

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA



OBTENÇÃO DE UMA MATRIZ DE
QUANTIZAÇÃO OTIMIZADA PARA IMAGENS DE
ULTRA-SONOGRAFIA E RESSONÂNCIA
MAGNÉTICA

JÚLIO CÉZAR COELHO

OUTUBRO

2005

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

OBTENÇÃO DE UMA MATRIZ DE QUANTIZAÇÃO
OTIMIZADA PARA IMAGENS DE ULTRA-SONOGRAFIA
E RESSONÂNCIA MAGNÉTICA

Dissertação apresentada por Júlio César Coelho à Universidade Federal de Uberlândia para obtenção do título de mestre em Engenharia Elétrica, aprovada em 28 de outubro de 2005 pela banca examinadora:

Professora Dr.^a Edna Lúcia Flôres, UFU (orientadora)

Professor Dr. Gilberto Arantes Carrijo, UFU (co-orientador)

Professor Dr. Antônio Cláudio Paschoarelli da Veiga, UFU

Professor Dr. Cláudio Afonso Fleury, UCG

OBTENÇÃO DE UMA MATRIZ DE QUANTIZAÇÃO
OTIMIZADA PARA IMAGENS DE ULTRA-SONOGRAFIA
E RESSONÂNCIA MAGNÉTICA

Júlio César Coelho

Dissertação apresentada por Júlio César Coelho à Universidade
Federal de Uberlândia como parte dos requisitos para obtenção
do título de mestre em Engenharia Elétrica.

Prof.^a Edna Lúcia Flôres, Dr.^a

Orientadora

Prof. Darizon Alves Andrade, PhD.

Coordenador do Curso de Pós-graduação

À Deus,
À minha família
À minha amada Carla.

“Uma coisa que aprendi ao longo
da vida: que toda nossa ciência,
comparada com a realidade, é
primitiva e infantil - e, ainda
assim, é a coisa mais preciosa
que nós temos.”

Albert Einstein

AGRADECIMENTOS

Agradeço à Deus pela vida.

À professora Edna Lúcia Flôres pela orientação, o companheirismo, a dedicação e o incentivo.

Ao professor Gilberto Arantes Carrijo a motivação e a colaboração.

Aos colegas Rodrigo, Leonardo, Arquimedes, João Manoel, Cleiber Silva que colaboraram na aquisição de conhecimento para a execução deste trabalho.

Aos amigos Tony e Dr. Eduardo Guilherme col aboradores fundamentais na aquisição e análise das imagens utilizadas neste trabalho.

Agradeço o apoio, o carinho e a compreensão do meu pai Jeová, da minha mãe Alda, do meu irmão Eduardo, das minhas irmãs Lucinéia e Juliana e da minha amada Carla.

A todos que colaboraram de forma direta ou indireta para a realização deste trabalho.

RESUMO

A quantidade de imagens médicas está crescendo, e em hospitais e clínicas onde o número de pacientes é grande, o espaço requerido para todas essas informações é muito grande, o que torna-se um problema. Uma das formas de diminuir esse problema é comprimir essas imagens. O objetivo deste trabalho é obter uma matriz de quantização para imagens de ultra-sonografia e ressonância magnética utilizando o método da média aritmética. Essa matriz deverá fornecer uma taxa de bits/pixel pequena e uma qualidade satisfatória para o diagnóstico médico dessas imagens. Para obter essa matriz de quantização foram encontradas quatro matrizes de quantização de Q2 a Q5, obtidas utilizando-se cinco imagens de ultra-sonografia e 454 imagens de ressonância magnética. Os resultados obtidos nos testes realizados neste trabalho mostraram que das matrizes de quantização de Q2 a Q5 obtidas nesta dissertação a que apresentou a menor taxa de bits/pixel foi a matriz Q5 com valor igual a 0,38081. Esta taxa é 29% menor do que a taxa de bits/pixel da matriz Q6 (matriz obtida por Hamamoto para imagens de ultra-sonografia) cujo valor é igual a 0,53638 e 37,55% menor do que a taxa de bits/pixel da matriz Q1 (matriz de quantização do padrão JPEG) cujo valor é igual a 0,61141. Como as diferenças percentuais das PSNRs das matrizes de Q2 a Q6 em relação a PSNR da matriz Q1 são menores do que 10% e como a opinião dos seis médicos consultados sobre a qualidade das imagens de ultra-sonografia e ressonância magnética reconstruídas utilizando as matrizes de Q1 a Q6 são satisfatórias para eles realizarem seus diagnósticos, pode-se considerar que essas imagens apresentam praticamente a mesma qualidade.

Portanto, a matriz de quantização Q_5 pode ser considerada como sendo a matriz de quantização otimizada para essas imagens, porque ela apresenta uma taxa de bits/pixel pequena e uma qualidade satisfatória.

Palavras Chaves: compressão de imagem, taxa de bits/pixel, relação sinal/ruído de pico, imagens de ultra-sonografia, imagens de ressonância magnética.

ABSTRACT

The amount of medical images is growing, and in hospitals and clinics, where the number of patients is high, the space requested for all those information is very big, and that becomes a problem. One of the ways to reduce this problem is to compress those images. The aim of this work is to obtain one quantization matrix for ultrasonic echo images and magnetic resonance using the arithmetic average method. This matrix should provide a small rate of bits/pixel and a satisfactory quality for the medical diagnosis of those images. In order to obtain that quantization matrix were found four quantization matrix from Q2 to Q5, obtained through the use of five ultrasonic echo images and 454 magnetic resonance images. The results obtained in the tests of this work showed that from the quantization matrix Q2 to Q5 obtained in this dissertation, one that presented the smallest rate of bits/pixel was the Q5 with value equals to 0,38081. This rate is 29 % smaller than the rate of bits/pixel of the matrix Q6 (matrix obtained by Hamamoto for ultrasonic echo images) whose value is equal to 0,53638 and 37,55 % smaller than the rate of bits/pixel of the matrix Q1 (quantization matrix of the pattern JPEG) whose value is equal to 0,61141. By the fact that the percentile differences of PSNRs of the matrix Q2 to Q6 in relation to PSNR of the matrix Q1 are smaller than 10 % and that the six consulted doctors' opinion on the quality of the ultrasonic echo images and resonance magnetic reconstructed using the matrix of Q1 to Q6 are satisfactory for them to accomplish their diagnoses, it can be considered that those images present the same quality. Therefore, the quantization matrix Q5 can be considered as being the quantization matrix

optimized for those images, because it presents a small rate of bits/pixel and a satisfactory quality.

Keywords: image compression, rate of bits/pixel, relation signal/noise of peak, ultrasonic echo image, resonance magnetic image.

Sumário

1	INTRODUÇÃO	1
1.1	Motivação	1
1.2	Levantamento Bibliográfico dos Trabalhos Desenvolvidos em Compressão de Imagens	3
1.3	Proposta deste Trabalho	6
1.4	Estrutura desta Dissertação	7
1.5	Considerações Finais deste Capítulo	8
2	COMPRESSÃO DE IMAGENS DIGITAIS	9
2.1	Introdução	9
2.2	Medidas de Fidelidade	13
2.3	Técnicas de Compressão	15
2.4	Técnicas de Compressão sem Perdas	15
2.4.1	Codificação por sequência repetitiva	16
2.4.2	Codificação por estatística	17
2.4.3	Codificação previsora sem perdas	20
2.4.4	Codificação por plano de bits	24
2.5	Técnicas de Compressão com Perdas	24
2.5.1	Codificação por truncagem de bloco	26

2.5.2	Codificação previsora com perdas	27
2.5.3	Codificação por transformada	28
2.5.4	Codificação em sub-bandas	32
2.5.5	Codificação por quantização vetorial	33
2.5.6	Codificação por fractais	33
2.6	Método de Compressão Joint Photographic Experts Group (JPEG)	35
2.7	Considerações Finais deste Capítulo	43
3	MÉTODO UTILIZADO	44
3.1	Introdução	44
3.2	Imagens Utilizadas	45
3.3	Método da Média Aritmética	53
3.4	Matriz de Quantização Obtida por Hamamoto	55
3.5	Considerações Finais deste Capítulo	56
4	RESULTADOS OBTIDOS	57
4.1	Introdução	57
4.2	Obtenção das Matrizes de Quantização	58
4.3	Resultados Obtidos	61
4.4	Conclusões	107
5	CONCLUSÕES E CONTRIBUIÇÕES	109
5.1	Conclusões	109
5.2	Contribuições deste Trabalho	112
5.3	Sugestões para Futuros Trabalhos	112

Lista de Figuras

1.1	Esquema de compressão e descompressão de uma imagem. . .	2
2.1	Principais técnicas de compressão sem perda.	16
2.2	Codificação RLE.	17
2.3	Codificação LZW.	20
2.4	Modelo de codificação previsora sem perdas: (a) codificador; (b) decodificador	22
2.5	Tipos de técnicas de compressão com perda.	26
2.6	Modelo de codificação previsora com perda (a) codificador (b) decodificador.	27
2.7	Modelo de codificação por transformada	29
2.8	Padrão de ordenação zigue-zague utilizado no método de com- pressão JPEG.	38
2.9	(a) quantizador; (b) desquantizador	39
2.10	Esquema de um quantizador uniforme.	42
2.11	Esquema de um quantizador não-uniforme.	43
3.1	(a) Imagem do fígado; (b) imagem do rim.	46
3.2	(a) Imagem da bexiga e do útero; (b) imagem do coração. . .	46

3.3	Imagem da próstata.	47
3.4	Imagens do abdome.	47
3.5	Imagens do abdome.	48
3.6	Imagens do abdome.	48
3.7	Imagens do abdome.	49
3.8	Imagens do encéfalo.	49
3.9	(a) Imagem do encéfalo; (b) imagem do joelho.	50
3.10	Imagens do joelho.	50
3.11	Imagens do joelho.	51
3.12	Imagens da coluna lombar.	51
3.13	(a) Imagem da coluna lombar; (b) imagem do tornozelo. . . .	52
3.14	Imagens do tornozelo.	52
3.15	Diagrama de blocos para obter uma matriz de quantização utilizando o método da média aritmética.	54
4.1	Imagem original de ultra-sonografia da bexiga e do útero. . . .	62
4.2	(a) Imagem de ultra-sonografia da bexiga e do útero recon- struída utilizando-se a matriz de quantização Q1; (b) imagem- erro.	62
4.3	(a) Imagem de ultra-sonografia da bexiga e do útero recon- struída utilizando-se a matriz de quantização Q2; (b) imagem- erro.	63
4.4	(a) Imagem de ultra-sonografia da bexiga e do útero recon- struída utilizando-se a matriz de quantização Q3; (b) imagem- erro.	63

4.5	(a) Imagem de ultra-sonografia da bexiga e do útero reconstruída utilizando-se a matriz de quantização Q4; (b) imagem-erro.	64
4.6	(a) Imagem de ultra-sonografia da bexiga e do útero reconstruída utilizando-se a matriz de quantização Q5; (b) imagem-erro.	64
4.7	(a) Imagem de ultra-sonografia da bexiga e do útero reconstruída utilizando-se a matriz de quantização Q6; (b) imagem-erro.	65
4.8	Gráfico da média da taxa de bits/pixel utilizando-se as matrizes de quantização de Q1 a Q6 para as cinco imagens de ultra-sonografia usadas neste trabalho.	66
4.9	Gráfico da média da PSNR utilizando-se as matrizes de quantização de Q1 a Q6 para as cinco imagens de ultra-sonografia usadas neste trabalho.	66
4.10	Imagem original do abdome.	69
4.11	(a) Imagem do abdome reconstruída utilizando-se a matriz de quantização Q1; (b) imagem-erro.	69
4.12	(a) Imagem do abdome reconstruída utilizando-se a matriz de quantização Q2; (b) imagem-erro.	70
4.13	(a) Imagem do abdome reconstruída utilizando-se a matriz de quantização Q3; (b) imagem-erro.	70
4.14	(a) Imagem do abdome reconstruída utilizando-se a matriz de quantização Q4; (b) imagem-erro.	71

4.15	(a) Imagem do abdome reconstruída utilizando-se a matriz de quantização Q5; (b) imagem-erro.	71
4.16	(a) Imagem do abdome reconstruída utilizando-se a matriz de quantização Q6; (b) imagem-erro.	72
4.17	Gráfico da média da taxa de bits/pixel utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens do abdome usadas neste trabalho.	73
4.18	Gráfico da PSNR utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens do abdome usadas neste trabalho.	73
4.19	Imagem original do encéfalo.	76
4.20	(a) Imagem do encéfalo reconstruída utilizando-se a matriz de quantização Q1; (b) imagem-erro.	76
4.21	(a) Imagem do encéfalo reconstruída utilizando-se a matriz de quantização Q2; (b) imagem-erro.	77
4.22	(a) Imagem do encéfalo reconstruída utilizando-se a matriz de quantização Q3; (b) imagem-erro.	77
4.23	(a) Imagem do encéfalo reconstruída utilizando-se a matriz de quantização Q4; (b) imagem-erro.	78
4.24	(a) Imagem do encéfalo reconstruída utilizando-se a matriz de quantização Q5; (b) imagem-erro.	78
4.25	(a) Imagem do encéfalo reconstruída utilizando-se a matriz de quantização Q6; (b) imagem-erro.	79
4.26	Gráfico da média da taxa de bits/pixel utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens do encéfalo usadas neste trabalho.	80

4.27 Gráfico da média da PSNR utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens do encéfalo usadas neste trabalho.	80
4.28 Imagem original do joelho.	83
4.29 (a) Imagem do joelho reconstruída utilizando-se a matriz de quantização Q1; (b) imagem-erro.	83
4.30 (a) Imagem do joelho reconstruída utilizando-se a matriz de quantização Q2; (b) imagem-erro.	84
4.31 (a) Imagem do joelho reconstruída utilizando-se a matriz de quantização Q3; (b) imagem-erro.	84
4.32 (a) Imagem do joelho reconstruída utilizando-se a matriz de quantização Q4; (b) imagem-erro.	85
4.33 (a) Imagem do joelho reconstruída utilizando-se a matriz de quantização Q5; (b) imagem-erro.	85
4.34 (a) Imagem do joelho reconstruída utilizando-se a matriz de quantização Q6; (b) imagem-erro.	86
4.35 Gráfico da média da taxa de bits/pixel utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens do joelho usadas neste trabalho.	87
4.36 Gráfico da PSNR utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens do joelho usadas neste trabalho. . .	87
4.37 Imagem original da coluna lombar.	90
4.38 (a) Imagem da coluna lombar reconstruída utilizando-se a matriz de quantização Q1; (b) imagem-erro.	90

4.39	(a) Imagem da coluna lombar reconstruída utilizando-se a matriz de quantização Q2; (b) imagem-erro.	91
4.40	(a) Imagem da coluna lombar reconstruída utilizando-se a matriz de quantização Q3; (b) imagem-erro.	91
4.41	(a) Imagem da coluna lombar reconstruída utilizando-se a matriz de quantização Q4; (b) imagem-erro.	92
4.42	(a) Imagem da coluna lombar reconstruída utilizando-se a matriz de quantização Q5; (b) imagem-erro.	92
4.43	(a) Imagem da coluna lombar reconstruída utilizando-se a matriz de quantização Q6; (b) imagem-erro.	93
4.44	Gráfico da média da taxa de bits/pixel utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens da coluna lombar usadas neste trabalho.	94
4.45	Gráfico da média da PSNR utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens da coluna lombar usadas neste trabalho.	94
4.46	Imagem original do tornozelo.	97
4.47	(a) Imagem do tornozelo reconstruída utilizando-se a matriz de quantização Q1; (b) imagem-erro.	97
4.48	(a) Imagem do tornozelo reconstruída utilizando-se a matriz de quantização Q2; (b) imagem-erro.	98
4.49	(a) Imagem do tornozelo reconstruída utilizando-se a matriz de quantização Q3; (b) imagem-erro.	98
4.50	(a) Imagem do tornozelo reconstruída utilizando-se a matriz de quantização Q4; (b) imagem-erro.	99

4.51	(a) Imagem do tornozelo reconstruída utilizando-se a matriz de quantização Q5; (b) imagem-erro.	99
4.52	(a) Imagem do tornozelo reconstruída utilizando-se a matriz de quantização Q6; (b) imagem-erro.	100
4.53	Gráfico da média da taxa de bits/pixel utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens do tornozelo usadas neste trabalho.	101
4.54	Gráfico da média da PSNR utilizando-se as matrizes de quantização de Q1 a Q6 para as imagens do tornozelo usadas neste trabalho.	101
4.55	Gráfico da média das taxas de bits/pixel utilizando-se as matrizes de quantização de Q1 a Q6 para todas as imagens de ultra-sonografia e ressonância magnética usadas neste trabalho.	105
4.56	Gráfico da média das PSNRs das matrizes de quantização de Q1 a Q6 para todas as imagens de ultra-sonografia e ressonância magnética usadas neste trabalho.	105

Lista de Tabelas

1.1	Recursos computacionais utilizados nesta dissertação.	7
2.1	Matriz de quantização de luminância típica do padrão JPEG.	37
2.2	Exemplo de um livro-código com 4 palavras-código	40
3.1	Matriz de quantização obtida por Hamamoto para imagens de ultra-sonografia.	56
4.1	Matriz de quantização Q2.	58
4.2	Matriz de quantização Q3.	59
4.3	Matriz de quantização Q4.	60
4.4	Matriz de quantização Q5.	60
4.5	Média dos resultados obtidos na compressão das cinco imagens de ultra-sonografia utilizadas neste trabalho.	67
4.6	Média dos resultados obtidos na compressão das imagens do abdome utilizadas neste trabalho.	74
4.7	Média dos resultados obtidos na compressão de imagens de encéfalo utilizadas neste trabalho.	81
4.8	Média dos resultados obtidos na compressão de imagens de joelho utilizadas neste trabalho.	88

4.9	Média dos resultados obtidos na compressão das imagens da coluna lombar utilizadas neste trabalho.	95
4.10	Média dos resultados obtidos na compressão das imagens de tornozelo utilizadas neste trabalho.	102
4.11	Média dos resultados obtidos na compressão de todas as im- agens de ultra-sonografia e ressonância magnética utilizadas neste trabalho.	106

Capítulo 1

INTRODUÇÃO

1.1 Motivação

A evolução da eletrônica, a popularização dos computadores e o crescimento da Internet, fizeram com que cada residência, comércio e indústria se tornassem uma fonte contínua de informações. Esses bilhões de fontes geram quantidades astronômicas de informações. O fluxo, o gerenciamento e o armazenamento dessas informações são problema, porque a Internet é um sistema distribuído e cada parte desse sistema tem seus proprietários, onde muitos deles não têm interesse em melhorar ou trocar seus equipamentos por equipamentos mais rápidos e de maior capacidade, para diminuir o congestionamento e o atraso da chegada de informações é preciso representá-las de forma mais eficiente, ou seja, com menor quantidade de bits, assim as informações chegam mais rápido ao destino e com menor possibilidade de congestionamento.

O processo de representar a informação de forma eficiente pela redução ou eliminação da redundância existente nela é chamado compressão. A compressão é realizada antes de transmitir ou armazenar uma imagem e posteriormente para a visualização, a imagem é reconstruída. O resultado da

descompressão é uma imagem idêntica ou próxima da imagem original. A Figura 1.1 mostra o processo de compressão e de descompressão de uma imagem.

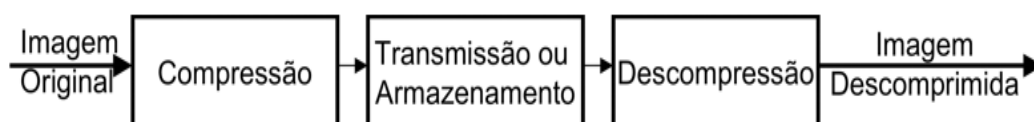


Figura 1.1: Esquema de compressão e descompressão de uma imagem.

Os três tipos de arquivos que mais congestionam a Internet, por causa de seus tamanhos são: arquivos de som, de vídeo e de imagem. A crescente utilização desses tipos de arquivos reforça a necessidade do uso do processo de compressão.

O comércio eletrônico cresce a cada dia, e nesse tipo de aplicação, arquivos de imagem são transmitidos via Internet, esses arquivos são fotos de produtos que estão disponíveis para venda, isto faz com que a imagem seja cada vez mais necessária e seu uso crescente exige o processo de compressão de imagens.

Outro tipo de arquivo com uso crescente são os arquivos de vídeos que são utilizados na Internet para comunicação (video-fone) e a transmissão de imagens de câmera de segurança. Eles reforçam mais a necessidade de utilizar a compressão de imagens.

Em aplicações onde é necessário o armazenamento de imagens, deve-se usar compressão de imagens, para reduzir os custos, fornecer maior rapidez ao processo e diminuir o espaço necessário para se armazenar tal quantidade de informações.

Outra área onde o uso do computador é bastante necessário é a medicina. Os computadores e aparelhos fornecem precisão e eficiência a essa área, trazem facilidades para os diagnósticos através dos aparelhos computadorizados que fornecem a imagem do corpo humano por dentro e por fora. Esta ligação entre a medicina e a eletrônica cresce a cada dia, e diagnósticos por imagem estão sendo usados para se saber realmente o que está acontecendo com o paciente.

Por causa de todos estes fatores, citados anteriormente neste item, a quantidade de imagens está crescendo, e em hospitais onde o número de pacientes é grande, o espaço requerido para todas essas informações deve ser muito grande, o que torna-se um problema. Uma das formas de diminuir esse problema é comprimir as imagens. Como citado por Hamamoto [1], a quantidade de imagens radiológicas capturadas por ano só nos Estados Unidos é da ordem de pentabytes. Este é um número considerável, e é necessário usar a compressão nessas imagens.

Este capítulo apresenta um levantamento bibliográfico dos trabalhos desenvolvidos em compressão de imagens, a proposta e a estrutura desta dissertação.

1.2 Levantamento Bibliográfico dos Trabalhos Desenvolvidos em Compressão de Imagens

Wallace [2] descreveu o padrão JPEG fornecendo uma visão geral do conjunto de operações que ocorrem na compressão e na descompressão desse

padrão.

Sherlock e outros [3] apresentaram uma estratégia para gerar tabelas de quantização JPEG ótimas que se aproximavam de uma taxa de compressão alvo. Eles usaram um modelo para expressar os coeficientes de quantização como função da taxa de compressão e de suas posições na tabela de quantização.

Kam e outros [4] desenvolveram um método de aproximação que garante a convergência em até oito iterações de um algoritmo iterativo para tabelas de quantização JPEG.

Neto [5] apresentou os aspectos da compressão sem perdas e sua importância na redução da grande quantidade de dados presentes nas imagens médicas digitais, viabilizando assim, o seu armazenamento e sua transmissão em redes de computadores.

Eskicioglu [6] tentou melhorar as medidas de qualidade incluindo modelos simples do sistema visual humano e ferramentas multi-dimensionais. Ele revisou de 1974 a 1999 os critérios de qualidade de imagens monocromáticas comprimidas.

Konstantinides e outros [7] apresentaram uma nova técnica para agudizar imagens comprimidas usando o padrão JPEG. A agudização da imagem foi atingida escalonando-se adequadamente cada elemento da tabela de quantização para realçar as características de alta frequência da imagem.

Fan e Queiroz [8] apresentaram um método para a estimação da máxima verossimilhança das tabelas de quantização do padrão JPEG.

Chang e outros [9] encontraram uma tabela de quantização para o codificador JPEG baseada no sistema visual humano. Com essa tabela, eles

obtiveram melhoramentos na PSNR de aproximadamente 0,2 dB a 2 dB, sem aumentar a complexidade do codificador e do decodificador.

Hamamoto [1] padronizou a tabela de quantização do padrão JPEG para imagens de ultra-sonografia utilizando o método da média aritmética. Os resultados obtidos por ele forneceram uma menor taxa de bits/pixel comparada com a do padrão JPEG, com a mesma qualidade das imagens.

Cruz e Ebrahimi [10] estudaram o padrão JPEG 2000 versus os padrões JPEG-LS, MPEG-4VTC e JPEG. Eles concentram o estudo na eficiência da compressão e na complexidade e funcionalidade desses padrões.

Arcangelo e Massino [11] propuseram um novo algoritmo para controlar o fator de compressão ao utilizar o padrão de compressão JPEG.

Ribas [12] estudou e comparou a compressão de imagens digitais utilizando transformadas e a técnica de quantização vetorial.

Fleury [13] desenvolveu a quantização vetorial classificada adaptada perceptivamente para melhorar os resultados obtidos ao utilizar a quantização vetorial de taxa variável.

Della Coleta [14] verificou o desempenho da transformada vetorial seguida da quantização vetorial utilizando o método de Linde-Buzo-Gray. Os resultados obtidos levaram-no a recomendar a utilização da transformada bidimensional com quadro de tamanho no mínimo igual a oito, obtendo taxas de compressão da ordem de 20:1 com qualidade da imagem reconstruída satisfatória.

1.3 Proposta deste Trabalho

A proposta deste trabalho é obter uma matriz de quantização para imagens de ultra-sonografia e ressonância magnética utilizando o método da média aritmética. Essa matriz deverá fornecer uma taxa de bits/pixel pequena e uma qualidade satisfatória para o diagnóstico médico dessas imagens. Para obter essa matriz serão encontradas algumas matrizes de quantização obtidas utilizando-se cinco imagens de ultra-sonografia e 454 imagens de ressonância magnética. Na obtenção dessas matrizes serão utilizadas as cinco imagens de ultra-sonografia, algumas imagens de ressonância escolhidas aleatoriamente e algumas imagens de ressonância magnética que caracterizem a diversidade desse conjunto de imagens. Em todas essas matrizes serão utilizados um fator de ajuste em seus coeficientes.

Das matrizes de quantização obtidas será escolhida a matriz que apresenta a menor taxa de bits/pixel e uma qualidade satisfatória para o diagnóstico médico das imagens de ultra-sonografia e ressonância magnética. Estas características da matriz escolhida serão comparadas com as mesmas características da matriz de quantização obtida por Hamamoto para imagens de ultra-sonografia [1] e também com as mesmas características da matriz de quantização do padrão JPEG [2]. Dessa comparação será respondida a pergunta: qual é a matriz de quantização otimizada para imagens de ultra-sonografia e ressonância magnética, a encontrada neste trabalho com uma taxa de bits/pixel pequena e uma qualidade satisfatória para o diagnóstico médico dessas imagens, a matriz de quantização obtida por Hamamoto para imagens de ultra-sonografia ou a matriz de quantização do padrão JPEG?

A Tabela 1.1 mostra os recursos computacionais utilizados nesta dis-

sertação.

Tabela 1.1: Recursos computacionais utilizados nesta dissertação.

CPU	Pentium III 933Mhz, 512MB RAM
Sistema Operacional	Windows Millennium
Linguagem de Programação	MatLab 6.5
Dispositivos de Saída	Monitor TCE 14" e Impressora HP PSC 1315

1.4 Estrutura desta Dissertação

Este capítulo apresenta um levantamento bibliográfico dos trabalhos desenvolvidos em compressão de imagens e a proposta deste trabalho.

O capítulo 2 descreve as medidas de fidelidade em compressão de imagens e as técnicas de compressão sem perdas e com perdas. Apresenta o padrão de compressão JPEG. E finalmente são realizadas considerações finais desse capítulo.

O capítulo 3 apresenta o método usado neste trabalho para obter a matriz de quantização utilizando a média aritmética e mostra a matriz de quantização obtida por Hamamoto para imagens de ultra-sonografia [1]. E finalmente são realizadas as considerações finais desse capítulo.

O capítulo 4 descreve os resultados obtidos nos testes realizados para obter a matriz de quantização para imagens de ultra-sonografia e ressonância magnética utilizando o método da média aritmética. Finalmente, são realizadas conclusões sobre os resultados obtidos neste teste.

Finalmente, o capítulo 5 apresenta as conclusões, as contribuições deste

trabalho, as sugestões para futuros trabalhos.

1.5 Considerações Finais deste Capítulo

Este capítulo apresentou um levantamento bibliográfico dos trabalhos desenvolvidos em compressão de imagens e a proposta deste trabalho.

Capítulo 2

COMPRESSÃO DE IMAGENS DIGITAIS

2.1 Introdução

A compressão de imagens é um conjunto de técnicas que visam reduzir a quantidade necessária de dados para representar uma imagem digital [15]. O processo de compressão consiste na remoção dos dados redundantes existentes na imagem, preservando a qualidade visual da imagem reconstruída o mais próximo possível da original. Quando a imagem reconstruída é idêntica à imagem original, considera-se que a compressão é sem perdas, caso contrário, a compressão é considerada com perdas.

Os fundamentos básicos para compressão iniciaram na década de 1940 com a teoria da Informação de C. E. Shannon [15] e, desde então, vem se desenvolvendo significativamente. Quando uma quantidade mínima de dados é necessária e suficiente para representar uma determinada quantidade de informação considera-se que não existe redundância de dados.

O conceito de redundância de dados é fundamental no estudo da compressão de imagens. A compressão é um problema de otimização que con-

siste em preservar a quantidade máxima de informação em uma quantidade mínima de dados.

A redundância de dados pode ser quantificada considerando n_1 e n_2 o número de unidades de transporte de informação em dois conjuntos de dados D_1 e D_2 , respectivamente, que representam a mesma informação. A redundância de D_1 relativa a D_2 é obtida pela Equação (2.1).

$$R_D = 1 - \frac{1}{C_R} \quad (2.1)$$

onde a taxa de compressão C_R é obtida pela Equação (2.2).

$$C_R = \frac{n_1}{n_2} \quad (2.2)$$

Pode-se observar na Equação (2.1) que, quanto maior o valor de C_R , maior a redundância da imagem original.

Pode-se verificar na Equação (2.2) que, quanto menor n_2 , maior a taxa de compressão.

Na prática, os valores de C_R são números positivos e os valores de R_D são menores do que um.

Em compressão de imagens existem três tipos de redundância de dados: redundância de codificação, redundância interpixel e redundância psico-visual [15]. Consegue-se a compressão de dados reduzindo ou eliminando uma ou mais dessas redundâncias.

Redundância de codificação

A redundância de codificação consiste no fato dos níveis de cinza de uma imagem serem codificados usando mais símbolos de codificação do que o necessário para representar cada nível de cinza. Geralmente, a redundância de codificação ocorre quando os códigos atribuídos a um conjunto de eventos são escolhidos sem levar em consideração as probabilidades dos eventos.

Redundância interpixel

A redundância interpixel está relacionada com as relações geométricas ou estruturais entre os objetos contidos na imagem. Dependendo do grau de correlacionamento entre os pixels adjacentes da imagem, o valor de qualquer pixel pode ser razoavelmente previsto com base nos valores dos seus pixels vizinhos. Essa redundância é chamada de redundância interpixel.

Redundância psico-visual

A redundância psico-visual está relacionada com o tratamento das informações visuais realizado pelo cérebro e pelo sistema visual. A visão permite a percepção e o reconhecimento de objetos à distância sem nenhum contato físico [16]. Segundo Kovacs [17], o sistema visual é o mais complexo, o mais estudado e o menos entendido entre os sistemas sensoriais. A percepção visual é um fenômeno de grande complexidade que depende de vários outros fatores além da luz refletida pelo objeto observado. Em síntese, o olho funciona como uma máquina fotográfica dotada de um sofisticado sistema de processamento que constrói a imagem percebida, em tempo real e de forma gradativa, com base na luz refletida pelo objeto e nas experiências

visuais acumuladas. Desta forma, o olho não responde com a mesma intensidade a todas as informações visuais recebidas. No processamento das informações para formação da imagem no cérebro, algumas informações são relativamente mais importantes do que as outras. As informações menos importantes podem ser eliminadas sem sacrificar significativamente a qualidade e a percepção da imagem. Essas informações são chamadas de psicovisualmente redundantes [15].

No primeiro estágio do processo visual, o observador procura características discriminadoras globais, como bordas ou regiões de texturas e as combinam em agrupamentos reconhecíveis. No segundo estágio, o cérebro correlaciona esses agrupamentos com o conhecimento prévio, com a finalidade de completar o processo de interpretação da imagem [17].

A redundância psico-visual está relacionada com a informação quantificável. A informação em si não é essencial para o processo visual normal, podendo por isso ser eliminada sem causar prejuízo significativo ao processo visual.

Quanto menos redundância mais eficiente é a compressão. Para calcular a eficiência na representação de uma imagem, utiliza-se a taxa de bits/pixel.

A taxa de bits/pixel (n) é a quantidade de bits que são necessários para representar um pixel. Essa taxa é obtido pela Equação (2.3).

$$n = \frac{n_t}{M \cdot N} \quad (2.3)$$

onde:

n_t - número total de bits da imagem.

M - número de colunas da imagem.

N - número de linhas da imagem.

Pode-se observar na Equação (2.3) que o número de bits/pixel depende das dimensões da imagem, e quanto menos bits são usados para representar uma imagem, menor é a taxa de bits/pixel, ou seja, quanto maior a compressão, menor é o número de bits/pixel.

Este capítulo descreve as medidas de fidelidade em compressão de imagens e as técnicas de compressão sem perdas e com perdas, apresenta o padrão de compressão JPEG, finalmente, são realizadas as considerações finais deste capítulo.

2.2 Medidas de Fidelidade

Como citado por Gonzalez [15], as técnicas utilizadas para compressão de imagens causam alguma degradação na imagem reconstruída. Para estimar o grau dessa degradação, existem duas classes de critérios, denominadas critério de fidelidade objetivo e critério de fidelidade subjetivo. O critério de fidelidade é objetivo quando o nível de perda da informação é expresso como função da imagem original e da imagem reconstruída.

A medida de fidelidade objetiva utilizada neste trabalho para verificar a qualidade obtida no processo de compressão é a relação sinal-ruído de pico (PSNR). Esta medida compara as imagens reconstruída e original utilizando como base o erro médio quadrático (MSE). Este erro é calculado pela Equação (2.4).

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (x_{i,j} - \hat{x}_{i,j})^2 \quad (2.4)$$

onde:

M - número de colunas da imagem

N - número de linhas da imagem

$x_{i,j}$ - valor de nível de cinza do pixel de coordenada i e j da imagem original.

$\hat{x}_{i,j}$ - valor de nível de cinza do pixel de coordenada i e j da imagem reconstruída.

Pode-se observar na Equação (2.4) que MSE é zero se a imagem reconstruída é igual a imagem original.

A relação sinal-ruído de pico (PSNR) é uma taxa em escala logaritma, que relaciona o nível máximo de cinza ao quadrado e o MSE. A PSNR é obtida pela Equação (2.5).

$$PSNR = 10 \cdot \log \left(\frac{(2^n - 1)^2}{MSE} \right) dB \quad (2.5)$$

onde:

n - número de bits/pixel.

Pode-se observar na Equação (2.5) que quanto menor o MSE, maior é a PSNR, mais próxima a imagem reconstruída está da imagem original.

O critério é considerado de fidelidade subjetivo quando a medida da qualidade da imagem reconstruída é feita por avaliações subjetivas de um observador humano. Para isto utiliza-se uma escala apropriada de notas absolutas ou comparações, lado a lado, das imagens originais $f(x, y)$ e reconstruída $\hat{f}(x, y)$.

Neste trabalho utiliza-se a PSNR como critério objetivo e a opinião de seis médicos como critério subjetivo para avaliar a qualidade da imagem reconstruída e para verificar a eficiência da compressão é usada a taxa de bits/pixel.

2.3 Técnicas de Compressão

As técnicas de compressão de imagens podem ser divididas em dois grupos: compressão sem perdas e compressão com perdas.

2.4 Técnicas de Compressão sem Perdas

A compressão de imagens sem perdas (ou compactação) é caracterizada pelo fato da imagem reconstruída ser idêntica a imagem original. Em diversas aplicações, por razões de ordem legal, por conveniência ou por exigência da aplicação, esta é a única técnica aceitável de compressão. Alguns tipos de imagens médicas, textos legais e códigos de programas são exemplos de aplicações que exigem compressão sem perdas. As técnicas tradicionais de compressão sem perdas suportam taxas de compressão relativamente baixas, da ordem de 10:1 no máximo, enquanto na compressão por fractal essas taxas podem atingir até a ordem de 10.000:1 [18, 19].

Um método simples e eficiente de compressão sem perdas é a codificação por tamanho variável. Nesta codificação são atribuídas as menores palavras-código (menores em quantidade de bits) aos níveis de cinza com maior probabilidade de ocorrência. Devido a isso, o código não é fixo, e seu

tamanho é variável. Com isso diminui-se a quantidade de bits necessária para o armazenamento de uma imagem e a transmissão torna-se mais rápida.

As principais técnicas de compressão sem perdas são: codificação por sequência repetida, codificação por estatística, codificação previsora sem perdas e codificação por plano de bits. A Figura 2.1 resume essas técnicas.

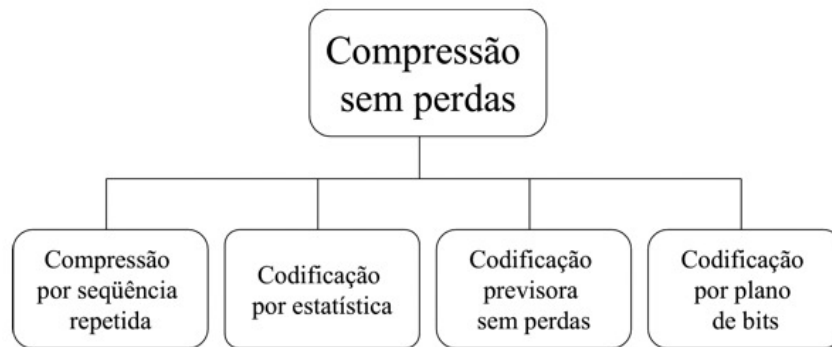


Figura 2.1: Principais técnicas de compressão sem perda.

2.4.1 Codificação por sequência repetitiva

A codificação por sequência repetitiva é a forma mais simples de obter a compressão de imagens sem perdas, explorando apenas a redundância da codificação. Um exemplo dessa técnica é a codificação Run-Lenght Encoding (RLE). Esta codificação consiste em identificar, na fonte original (entrada), cada sequência de valores repetidos e substituí-la, na saída codificada (destino), por um par de valores da forma (valor repetido, número de repetições). A decodificação RLE faz o processo inverso, substituindo cada ocorrência do par (valor repetido, número de repetições) pelo valor repetido quantas vezes

forem as unidades do número de repetições. A Figura 2.2 mostra essa técnica.

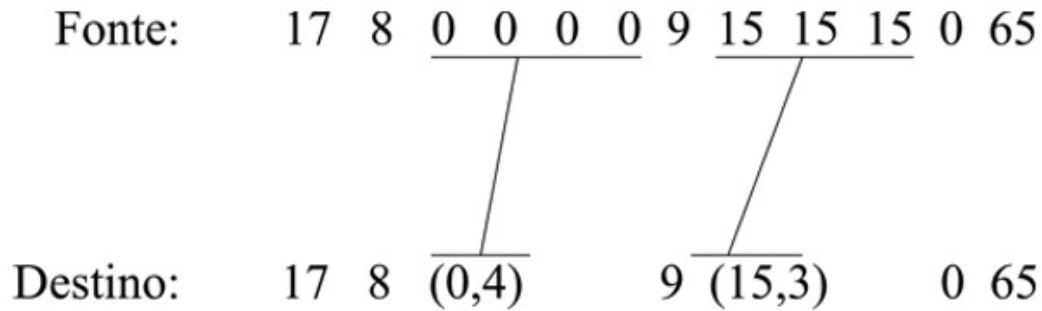


Figura 2.2: Codificação RLE.

2.4.2 Codificação por estatística

A codificação por estatística é caracterizada pela necessidade do levantamento das probabilidades dos símbolos da fonte. Para explorar a redundância de codificação, codifica-se os valores de níveis de cinza atribuindo-se palavras-código mais curtas aos símbolos mais prováveis. A Teoria da Informação fornece os subsídios necessários para a representação e a manipulação da informação quantitativamente. A premissa fundamental, dessa teoria diz que a geração da informação pode ser modelada como um processo probabilístico, no qual um evento aleatório E , com probabilidade $P(E)$, contém $I(E)$ unidades de informação. $I(E)$ é obtido pela Equação (2.6).

$$I(E) = \log \frac{1}{P(E)} = -\log P(E) \quad (2.6)$$

A entropia ($H(z)$) de uma fonte de informação A , com alfabeto $A = \{a_1, a_2, \dots, a_j\}$ e probabilidade dos símbolos $\{P(a_1), P(a_2), \dots, P(a_j) | \sum_{i=1}^j a_i = 1\}$, é definida pela Equação (2.7).

$$H(z) = \sum_{i=1}^j P(a_i) \log P(a_i) \quad (2.7)$$

onde:

z é o vetor $[P(a_1), P(a_2), \dots, P(a_j) | \sum_{i=1}^j a_i = 1]^T$.

O conjunto finito definido pelo par (A, z) descreve completamente a fonte de informação A .

Definido um modelo de fonte particular, calcula-se a entropia da imagem com base nesse modelo, de acordo com uma probabilidade pré-definida [15].

As codificações de Huffman, aritmética e Lempel-Ziv-Welch (LZW) são exemplos de codificação por estatística. As principais características de cada uma dessas codificações são:

- **Codificação de Huffman:** esta técnica foi criada em 1951 por D. A. Huffman. Ela foi a pioneira em explorar a codificação por redundância mínima para compressão de dados. Esta codificação sugere que, conhecendo-se a frequência de cada símbolo em um conjunto de dados, pode-se representar os símbolos de tal forma que os dados necessitem de menos espaço. Para isso, codificam-se os símbolos que ocorrem com maior frequência com menos bits do que aqueles que ocorrem com

menor frequência [20].

- **Codificação aritmética:** nesta técnica, toda uma sequência de símbolos-fonte é atribuída a uma única palavra do código aritmético. A própria palavra-código define um sub-intervalo do intervalo unitário $[0,1]$ de números reais. À medida em que o número de símbolos da palavra aumenta, o sub-intervalo usado para representá-la reduz, e o número de bits da informação necessário para representá-la aumenta. Cada símbolo-fonte reduz o tamanho do sub-intervalo de acordo com sua probabilidade de ocorrência [15].
- **Codificação Lempel-Ziv-Welch (LZW):** esta técnica foi desenvolvida por Abraham Lempel e Jacob Ziv, aperfeiçoada por Terry A. Welch, e publicada em 1977. Em 1985 foi patenteada pela Sperry Corporation. A compressão LZW usa um dicionário ou uma tabela de códigos para armazenar as sequências de símbolos que ocorrem com maior frequência. Em geral, a tabela de códigos (ou dicionário) é constituída de 4096 entradas. Neste caso, os dados codificados consistem de códigos de 12 bits, cada um referenciando uma das entradas da tabela. No início do processo de compressão, é montada a tabela de códigos com as 256 primeiras entradas (0-255) inicializadas com valores idênticos (0-255); o restante da tabela (256-4095) é inicializada com brancos. A codificação LZW usa as entradas 256-4095 para representar sequências de bytes [21]. Por exemplo, o código 257 pode representar 145 201 73; cada vez que o algoritmo de compressão encon-

tra essa sequência, na fonte, o código 257 é colocado na saída (destino), como mostrado na Figura 2.3.

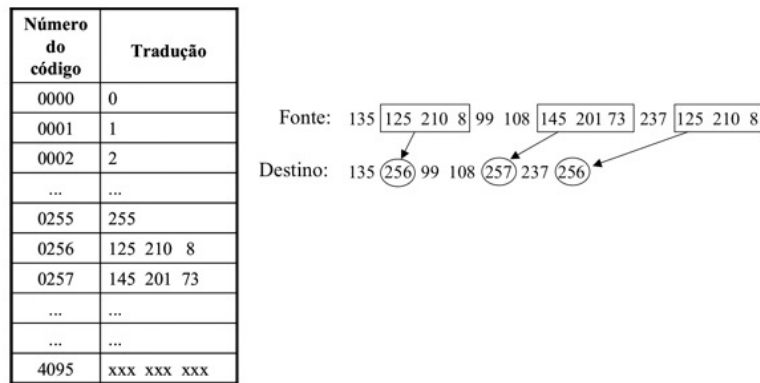


Figura 2.3: Codificação LZW.

A descompressão é realizada fazendo a tradução direta da tabela de códigos.

2.4.3 Codificação previsora sem perdas

A codificação previsora sem perdas é baseada na redundância inter-pixels de pixels pouco espaçados. Para isto, extrai-se e codifica-se apenas a informação definida em cada pixel como a diferença entre o valor do pixel real e um valor previsto daquele pixel [15]. Tanto o codificador como o decodificador possui um predictor idêntico. À medida em que a codificação (ou decodificação) avança sobre a imagem fonte f_n , o predictor calcula um valor antecipado daquele pixel em função dos valores de um determinado número de pixels anteriores. Esse valor estimado é arredondado para o inteiro mais próximo, representado por $\hat{f}_n$, usado para formar o erro de previsão, obtido

pela Equação (2.8).

$$e_n = f_n - \hat{f}_n \quad (2.8)$$

O erro e_n é usado para gerar o próximo elemento dos dados codificados f_n , utilizando uma codificação de tamanho variável.

O processo de decodificação reconstrói f_n , a partir do código de tamanho variável, realizando a operação inversa, como mostrado na Equação (2.9).

$$f_n = e_n + \hat{f}_n \quad (2.9)$$

A Figura 2.4 [15], mostra a codificação e a decodificação previsora sem perdas.

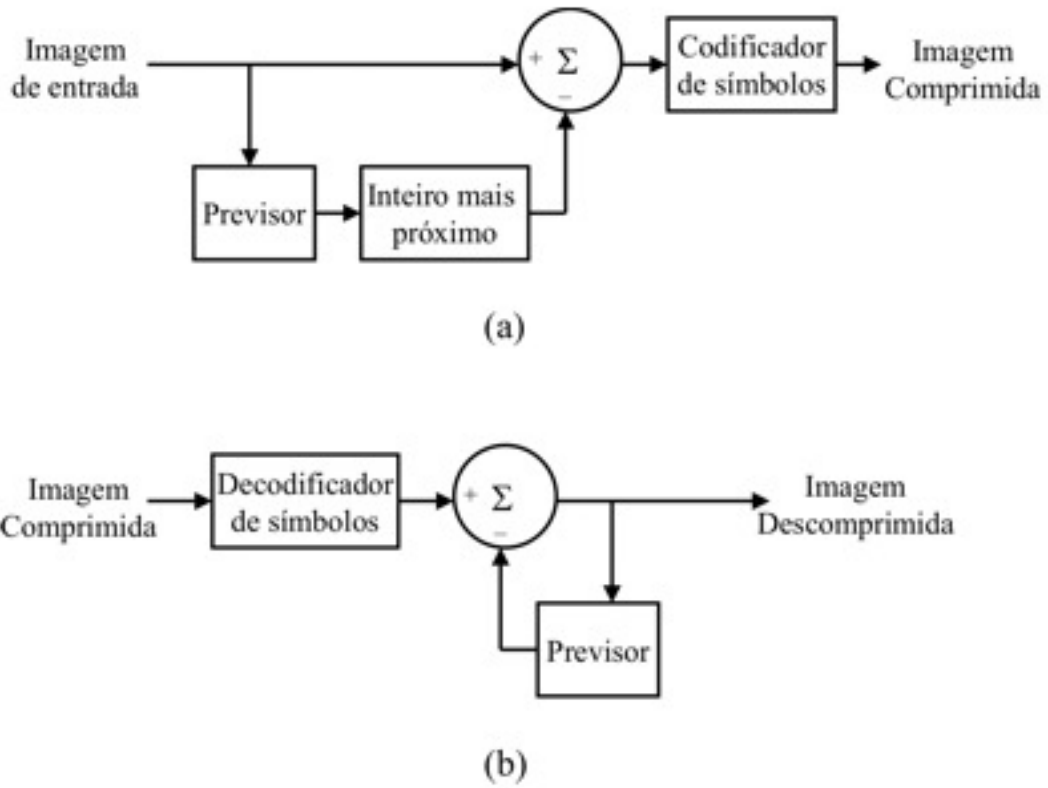


Figura 2.4: Modelo de codificação preditiva sem perdas: (a) codificador; (b) decodificador

O exemplo mais comum de codificação preditiva sem perdas é a codificação Differential Pulse Code Modulation (DPCM). Nesta situação o valor de cada pixel da imagem original f_n , com exceção do primeiro e do último pixel, é estimado com base no valor dos pixels vizinhos para obter a imagem prevista $\hat{f}_n$. A imagem diferencial (ou erro de previsão d_n) é obtida calculando a diferença entre os valores f_n e $\hat{f}_n$ [12], pela Equação (2.10).

$$d_n = f_n - \hat{f}_n \quad (2.10)$$

Os valores dos pixels da imagem diferencial d_n são menores do que os valores dos pixels da imagem original f_n , assim, cada pixel de d_n necessita de um número menor de bits do que cada pixel de f_n . Este fato garante que a imagem inteira d_n seja muito menor do que a imagem original f_n .

A decodificação DPCM é realizada adicionando o erro de previsão d_n ao valor previsto da amostra $\hat{f}_n$, como mostrado na Equação (2.11).

$$f_n = d_n + \hat{f}_n \quad (2.11)$$

onde $\hat{f}_n$ pode ser calculado usando métodos locais, globais e adaptativos. Na maioria dos casos, uma combinação linear dos m pixels anteriores é uma boa previsão para $\hat{f}_n$, como mostrado na Equação (2.12).

$$\hat{f}_n = \text{round} \left[\sum_{i=1}^m \alpha_i f(x, y - 1) \right] \quad (2.12)$$

onde m é a ordem do previsor linear, round é uma função de arredondamento para o inteiro mais próximo, e os α_i , para $i = 1, 2, \dots, m$ são os coeficientes de previsão.

2.4.4 Codificação por plano de bits

A técnica de codificação por plano de bits consiste em decompor a imagem original (níveis de cinza ou colorida) em tantas imagens binárias quantos forem os níveis de cinza ou de cor. Por exemplo, uma imagem colorida (padrão RGB), com 256 níveis para cada cor (8 bits), pode ser decomposta em 768 (3 x 256) planos de 1 bit, com as mesmas dimensões da imagem original. Cada imagem binária resultante da decomposição pode ser codificada utilizando-se uma técnica de codificação sem perdas.

2.5 Técnicas de Compressão com Perdas

A codificação com perdas (ou compressão) é caracterizada pelo fato que a imagem reconstruída sofre uma perda tolerável da informação; em compensação, proporciona taxas de compressão maiores do que a codificação sem perdas. Essa técnica é útil em muitas aplicações nas quais uma certa quantidade de erro é aceitável em benefício de uma maior taxa de compressão, como por exemplo, transmissão de televisão, video-conferência, transmissão por fac-símile e por Internet.

Em geral, as técnicas tradicionais de compressão com perdas obtém imagens reconstruídas de boa qualidade visual com taxas de compressão em torno de 20:1 [12]; excepcionalmente, técnicas mais sofisticadas, como fractais, podem atingir taxas bem altas, da ordem de 10.000:1 [18, 19].

Codificação visualmente sem perdas

Uma imagem reconstruída é chamada visualmente sem perdas se um observador humano é incapaz de perceber qualquer diferença entre a imagem original e a reconstruída. Uma imagem codificada sem perdas é, por definição, também visualmente sem perda [16].

Nas aplicações de compressão com perdas, a taxa de compressão e a qualidade visual são dois fatores antagônicos, isto é, quanto menor a taxa de compressão, melhor é a qualidade visual, e vice-versa. Na prática, o processo de compressão estabelece o compromisso da relação entre esses fatores de acordo com as exigências de cada aplicação. Por exemplo, na compressão de imagens médicas, onde a compressão deve ser visualmente sem perdas, a qualidade visual tem relevância maior do que a taxa de compressão.

A maioria dos sistemas de compressão com perdas é híbrida, combinando a aplicação de uma das técnicas de codificação com perdas na primeira fase; seguida da aplicação de uma das técnicas sem perdas na segunda fase, e da codificação de entropia na última [12].

Ao longo dos anos, foram desenvolvidas diversas técnicas de compressão com perdas, como mostrado na Figura 2.5.

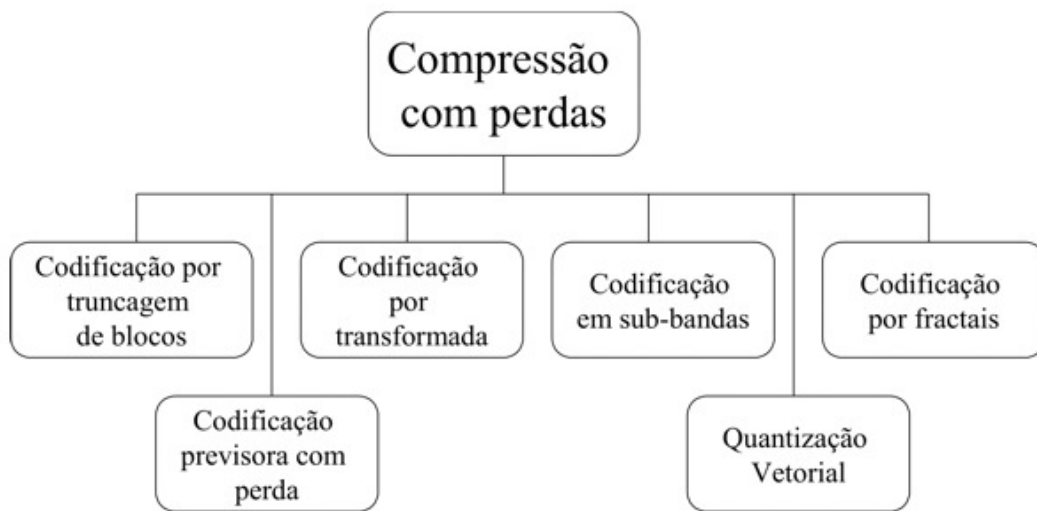


Figura 2.5: Tipos de técnicas de compressão com perda.

2.5.1 Codificação por truncagem de bloco

Na codificação por truncagem de blocos a imagem original é decomposta em blocos 4 x 4 pixels. Para cada bloco são calculados o valor médio ($\bar{x}$) dos pixels (usado como limiar) e o desvio padrão (σ) [12].

A quantização de cada bloco é feita atribuindo-se zero aos valores dos pixels menores do que o limiar $\bar{x}$ e um para os demais valores.

A decodificação de cada bloco é obtida com base no valor médio ($\bar{x}$) do pixels, no valor do desvio padrão (σ) e nas quantidades de zeros e uns resultantes da quantização.

2.5.2 Codificação preditora com perdas

A codificação preditora com perdas também explora a redundância interpixels, como a codificação preditora sem perdas vista na seção 2.4.3 deste capítulo.

Os componentes básicos da codificação preditora com perdas são mostrados na Figura 2.6, além da codificação dos símbolos e do preditor já vistos na seção 2.4.3 deste capítulo, inclui-se um quantizador, cuja função é mapear o erro de previsão e_n em um intervalo limitado de saída, representado por $\hat{e}_n$, que estabelece a quantidade de compressão e de distorção associada à codificação [15].

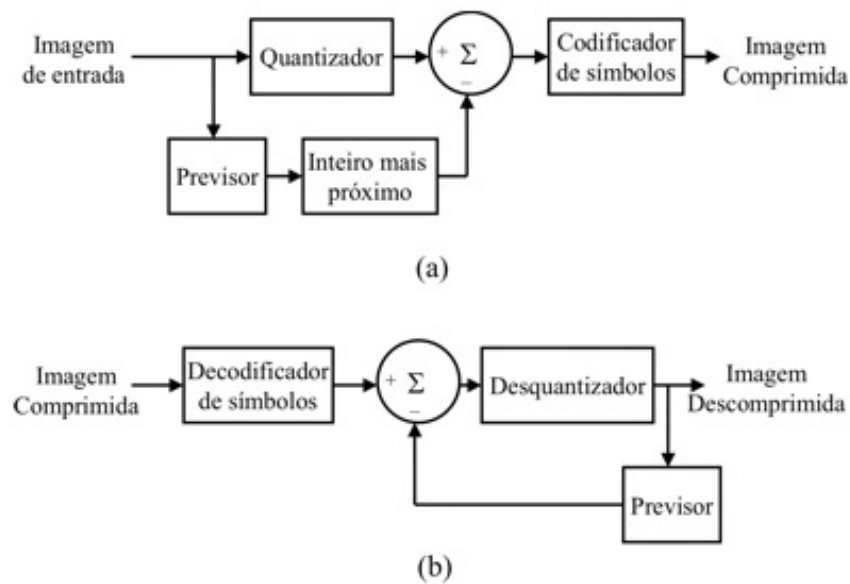


Figura 2.6: Modelo de codificação preditora com perda (a) codificador (b) decodificador.

As previsões geradas pelo codificador e pelo decodificador devem ser

equivalentes. Para isto, coloca-se o previsor com perdas do codificador dentro de um laço de retro-alimentação, em que sua entrada, representada por $\dot{f}_n$, é gerada como função das previsões usadas ($\hat{f}_n$) e dos erros quantizados correspondentes e_n . Ou seja,

$$\dot{f}_n = \dot{e}_n + \hat{f}_n \quad (2.13)$$

onde ($\hat{f}_n$) foi definida na Equação (2.12).

2.5.3 Codificação por transformada

A codificação por transformada utiliza uma ferramenta matemática que possibilita transportar uma matriz de pixels (imagem) do domínio espacial, geralmente, para o domínio da frequência, permitindo a concentração da energia do sinal em poucos coeficientes da matriz transformada. Para tal, utiliza-se uma transformada linear reversível que associa uma imagem a um conjunto de coeficientes da transformada, que são quantizados e codificados. Um número significativo desses coeficientes têm pequenas magnitudes, podendo ser descartado sem causar distorção significativa da imagem. Quanto maior a capacidade da transformada de compactar informações em poucos coeficientes, melhor é o seu poder de compressão.

Existem diversas transformadas que podem ser aplicadas, cada uma com suas peculiaridades, sendo comum o uso da transformada cosseno discreta (DCT), Fourier (DFT), Karhunen-Lòeve (KLT), Walsh/Hadamard e Haar. O critério de escolha de uma transformada em particular depende da

quantidade de erro de reconstrução tolerado e dos recursos computacionais disponíveis.

A KLT é uma ferramenta que apresenta os melhores resultados para compressão de imagens, mas o seu custo computacional diminui a viabilidade de uso dessa transformada. A DCT também apresenta bons resultados para compressão com custo computacional razoável. Essa transformada foi escolhida para o padrão de compressão de imagens Joint Photographic Experts Group (JPEG), sendo assim ela se tornou a transformada mais utilizada por causa da larga utilização desse padrão, principalmente na Internet.

A Figura 2.7 mostra o modelo de codificação e decodificação por transformada.

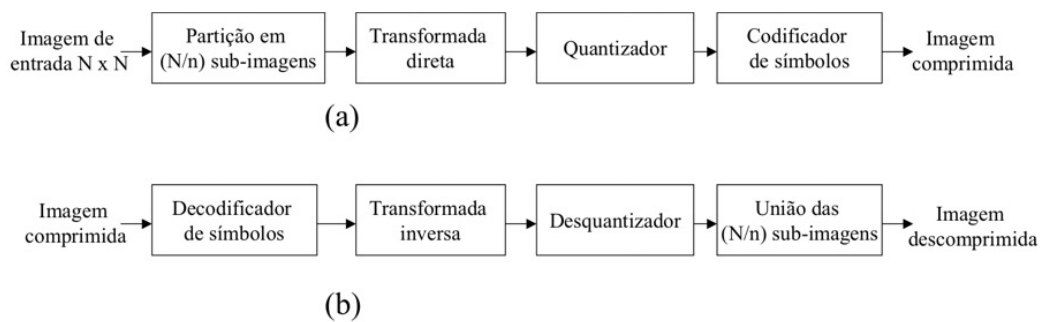


Figura 2.7: Modelo de codificação por transformada: (a)codificador; (b)decodificador

O processo de codificação consiste de quatro componentes básicos: partição em sub-imagens, transformada direta, quantização e codificação de símbolos.

A Figura 2.7(a) mostra esses componentes.

A partição em sub-imagens divide a imagem $N \times N$ em blocos quadra-