressão do bloco, pois nesses casos a compressão pode ser máxima sem produzir perdas notórias de alta frequência e a conseguinte redução da taxa de bits.

Implementou-se o algoritmo proposto para vencer essas deficiências específicas mencionadas previamente, obtendo-se um resultado com respeito à Degradação qualificada na escala subjetiva de 5 graus com as restrições da secção 4.5.7.1, no nível 4 a 5 e com respeito à Qualidade. Devido à baixa resolução da tela das estações de trabalho (256 tons), pode-se estabelecer que ficou também entre o nível 4 e 5, ou seja, a Qualidade é de boa a excelente somente para imagens que tem um comportamento Markoviano, espaço onde fica a grande maioria das imagens reais. Este comportamento foi totalmente comprovado

através das avaliações objetivas do nível Sinal-Ruído cujos resultados foram da ordem dos 40 db e através dos Histogramas e Entropia Acumuladas onde a diferença foi mínima para as imagens de comportamento Markoviano. O caso de exceção foi a Imagem Babuíno, onde a avaliação objetiva castigou severamente a Relação Sinal-Ruído, cujo valor caiu antes do quantizador de 34.08 db para 26.83 db, valor considerado baixo do ponto de vista de Qualidade, embora a Degradação subjetiva da Escala de 5 graus acuse o nível de 4 a 5, quase-imperceptível, devido ao comportamento do S.V.H. neste tipo de imagem. A perda da Qualidade da avaliação objetiva apenas se justifica pelo fato da projeção da seleção de coeficientes transformados estar baseado no comportamento Laplaciano dos coeficientes transformados pela DCT de duas dimensões para imagens do tipo Markoviano e, pelo fato dos valores das tabelas de ponderação na quantização dos coeficientes transformados estarem baseados no comportamento psico-visual para o S.V.H. de imagens reais tipo Markovianas, portanto era de se esperar que o resultado para o caso do Babuíno, que saísse do modelamento estatístico mencionado.

No que diz respeito à precisão de codificação, pode-se resumir que o uso do algoritmo proposto no caso de imagens de comportamento Markoviano com variações suaves como foi o caso de Lena, usando o controle do Desvio Padrão (teste 1.) para a seleção do tamanho de bloco, não produz grande variação quantitativa da distribuição de blocos e por consequência da precisão de codificação. Neste caso o número de diferentes blocos passa de 6670 para 5944 e por sua vez a precisão de codificação passa de 2.2619 para 2.1078 bits/EI, o que significaria, sem considerar o acréscimo de eficiência de codificação fornecidos pela codificação entrópica e a codificação DPCM do sistema final híbrido, (ver figura(5.1)), passar de uma taxa de 17.78 Mbit/s para uma de 16.5150 Mbit/s, considerando um quadro SMPTE de 512x512 com 30 repetições por segundo. Agora, no caso de imagens não Markovianas como no caso do Babuíno, o algoritmo proposto cumpre um eficiente objetivo, já que o número de diferentes blocos passa de 10702 para 4759, significando que a precisão de codificação se reduz de 3.10 para 1.8650 bits/EI com as mesmas considerações prévias, o que significa finalmente que a taxa de bits cai de 24.37 Mbit/s para 14.66 Mbit/s. Do ponto de vista objetivo, acontece uma redução dos fatores de qualidade (Histogramas, Entropia Acumulada e Relação Sinal-Ruído). Não obstante, do ponto de vista subjetivo do S.V.H. a degradação é imperceptível.

Conclui-se que, se o sistema fosse implementado usando um formato de menor

resolução por quadro, caso do CIF (Formato Intermediário Comum), tabela (3.2), onde seria compatível com os sistemas NTSC e PAL, este algoritmo daria um excelente resultado no canal de distribuição secundária da CCITT, "Broadband ISDN (B-ISDN)", visto que é possível ter uma imagem de alta qualidade a uma taxa de 17 Mbit/s (ou menor) que é o requerimento deste canal futuro [40].

Uma possível estrutura aplicada como codificador DPCM/Transformada de tipo Híbrido é apresentada na figura (5.1) usando predição inter-quadro e cuja avaliação será dependente do algoritmo proposto.

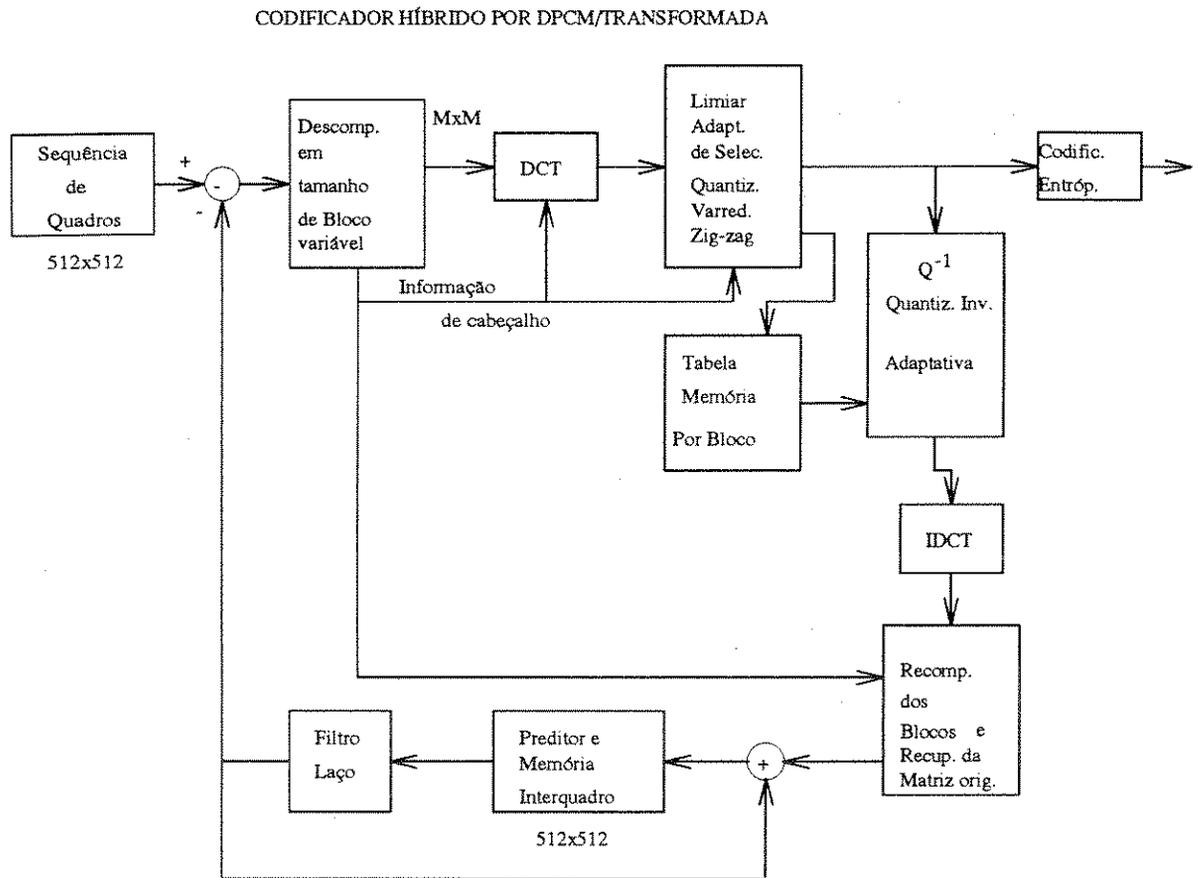


Figura 5.1: Codificador DPCM/Transformada Híbrido

Ainda que a primeira transformação do espaço de cor é uma parte do processo de remoção de redundâncias, este trabalho supõe as componentes de primeira transformação:

Y , C_R e C_B da referência [1], e o algoritmo proposto nesta dissertação tem sido avaliado para o sinal mais crítico, neste caso o sinal de Luminância onde o S.V.H. é mais sensível à degradação do que os sinais componentes da cor. Não obstante este algoritmo proposto é independente do domínio ou espaço de cor, portanto pode ser usado para comprimir os dados dos diferentes espaços, tais como: RGB, $YC_R C_B$, e CMYK.

Contudo, a melhor compressão resulta quando os sinais componentes têm menos correlação, como no caso Y , C_R e C_B , onde a maior quantidade de informação encontra-se concentrada na Luminância e em menor grau nos sinais de Crominância.

Finalmente, como uma extensão da aplicação deste algoritmo, utilizou-se no canal usado pelo sinal componente de Luminância Y , os sinais primários das imagens de Teste: sm01r, sm01g, sm01b, sm15r, sm15g e sm15b, processando-as em forma individual e comprovou-se que os resultados nos Histogramas, Entropias Acumuladas e relações Sinal-Ruído são exatamente os mesmos tanto para o sinal de Luminância bem como para os sinais primários de cor R,G e B, apresentando como resultado nas imagens “.ras” e “.gif” com apenas o acréscimo de uma ou duas tonalidades com respeito as originais do mesmo tipo, (apêndice (f)).

Por último, é importante frisar que, com este trabalho, pretende-se apenas dar uma contribuição para o desenvolvimento deste campo de compressão de imagens via transformada cosseno para transmissão e permitir a abertura de novas frentes de pesquisa nesta especialidade de vídeo bastante ampla.

Apêndice a

**GLOSSÁRIO DE TERMOS
USADOS**

Segue abaixo, uma relação dos termos usados nesta dissertação:

- C_R = O sinal Diferença de cor vermelho.
 C_B = O sinal Diferença de cor azul.
CCF = Código de Comprimento Fixo.
CCIR = Comité Consultivo de Radio Internacional.
CCITT = Comité Consultivo de Telefone e Telegrafo Internacional.
CIF = Formato Intermediário Comum.
CEC = Comité de Países Europeos.
DCT-2D = Transformada Cosseno Discreta em duas Dimensões.
EI = E.I. = Elementos de Imagem = "pixel".
JPEG = Joint Photographic Experts Group.
KLT = Transformada de Karhunen e Loeve.
MCPD = DPCM = Modulação por Código de Pulso Diferencial.
MPEG = Motion Picture Experts Group.
NTSC = Comité de Sistemas de Television Nacional.
PAL = Linhas de Fase Alternadas.
RISC = Computador de conjunto de instruções reduzidas.
SECAM = Sequências com Memórias.
SMPTE = Society of Motion Pictures and Television Engineers.
SVH = Sistema Visual Humano.
VLC = Codificação de Comprimento Variável.
VLSI = Integração a muito grande escala.
Y = O sinal Luminância.
 f_{sc} = Frequência sub-portadora
Broadcasting = Transmissão profissional.
Super-Nyquist = Frequência Superior à determinada pelo Teorema de Nyquist.

Apêndice b

LISTA DE TRABALHOS PRODUZIDOS E RELACIONADOS.

Segue abaixo, uma relação dos resumos dos trabalhos produzidos desta dissertação:

No período dos anos 1992 e 1993 se apresentaram os seguintes trabalhos em Congressos e Seminários com o linguagem correspondente:

Estudo de uma Codificação de Imagens Estáticas por Transformadas DCT-2D.

Apresentado no SBT, Telemo92. 1992 Brasília.

Este estudo apresenta uma codificação eficiente por transformada DCT-2D (Transformada Cosseno Discreta em duas dimensões) associada a métodos adicionais de processamento digital. Mediante o uso de codificação DCT é possível tirar redundâncias e compactar a energia dos coeficientes transformados de maior variância; podendo ser adaptativa pela mudança do critério de seleção, quantização dos coeficientes e alocação de bits para permitir um mínimo espaço de armazenamento em memória.

Estudio de Compresión de Imagen usando Codificación DCT-2D

Apresentado no Seminário de Comunicações e Computação no México, organizado pela IEEE. Seção México. Novembro 1992.

Este estudio presenta una codificación eficiente por transformada DCT-2D (Transformada Coseno Discreta en dos dimensiones) asociada a métodos adicionales de procesamiento digital. Mediante el uso de codificación DCT es posible eliminar redundancias en memoria gracias a un esquema adaptativo a dos parámetros para la cuantización y localización de bits por bloques. Finalmente se obtiene una precisión de codificación por bloque de 2 bit por muestra con el mejor compromiso entre la calidad de imagen, la complejidad del sistema y la codificación eficiente.

Compresión de Imagen Digital por Transformada DCT-2D usando Segmentación de Bloque Variable Adaptativo.

Apresentado no X Congreso Chileno de Ingeniería Eléctrica. Valdivia. 1993

Este trabajo enfatiza los defectos del algoritmo de codificación usando Transformada Coseno Discreta (DCT) convencional que trata todas las áreas segmentadas de la

imagen en bloques fijos en forma indiscriminada produciendo una calidad no uniforme para diferentes contenidos de la imagen.

Se presenta un esquema de bloques variable adaptativo al contenido de la imagen a través de herramientas estadísticas de primer y segundo orden que determinan un resultado promisorio pensando en la aplicabilidad del procesamiento paralelo intenso y los resultados de las evaluaciones hechas para el mejor compromiso entre el número de bits por elemento de imagen (E.I.) y la calidad de la escena recuperada.

Apêndice c

LISTAS DE INTERESSE.

Lista de Figuras

1.1	Codificação por Transformada DCT do sinal de vídeo em componentes . . .	3
2.1	Codificação Preditiva básica	10
3.1	Princípio de Formato Intermediário Comum que permite a conexão direta entre países com Sistemas NTSC e PAL. Conversão para CIF é possível com filtragem e amostragem apropriados.	17
3.2	Processamento básico da imagem espacial submetida a amostragem e quantização	20
3.3	Espectro de frequência de Fourier da imagem amostrada $I_s(x, y)$. X_{fx} e Y_{fy} são respectivamente os ciclos horizontais e verticais. O intervalo de baixa frequência é a imagem $\bar{I}(x, y)$ original de banda limitada. Fora estão as frequências imagem.	25
3.4	Espectro de Frequência da imagem amostrada $I_s(x, y)$ quando a imagem original $\bar{I}(x, y)$ contém componentes de frequência cujos valores estão sobre a metade da frequência da amostragem. Componentes na zona sombreada são afetadas pelas frequências imagem. Esta distorção é denominada “pre-aliasing” e é devido a insuficiência de pré - filtragem na imagem original $\bar{I}(x, y)$	26
3.5	Pós-filtragem não ideal. A linha segmentada indica a extensão da frequência de um filtro sub-ótimo $h_d(u, v)$. A linha sólida indica o filtro ótimo. Componentes de frequências “ alias” aparecem nas áreas sombreadas na imagem reconstruída. Esta distorção é denominada “post-aliasing”, e é devido a pós-filtragem inadequada de $I_s(x, y)$	28

3.6	Exemplo de uma borboleta	39
3.7	Duas soluções de 4 pontos	40
3.8	Especificação externa do algoritmo da DCT	42
3.9	Análise do não casamento da IDCT para um esquema híbrido de codificação contendo duas transformadas no laço	45
3.10	Análise do não casamento da IDCT para um esquema híbrido de codificação não contendo transformadas no laço	48
3.11	Quantização de Variáveis	55
3.12	Função Densidade de Probabilidade do sinal de entrada nos intervalos de decisão com igual ocorrência	58
3.13	Característica não linear de quantização	59
3.14	Característica de quantização uniforme.	60
4.1	Esquema da utilização das sinais de imagem em componentes no algoritmo empregado nesta dissertação	66
4.2	Algoritmo Estatístico de Bloco Variável Adaptativo Empregado.	69
4.3	Estatística por bloco de 16x16, 8x8, 4x4.	70
4.4	Codificador convencional	71
4.5	Segmentação de tamanho de blocos 1x1, 4x4, 8x8 y 16x16 para a imagem padrão SMPTE01Y.	72
4.6	Segmentação de tamanho de blocos 1x1, 4x4, 8x8 y 16x16 para a imagem padrão SMPTE15Y.	73
4.7	Imagens de Teste usadas nesta dissertação: smpte01y, smpte15y, Lena e Babuíno	76
4.8	Sinal de Teste CCIR II ou Multiburst	78
4.9	Sinal de Teste CCIR II, na tela	79
4.10	Avaliação e seleção ótima de coeficientes.	81
4.11	Estruturas de Coeficientes Transformados em blocos de 16x16 (a).	83
4.12	Estruturas de Coeficientes Transformados em blocos de 16x16 (b).	84

4.13	Matriz de Funções Bases da DCT 4x4	85
4.14	Estruturas de Geometrias Simples.	86
4.15	Comportamento do Desvio Padrão na segmentação de 16x16 na imagem smpte01y.	90
4.16	Comportamento da Média na segmentação de 16x16 na imagem smpte01y.	92
4.17	Comportamento da Variância na segmentação de 16x16 na imagem smpte01y.	93
4.18	Comportamento da Energia Total na segmentação de 16x16 na imagem smpte01y.	94
4.19	Comportamento do Desvio Padrão na segmentação de 16x16 na imagem smptel5y.	95
4.20	Comportamento da Média na segmentação de 16x16 na imagem smptel5y.	96
4.21	Comportamento do Desvio Padrão na segmentação de 16x16 na imagem Lena.	98
4.22	Comportamento da Média na segmentação de 16x16 na imagem Lena.	99
4.23	Comportamento do Desvio Padrão na segmentação de 16x16 na imagem Babuíno.	100
4.24	Comportamento da Média na segmentação de 16x16 na imagem Babuíno.	102
4.25	Comportamento do Desvio Padrão na segmentação de 16x16 nas diferentes imagens de teste.	103
4.26	Comportamento da Média na segmentação de 16x16 nas diferentes imagens de teste.	105
4.27	Comportamento do Desvio Padrão na segmentação de 8x8 na imagem smpte01y.	106
4.28	Comportamento da Média na segmentação de 8x8 na imagem smpte01y.	108
4.29	Comportamento do Desvio Padrão na segmentação de 8x8 na imagem smptel5y.	109
4.30	Comportamento da Média na segmentação de 8x8 na imagem smptel5y.	110
4.31	Comportamento do Desvio Padrão na segmentação de 8x8 na imagem Lena.	112
4.32	Comportamento da Média na segmentação de 8x8 na imagem Lena.	113
4.33	Comportamento do Desvio Padrão na segmentação de 8x8 na imagem Babuíno.	114
4.34	Comportamento da Média na segmentação de 8x8 na imagem Babuíno.	116

4.35	Comportamento do Desvio Padrão na segmentação de 8x8 nas diferentes imagens de teste.	117
4.36	Comportamento da Média na segmentação de 8x8 nas diferentes imagens de teste.	118
4.37	Comportamento do Desvio Padrão na segmentação de 4x4 nas diferentes imagens de teste.	120
4.38	Comportamento da Média na segmentação de 4x4 nas diferentes imagens de teste.	121
4.39	Codificação por Transformada adaptativa com segmentação de tamanho de bloco variável.	124
4.40	Segmentação de tamanho de bloco por junção ou união de quatro sub-blocos adjacentes.	125
4.41	Imagem segmentada por Regiões e sua estrutura em árvore de quarto grau.	129
4.42	Oct-tree de nós terminais escuros e sua representação	130
4.43	Árvore de quarto grau correspondente à figura (4.39)	131
4.44	Localização das faixas verticais de frequência do sinal Multisalva(pontos horizontais simétricos e harmônicos às funções bases de imagem da transformada), no domínio da transformada DCT-2D	133
4.45	Distribuição da energia dos coeficientes usando a equação 3.88, variando os valores das coordenadas U e V.	134
4.46	Controle de seleção adaptativa ao Desvio Padrão aplicado à amostragem por zona de coeficientes. Mostra-se a diferença entre a seleção triangular e a seleção por zona Laplaciana para um bloco de 4x4. Neste caso usa-se $k=-3$	135
4.47	Processo de quantização empregando Tabela.	135
4.48	Tabelas de quantização baseadas em funções de ponderação projetadas para o Sistema Visual Humano.(Psychovisual Weighting Functions).	137
4.49	Resumo do processamento a nível de blocos segmentados e suas avaliações finais que requerem um intenso uso de facilidades computacionais considerando e não considerando o teste (1.).	139
4.50	Esquema de varredura em zig-zag para um bloco de 4x4.(Método de Parsons-Tescher)	141

4.51	Plano de alocação de bits para os diferentes tamanhos de blocos.	142
4.52	Distribuição de Blocos por Controle Estatístico para imagem smpte01y. . .	146
4.53	Distribuição de Blocos por Controle Estatístico para imagem smpte15y. . .	147
4.54	Distribuição de Blocos por Controle Estatístico para imagem Lena.	148
4.55	Distribuição de Blocos por Controle Estatístico para imagem Babuíno. . . .	149
4.56	Comparação da Imagem Lena, esquerda, apenas com o teste de segmentação usando a Média; Direita, usando o teste 1. do Desvio Padrão empregado nesta dissertação.	151
4.57	Histograma da Imagem recuperada smpt01y, apenas comprimida pelo algoritmo empregado e comparada com a original.	154
4.58	Entropia Acumulada da Imagem recuperada smpt01y, apenas comprimida adaptativamente com os parâmetros estatísticos do algoritmo empregado e comparada com a original.	155
4.59	Histograma da Imagem recuperada smpt15y, apenas comprimida pelo algoritmo empregado e comparada com a original.	156
4.60	Entropia Acumulada da Imagem recuperada smpt15y, apenas comprimida adaptativamente com os parâmetros estatísticos do algoritmo empregado e comparada com a original.	157
4.61	Histograma da Imagem recuperada Lena, apenas comprimida pelo algoritmo empregado e comparada com a original.	158
4.62	Entropia Acumulada da Imagem recuperada Lena, apenas comprimida adaptativamente com os parâmetros estatísticos do algoritmo empregado e comparada com a original.	159
4.63	Histograma da Imagem recuperada Babuíno, apenas comprimida pelo algoritmo empregado e comparada com a original.	160
4.64	Entropia Acumulada da Imagem recuperada Babuíno, apenas comprimida adaptativamente com os parâmetros estatísticos do algoritmo empregado e comparada com a original.	161
4.65	Comparação das Entropias acumuladas para a imagem recuperada smpte01y, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.	163

4.66	Comparação das Entropias acumuladas para a imagem recuperada smpte15y, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.	164
4.67	Comparação das Entropias acumuladas para a imagem recuperada Lena, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.	165
4.68	Comparação das Entropias acumuladas para a imagem recuperada Babuíno, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.	166
4.69	Comparação dos Histogramas para a imagem recuperada smpte01, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.	168
4.70	Comparação dos Histogramas para a imagem recuperada smpte15, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.	169
4.71	Comparação dos Histogramas para a imagem recuperada Lena, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.	170
4.72	Comparação dos Histogramas para a imagem recuperada Babuíno, usando o teste da Média e o teste do Desvio Padrão empregado nesta dissertação e comparadas com a original.	171
4.73	Comparação das Imagens de Teste originais (esquerda) com as recuperadas (direita), parte superior: smpte01y e parte inferior: smpte15y.	174
4.74	Comparação das Imagens de Teste originais (esquerda) com as recuperadas (direita), parte superior: Lena e parte inferior: Babuíno.	175
4.75	Sinal Multisalva recuperada, com mínima degradação no intervalo das altas frequências.	176
4.76	Sinal de Teste CCIR II, na tela	177
5.1	Codificador DPCM/Transformada Híbrido	183
d.1	Forma de aplicação dos programas desenvolvidos.	201

e.1	Cabeçalho de sincronismo de quadro.	203
f.1	Comparação das imagens coloridas, smpte01(acima) e smpte15(abaixo), originais(esquerda) e recuperadas(direita).	205

Lista de Tabelas

3.1	Parâmetros CCIR 601	15
3.2	Comparação do formato CIF com NTSC e PAL	17
3.3	Comparação do número de blocos em um quadro com formato SMPTE	51
4.1	Características subjetivas das quatro imagens de teste. * Ver secção 3.8	75
4.2	Características estatísticas das quatro imagens de teste	75
4.3	Relação Sinal/Ruído da Imagem smpte15y recuperada utilizando apenas 25 % dos coeficientes transformados	87
4.4	Relação Sinal/Ruído para as segmentações de bloco fixo 4x4, 8x8 e 16x16 com quantização uniforme e em torno de 40% de coeficientes selecionados.	87
4.5	Resumo dos Valores limites de Intervalos Adequados do Desvio Padrão usado para o Controle da segmentação das Imagens de Teste.	122
4.6	Resumo dos Valores limites de Intervalos Adequados da Média usado para o Controle da segmentação das Imagens de Teste.	123
4.7	Distribuição de blocos por imagem de teste. Onde, CDP = com Desvio Padrão e SDP= sem Desvio Padrão.	150
4.8	Precisão de codificação. (Unidade: bits/E.I.)	150
4.9	Cabeçalho da árvore incluído na precisão de codificação. (Unidade: $\times 10^{-2}$ bits/E.I.)	152
4.10	Características subjetivas da qualidade em função da compressão usada pela JPEG.	152

4.11	Características estatísticas das quatro imagens de teste recuperadas antes e depois do quantizador, submetidas ao controle dos parâmetros estatísticos, Média e Desvio Padrão	167
4.12	Quantidade de tonalidades reproduzida na tela da estação de trabalho de acordo com a quantização de tons de cinza permitida e uma comparação do comportamento da Relação Sinal-Ruído para a imagens de teste recuperadas.	172
4.13	Escala de 5 graus para medida subjetiva de qualidade ou degradação da imagem.	173

Apêndice d

APLICAÇÃO DOS PROGRAMAS

A continuação segue a forma de aplicação dos programas utilitários desenvolvidos (figura(d.1)):

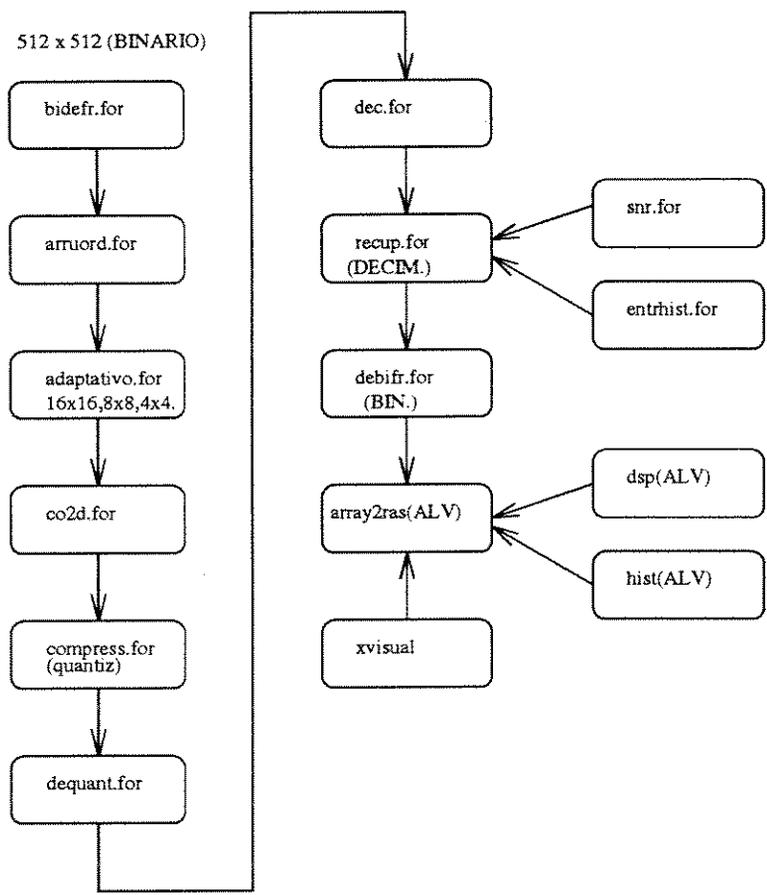


Figura d.1: Forma de aplicação dos programas desenvolvidos.

Apêndice e

**CABEÇALHO DE
SINCRONISMO DE QUADRO**

A continuação apresenta-se o cabeçalho de quadro presente no início de quadro. Este cabeçalho tem por objetivo especificar as características das fonte de imagem, os componentes no quadro, o fator de amostragem de cada componente, e selecionar a tabela de quantização para ser usada em cada tamanho de bloco e para cada componente.

Os marcadores e parâmetros são amostrados na figura (e.1)

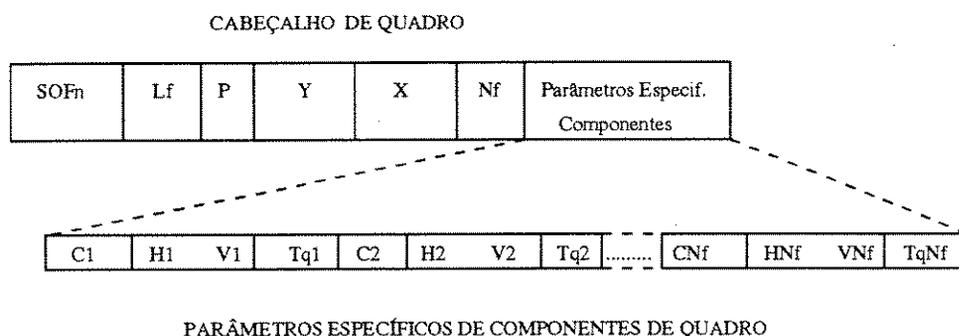


Figura e.1: Cabeçalho de sincronismo de quadro.

O significado de cada de cada termo é:

SOFn = marcador de início de quadro, da o començo dos parâmetros de quadro. A letra n, indica se o processo de codificação é de linhabase sequêncial ou qualquer outro tipo de codificação entrópica.

Lf = especifica o comprimento do cabeçalho de quadro.

P = especifica a precisão em bits, para as amostras das componentes no quadro.

Y = especifica o número de linhas da fonte de imagem.

X = especifica o número de amostras por linha da fonte de imagem.

Nf = especifica o número de fuentes de imagem componentes no quadro.

Ci = especifica um identificador único ao i-ésimo componente na sequêncial de parâmetros de especificação de componentes de quadro.

Hi = fator de amostragem horizontal.

Vi = fator de amostragem vertical.

Tqi = seletor de tabela de quantização.

Apêndice f

IMAGENS “.gif” E “.ras”

A continuação apresenta-se uma comparação das imagens SMPTE originais e recuperadas do ponto de vista de Cores(figuraf.1).

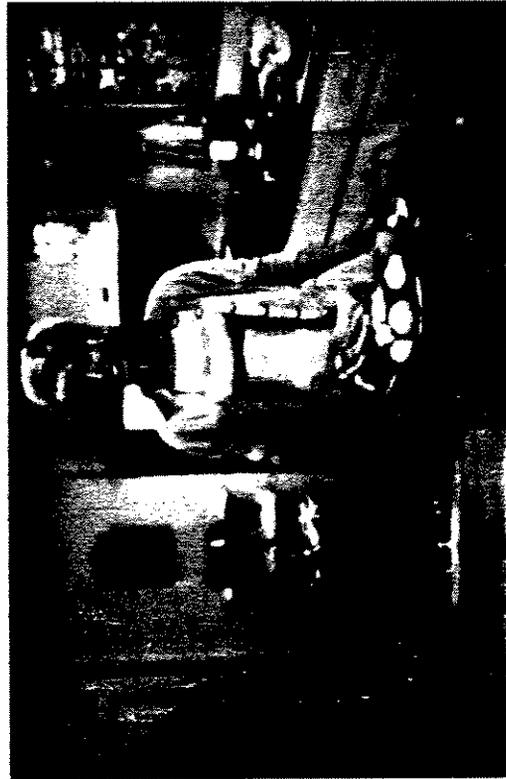
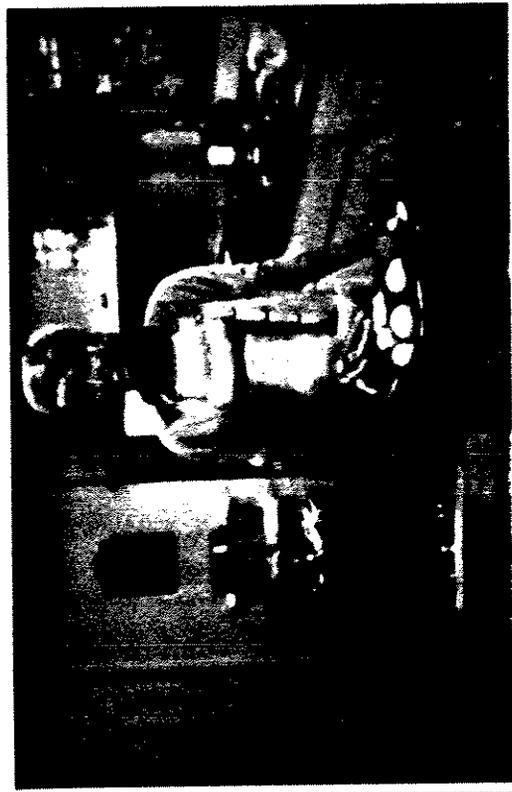


Figura f.1: Comparação das imagens coloridas, smpte01(acima) e smpte15(abaixo), originais(esquerda) e recuperadas(direita).

Apêndice g

ARREDONDAMENTO E EXATIDÃO DA IDCT.

Especificações da IDCT. CCITT SG XV

1. Gerar em forma aleatória valores inteiros de dados de EI no intervalo $-L$ a $+H$ de acordo a geradores aleatórios específicos. Arrumar em blocos de 8×8 considerando conjuntos individuais de 8 números consecutivos em fila.
2. Para cada bloco de 8×8 , a DCT deve ser calculada usando 64 bit com exatidão de ponto flutuante.
3. Arredondar os 64 bit de número de ponto flutuante ao mais próximo inteiro. Logo cortar ao intervalo -2048 a $+2047$. Esse valor será a entrada de 12 bits para a Transformada Inversa.
4. Aplicar a IDCT usando pelo menos a exatidão de 64 bit de ponto flutuante. Arredondar os EI resultantes ao mais próximo inteiro e cortar ao intervalo -256 a $+255$. Esses blocos de EI de 8×8 são a referencia da saída da IDCT.
5. Para cada bloco 8×8 de 12 bit que usa o chip da IDCT proposto, cortar a saída a -256 a $+255$.
6. Os requerimentos são:
 - Para qualquer EI o erro Máximo não deve superar 1 em magnitude.
 - Para qualquer EI o erro quadrático médio não deve superar 0.06.
 - Em total, o erro quadrático médio não deve superar 0.02.
 - Para qualquer EI o erro médio não deve superar 0.015 em magnitude.
 - Em total, o erro médio não deve superar 0.0015 em magnitude.
7. Todos os zeros de entrada devem produzir zeros de saída.

BIBLIOGRAFIA

- [1] MARTINI, L.C. ; *Proposta de um Sistema de Codificação MCPD em 34 Mbit/s para Processamento de Sinais de TV em Componentes Y, C_R, C_B* , Tese de Doutorado UNICAMP, Dez. 1989.
- [2] YABU-UTI, J.B.T; *Redução da Taxa de bits para transmissão do sinal de TV composto PAL-M*, Tese de Doutorado UNICAMP, Fev. 1983.
- [3] TIE,C.F. & HU,H.C.; *Real Time Chips applied to Multimedia*, Proc. in Signal Processing 1989, 149, pp 182-188.
- [4] PLOMPEN, R.H.J.M. ; *Motion Video Coding for Visual Telephony*, PTT Research Neher Laboratorium. 1989.
- [5] NETRAVALI, A.N. & LIMB,J.O.; *Picture Coding : A Review*,Proc.IEEE, Vol.68 No.3,pp.366-406,Mar.1980.
- [6] FERNÁNDEZ,G.S. & MARTINI,L.C.; *Técnicas de Processamento Digital de Imagens*, Publicação Interna. IA-348.Dez.1990.
- [7] CLARKE, R.J.; *Transform Coding of Images* . Academic Press. 1985.
- [8] NETRAVALI, A.N. & HASKELL, B.G.; *Digital Pictures, Representation and Compression*. Plenum Press. 1989
- [9] JUNIOR, A.D. & MARTINI, L.C.; *Comparação entre a Codificação de Imagens empregando Amostragem a Frequência Sub-Nyquist e a Taxas Sub-Nyquist*, Tese de Mestrado UNICAMP, Jul. 1992.
- [10] HALL, C.F. & ANDREWS,H.C.; *Digital Color Image Compression in a perceptual Space*, Proc. SPIE 1978, 149, pp 182-188.

- [11] **HALL, C.F.**; *Perceptual Coding in the Fourier Transform Domain*, NTC Convention Record, 1980, pp 36.1.1-36.1.7.
- [12] **MARAKAN,H., HALOM,Y. & YAMAMOTO,H.**; *Comparison between DPCM and Hadamard transform coding in the composite coding of the NTSC color TV signal*, IEEE Trans.Comm., Vol. COM-30, 1982, pp 469 - 479.
- [13] **REININGER,C. & GIBSON,J.B.**; *Distribution of the two-dimensional DCT coefficients for images*, IEEE Trans.Inf.Theory, Vol. IT-14, 1983, pp.650-661.
- [14] **MANNOS,J. & SAKRINSON,D.**; *The Effects of a Visual Fidelity Criterion on the Encoding of Images*, IEEE Trans.1974,IT-20, pp 525 - 236.
- [15] **CLARKE, R.J. & TECH,B.**; *Spectral Response of the Discrete Cosine and Walsh Hadamard Transforms*, IEEE Proc., Vol 130, N.4, June 1983, pp 309-313.
- [16] **ELLIOT,D.F. & RAO,K.R.**; *Fast Transforms; Algorithms, Analysis, Applications*. Academic Press,1982.
- [17] **CHENG-TIEÇ.**; *Transform Coding of Digital Images Using Variable Block Size DCT with Adaptive Thresholding and Quantization*, Applications of Digital Image Processing XIII, 1990.SPIE Vol.1349, pp 43-54.
- [18] **JAYANT,N.S. & NOLL,P.**; *Digital Coding of Waveforms: Principles and Applications to Speech and Video*,Prentice Hall,1984.
- [19] **PRATT,W.K.**; *Digital Image Processing*, John Wiley,1978. Interscience publication.
- [20] **FERNÁNDEZ,G.S. & MARTINI,L.C.**; *Estudo de uma Codificação de Imagens Estáticas por Transformada DCT-2D*, TELEM092.X-SBT.V-SBMO, 1992. Vol.I ,pp 479-484.
- [21] **PRATT,W.K. & CHEN,W.H.**; *Slant Transform Image Coding*, IEEE Trans. Communications, Vol. COM-22, pp 1075-1093. 1974.
- [22] **FERNÁNDEZ,G.S. & MARTINI,L.C.**; *Estudio de Compresión de Imagen usando Codificación DCT-2D*, ROC&C92-IEEE,1992. Mexico.Vol.I,pp 4-12,4-16.
- [23] **SEYMOUR LIPSCHUTZ.**; *Theory and Problems of Linear Algebra*, McGraw-Hill,1968.

- [24] CHEN, A., SMITH, C.H. & FRALICK, S.C.; *A fast Computational Algorithm for the Discrete Cosine Transform*, IEEE Transactions on Communications, 25 : 1004-1009, 1977.
- [25] LEE, B.G.; *A new algorithm to compute the discrete cosine transform*, IEEE Transactions on Acoustics, Speech and Signal Processing, 32, December 1985.
- [26] VETTERLI, M. & NUSSBAUMER, H.J.; *Simple fft e dct algorithms with a reduced number of operations*, Signal Processing.6(4). July 1984.
- [27] MILT,L.; *Fast Compression Algorithm Simplifies JPEG Silicon*, Electronic Design. Oct.1991.
- [28] WINTZ,P.A.; *Transform Picture Coding*,IEEE Transactions on Information Theory, IT-17,pp. 180-197. March 1971.
- [29] LATHI,B.P.; *Randoms Signals and Communication Theory*, International Texbook Company.1968.
- [30] PAPOULIS,A.; *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York.1965.
- [31] HUFFMAN,D.; *A Method for Construction of Minimum Redundancy Codes*, In Proceeding of the IRE, Sept. 1952.
- [32] YUZO,Y. ET AL.; *Imagens Digitalizadas da SMPTE Armazenadas em Fitas Magnéticas*, Contrato TELEBRÁS 208/87, Transmiss ao Digital RT-153, Julho . 1987.
- [33] SUN,H. & GOLDBERG,M.; *Image Sequence Coding by Adaptive Vector Quantization with Variable Block Size*, Journal of Visual Comm. and Image Repres.1991. Vol.2. No.1, 32, pp. 91-99.
- [34] BHASKAR,S.K. & AZRIEL,R.; *Parallel Processing of Regions Represented by Linear Quadtrees*, Comp. Vision,Graphics, and Image Processing, Academic Press,Inc.1988. Vol.42, No.3, pp. 371-380.
- [35] CHEN,H.H. & HUANG,T.S.; *A survey of Construction and Manipulation of Octrees*, Comp. Vision,Graphics, and Image Processing, Academic Press,Inc.1988. Vol.43, No.3, pp. 409-431.
- [36] GARGANTINI,I.; *An efective way to represent Quadtrees*, Comm. ACM 25, 1982, pp. 905-910.

- [37] **ROSENFELD,A. & KAK,A.C.**; *Digital Picture Processing*, 2nd. ed., Vol. 2, Chap.11, Academic Press,Inc. New York,1982.
- [38] **WALLACE,G.K.**; *The JPEG Still Picture Compression Standard*, IEEE Trans. on Consumer Electronics, Vol. 38. No1. pp 18-34.Feb.1992.
- [39] **KAI,H. & YENG-HENG,C.**; *Partitioned Matrix Algorithms for VLSI Arithmetic Systems*,IEEE Transactions on Computers, Vol.C-31,pp. 1215-1224. December 1982.
- [40] **HENOT,J.P.**; *17 Mbit/s Algorithm for Secondary Distribution of 4.2.2 Video Signals.*, Visual Communications and Image Processing IV. SPIE, Vol. 1199. pp 1590-1598.1989.