VITOR HITOSHI TSUJIGUCHI

IDENTIFICAÇÃO DA CORRELAÇÃO ENTRE AS
CARACTERÍSTICAS DAS IMAGENS DE
DOCUMENTOS E OS IMPACTOS NA FIDELIDADE
VISUAL EM FUNÇÃO DA TAXA DE COMPRESSÃO

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do Título de Mestre em Engenharia Elétrica.

São Paulo
2011

VITOR HITOSHI TSUJIGUCHI

IDENTIFICAÇÃO DA CORRELAÇÃO ENTRE AS
CARACTERÍSTICAS DAS IMAGENS DE
DOCUMENTOS E OS IMPACTOS NA FIDELIDADE
VISUAL EM FUNÇÃO DA TAXA DE COMPRESSÃO

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do Título de Mestre em Engenharia Elétrica.

Área de Concentração:
Sistemas Digitais

Orientador:
Edson Satoshi Gomi

São Paulo
2011

Este exemplar foi revisado e alterado em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 09 de Dezembro de 2011.

Assinatura do autor _____

Assinatura do orientador _____

FICHA CATALOGRÁFICA

Tsujiguchi, Vitor Hitoshi

**Caracterização e compressão de imagens de documentos em bibliotecas digitais / V.H. Tsujiguchi. -- ed. rev. -- São Paulo, 2011.
119 p.**

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Computação e Sistemas Digitais.

1.Processamento digital de imagens 2.Documentos eletrônicos 3.Análise de ondaletas I. Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Computação e Sistemas Digitais II.t.

Dedico este trabalho à Kátia Miyuki
Tsuji, minha eterna companheira.
Tenho uma imensa gratidão pelo apoio
incondicional durante essa longa jornada.
Com muito carinho.

AGRADECIMENTOS

Agradeço aos meus pais, James e Helena, e meu irmão Takeshi, pelo apoio, torcida e confiança depositados em todos os momentos.

Agradeço ao meu orientador, Prof. Dr. Edson Gomi, por exigir de mim, em todos os momentos, um pouco mais do que eu acreditava ser capaz de alcançar. Graças a esses ensinamentos, tive a oportunidade de amadurecer como engenheiro, pesquisador e sobretudo como pessoa.

Agradeço ao Prof. Dr. Pedro Puntoni, que mesmo estando sempre muito ocupado, ofereceu-me conselhos e dicas em momentos cruciais nessa jornada.

Agradeço ao Prof. Dr. Hae Yong Kim pelas poucas, porém interessantes discussões sobre processamento de imagens durante meu mestrado.

Agradeço aos meus colegas de trabalho da Brasiliana Digital pelo companheirismo e pela troca de experiências enquanto desenvolvemos o incrível projeto da Brasiliana Digital.

Agradeço à minha "galerinha" do BSE pelos muitos momentos marcantes que passamos juntos. Em especial ao Gustavo, Gabi, Felipe A., Juliano, Felipe P., Filipe Viking, Veronica, Nayara, Marina, Helena, Aline e Fábio, pela torcida e pelo incrível trabalho realizado comigo na Brasiliana.

Agradeço aos colegas do Laboratório Knoma, com quem aprendi muito e compartilhei ótimos momentos.

Agradeço à família Tsuji, sr. Kendi, Gustavo, Karina e principalmente à dona Keiko Tsuji pelo incentivo, reflexões e pelas as maravilhosas refeições que tive o prazer de desfrutar.

Agradeço aos meus colegas engenheiros elétricos da Poli-USP, pela companhia nas comemorações e nos happy-hours animados.

Agradeço à minha coach, Candida Sevilhano, por todas as lições e palavras de sabedoria transmitidas em nossas sessões. Seu apoio foi fundamental para a conclusão desse trabalho.

Agradeço à bibliotecária Ana Maria pela ajuda na revisão bibliográfica deste trabalho e também pelos divertidos bate-papos em sua biblioteca na engenharia elétrica.

Tenho uma gratidão especial ao professor Istvan Jancsó (in memoriam), quem tenho como modelo de profissionalismo e de um verdadeiro professor.

Expresso minha igual admiração ao dr. José Mindlin (in memoriam), pelo seu gesto de generosidade ao doar sua biblioteca para a USP. Seu nome e seus ideais serão eternamente lembrados.

RESUMO

Imagens de documentos são documentos digitalizados com conteúdo textual. Estes documentos são compostos de caracteres e diagramação, apresentando características comuns entre si, como a presença de bordas e limites no formato de cada caractere. A relação entre as características das imagens de documentos e os impactos do processo de compressão com respeito à fidelidade visual são analisadas nesse trabalho. Métricas objetivas são empregadas na análise das características das imagens de documentos, como a medida da atividade da imagem (IAM) no domínio espacial dos pixels, e a verificação da medida de atividade espectral (SAM) no domínio espectral. Os desempenhos das técnicas de compressão de imagens baseada na transformada discreta de cosseno (DCT) e na transformada discreta de Wavelet (DWT) são avaliados sobre as imagens de documentos ao aplicar diferentes níveis de compressão sobre as mesmas, para cada técnica. Os experimentos são realizados sobre imagens digitais de documentos impressos e manuscritos de livros e periódicos, explorando texto escritos entre os séculos 16 ao século 19. Este material foi coletado na biblioteca Brasileira Digital (www.brasiliana.usp.br), no Brasil. Resultados experimentais apontam que as medidas de atividade nos domínios espacial e espectral influenciam diretamente a fidelidade visual das imagens comprimidas para ambas as técnicas baseadas em DCT e DWT. Para uma taxa de compressão fixa de uma imagem comprimida em ambas técnicas, a presença de valores superiores de IAM e níveis menores de SAM na imagem de referência resultam em menor fidelidade visual, após a compressão.

Palavras-chaves: Compressão de imagens; Documentos digitalizados; Bibliotecas digitais; Brasileira-Biblioteca digital; Imagens de documentos; Qualidade da imagem

ABSTRACT

Document images are digitized documents with textual content. These documents are composed of characters and their layout, with common characteristics among them, such as the presence of borders and boundaries in the shape of each character. The relationship between the characteristics of document images and the impact of the compression process with respect to visual fidelity are analyzed herein. Objective metrics are employed to analyze the characteristics of document images, such as the Image Activity Measure (IAM) in the spatial domain, and assessment of Spectral Activity Measure (SAM) in the spectral domain. The performance of image compression techniques based on Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) are evaluated from document images by applying different compression levels for each technique to these images. The experiments are performed on digital images of printed documents and manuscripts of books and magazines, exploring texts written from the 16th to the 19th century. This material was collected in the Brasiliana Digital Library in Brazil. Experimental results show that the activity measures in spatial and spectral domains directly influence the visual fidelity of compressed images for both the techniques based on DCT and DWT. For a fixed compression ratio for both techniques on a compressed image, higher values of IAM and low levels of SAM in the reference image result in less visual fidelity after compression.

Keywords: Image compression; Digitized documents; Digital libraries; Brasiliana-Digital library; Document images; Image quality

CONTEÚDO

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

1	Introdução	18
1.1	Motivação	33
1.2	Objetivo	35
1.3	Metodologia da pesquisa	35
1.4	Contribuições esperadas	36
1.5	Organização do Trabalho	36
2	Imagens digitais	38
2.1	Categorias de imagens para testes	41
2.2	Métodos para caracterização de imagens	44
3	Compressão de Imagens	47
3.1	Métodos de Compressão sem Perdas	50
3.1.1	Métodos estatísticos	50
3.1.2	Métodos baseados em dicionários	51
3.1.3	Formatos de compressão de imagens sem perdas	53

3.2	Métodos de Compressão com Perdas	54
3.2.1	Transformada Discreta de Cosseno	55
3.2.2	Transformada Discreta de Wavelets	60
3.2.3	Padrão de compressão JPEG2000	62
3.2.3.1	Compressão JPEG2000 com perdas	63
3.2.3.2	Compressão JPEG2000 sem perdas	64
3.3	Métricas de compressão	65
3.3.1	Métricas objetivas de qualidade visual	66
3.3.2	Métricas subjetiva de qualidade visual	70
4	Análise do impacto da taxa de compressão sobre a fidelidade visual	72
4.1	Medição das imagens de documentos	75
4.2	Análise do impacto da taxa de compressão das imagens	82
4.3	Análise das características das imagens	111
4.4	Correlações entre compressão e características das imagens	116
5	Considerações finais	120
5.1	Trabalhos futuros	124
	Referências	126
	Apêndice A – Códigos-fontes dos algoritmos utilizados	135
	Apêndice B – Imagens utilizadas no trabalho	138

LISTA DE FIGURAS

1	World Digital Library [WDL 2010].	21
2	Projeto Gutenberg [Gutenberg 2010].	22
3	Biblioteca Digital Gallica [Gallica 2010]	23
4	Página inicial da Brasileira Digital	24
5	Apresentação do documento digital	25
6	Visualização de imagens	25
7	Projeto Google Books [Google 2010]	26
8	Imagem Lena, com zoom em 300%	28
9	Imagem Lena, com artefatos de blocos	29
10	Imagem Lena, com artefatos borrados	29
11	Imagem de documento, com zoom em 300%	30
12	Imagem de documento, com artefatos de blocos	30
13	Imagem de documento, com artefatos borrados	31
14	Duas versões de imagens de um documento da Brasileira Digital [Cunha BBD].	34
15	Variáveis de uma imagem digital	39
16	Representação gráfica e sob a forma de matriz de uma imagem digital	40
17	Imagens naturais em grayscale	42

18	Imagens de impressões digitais, capturadas por um sensor óptico de baixo custo.	43
19	Imagens médicas, em grayscale.	43
20	Imagens de documentos digitalizados.	44
21	Sistema de compressão e transmissão de dados	48
22	Bloco correspondente ao codificador	49
23	Bloco correspondente ao decodificador	50
24	Sequência zig-zag de codificação dos coeficientes quantizados . . .	59
25	Manuscrito século XVIII	
	[Sá BBD]	76
26	Livro século XVI [Staden BBD]	77
27	Livro século XVII [The Dutch Survey BBD]	78
28	Livro século XVIII [Vicente BBD]	79
29	Livro século XIX [Vieira BBD]	80
30	Imagens sintéticas incluída entre as imagens analisadas	81
31	Compressões da imagem de documento do século XIX [Castro BBD], onde a) representa a imagem de referência (sem compressão) e b) e c) representam as imagens comprimidas a uma taxa de compressão de 0.88 ($CR = 88\%$), nos formatos de imagem JPEG e JPG2000, respectivamente e d) e e) representam as imagens comprimidas à taxa de compressão de 0.98 ($CR = 98\%$)	84
32	Compressões da imagem de documento do século XVIII [Garcia BBD]	85
33	Compressões da imagem de documento do século XVII [Moreau BBD]	86

34	Compressões da imagem de documento do século XVI [Alvarez BBD]	87
35	Compressões da imagem de documento manuscrito do século XIX [Morais BBD]	88
36	Gráfico de CR por PSNR, para imagens de documento JPEG	92
37	Gráfico de CR por PSNR, para imagens de documento JPG 2000	92
38	Gráfico de CR por AD, para imagens de documento JPEG	92
39	Gráfico de CR por AD, para imagens de documento JPG 2000	93
40	Gráfico de CR por LMSE, para imagens de documento JPEG	93
41	Gráfico de CR por LMSE, para imagens de documento JPG 2000	93
42	Gráfico de CR por NK, para imagens de documento JPEG	94
43	Gráfico de CR por NK, para imagens de documento JPG 2000	94
44	Gráfico de CR por MD, para imagens de documento JPEG	94
45	Gráfico de CR por MD, para imagens de documento JPG 2000	95
46	Gráfico de CR por NAE, para imagens de documento JPEG	95
47	Gráfico de CR por NAE, para imagens de documento JPG 2000	95
48	Gráfico de CR por SC, para imagens de documento JPEG	96
49	Gráfico de CR por SC, para imagens de documento JPG 2000	96
50	Gráfico de CR por SSIM, para imagens de documento JPEG	96
51	Gráfico de CR por SSIM, para imagens de documento JPG 2000	97
52	Compressões da imagem natural Lena	99
53	Compressões da imagem natural Baboon	100
54	Compressões da imagem natural Barbara	101

55	Compressões da imagem sintética Vertical	102
56	Imagem de documento do século XVII [The Dutch Survey BBD], sem compressão	104
57	Imagem de documento do século XVII, onde a) representa a ima- gem de referência (sem compressão) e b) e c) representam as ima- gens comprimidas a uma taxa de compressão de 0.88 ($CR = 88\%$), nos formatos de imagem JPEG e JPG2000, respectivamente . . .	105
58	Imagem de documento do século XVII, onde a) representa a ima- gem de referência (sem compressão) e b) e c) representam as ima- gens comprimidas a uma taxa de compressão de 0.98 ($CR = 98\%$), nos formatos de imagem JPEG e JPG2000, respectivamente . . .	105
59	CR x PSNR	106
60	CR x LMSE	108
61	CR x SSIM	108
62	Gráfico de CR por PSNR, para diversas imagens JPEG	109
63	Gráfico de CR por PSNR, para diversas imagens JPG 2000	109
64	Gráfico de CR por LMSE, para diversas imagens JPEG	109
65	Gráfico de CR por LMSE, para diversas imagens JPG 2000	110
66	Gráfico de CR por SSIM, para diversas imagens JPEG	110
67	Gráfico de CR por SSIM, para diversas imagens JPG 2000	110
68	Imagem de documento do século XVII [The Dutch Survey BBD], sem zoom	112
69	IAM e SAM de imagens de documentos	114

70	IAM e SAM de imagens naturais	115
----	---	-----

LISTA DE TABELAS

1	Quantidade de bits de imagens digitais	40
2	Quadro 8x8 de pixels, extraído de [Wallace 1992]	57
3	Quadro após DCT, extraído de [Wallace 1992]	57
4	Quadro 8x8 de pixels reconstruído sem quantização	58
5	Quadro de quantização, extraído de [Wallace 1992]	58
6	Quadro de coeficientes quantizados e desnormalizados, extraído de [Wallace 1992]	59
7	Quadro 8x8 reconstruído, extraído de [Wallace 1992]	60
8	Grau de percepção da qualidade visual	71
9	Tabela de referência com as imagens de documentos utilizadas . .	73
10	Legenda das métricas objetivas utilizadas para averiguar a fide- lidade visual após a compressão	89
11	Compressão da imagem de documento do século XIX, no formato JPEG	89
12	Compressão da imagem de documento do século XIX, no formato JPG 2000	89
13	Compressão da imagem de documento do século XVIII, no formato JPEG	89
14	Compressão da imagem de documento do século XVIII, no formato JPG 2000	90

15	Compressão da imagem de documento do século XVII, no formato JPEG	90
16	Compressão da imagem de documento do século XVII, no formato JPG 2000	90
17	Compressão da imagem de documento do século XVI, no formato JPEG	90
18	Compressão da imagem de documento do século XVI, no formato JPG 2000	91
19	Compressão da imagem de documento manuscrito do século XIX, no formato JPEG	91
20	Compressão da imagem de documento manuscrito do século XIX, no formato JPG 2000	91
21	Compressão da imagem Lena, no formato JPEG	99
22	Compressão da imagem Lena, no formato JPG 2000	100
23	Compressão da imagem Baboon, no formato JPEG	101
24	Compressão da imagem Baboon, no formato JPG 2000	101
25	Compressão da imagem Barbara, no formato JPEG	102
26	Compressão da imagem Barbara, no formato JPG 2000	102
27	Compressão da imagem Vertical, no formato JPEG	103
28	Compressão da imagem Vertical, no formato JPG 2000	103
29	Resultados da compressão do documento do século XVII, em JPEG e JPG 2000, para $CR = 0.88$ e $CR = 0.98$	104
30	Características do documento do século XVII	112

31	Características de algumas imagens analisadas	113
32	Tabela de referência das imagens de documentos analisadas	117
33	Tabela comparativa para imagens em JPG	118
34	Tabela comparativa para imagens em JPG 2000	119

LISTA DE ABREVIATURAS E SIGLAS

OCR	Optical Character Recognition – Reconhecimento óptico de caracteres
DCT	Discrete Cosine Transform – Transformação discreta de cossenos
DFT	Discrete Fourier Transform – Transformação discreta de Fourier
DWT	Discrete Wavelets Transform – Transformação discreta de wavelets
CWT	Continuous Wavelets Transform – Transformação contínua de wavelets
IDCT	Inverse Discrete Cosine Transform – Transformação inversa discreta de cossenos
IAM	Image Activity Measure – Medição da atividade da imagem
SAM	Spectral Activity Measure – Medição da atividade espectral
PSNR	Peak Signal-to-Noise Ratio – Pico da razão sinal-ruído
CR	Compression-Rate – Taxa de compressão
MSE	Mean Square Error – Erro médio quadrático
LMSE	Laplacian Mean Square Error – Erro médio quadrático laplaciano
SSIM	Structural Similarity Index – Índice de similaridade estrutural
AD	Average Difference – Diferença média
MD	Maximum Difference – Diferença máxima
NK	Normalized Cross-Correlation – Correlação cruzada normalizada
NAE	Normalized Absolute Error – Erro absoluto normalizado

SC	Structural Content – Conteúdo estrutural
HVS	Human Visual System – Sistema visual humano
MOS	Mean Opinion Score – Pontuação da opinião média
FR	Full Reference – Referência completa
RR	Reduced Reference – Referência reduzida
NR	No Reference – Sem referência
RCT	Reversible Component Transform – Transformação reversível de componente
ICT	Irreversible Component Transform – Transformação irreversível de componente
JPEG	Joint Photographic Experts Group – Grupo de especialistas em articulação fotográfica
GIF	Graphics Interchange Format – Formato para intercâmbio de gráficos
TIFF	Tagged Image File Format – Formato de arquivos de imagens marcados
W3C	World Wide Web Consortium – Consórcio da Rede Mundial de Computadores
ISO	International Organization for Standardization – Organização internacional para padronização
IEC	International Electrotechnical Commission – Comissão eletrotécnica internacional
WSQ	Wavelet Scalar Quantization – Wavelet de Quantização Escalar
MSVQ	Multiple Stages Vector Quantization – Quantização vetorial de múltiplos estágios

FBI	Federal Bureal of Investigation – Departamento federal de investigação
FVC2000	Fingerprint Verification Competition 2000 – Competição de verificação de impressão digital do ano 2000
BBD	Biblioteca Brasileira Digital
RLE	Run-Lenght Encoding – Codificação de comprimento corrido
CDF 9/7	Cohen-Daubechies-Feauveau wavelet 9/7 – Wavelet 9/7 de Cohen-Daubechies-Feauveau

1 INTRODUÇÃO

Segundo a Digital Library Federation [DLF 1998], bibliotecas digitais são organizações que proporcionam os recursos necessários para selecionar, estruturar, distribuir, preservar a integridade e assegurar a persistência através do tempo de coleções de recursos digitais, de forma que estes estejam disponíveis para uma comunidade alvo definida. Ian H. Witten [Witten et al. 2000] define biblioteca digital como um repositório de coleções de objetos digitais. Esse repositório é organizado e disposto de mecanismos de busca, de seleção e de apresentação do conteúdo dessas coleções, composta por arquivos sob a forma de texto, imagens, áudio e vídeo são apresentados como objetos digitais. Rich Gazan [Gazan 2005] associa bibliotecas digitais a entidades sociais, um local onde são registradas histórias sobre a cultura, costumes, localidades e épocas, através dos itens que compõem as coleções digitais.

Bibliotecas digitais permitem o acesso ao conteúdo digital de uma forma mais ampla que as bibliotecas tradicionais, uma vez que rompe barreiras físicas e limites de disponibilidade. É comum, inclusive, a existência de portais de bibliotecas digitais, onde o conteúdo é agregado e distribuído por uma única entidade, independente da localidade física dessas coleções de objetos digitais.

Além das bibliotecas digitais, diversas instituições e órgãos corporativos também estão convertendo suas coleções de documentos físicos em cópias digitais, dada a agilidade do acesso remoto aos objetos digitais e ao baixo custo de

armazenamento digital, comparado ao equivalente para armazenamento físico [Aung, Ng e Poh 2009]. Assim como Lu [Lu, Zhang e Tan 2004], caracterizamos documentos digitais como material acessível por recursos computacionais, cuja fonte de informação dominante está em seu conteúdo textual, na forma de redação manuscrita ou impressa. Pilhas de documentos em papel podem ser arquivadas em estantes e depósitos, exigindo métodos eficientes de gerenciamento, tanto na recuperação das informações quanto sobre as dimensões físicas ocupadas. Esta mesma pilha de documentos quando digitalizada, permite que seu conteúdo seja armazenado em mídia digital, ocupando menos espaço nas prateleiras; suas informações são distribuídas rapidamente e oferecendo inúmeros acessos simultâneos ao mesmo documento. Outra grande vantagem da digitalização de documentos é a preservação de sua integridade visual, sob a forma digital; isto é, permite que se guarde um retrato do documento físico em um dado instante, assegurando que seu conteúdo possa ser extraído mesmo que o documento original seja degradado ou perdido.

Em seu trabalho sobre a preservação da herança cultural no mundo virtual, Ruan Jianhai [Jianhai e McDonough 2009] classifica documentos digitais em duas categorias: documentos nativos digitais e documentos digitalizados. Documentos nativos digitais são produtos produzidos diretamente no computador, através de programas de processamento de texto, como Microsoft Office Word, OpenOffice Writer, LyX e outros editores de texto. Grande parte das teses e dissertações acadêmicas produzidas recentemente são documentos nativos digitais e podem ser gerenciadas por Sistemas Eletrônicos de Teses e Dissertações (ETD), de forma que sua contribuição de pesquisa pode ser disponibilizada rapidamente, no repositório de sua instituição, onde é realizado o gerenciamento e disponibilização deste material na internet. Atualmente, diversas universidades e instituições adotam esse modelo de apresentação do material acadêmico produzido, como a Biblioteca

Digital Brasileira de Teses e Dissertações [BDTD 2010], coordenado pelo Instituto Brasileiro de Informação em Ciências e Tecnologia (IBICT), a Biblioteca de Teses e Dissertações da USP [USP 2010], a Networked Digital Library of Theses and Dissertations [NDLTD 2010], entre outras.

Documentos digitalizados correspondem aos documentos que se apresentavam sob forma de papel e foram convertidos para sua forma digital. O processo de digitalização ocorre com o auxílio de equipamentos de captura, como scanners e máquinas fotográficas, obtendo a imagem dos documentos físicos. Diversas bibliotecas digitais miram a digitalização e disponibilização de documentos impressos históricos e modernos [Marinai, Marino e Soda 2006]. As coleções de documentos históricos apresentam grande importância cultural e científica [Kavallieratou e Stathis 2006], por conterem um material fundamental sobre os registros da vida humana e sua história [Ding et al. 2004]. Este material é geralmente composto de jornais, revistas, livros e documentos manuscritos, tendo exemplares que podem datar séculos desde sua criação. Visando a preservação desta herança cultural, a digitalização destes documentos permite separar a deterioração do material físico da deterioração do conteúdo da informação [Agosti, Ferro e Orio 2005]. Desta forma, apresenta-se uma alternativa de conservação da aparência e do conteúdo de uma obra física com sua contraparte digital, permitindo também a utilização dos avanços da tecnologia no tratamento de imagens para otimização da apresentação deste material digital. Vale ressaltar que a qualidade das cópias digitais estão intimamente ligadas à qualidade visual dos documentos originais, que podem apresentar alguns tipos de degradações, devido à deterioração do material causada pelo tempo, formas de manipulação e condições de armazenamento, como umidade e temperatura [Drira 2006], além das condições de geração e conservação e até da digitalização desses documentos [Zheng, Li e Doermann 2004].

A UNESCO (United Nations Educational, Scientific and Cultural Organization) [UNESCO 2010] é uma das principais instituições contribuintes para a preservação da herança cultural mundial sob a forma digital. Em parceria com diversas instituições, foi criada a World Digital Library (WDL) [WDL 2010], uma biblioteca digital reunindo acervos históricos de diversos países no mundo, apresentando seu conteúdo em mais de dez idiomas. Na figura 1, encontram-se algumas ilustrações sobre a interface com o usuário.

(a) Página inicial

(b) Apresentação do documento digital

Figura 1: World Digital Library [WDL 2010].

Outra biblioteca digital responsável por disponibilizar uma vasta coleção de documentos digitais, é o Projeto Gutenberg [Gutenberg 2010]. Este acervo digital apresenta desde documentos digitalizados até livros nativos digitais, onde grande parte destes objetos digitais são apresentados na forma de texto puro em código ASCII, como pode ser observado na figura 2 [Gardner 2006].

(a) Página inicial

(b) Apresentação do documento digital

Figura 2: Projeto Gutenberg [Gutenberg 2010].

A seguir, apresentamos como são disponibilizados os documentos digitais em algumas bibliotecas digitais mundialmente conhecidas. A biblioteca nacional

francesa disponibiliza livros digitalizados através da Biblioteca Digital Gallica [Gallica 2010] (vide figura 3).

(a) Página inicial

(b) Apresentação do documento digital

Figura 3: Biblioteca Digital Gallica [Gallica 2010]

A biblioteca Brasileira Digital [BBD 2010], é desenvolvida pela Biblioteca Guita e José Mindlin, com o o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp), por meio do processo 07/59783-3 de auxílio à pesquisa. Grande parte do acervo digital é composto pela doação de parte da biblioteca de

José Mindlin. Com pouco mais de um ano desde o primeiro lançamento, já foram publicados mais de 1000 livros e 3000 objetos digitais em seu acervo digital. Nas figuras 4, 5 e 6, apresentamos algumas imagens do site da biblioteca digital.

Figura 4: Página inicial da Brasileira Digital

Figura 5: Apresentação do documento digital

Figura 6: Visualização de imagens

O projeto Google Books [Google 2010], da organização Google, oferece serviços de digitalização e disponibilização de documentos digitais em seu repositório, adicionando registros de metadados e indexando-os em seu mecanismo de busca.

Sua interface com o usuário é apresentada na figura 7.

(a) Página inicial

(b) Apresentação do documento digital

Figura 7: Projeto Google Books [Google 2010]

Apesar do espaço de armazenamento de documentos digitais representar uma quantidade muito menor de material do que o ocupado por documentos físicos nas estantes de uma biblioteca tradicional, ainda apresenta algumas inconveniências quanto ao seu armazenamento e acesso. As instituições e bibliotecas digitais utilizam equipamentos de captura em alta resolução durante a digitalização, ge-

rando arquivos com grande riqueza de detalhes, o que resulta em uma maior quantidade de informação a ser armazenada. Usuários de bibliotecas digitais geralmente têm acesso ao conteúdo digital por meio de transferência de arquivos via conexão de internet. Uma implicação natural disso é o fato de grandes arquivos resultarem em maior tempo de transferência, causando um impacto sobre a usabilidade do sistema. Neste cenário, torna-se interessante a aplicação de técnicas de compressão de imagens para a redução da quantidade de bits necessários para a representação de uma imagem digital, mantendo níveis adequados de fidelidade e qualidade da imagem apresentada [Khalifa e Dlay 1998].

Outro desafio enfrentado em bibliotecas digitais é a obsolescência das tecnologias de software e hardware para representação dos documentos digitais [Heminger e Robertson 1998]. Atualmente é papel da instituição ou biblioteca digital decidir sobre os formatos e padrões de apresentação dos documentos digitais, tendo em vista questões como portabilidade, flexibilidade e compatibilidade, uma vez que ainda não há padrões universalmente regulamentados.

Imagens de diferentes naturezas requerem formas específicas de tratamento, tendo em vista seus usos finais. Para documentos com conteúdo textual, a aplicação de técnicas de ajuste de nitidez auxilia na legibilidade de seu conteúdo. Entretanto, a aplicação dessa mesma técnica pode degradar a aparência de logotipos de empresas e fotografias de pessoas e paisagens; nessas imagens de logotipos e artes gráficas de webdesign, normalmente se utiliza uma paleta de cores simplificada, enquanto fotos podem ser apresentadas com uma resolução visual reduzida, caso seja ajustada a resolução dos bits, uma vez que há pouca perda visual nesse processo [Lins et al. 2009]. Logo, percebe-se que para diferentes tipos de imagens, encontram-se uma série de ferramentas e técnicas adequadas para seu processamento e sua apresentação.

No caso de compressão de imagens com perdas, é comum observar a presença

de artefatos visíveis sobre a imagem comprimida. Há dois tipos de artefatos visuais que geralmente são encontrados em imagens comprimidas com perdas: artefatos de blocos (blocking artifacts) e artefatos de borrões (blurring artifacts). A figura 8, apresenta uma mesma imagem original submetida a duas formas de compressão com perdas, evidenciando cada um dos tipos de artefatos nas imagens: artefatos de blocos (figura 9) e artefatos de borrões (figura 10) .

Figura 8: Imagem Lena, com zoom em 300%

Figura 9: Imagem Lena, com artefatos de blocos

Figura 10: Imagem Lena, com artefatos borrados

Ambas as imagens foram comprimidas a uma taxa de 64 vezes em relação à imagem original. A imagem de teste Lena representa uma imagem natural de

pessoa. Aplicando-se o mesmo experimento para a imagem de um documento digitalizado (ver figura 11), obtém-se os seguintes exemplares (figuras 12 e 13):

Figura 11: Imagem de documento, com zoom em 300%

Figura 12: Imagem de documento, com artefatos de blocos

Figura 13: Imagem de documento, com artefatos borrados

Pode-se verificar que, para as mesmas taxas de compressões utilizadas na imagem Lena, a imagem de documento apresenta maior quantidade de artefatos visuais nas imagens comprimidas. Devido à presença de muitas bordas ("edges") no formato das letras que compõem a imagem, ocorre um aumento perceptível da quantidade de artefatos nessas regiões. Isso ilustra como imagens de diferentes naturezas apresentam comportamentos distintos durante a compressão de imagens com perdas.

Um exemplo de pesquisa sobre técnicas adequadas para compressão de imagens de natureza específica é o caso das imagens digitalizadas de impressões digitais. O banco de dados de impressão digital do FBI, que armazena impressões digitais desde 1924, apresenta mais de 200 milhões de registros, exigindo buscas diárias ao procurar os registros de qualquer suspeito. Segundo Chris Brislawn [Brislawn 2002], levava-se aproximadamente três horas para se realizar uma determinada transferência diária entre os repositórios de arquivos. Em

1994, Bradley e Brislawn apresentaram um método de compressão de imagens de impressões digitais, o Wavelet de Quantização Escalar (WSQ), a partir das técnicas de Transformada Discreta de Wavelets (DWT), obtendo taxas de compressão de 20:1 [Bradley e Brislawn 1994]. Pode-se utilizar as técnicas empregadas no JPEG2000 para compressão de imagens digitais, obtendo altas taxas de compressão, porém não preserva tantos detalhes dos contornos quanto a WSQ [Figueroa-Villanueva, Ratha e Bolle 2003]. Em 2006, Esakkirajan et al. apresentou outra técnica de codificação para impressões digitais, baseada na transformada de contourlet e na quantização vetorial de múltiplos estágios (MSVQ) [Esakkirajan et al. 2006]. As pesquisas persistem para esse tipo específico de imagem, buscando sempre preservar os detalhes específicos para reconhecimento dos padrões exigidos nas buscas do FBI e maximizar a relação taxa de compressão e perda visual de detalhes.

Na área de compressão de imagens de documentos, essa particular questão foi apontada por Witten et al. em 1994, ao apresentar a alternativa de compressão de imagens textuais realizando o reconhecimento óptico de caracteres (OCR) do texto e armazenando apenas os códigos ASCII relativo ao conteúdo, junto com alguma informação sobre suas coordenadas na página original [Witten et al. 1994]. Com a utilização de técnicas de compressão de dados, como compressão por métodos de dicionário (Dictionary Compression) ou por modelos estatísticos, como a codificação de Huffman, obtém-se taxas de compressão superiores para esse tipo de documentos [Solomon 2004]. Entretanto, o reconhecimento dos caracteres não é completamente acurado em muitos documentos, principalmente nos documentos antigos, que estão sujeitos a degradações decorrentes do tempo e das próprias condições de armazenamento; a maioria das bibliotecas digitais adotam a apresentação desses documentos apenas como imagens de documentos [Bourgeois et al. 2004]. Desta forma, neste trabalho é realizado um estudo

sobre a aplicação de técnicas de compressão que exploram as características e propriedades distintas das imagens de documentos, de modo a não depender diretamente da taxa de acertos no reconhecimento de caracteres de um texto.

1.1 Motivação

Bibliotecas digitais, em geral, oferecem versões em fac-símile ou em versões modificadas, com a aplicação de filtros que melhoram a legibilidade de um documento digitalizado. As figuras 14(a) e 14(b) apresentam esses dois tipos de documentos, um processado para facilitar a impressão e apresentação e outro fac-símile do documento físico, respectivamente. A versão fac-símile, normalmente, é utilizada por um público-alvo interessado em estudar as minúcias do documento original, como condições de conservação, vestígios da impressão ou a forma do material. Para disponibilizar qualquer um destes dois tipos de materiais digitais pela internet e para armazená-los em seus servidores, as bibliotecas digitais buscam encontrar formas de apresentar estas cópias digitais com menor tamanho possível, mantendo um nível aceitável de legibilidade [Aung, Ng e Poh 2009], com foco na usabilidade e acesso do usuário.

(a) Original em fac-símile

(b) Tratado para impressão

Figura 14: Duas versões de imagens de um documento da Brasiliana Digital [Cunha BBD].

Neste trabalho, utilizou-se como base de estudos o acervo da biblioteca Brasiliana Digital [BBD 2010], composta de uma vasta coleção de livros da temática brasiliana, que datam do século XVI ao início do século XX. Esta biblioteca foi escolhida por estar em processo de desenvolvimento, tendo inclusive o presente mestrando como um dos participantes ativos em seu desenvolvimento; a biblioteca apresenta uma grande amostra de documentos para realização dos experimentos científicos e apresentou a motivação dos desafios de representação e compressão dos documentos. Grande parte dos documentos digitais disponíveis apresentam tamanho elevado de armazenamento, que tornam o acesso lento para o usuário, além de consumirem considerável carga no espaço dos servidores deste projeto. Até o momento da escrita deste trabalho, a Brasiliana Digital oferece apenas as versões tratadas dos documentos, com foco na facilidade de leitura e impressão para o usuário final, excluindo uma versão fac-símile, sem interferência nas cores

e aspectos do documento original digitalizado.

1.2 Objetivo

O objetivo deste trabalho de pesquisa é identificar a correlação entre as características das imagens de documentos e os impactos na fidelidade visual em função da taxa de compressão.

1.3 Metodologia da pesquisa

Para alcançar os objetivos deste trabalho de pesquisa, foi estabelecida uma divisão de atividades:

- Coleta e caracterização de imagens digitais: Nessa etapa, busca-se coletar e analisar uma amostra diversificada de imagens, com o objetivo de identificar um conjunto de propriedades específicas, para distinguir os documentos digitais dos demais tipos de imagens existentes;
- Avaliação do desempenho de compressão de imagens: Realizar uma análise do desempenho das técnicas de compressão com perdas e sem perdas para imagens de documentos, avaliando os resultados obtidos através de métricas objetivas e subjetivas;
- Criação de laboratório para experimentações: Elaboração de um ambiente para a execução dos experimentos e posterior coleta e avaliação dos resultados;
- Correlação entre os resultados obtidos: Encontrar as características que comprovem os fatores limitantes na capacidade de compressão de imagens de documentos, utilizando os resultados obtidos das atividades anteriores.

1.4 Contribuições esperadas

Este trabalho propõe contribuir com uma investigação sobre a compressão de imagens de documentos digitalizados, utilizamos abordagem de caracterizar esse tipo de imagem em relação aos demais tipos de imagens (naturais, sintéticas, impressões digitais, mapas, etc).

Uma biblioteca digital geralmente é composta de centenas a milhões de imagens de documentos digitalizados. Durante a construção e a manutenção de uma biblioteca digital, aspectos como apresentação do acervo digital e quantidade de dados a serem armazenados tornam-se rapidamente fatores críticos e decisivos. Esse trabalho propõe apresentar resultados da análise do comportamento das imagens de documentos frente a diferentes métodos e taxas de compressão, tanto no aspecto visual quanto na quantidade de compressão alcançada.

Com os resultados desse trabalho de pesquisa, permite-se vislumbrar caminhos para uma investigação posterior sobre formas de melhorar o desempenho das técnicas atuais de compressão, além da perspectiva de desenvolvimento de novos métodos de medição da compressão.

1.5 Organização do Trabalho

Neste capítulo foi apresentada uma breve introdução sobre bibliotecas digitais, com ênfase na importância na preservação da herança cultural digital. Também constata uma definição simples de documentos digitais, identificando suas naturezas e alguns exemplos de seus usos. Por fim situa o trabalho no atual estado da arte em compressão de imagens, seguido pela motivação e do objeto de estudos – o acervo da biblioteca Brasileira Digital – caracterizando o objetivo desta dissertação.

O segundo capítulo apresenta uma análise sobre imagens digitais, identificando as características de diversos tipos de imagens, como paisagens, fotos de pessoas, impressões digitais, imagens médicas e documentos de texto digitalizado. Para a análise de documentos digitais, são exploradas algumas métricas para realizar esta avaliação. Por fim, é selecionado um conjunto de imagens de documentos digitais para a comparação dos métodos de compressão de imagens.

O terceiro capítulo aborda as técnicas de compressão de imagens com e sem perdas. São apresentadas alguns métodos de compressão de imagens, abordando seus aspectos teóricos e funcionalidades práticas para imagens de documentos.

O quarto capítulo inicialmente descreve o laboratório de experimentos para a aplicação das análises sobre os objetos de estudo. Em seguida, relata o resultado dos experimentos e as análises sobre os resultados obtidos..

Concluindo o trabalho, no quinto capítulo são feitas as considerações finais e sugestões para trabalhos futuros.

2 IMAGENS DIGITAIS

Uma imagem digital pode ser descrita por uma matriz bidimensional de elementos de imagem (picture elements). Cada elemento de imagem denomina-se pixel ou pel e representa a intensidade de luz emitida em sua determinada posição nesta matriz. A escala de valores para as representações de cada pixel define alguns tipos de imagens [Solomon 2004]:

- Imagens monocromáticas ou binárias: Cada pixel apresenta um de dois valores possíveis, normalmente branco ou preto;
- Imagens em tons de cinza (grayscale): Cada pixel apresenta um de n valores, variando de 0 a $n - 1$, resultando em 2^n tonalidades de cinza;
- Imagens coloridas: Cada pixel é representado pela composição de m componentes, cada componente apresentando um de n_i valores, variando de 0 a $n_i - 1$. Tipicamente, são adotados os padrões RGB, HSI e YCrCb, com 3 componentes cada.

Para se compor uma imagem digital, deve-se especificar a quantidade de pixels nas linhas e colunas da matriz bidimensional e seus valores individuais. Esta imagem pode ser decrita como uma função $f(x, y)$, cujas coordenadas espaciais (x, y) representam a amostragem da imagem e o valor de $f(x, y)$ em cada ponto representa a amplitude de um sinal – por exemplo, intensidade de brilho de cinza –, sendo limitada pelo nível de quantização da imagem. Os valores de amostragem e

quantização podem ser representados nos conjuntos $\mathbb{Z}$ e $\mathbb{R}$, conjuntos dos números inteiros e reais, respectivamente [Gonzales e Woods 2000]. O cálculo do número de bits necessários para representar uma imagem digital depende dos valores do número de linhas (N), o número de colunas (M), o número de componentes (C) e o número de valores de cada componente (G_c). Tipicamente, os valores de N , M e G_c são representados em potências inteiras de dois, isto é:

$$N = 2^n, \quad M = 2^m \quad (2.1)$$

e

$$G_c = 2^{g_c} \quad (2.2)$$

onde G é o número de tonalidades do componente C . A figura 15 apresenta a localização destes valores dentro de uma imagem digital.

Figura 15: Variáveis de uma imagem digital

O número, b , de bits necessários para representar uma imagem digital é obtido na equação 2.3:

$$b = m \cdot n \cdot \sum_c g_c \quad (2.3)$$

A tabela 1 apresenta uma comparação do tamanho de bits necessários para representar uma mesma imagem, com variações nas dimensões e nas tonalidades.

Tabela 1: Quantidade de bits de imagens digitais

Tipo de Imagem	Linhas(M)	Colunas(N)	Tonalidades (g_c)	Bits
P & B	256	256	1 bit	65536
Grayscale	256	256	8 bits	524288
Grayscale	512	512	8 bits	2097152
Colorida	256	512	20 bits	2621440
Colorida	512	512	24 bits	6291456

A figura 16 apresenta os valores de uma pequena região da imagem em termos de matriz de pixels, para uma imagem em grayscale com 8 bits de tonalidades de cinza.

Figura 16: Representação gráfica e sob a forma de matriz de uma imagem digital

A resolução – grau de detalhamento – de uma imagem é altamente dependente dos parâmetros de amostragem e quantização. Quanto maiores os valores

desses parâmetros, a imagem digital apresentará maiores níveis de detalhamento, em relação à imagem real. Porém, conforme aponta a equação 2.3, o ajuste desses parâmetros implica em um aumento da quantidade de bits necessários para a apresentação e armazenamento do arquivo [Gonzales e Woods 2000]. Para encontrar um equilíbrio entre a resolução e o tamanho do arquivo, a aplicação de técnicas adequadas de compressão de imagens tem papel fundamental. Para a associação de uma imagem digital e seu tamanho físico, no caso de uma impressão por exemplo, geralmente é utilizada a medida pontos por polegada (DPI).

2.1 Categorias de imagens para testes

Neste trabalho, foi adotada a seguinte classificação das imagens digitais:

- Imagens naturais, como fotografias de pessoas ou paisagens;
- Imagens sintéticas, como banners e logotipos, produzidas no computador;
- Imagens médicas, por exemplo radiografias e ultra-som;
- Imagens textuais, como manuscritos e textos de jornal e livros;
- Imagens de impressão digital (fingerprint images).

Imagens de diferentes categorias tendem a apresentar diferentes características no domínio espacial e espectral [Vemuri e Saha 2000]. Logo, a distinção dessas propriedades em ambos os domínios se apresenta como um bom ponto de partida para a caracterização de uma imagem de documento. Na literatura e na internet se encontram inúmeras imagens digitais para experimentação para análise de imagens. Neste trabalho, utilizamos imagens de diversas categorias, como imagens naturais, médicas, de impressões digitais e de textos digitalizados, coletadas nas variadas bases de imagens disponíveis.

Para imagens naturais, foram utilizadas as imagens Lena, Goldhill, Peppers, Baboon e Bird, coletadas da Universidade de Waterloo, Canada [WFCAG 2010]. As imagens coletadas estão em grayscale, com dimensões de 512 por 512 pixels, com 8 bits de níveis de cinza e são apresentadas na figura 17.

(a) Lena

(b) Baboon

(c) Peppers

(d) Goldhill

(e) Bird

Figura 17: Imagens naturais em grayscale

As imagens de impressões digitais foram obtidas no FVC2000, utilizadas na competição de verificação de impressões digitais [Maio et al. 2002]. Apresentamos na figura 18 a categoria DB1 de impressões digitais capturadas por um sensor óptico de baixo custo, contendo 300x300 pixels, com 8 bits de profundidade em grayscale.

(a) Tb01 (b) Tb02 (c) Tb03 (d) Tb04 (e) Tb05

Figura 18: Imagens de impressões digitais, capturadas por um sensor óptico de baixo custo.

Para imagens médicas, foram coletadas [MIS 2010] cinco imagens de diferentes categorias médicas: uma radiografia computadorizada (CR) do peito, uma tomografia computadorizada (CT) do tornozelo, uma ressonância magnética (MR) da cabeça, radiografias do cólon e do quadril (OT) (figura 19).

(a) Peitoral (b) Tornozelo (c) Cabeça (d) Cólon (e) Quadril

Figura 19: Imagens médicas, em grayscale.

A amostra de testes para documentos digitalizados (figura 20) foram extraídos de diversas bases, abrangendo um grande número de tipos de documentos. Há versões digitais de manuscritos, jornais, textos recentes e livros antigos, dos séculos XVI e XVII, extraídos do acervo da biblioteca Brasileira Digital.

Figura 20: Imagens de documentos digitalizados.

2.2 Métodos para caracterização de imagens

Zhaohui Cui em seu artigo [Cui et al. 2008] procurou uma relação entre as características de uma imagem digital e a qualidade da percepção visual (PSNR), quando submetida à compressão JPEG2000, realizando uma série de métricas sobre uma coleção de imagens diferentes. Grande parte dessas métricas são baseadas e reformuladas a partir do artigo de Saha e Vemuri [Saha e Vemuri 2000]. Neste trabalho, utilizamos estas métricas para classificar as imagens e permitir

especular sobre o resultado da compressão destas imagens:

- Média: representa o valor médio de intensidade de brilho da imagem:

$$m = \frac{1}{MN} \sum_{j=1}^M \sum_{k=1}^N f(j, k) \quad (2.4)$$

- Variância: definida como uma medida da dispersão da imagem. é representada por:

$$\sigma^2 = \frac{1}{MN-1} \sum_{j=1}^M \sum_{k=1}^N (f(j, k) - m)^2 \quad (2.5)$$

- Coeficiente de Variação: medida de dispersão relativa. é dada por:

$$\eta = \frac{\sigma}{m} \quad (2.6)$$

adota-se $f(j, k)$ o valor de intensidade de brilho de cada pixel e M e N são as dimensões em pixels da imagem.

De forma a complementar as métricas de Cui et al., pode-se verificar o comportamento de uma imagem no domínio espacial e no domínio espectral. Saha e Vemuris apresentam uma medida para a entropia no domínio espacial, chamada Image Activity Measure (IAM) [Vemuri e Saha 2000], que é descrito pela equação 2.7:

$$IAM = \frac{1}{MN} \left[\sum_{i=1}^{M-1} \sum_{j=1}^N |f_{i,j} - f_{i+1,j}| + \sum_{i=1}^M \sum_{j=1}^{N-1} |f_{i,j} - f_{i,j+1}| \right] \quad (2.7)$$

A atividade espacial da imagem está associada à distribuição dos pixels na imagem. Quanto maior o grau de detalhamento de uma imagem, como grande variação na intensidade dos valores de pixels vizinhos em uma determinada área, maior é a atividade espacial da imagem. De forma análoga, quanto mais estável é a dispersão de valores, com pouca variação entre pixels vizinhos, menor é a atividade da imagem.

Para se verificar as características espectrais de uma imagem, é necessário realizar inicialmente uma transformação da imagem do domínio espacial $f(x, y)$ para o domínio da frequência $F(s)$. Para realizar essa transformação, emprega-se a Transformada Discreta de Fourier (DFT) em duas dimensões.

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp \left[-j2\pi \left(\frac{ux}{M} + \frac{vy}{N} \right) \right] \quad (2.8)$$

onde u e v são os coeficientes discretos dos respectivos valores de x e y , na equação 2.8 .

Obtida a matriz de frequências, $F(s)$, obtém-se a as propriedades espectrais da imagem através da spectral activity measure (SAM):

$$SAM = \frac{\frac{1}{M.N} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} |F(j, k)|^2}{\left[\prod_{j=0}^{M-1} \prod_{k=0}^{N-1} |F(j, k)|^2 \right]^{\frac{1}{M.N}}} \quad (2.9)$$

De acordo com Ferreira [Ferreira 2002], SAM é uma medida de previsibilidade da redundância da imagem. No campo espectral, é possível perceber aspectos da redundância na distribuição de pixels sobre a imagem. A transformada discreta de cosseno (DCT), empregada na compressão de imagens no formato JPEG, utiliza desse aspecto ao converter a imagem ao campo espectral, para realizar a eliminação de redundâncias.

A métrica SAM é representada pelo quociente entre as médias aritmética e geométrica do valor absoluto dos coeficientes da Transformada Discreta de Fourier (DFT) (vide equação 2.8) da função de entrada. Valores elevados de SAM indicam menor entropia e uma menor distorção quadrática média [Jayant e Noll 1984].

3 COMPRESSÃO DE IMAGENS

Documentos digitalizados são apresentados como uma sequência de imagens digitais, o que gera um grande volume de informação e exige que os sistemas computacionais apresentem grande capacidade de armazenamento, além das demandas na distribuição e apresentação dos arquivos. Um recurso utilizado para otimizar essas demandas de armazenamento, distribuição e apresentação é a aplicação de técnicas de compressão de dados, mais especificamente, em compressão de imagens.

A compressão de dados é alcançada ao reduzir a redundância de um arquivo. Exemplos históricos de codificação de dados que se utilizou da compressão e da redundância dos símbolos para transmissão de informação é o código Braille e o código Morse [Solomon 2004]. Existem quatro tipos de redundâncias encontrados em imagens [Conci, E. e R. 2008]:

- Redundância de codificação de tons ou cor: quando são utilizados mais símbolos de codificação do que o necessário. Por exemplo, uma imagem em preto e branco que foi codificada usando uma tabela de 256 tons de cinza;
- Redundância espacial: também chamada de redundância inter-pixel, é resultante das relações estruturais ou geométricas dos objetos da imagem. Por exemplo, quando os objetos presentes em uma imagem podem ser descritos apenas por seu contorno;

- Redundância espectral: ocorre quando valores espectrais na mesma posição em faixas diferentes na matriz de pixels são correlacionados. Essa redundância é explorada na compressão com perdas do JPEG, apresentada a seguir;
- Redundância psicovisual: Relaciona o fato de o sistema visual humano (HVS) não responder com a mesma sensibilidade a todos os estímulos visuais. Por exemplo, o ser humano não consegue detectar pequenas variações de crominância.

Existem dois tipos de compressão de dados: compressão sem perdas (lossless), onde toda a informação é transmitida e pode ser fielmente reproduzida após sua descompressão, e a compressão com perdas (lossy), também chamada de destrutiva ou irreversível, uma vez que parte da informação é descartada para aumentar a efetividade da taxa de compressão.

Normalmente, são utilizadas diversas técnicas de compressão combinadas, de modo a formar um sistema prático de compressão de imagens. Um sistema típico de compressão de imagens é composto de dois blocos estruturais distintos – um codificador e um decodificador – ligados por um canal de comunicação. Uma imagem de entrada $f(x,y)$ é alimentada no codificador, onde é criado um conjunto de símbolos a partir dos dados de entrada; em seguida esse conjunto de símbolos são transmitidos ao longo do canal e, por fim, a representação codificada é alimentada no decodificador, onde uma imagem de saída reconstruída $f'(x,y)$ é gerada [Gonzales e Woods 2000]. A figura 21 representa essa estrutura de blocos, para a transmissão de um sinal.

Figura 21: Sistema de compressão e transmissão de dados

A redução ou eliminação das redundâncias na imagem de entrada é realizada no codificador. De acordo com a exigência da situação – por exemplo, requisitos específicos de fidelidade da imagem resultante – nessa etapa são empregadas as técnicas adequadas de codificação. Em geral, a etapa de codificação pode ser modelada através de uma série de três operações independentes: Mapeador, Quantizador e Codificador de símbolos (vide figura 22). Os modelos de codificação baseados em transformadas utilizam essas três estruturas [Ferreira 2002].

Figura 22: Bloco correspondente ao codificador

No primeiro estágio do processo de codificação, o mapeador transforma os dados de entrada da imagem em outro formato, de modo a reduzir as redundâncias entre pixels; as transformadas desempenham um papel importante nesta etapa, uma vez que são obtidas evidências de redundâncias em diferentes perspectivas, como no domínio espacial e no domínio espectral. Em seguida, é aplicada a quantização dos dados, onde há uma redução na acurácia dos dados obtidos do mapeador, de acordo com algum critério de fidelidade pré-estabelecido [Gonzales e Woods 2000]. Nesta etapa, ocorrem perdas de dados em relação à imagem original, representando uma operação irreversível. Desta forma, esta etapa é omitida nos casos de compressão sem perdas. Por fim, o codificador de símbolos gera um código de comprimento arbitrário para representar a saída do quantizador, de modo a otimizar a transmissão dos dados codificados.

Figura 23: Bloco correspondente ao decodificador

O decodificador apresenta apenas dois componentes, de forma inversa às etapas de codificação, conforme apresentado na figura 23. Primeiro é realizada a decodificação dos símbolos e em seguida o mapeamento inverso, de modo a obter a imagem de saída reconstruída. Como a etapa de quantização apresenta perdas irreversíveis de dados, não é possível incluir um bloco quantizador inverso no modelo.

3.1 Métodos de Compressão sem Perdas

Técnicas de compressão sem perdas apresentam métodos ditos universais, uma vez que não exigem conhecimento prévio do dado a ser comprimido. O arquivo original é analisado como uma sequência de bits, realizando a eliminação de redundância sequencialmente. As técnicas de compressão de imagens sem perdas utilizam um de dois conceitos: mapeamento e codificação de símbolos. A primeira identifica os símbolos utilizados para descrever o arquivo de dados, enquanto o segundo associa uma codificação específica para esse símbolo, de modo a modificar a quantidade de informação necessária para descrever estes dados. A seguir, apresentaremos alguns métodos de compressão de imagens sem perdas.

3.1.1 Métodos estatísticos

Uma abordagem simples de compressão de imagens sem perdas é realizar eliminação apenas sobre a redundância de codificação. Baseadas nos conceitos fundamentais da teoria da informação, especialmente na quantificação de infor-

mação, encontram-se as técnicas de codificação por tamanho variável. Considerando que alguns dados carregam informações mais relevantes que outros, suas representações também podem ser diferenciadas em relação aos menos relevantes. Entre as técnicas mais populares para a redução de redundância de codificação apresentamos a codificação de Huffman e a Run-Length Encoding (RLE).

O método de Huffman gera uma lista de todos os símbolos do alfabeto em ordem decrescente de suas probabilidades de ocorrência. Com base numa árvore de ocorrências, são distribuídos pesos para a representação dos símbolos na árvore, de forma que os símbolos mais frequentes necessitem de menos informação para serem representados [Solomon 2004]. Por isso leva o nome de codificação por tamanho variável, uma vez que esta técnica apresenta diferentes tamanhos de representação dos símbolos dos dados, de forma a otimizar sua apresentação.

Outro método de diminuição da redundância de codificação é o Run-Length Encoding (RLE). A ideia básica está em codificar cada grupo contíguo de informações, de acordo com a quantidade de aparições deste grupo, em uma leitura em varredura por linhas. Se a unidade de informação d ocorre n vezes consecutivas no mapeamento do objeto, pode-se substituir a representação de d repetidas n vezes por um único par de símbolos nd . Desta forma, a string 00001111 poderia ser representada por @40@41, reduzindo neste caso a transmissão de 8 bits para 6 bits. Nesta codificação de RLE, foi utilizado o símbolo "@" para representar o caractere de escape, indicando que se inicia uma sequência de n vezes a informação d .

3.1.2 Métodos baseados em dicionários

Os métodos baseados em dicionários consistem em selecionar conjuntos de símbolos recorrentes em um texto e utilizar outros símbolos para representá-los, criando um dicionário de símbolos. Essas técnicas também são baseadas na

teoria da informação, utilizando principalmente os fundamentos sobre entropia dos dados. Dado um conjunto de n símbolos, um compressor baseado em dicionário consegue comprimir este conjunto para nH bits, onde H representa a entropia deste conjunto. Estes dicionários podem ser gerados de forma estática ou de forma adaptativa, sendo este último mais utilizado, tanto pelo desempenho na compressão quanto na complexidade computacional. Uma das técnicas mais representativas desses métodos foi desenvolvida inicialmente por Jacob Ziv e Abraham Lempel, chamado método de compressão LZ, que apresentou diversas adaptações ao longo do tempo [Solomon 2004].

A compressão LZ se fundamenta em promover a modelagem e codificação dos dados em conjunto, identificando e removendo os diferentes tipos de redundância. A base da modelagem e codificação está em substituir os símbolos repetidos que se sucedem na fonte de dados por códigos únicos que compõem um dicionário de dados, com comprimentos menores ou iguais aos símbolos que eles representam. Inicialmente, a tabela de códigos é constituída apenas por símbolos isolados. À medida que são lidos novos conjuntos, a lista de símbolos é atualizada – utilizando subsequências temporárias – de forma a codificar de forma eficiente todos os símbolos únicos do dicionário. Um fator limitante que pode ocorrer é o tamanho da tabela do dicionário se tornar maior que o tamanho do arquivo original, necessitando-se arbitrar um limite máximo para entrada de códigos [Monteiro 1996].

Os métodos de compressão por dicionário criados por Jacob Ziv e Abraham Lempel (LZ), receberam diversas adaptações de acordo com a aplicação a ser utilizada. Apenas para citar alguns exemplos dessas variações: LZ77 [Ziv e Lempel 1977] e LZ78 [Ziv e Lempel 1978], criados pelo próprio Jacob Ziv; LZW [Welch 1984], desenvolvida por Terry Welch em 1984, tornou-se uma das variações mais populares destes métodos, porém seu uso permaneceu durante muito

tempo sob a licença de uma patente (número da patente nos EUA – 4.558.302, recebida em 10 de dezembro de 1985), embora esta esteja atualmente expirada¹; e a LZW [Bloom 1996], de Charles Bloom em 1996, com predição dos valores do dicionário.

3.1.3 Formatos de compressão de imagens sem perdas

Entre os formatos populares de apresentação de imagens digitais que utilizam técnicas de compressão sem perdas, encontram-se GIF, JPEG2000 sem perdas, TIFF e PNG. Nesta seção, apresentamos apenas os formatos GIF, TIFF e PNG, já que o formato JPEG2000 sem perdas é abordado na sessão seguinte, juntamente com sua contraparte que apresenta perdas.

O formato Tagged Image File Format, também conhecido como TIFF, está atualmente sob controle da Adobe Systems. É um formato muito utilizado no armazenamento de imagens escaneadas em alta resolução, suportando até 72 bits de tonalidades de cores (RGB) e arquivos de imagens de até 4 GiB (4.294.967.296 bytes). Assim como o formato GIF, utiliza o método LZW para compressão sem perdas de imagens, além de permitir que se armazene mais de uma imagem dentro de um mesmo arquivo no formato TIFF [Systems 1992].

O formato de arquivo de imagem GIF é um acrônimo para Graphics Interchange Format, desenvolvido pela CompuServe Information Services em 1987. Este formato é limitado a um espectro de cores de até 256 cores ou em tonalidades de cinza. Para a compressão dos dados da imagem, utiliza o método LZW [Jackson e Hannah 1993]. Seu funcionamento baseia-se em uma varredura linha a linha, buscando eliminar redundâncias através de correlação espacial.

Pelo fato do formato GIF utilizar métodos de compressão protegidos por patentes, iniciou-se o projeto de criação de um formato de representação de imagens

¹O método de compressão LZW [Solomon 2004].

com compressão sem perdas livre de patentes, flexível e com taxa de compressão superior ao GIF.

Em outubro de 1996, foi inaugurado o formato de imagens PNG, um formato de arquivos baseado em software livre e recomendado pela W3C [W3C 1996]. O formato PNG utiliza uma variação do método LZ 77, denominada LZH ou Deflate, seguido da compressão por codificação de Huffman [Conci, E. e R. 2008]. Diferente do formato GIF, o PNG suporta até 48 bits por pixel para imagens coloridas (RGB), além de um canal alfa para transparência, com até 256 níveis de transparência.

3.2 Métodos de Compressão com Perdas

As técnicas de compressão de imagens sem perdas utilizam a redução da redundância na representação de uma imagem para sua compactação. Nesse tipo de compressão, toda informação sobre a imagem original pode ser recuperada fielmente. Outra abordagem para a compressão é reduzir a quantidade de informações irrelevantes de uma imagem. Tomando como exemplo uma figura em tons de cinza, com 256 níveis distintos de representação de cinza: pessoas conseguem distinguir com precisão por volta de 40 níveis de cinza, o que permite que esta figura apresente ruídos e imprecisões sem afetar drasticamente as condições de observação a olho nu [Phillips 2000]. A redução baseada na irrelevância de informação é uma das práticas adotadas para se realizar a compressão com perdas.

Em geral, os pixels vizinhos de uma imagem encontram-se altamente correlacionados, isto é, o valor de um pixel está ligado aos valores de seus vizinhos. Isso implica em uma grande taxa de redundância entre pixels, o que representa uma possibilidade interessante de compressão. Quando conseguimos encontrar uma re-

apresentação dessa imagem em que esses valores encontram-se descorrelacionados, obtemos também a eliminação da redundância e a imagem pode ser considerada comprimida [Solomon 2004]. As técnicas de compressão de imagens com perdas utilizam essa técnica de transformação. Nas próximas seções, apresentamos duas técnicas de transformadas frequentemente empregadas na compressão de imagens: Transformada Discreta de Cosseno (DCT) e Transformada Discreta de Wavelets (DWT).

3.2.1 Transformada Discreta de Cosseno

A transformada discreta de cosseno (DCT) permite transformar um sinal de dados altamente correlacionados no domínio espacial em uma representação onde a energia está concentrada em apenas alguns coeficientes. Durante a transformação, os dados de entrada são convertidos em coeficientes da transformada de cosseno, sendo divididos em componentes de baixa frequência e alta frequência. A DCT é baseada em valores reais, ao contrário da DFT – que apresenta valores complexos em seus coeficientes – o que permite que obter uma melhor aproximação do sinal com menor quantidade de coeficientes [Saha 2000], além de não assumir que os dados a serem transformados sejam periódicos, como a DFT.

A transformada discreta de cosseno (DCT) para duas dimensões é definida pela equação 3.1 [Solomon 2004]:

$$G_{ij} = \frac{2}{\sqrt{mn}} C_i C_j \sum_{x=0}^{n-1} \sum_{y=0}^{m-1} p_{xy} \cos \left[\frac{(2y+1)j\pi}{2m} \right] \cos \left[\frac{(2x+1)i\pi}{2n} \right] \quad (3.1)$$

onde

$$C_t = \begin{cases} \frac{1}{\sqrt{2}} & \text{se } t = 0, \\ 1 & \text{se } t > 0, \end{cases} \text{ para } t = 0, 1, 2, \dots, n-1$$

Esta transformada recebe como entrada um conjunto de n valores de pixels, e a saída é um conjunto de n coeficientes da transformada de cosseno.

A função inversa da transformada discreta bidimensional do cosseno, chamada de IDCT, reconstrói os coeficientes da transformada para os pixels no domínio espacial:

$$p_t = \frac{2}{\sqrt{mn}} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} C_i C_j G_{ij} \cos \left[\frac{(2x+1)i\pi}{2n} \right] \cos \left[\frac{(2y+1)j\pi}{2m} \right] \quad (3.2)$$

onde

$$C_t = \begin{cases} \frac{1}{\sqrt{2}} & \text{se } t = 0, \\ 1 & \text{se } t > 0, \end{cases} \text{ para } t = 0, 1, 2, \dots, n-1$$

O padrão JPEG utiliza a DCT em conjunto com um bloco quantizador e com um bloco codificador para obter a compressão com perdas da imagem a ser representada. No caso de imagens coloridas, o JPEG converte a escala de cores para YCrCb, uma vez que o olho humano é mais sensível a pequenas mudanças em luminância do que na contraparte crominante; desta forma, o padrão JPEG realiza a maior parte de perda de dados nos campos de crominância, preservando maior parte dos dados no campo da luminância.

Tanto para imagens coloridas quanto para grayscale, o padrão JPEG subdivide a imagem original em blocos de 8x8 pixels, dividindo a aplicação da DCT sobre a imagem. Sobre cada bloco, aplica-se a DCT, gerando diversos blocos de 64 coeficientes, no domínio espectral. A equação 3.3 representa a DCT para esses

64 blocos, é derivada da equação 3.1 [Wallace 1992].

$$G_{ij} = \frac{1}{4}C_iC_j \sum_{x=0}^7 \sum_{y=0}^7 p_{xy} \cos \left[\frac{(2y+1)j\pi}{16} \right] \cos \left[\frac{(2x+1)i\pi}{16} \right] \quad (3.3)$$

A seguir, apresentamos um exemplo de aplicação da transformada do cosseno sobre um quadro de 8x8 pixels em tons de cinza, na tabela 2. A tabela 3 apresenta os coeficientes obtidos após a transformada. Logo em seguida, aplicamos sua anti-transformada (equação 3.2) para a reconstrução da matriz original na tabela 4.

Tabela 2: Quadro 8x8 de pixels, extraído de [Wallace 1992]

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	160	159	159	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

Tabela 3: Quadro após DCT, extraído de [Wallace 1992]

235.6	-1.0	-12.1	-5.2	2.1	-1.7	-2.7	1.3
-22.6	-17.5	-6.2	-3.2	-2.9	-0.1	0.4	-1.2
-10.9	-9.3	-1.6	1.5	0.2	-0.9	-0.6	-0.1
-7.1	-1.9	0.2	1.5	0.9	-0.1	0.0	0.3
-0.6	-0.8	1.5	1.6	-0.1	-0.7	0.6	1.3
1.8	-0.2	1.6	-0.3	-0.8	1.5	1.0	-1.0
-1.3	-0.4	-0.3	-1.5	-0.5	1.7	1.1	-0.8
-2.6	1.6	-3.8	-1.8	1.9	1.2	-0.6	-0.4

Pode-se perceber que, ao aplicarmos a transformada e, logo em seguida, sua anti-transformada sobre um sinal, conseguimos reconstruir de forma bastante fiel

Tabela 4: Quadro 8x8 de pixels reconstruído sem quantização

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	160	159	159	159
159	160	161	162	162	155	155	155
161	161	161	161	160	158	157	157
162	162	161	163	163	157	157	157
163	162	161	161	163	158	158	157

sua estrutura. As pequenas diferenças entre as duas imagens são provenientes das limitações de precisão durante as transformações. Após a DCT, o padrão JPEG realiza uma etapa de quantização dos elementos antes da anti-transformação, obtendo assim uma taxa maior de compressão sobre os coeficientes restantes, ao custo de eliminar dados e não permitir a reconstrução fiel do objeto inicial.

O JPEG utiliza uma tabela de quantização de 64 elementos – criada pelo próprio grupo da Joint Photographic Expert Group – e que serve como base para a quantização da matriz de coeficientes da transformada de cosseno [Saha 2000]. A tabela 5 apresenta a tabela padrão de quantização da International JPEG Group (IJG) [IJG 2010], cuja documentação e pacotes de aplicativos são disponíveis gratuitamente [Kornblum 2008].

Tabela 5: Quadro de quantização, extraído de [Wallace 1992]

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	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

Cada elemento da matriz de coeficientes é dividido pelo elemento correspondente na tabela de quantização e o resultado é arredondado para o inteiro mais

próximo. Em seguida, o resultado é desnormalizado, multiplicando cada elemento novamente pela própria tabela de quantização, gerando a nova tabela com valores quantizados (tabela 6). Esta etapa corresponde ao bloco Quantizador na figura 22 e onde ocorre o efeito da perda de dados na compressão.

Tabela 6: Quadro de coeficientes quantizados e desnormalizados, extraído de [Wallace 1992]

240	0	-10	0	0	0	0	0
-24	-12	0	0	0	0	0	0
-14	-13	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Após a quantização ocorre o terceiro processo do codificador JPEG. Os coeficientes da nova tabela são ordenados sequencialmente em uma organização em "zig-zag"(vide figura 24). Esta ordenação facilita a codificação da entropia ao agrupar os coeficientes de baixas frequências antes dos coeficientes de alta frequência [Saha 2000]. Desta forma, é possível obter uma compressão sem perdas durante a etapa de codificação. O padrão JPEG utiliza a codificação de Huffman nesta etapa.

Ao descomprimir o arquivo no padrão JPEG para visualização, aplicam-se as decodificação e a anti-transformada IDCT, segundo a equação 3.2. Na tabela 7, verifica-se o resultado após a descompressão. Nota-se que há diferença entre os valores, em relação à tabela 2, decorrentes das perdas da compressão.

Tabela 7: Quadro 8x8 reconstruído, extraído de [Wallace 1992]

144	146	149	152	154	156	156	156
148	150	152	154	156	156	156	156
155	156	157	158	158	157	156	155
160	161	161	162	161	159	157	155
163	163	164	163	162	160	158	156
163	164	164	164	162	160	158	157
160	161	162	162	162	161	159	158
158	159	161	161	162	161	159	158

3.2.2 Transformada Discreta de Wavelets

De forma análoga à transformada discreta de Fourier (DFT) e à transformada discreta de cosseno (DCT), a transformada de wavelets decompõe o sinal em subconjuntos apropriados de funções de base. Wavelets também podem ser denominadas small waves ou ondaletas. Como o próprio nome sugere, são ondas, oscilam e podem ser representadas por funções definidas sobre um intervalo finito. A ideia básica da transformada de wavelet é representar uma função arbitrária $f(u)$ como uma superposição de um conjunto de wavelets, derivadas por dilatações, contrações e translações de uma wavelet-mãe particular. A wavelet-mãe é uma função contínua real ou complexa que satisfaz as seguintes propriedades [Solomon 2004]:

1. A área total sobre a curva da função é zero:

$$\int_{-\infty}^{\infty} \psi(u) du = 0 \quad (3.4)$$

e

2. É do tipo quadrado integrável:

$$\int_{-\infty}^{\infty} |\psi(u)|^2 du = 1 \quad (3.5)$$

Uma wavelet é caracterizada como uma função de duração finita com valor médio nulo. Existe um número infinito de funções que satisfazem essas propriedades, ou seja, são candidatas a wavelets-mães.

Independente da escolha da wavelet-mãe, a transformada contínua de wavelet (CWT) é representada como o produto interno da função analisada e o conjunto de wavelets derivados da wavelet-mãe, conforme apresentada na equação 3.6:

$$F(a, b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{\infty} f(t) \psi_{a,b}(t) dt \quad (3.6)$$

nesta equação a e b representam os parâmetros de escala (contração ou dilatação) e de localização (deslocamento), respectivamente. A função $\psi_{a,b}(t)$ representa os valores da wavelet-mãe na transformada para os parâmetros a e b , segundo a fórmula 3.7:

$$\psi_{a,b}(u) = \frac{1}{\sqrt{a}} \psi\left(\frac{u-b}{a}\right) \quad , \quad a \neq 0, b \in \mathbb{R} \quad (3.7)$$

Pode-se perceber que a função de entrada $f(u)$ é representada por uma função de saída de duas variáveis $F(a, b)$, após a transformada. Essa é uma das propriedades da transformada de wavelets, representando simultaneamente a fase e a translação do sinal de entrada em uma mesma função. As técnicas de compressão de imagens baseadas em transformadas de wavelets se baseiam nessa representação bidimensional para a detecção e eliminação de redundâncias.

A transformada de wavelet consiste na obtenção de coeficientes de escala

aproximada e de escala detalhada de um mesmo sinal de entrada. Os coeficientes de escala aproximada são representados pelos valores de baixa frequência desse sinal, enquanto os coeficientes de escala detalhada localizam-se nos valores de alta frequência. A transformada de wavelet utiliza essa propriedade para realizar uma análise em multiresolução, técnica útil sobretudo para a compressão de imagens, onde podem ser obtidas diversos níveis de resolução de uma mesma imagem. Pode-se representar esta situação como a aplicação de um filtro passa-baixa e um filtro passa-alta sobre a função de entrada, coletando os coeficientes obtidos. Sobre os valores de escala aproximada (baixas frequências) obtidos, aplica-se novamente esta transformada, gerando um segundo nível de escala aproximada e de escala detalhada. Aplicando essa transformação sucessivamente, obtém-se diversas resoluções da mesma função de entrada, segmentados sempre em uma versão aproximada e outra versão complementar com os detalhes.

Através desta abordagem da transformada de wavelets para a obtenção de coeficientes de um filtro passa-alta (H) e um filtro passa-baixa (L), é introduzido o conceito de bancos de filtros para representar a transformada. Os coeficientes do filtro passa-baixa (L) são representados pela função $\theta(t)$ e os coeficientes do filtro passa-alta (H) são representados pela função $\psi(t)$. Após a geração desses coeficientes é realizada uma sub-amostragem em cada filtro, eliminando os coeficientes ímpares de cada um. Durante a síntese dos sinais, na transformada inversa de wavelets, é realizada uma super-amostragem, antes de se passar pelos filtros passa-alta e passa-baixa novamente.

3.2.3 Padrão de compressão JPEG2000

Apresentado em meados de 2000 como um novo padrão internacional de compressão de imagens, desenvolvido de forma conjunta pela International Organization for Standardization (ISO), International Electrotechnical Commission (IEC)

e a JPEG (Joint Photographic Experts Group) [Acharya 2005].

O padrão JPEG2000 é baseado na transformação discreta de wavelets (DWT). Apresenta diversas funcionalidades como as opções de compressão de imagens com perdas e sem perdas, transmissão progressiva da imagem, segundo critérios de qualidade, resolução, componentes ou localizações espaciais, e codificação por região de interesse [Marcellin et al. 2000]. Usando a propriedade de multiresolução da transformada de wavelets, é possível comprimir o arquivo uma vez e descomprimí-lo de inúmeras formas, sob os diferentes níveis de resolução codificados.

A compressão JPEG2000 é realizada em uma série de etapas. Inicialmente, é feita o mapeamento da imagem de entrada, seguido da transformada discreta de wavelet, quantização dos coeficientes e por fim a codificação de entropia. Para a descompressão do arquivo, são efetuados os passos inversos, excetuando o passo de quantização, da mesma forma que o padrão JPEG. Uma grande diferença entre esses dois padrões é o fato do JPEG2000 oferecer compressão com perdas e sem perdas, enquanto o JPEG permite apenas uma compressão com perdas. Para cada tipo de compressão, são executados algoritmos diferentes no JPEG2000.

3.2.3.1 Compressão JPEG2000 com perdas

Na compressão com perdas, utiliza-se a técnica de transformação irreversível de componente (ICT) para o mapeamento dos valores dos pixels da imagem de entrada. Inicialmente, transforma-se os valores de cada pixel sem sinal (isto é, que assumem valores de $[0, 2^{s-1}]$) para valores com sinal (de $[-2^{s-1}, 2^{s-1} - 1]$). Em seguida, realiza o mapeamento dos valores de pixels segundo a equação 3.10 [Solomon 2004]:

$$Y_0(x, y) = 0.299I_0(x, y) + 0.587I_1(x, y) + 0.144I_2(x, y), \quad (3.8)$$

$$Y_1(x, y) = -0.16875I_0(x, y) - 0.33126I_1(x, y) + 0.5I_2(x, y), \quad (3.9)$$

$$Y_2(x, y) = 0.5I_0(x, y) - 0.41869I_1(x, y) + 0.08131I_2(x, y) \quad (3.10)$$

onde, Y_0 , Y_1 e Y_2 correspondem aos três componentes de uma imagem colorida. No caso da imagem de entrada estar no formato RGB, verifica-se que o ICT é similar à representação dos componentes de cor YCrCb.

Após o mapeamento, ocorre a transformação discreta de wavelet. No caso da compressão com perdas, o JPEG2000 utiliza a família de wavelet CDF (9,7) de ponto flutuante.

A quantização dos coeficientes da transformada ocorre por nível de resolução, de acordo com o nível de compressão aplicado. Desta forma, é possível especificar a taxa de compressão desejada e a tabela de quantização é gerada de acordo. Por fim, a codificação dos coeficientes resultantes utiliza uma codificação MQ, classificando os coeficientes de acordo com sua dispersão estatística, relacionando os símbolos mais prováveis (MPS) e os menos prováveis (LPS), para aplicar uma codificação de estimativa probabilística de forma similar à técnica empregada na compressão binária JBIG e JBIG2 [Reavy e Boncelet 2001].

3.2.3.2 Compressão JPEG2000 sem perdas

A compressão sem perdas do JPEG2000 ocorre de forma análoga à compressão com perdas, com pequenas variações sobre os algoritmos utilizados. No mapeamento dos valores dos pixels da imagem de entrada, é empregada a transformação reversível de componente (RCT). Do mesmo modo que o ICT, transforma-se os valores de cada pixel sem sinal para valores com sinal. Já o mapeamento para o RCT ocorre segundo a equação 3.13 [Solomon 2004]:

$$Y_0(x, y) = \frac{I_0(x, y) + 2I_1(x, y) + I_2(x, y)}{4}, \quad (3.11)$$

$$Y_1(x, y) = I_2(x, y) - I_1(x, y), \quad (3.12)$$

$$Y_2(x, y) = I_0(x, y) - I_1(x, y) \quad (3.13)$$

Na transformação discreta de wavelets da compressão sem perdas, o JPEG2000 utiliza a família de wavelet CDF (5,3) de valores inteiros.

Por ser uma compressão reversível, a etapa de quantização ocorre com os valores unitários, ou seja, sem alteração sobre os coeficientes da transformação. Por fim, aplica-se as mesmas regras de codificação dos coeficientes utilizadas na compressão com perdas.

3.3 Métricas de compressão

Existem diversas técnicas de compressão de imagens com perdas e sem perdas. Para que seja possível avaliar a eficiência de cada técnica de compressão, é necessário adotar uma base adequada de métricas de compressão, para ambas as categorias de compressão de imagens.

A avaliação de técnicas de compressão de imagens sem perdas se baseia principalmente na taxa de compressão (equação 3.14) e no tempo de execução, uma vez que a qualidade da imagem é preservada durante o processo. Nas técnicas de compressão de imagens com perda, no entanto, a avaliação é mais complexa, devido às consequências associadas à perda de informação durante a compressão [Mrak, Grgic e Grgic 2003]. Para as técnicas de compressão com perdas, recomenda-se o emprego de avaliação objetivas e subjetivas para medir a qualidade gráfica das imagens comprimidas.

Nas pesquisas sobre métricas da avaliação da qualidade, é comum o emprego

de métricas de fidelidade de sinal para a realização da avaliação da qualidade visual das imagens comprimidas. A fidelidade do sinal, por sua vez, é medida com através da avaliação em relação a um sinal de referência, onde assume-se qualidade máxima [Sheikh e Bovik 2006]. Com isso, pode-se verificar a quantidade de artefatos introduzidos e o ruído obtido no impacto visual de uma imagem processada. Métodos de avaliação da qualidade visual que utilizem da imagem de referência para comparação são chamados métodos de referência completa (Full Reference – FR).

Em alguns casos, entretanto, trabalha-se com imagens processadas de imagens que por sua vez também apresentam artefatos ou distorções, ou seja, a imagem de referência não possui qualidade máxima, muitas vezes decorrente de ser meramente uma versão processada de uma imagem de referência que não está mais disponível [Rehman e Wang 2010]. Nestes casos, empregam-se métricas de avaliação de qualidade de referência reduzida (Reduced Reference – RR), analisando em geral apenas algumas propriedades da imagem, como efeitos nas bordas ou ruídos excessivos [Haddad et al. 2010]. Por fim, existem diversas pesquisas sobre o desenvolvimento de métricas para quando não há imagem de referência [Turaga, Chen e Caviedes 2002] [Tong et al. 2004] e [Zhang e Le 2010], chamadas métricas sem referência (No Reference – NR) ou métrica de qualidade cega (Blind Image Quality Assessment), onde trabalha-se apenas sobre a imagem observada, buscando identificar os artefatos inclusos, como efeitos de blocos ou de borrões nas imagens.

3.3.1 Métricas objetivas de qualidade visual

Neste trabalho, são empregadas apenas métricas de referência completa (FR), uma vez que todas as imagens analisadas possuem uma imagem de referência que é utilizada como base para os processamentos. No caso de imagens de documen-

tos digitalizados, essas imagens de referência correspondem à imagem original digitalizada (master), geralmente armazenada também para fins de preservação digital do documento. A seguir apresentam-se as métricas objetivas de avaliação de qualidade visual de referência completa.

Taxa de Compressão – Compression Rate (CR)

$$CR = \frac{tamOriginal - tamComprimida}{tamOriginal} \times 100\% \quad (3.14)$$

De acordo com Dung [Dung et al. 1998], uma das técnicas de avaliação objetiva mais utilizadas é a medida da razão sinal-ruído (PSNR), apresentada na equação 3.16. Essa métrica, por sua vez, baseia-se na medida do erro quadrático médio (MSE) (equação 3.15)[Mrak, Grgic e Grgic 2003].

Nas equações a seguir, M e N representam a quantidade de pixels na horizontal e na vertical, respectivamente; $f_{j,k}$ o valor do pixel nas coordenadas (j, k) da imagem de referência e $f'_{j,k}$ o valor do pixel nas coordenadas (j, k) da imagem comprimida.

Erro Médio Quadrático – Mean Square Error (MSE)

$$MSE = \frac{1}{MN} \sum_{j=1}^M \sum_{k=1}^N \left(f_{j,k} - f'_{j,k} \right)^2 \quad (3.15)$$

Pico da Razão Sinal-Ruído – Peak Signal-to-Noise Ratio (PSNR)

$$PSNR = 10 \log \frac{(2^n - 1)^2}{MSE} = 20 \log \frac{2^n - 1}{\sqrt{MSE}} \quad (3.16)$$

onde n é o número de bits para representar as tonalidades da imagem resultante. No caso da imagem digital ser representada com 256 tonalidades em grayscale ($n = 8$), o valor de PSNR é igual a $20 \log \frac{255}{\sqrt{MSE}}$.

Além destas métricas, utilizamos neste trabalho algumas outras métricas objetivas, apresentadas em: 3.20, 3.22, 3.23 [Mrak, Grgic e Grgic 2003] e em 3.21 [Rani, Bansal e Bansal 2009], 3.18 [Nirmala, Dandapat e Bora 2007].

Diferença Média – Average Difference (AD)

$$AD = \frac{1}{MN} \sum_{j=1}^M \sum_{k=1}^N \left([f_{j,k} - f'_{j,k}] \right) \quad (3.17)$$

Erro Médio Quadrático Laplaciano – Laplacian Mean Squared Error (LMSE)

$$LMSE = \frac{\sum_{j=1}^M \sum_{k=1}^N \left[O(f(j,k)) - O(f'(j,k)) \right]^2}{\sum_{j=1}^M \sum_{k=1}^N \left[O(f(j,k)) \right]^2} \quad (3.18)$$

onde, $O(f(j,k))$ é dado pela equação 3.19:

$$O(f(j,k)) = f(j+1,k) + f(j-1,k) + f(j,k+1) + f(j,k-1) \quad (3.19)$$

Correlação Cruzada Normalizada – Normalized Cross-Correlation(NK)

$$NK = \frac{\sum_{j=1}^M \sum_{k=1}^N [f_{j,k} \cdot f'_{j,k}]}{\sum_{j=1}^M \sum_{k=1}^N [f_{j,k}]^2} \quad (3.20)$$

Diferença Máxima – Maximum Difference (MD)

$$MD = \max \left(|x(m,n) - x'(m,n)| \right) \quad (3.21)$$

Erro Absoluto Normalizado – Normalized Absolute Error (NAE)

$$NAE = \frac{\sum_{i=1}^M \sum_{j=1}^N [|f_{j,k} - f'_{j,k}|]}{\sum_{i=1}^M \sum_{j=1}^N [f_{j,k}]^2} \quad (3.22)$$

Conteúdo Estrutural – Structural Content(SC)

$$SC = \frac{\sum_{i=1}^M \sum_{j=1}^N [f_{j,k}]^2}{\sum_{i=1}^M \sum_{j=1}^N [f'_{j,k}]^2} \quad (3.23)$$

Diversos autores [Mrak, Grgic e Grgic 2003] [Eskicioglu e Fisher 1995]. , inclusive o próprio artigo de Dung [Dung et al. 1998], afirmam que estas métricas objetivas não correlacionam muito bem com as avaliações subjetivas sobre a qualidade de imagens.

Conforme Chandler e Hermami [Chandler e Hemami 2005], diversas técnicas de compressão de imagens exploram o fato do sistema visual humano (HVS) ser um sensor imperfeito, uma vez que é possível distinguir uma categoria conhecida como visualmente sem perdas.

Zhou Wang et al [Wang et al. 2004] propuseram uma métrica objetiva de qualidade de imagem, chamada de Structural SIMilarity Index (SSIM), para comparar padrões locais na intensidade dos pixels entre a imagem comprimida e a imagem original, uma vez que o sistema visual humano é mais sensível a diferenças estruturais locais do que a pequenas variações na intensidade de brilho de uma imagem. A SSIM é definida como:

$$SSIM(x, y) = f(l(x, y)^\alpha, c(x, y)^\beta, s(x, y)^\gamma) \quad (3.24)$$

onde $l(x, y)$ representa a Comparação Luminosa, $c(x, y)$ é a Comparação de Contraste e $s(x, y)$ a Comparação Estrutural. Esses valores podem ser obtidos por:

$$l(x, y) = \frac{(2\mu_x\mu_y + C_1)}{(\mu_x^2 + \mu_y^2 + C_1)} \quad (3.25)$$

$$c(x, y) = \frac{(2\sigma_x\sigma_y + C_2)}{(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3.26)$$

$$s(x, y) = \frac{(2\sigma_{xy} + C_3)}{(\sigma_x\sigma_y + C_3)} \quad (3.27)$$

Os valores de $\alpha, \beta, \gamma > 0$ representam os pesos de cada comparação. Neste trabalho, foi adotado $\alpha, \beta, \gamma = 1$ para a análise do SSIM das imagens, compostas por janelas de tamanho de 8x8 pixels e pode-se reescrever a equação 3.24 como [Nirmala, Dandapat e Bora 2007]:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3.28)$$

onde $C_1, C_2, C_3 \ll 1$ são constantes utilizadas para evitar instabilidades quando $\mu_x^2 + \mu_y^2$ se aproximam de zero. Adotou-se $C_1, C_2 = 0$ durante os testes realizados.

3.3.2 Métricas subjetiva de qualidade visual

A avaliação subjetiva do resultado da compressão com perdas de uma imagem está diretamente relacionada com a interpretação realizada pelo sistema visual humano (HVS).

Uma das técnicas mais utilizadas na avaliação subjetiva é a avaliação baseada na percepção, quantificada pelo Pontuação da Opinião Média – Mean Opinion Score (MOS). Esta técnica envolve a participação de pessoas voluntárias, com

familiaridade no assunto de processamento de imagens, para a coleta dos resultados sobre a comparação da experiência visual com as imagens comprimidas, em relação à imagem original. Utiliza-se uma abordagem de estímulo duplo, usando a imagem original como referência e comparando com a imagem comprimida. A resposta é fornecida sob a forma de voto, com cinco níveis de percepção de qualidade visual [Grgic, Grgic e Mrak 2004], conforme a tabela 8:

Tabela 8: Grau de percepção da qualidade visual

Grau de Percepção	Percepção sobre a imagem
5	Imperceptível, fiel à imagem original
4	Perceptível, mas possível distinguir detalhes
3	Ligeiramente difícil de distinguir detalhes na imagem
2	Difícil de distinguir detalhes na imagem
1	Imagem irreconhecível

Ao fim dos testes, a medida de MOS da imagem de teste é calculada pela equação 3.29:

$$MOS = \sum_{i=1}^5 ip(i) \quad (3.29)$$

Na equação 3.29, i representa o nível de percepção e $p(i)$ indica a probabilidade do nível. Neste trabalho, não foram utilizadas métricas subjetivas da qualidade visual dos arquivos, apenas mantendo o foco nas métricas objetivas, para uso prático em um repositório de documentos digitais.

4 ANÁLISE DO IMPACTO DA TAXA DE COMPRESSÃO SOBRE A FIDELIDADE VISUAL

Para a realização da análise do impacto da taxa de compressão sobre a fidelidade visual, foram selecionados programas adequados para essa tarefa. Configura-se então um ambiente de simulação que apresente recursos computacionais que dêem suporte aos algoritmos e operações aritméticas utilizadas, com o intuito de se obter resultados consistentes com a avaliação teórica. Nesta seção, inicialmente é apresentado o laboratório de experimentação empregado e, em seguida, a análise dos resultados obtidos no trabalho de pesquisa de caracterização e compressão de imagens de documentos digitais.

Os testes foram realizados sobre diversas obras produzidas desde o século XVI até o século XIX, entre documentos impressos e documentos manuscritos, cuja relação é apresentada na tabela 9. Muitos destes documentos antigos apresentam efeitos de degradação causados pelo tempo e condições de conservação, cujo efeito visual torna-se ainda mais acentuado quando há introdução de ruídos durante a compressão com perdas. Para fins de comparação dos resultados apresentados pelas imagens de documentos, utilizou-se também imagens naturais e sintéticas, analisando assim o impacto causados pelas características distintas entre esses tipos de imagens. Todas as imagens utilizadas encontram-se no CD anexo à dissertação e são descritas no apêndice B, ao final do texto.

Este trabalho teve como foco o uso de técnicas de processamento de imagens para analisar o comportamento de imagens de documentos sob diferentes níveis de

Tabela 9: Tabela de referência com as imagens de documentos utilizadas

Título	Referência
Irmandade da Virgem Sra. do Rozário dos Pretos	[Sá BBD]
Physiologia das paixões e afecções, (...)	[Moraes BBD]
Historia de las cosas de Ethiopia	[Alvarez BBD]
Nova Lusitania , historia da guerra brasilica	[Freyre BBD]
C est la dedution du sumptueux ordre...	[Rouen BBD]
Historia dell'India America detta altramente Francia Antartica	[Thevet BBD]
Klare en Waarachtige Beschryving van de leste	[Moreau BBD]
Conquista espiritual	[Montoya BBD]
Sermoens do P. Antonio Vieira (Volume 05)	[Vieira BBD]
De saluberrima potione Chahve	[Banesio BBD]
Reflexões sobre a vaidade dos homens (...)	[Eca BBD]
Caramurú : poema epico do descobrimento da Bahia	[Duraõ BBD]
Origen de los indios de el nuevo mundo, e indias occidentales	[Garcia BBD]
Diccionario dos termos technicos de Historio Natural	[Vandelli BBD]
Auto da barca do inferno	[Vicente BBD]
Le pilote du Bresil - 1a ed	[Roussin BBD]
Cartas Botanicas	[Rousseau BBD]
Hans Staden: suas viagens e captiveiro entre os selvagens	[Staden BBD]
Almanaque do portugal luzo-brazileiro para 1818	[Castro BBD]

compressão. Para corroborar as abordagens teóricas apresentadas, foi empregado um laboratório de experimentação, com uma coleção diversificada de imagens de testes, além de um conjunto de programas específicos para esse fim. Para grande parte dos cálculos foram utilizados os programas MatLab 9.0 e Simulink [MATLAB 2010]; para integrar os diversos cálculos e replicar para as diversas imagens analisadas foram geradas rotinas automatizadas, de forma a acelerar a coleta dos resultados.

Entre os métodos de compressão abordados neste trabalho está a aplicação dos métodos de transformada de cosseno (DCT) e da transformada de wavelets (DWT) utilizados nos arquivos nos formatos JPEG [Wallace 1992] e JPG 2000 [Skodras, Christopoulos e Ebrahimi 2001], respectivamente. As imagens em JPEG são geradas a partir do programa lib-jpeg [JPG 2010]. Para realizar a compressão e conversão dos arquivos de imagem no formato JPG 2000, é necessário

a utilização ou construção de um conversor adequado. Neste trabalho, utiliza-se o kit de ferramentas do programa Jasper [JasPer 2010], desenvolvido por Michael D. Adams e baseado no modelo de código aberto. Este programa permite converter arquivos de imagem de diversos formatos (JPEG, BMP, PNM) para o formato JPG 2000 e vice-versa. Este programa é integrado a um sistema de processamento de imagens, facilitando a conversão entre os demais formatos de imagens disponíveis.

Nas simulações foram utilizadas técnicas de processamento de imagens, principalmente na conversão de imagens para outros formatos. Para realizar esse processamento das imagens, o software Imagemagick [IM 2010] foi empregado. Este programa, de código aberto, foi projetado para criar, editar e manipular imagens digitais; possui extensas bibliotecas sobre diversos formatos de imagens, apresentando o próprio programa Jasper em uma das bibliotecas disponíveis no Imagemagick [JasPer 2010]. As imagens em JPEG podem apresentar níveis de qualidade entre de 10 a 100, sendo 10 a qualidade visual mais inferior e maior taxa de compressão. As imagens em JPG 2000 são geradas pelo Jasper, realizando-se taxas de compressão de 1:4, 1:8, 1:16, 1:32 e 1:64, todos com blocksize de 32x32, de forma a obter taxas de compressão compatíveis com as imagens em JPEG, para comparação.

Devido à grande quantidade de operações aritméticas complexas envolvidas nos experimentos, optou-se por utilizar um conjunto de programas adequados para o laboratório, tanto do ponto de vista de desempenho quanto de funções pré-configuradas e possibilidade de ganho de escalabilidade. Para efetuar os cálculos aritméticos e algoritmos complexos, utilizamos o sistema de aplicativos MATLAB e Simulink [MATLAB 2010].

Para realizar a integração dessas ferramentas, utilizou-se algoritmos de processamento em lote, de forma a automatizar a sequência de passos na simulação

de cada imagem estudada. No próprio ambiente Linux, esses scripts de atividades foram desenvolvidos utilizando comandos da linguagem de programação Shell Script.

4.1 Medição das imagens de documentos

Todos os arquivos de imagem de entrada utilizados neste trabalho estão no formato *bmp*, seguindo a padronização nas pesquisas sobre imagens digitais. Para a utilização de aplicativos para representação da matriz de pixels foi utilizado o programa MatLab, com a função *imread*, que recebe uma imagem de entrada e armazena os valores dos pixels em uma matriz em código ASCII. Desta forma, é possível trabalhar apenas com matrizes de valores reais, onde serão aplicadas as equações apresentadas nos capítulos 2 e 3.

Um total de 456 imagens no formato JPEG e 380 imagens no formato JPG 2000 foram utilizadas nos experimentos, geradas a partir de 76 imagens de referência, entre documentos antigos, imagens naturais e sintéticas. O conjunto de imagens de documentos selecionado buscou apresentar diferentes formas de escrita (materiais manuscritos e impressos), textura de papéis, tipos de documentos (folhetos, livros e jornais) de diferentes épocas, publicados entre os séculos XVI e XIX. A tabela 9 apresenta a listagem dos documentos utilizados. No apêndice B, ao final da dissertação, há uma descrição da organização do CD, anexo ao trabalho. As figuras 25, 28, 27, 29 apresentam algumas imagens de documentos antigos analisadas.

Figura 25: Manuscrito século XVIII[Sá BBD]

gezeygte dinge/nit möchlich sein/geglaubt werden/vnd doch
so diese dinge bei verstendigen leuthen vorbracht/vnd erwas
gen werden/vor die gewiste vnd bestendigste ding geacht wer
den/vnd sich auch also erfinden.

Diß merck aus eynem oder.ij. Exempel so aus der Astro
nomi genommen/oder gezogen werden. Wir leuthe so wir
hie vmb Deutschlandt/oder nahe dabei wonen/wissen aus
lang herbrachter erfahrung/wie lang der Winter/der Sommer/
samt den andern zweien jars zeiten/herbst vnd lantz weren.
Item wie lang oder kürtz/der lengste tag im Sommer/vnd
der kürzte tag im winter/vnd so mit der nacht zu achten sei.

Wan nun gesagt wirt/das etzliche Ort in der welt sein
da die son in einem halben jar nit vnder gehe/vnd der lengste
tag bei den selben leuten vj. Monat/das ist eyns halben jars
langt sei/vnd her widderumb die lengste nacht auch vj. Mo
nat/odder eyn halb jar langt sei. Item das örer im der
welt funden werden/do in eynem jar/die quatuor tempora/
das ist die vier zeit des jars doppel sein. Also das zween wins
ter/zwen Sommer in einem jar gewislich da vorhanden sein.

Item das die Son samt andern sternenn wie kleyne sie vns
hie sein düncken/doch der kleinste stern im hymel grösser sei/
dann die ganze Erde/vnd der dinge vnzelich viel.

Wan nun der gemeyn Wan diese dinge höret/veracht ers
zum höchsten/gibt jm keynen glauben/vnd acht es sein din
ge/die vnmöglich sein. Diweil aber diese Natürliche dinge
bei den Astronomis der massen dargethan werden/das die
verstendigen der kunst/hieran nicht zweiffeln.

So muß derhalben nit volgen/diweil der gemeyn hauff
diese dinge vnwar helt/das es eben also sein muß/vnd wie
vbel würde die kunst der Astronomi stehen/wan sie dise him
lische corpora/nit künnte demonstrieren vnd anzeygen aus ge
wissen

Figura 26: Livro século XVI [Staden BBD]

(15)

Share. It is seated in the Persian Gulph, opposite to *Combrene*, on the maine land of *Persia*, and in respect of the commodious scituation, it stands fitly for the traf-
fique of *Arabia*, *Persia*, and *India*. Of this Iland the Ara-
bians vse to say.

Si terrarum orbis quasqua patet annulus esset,

Illius Ormusium gemma decusq; foret.

Which we may English thus:

If the world a ring were, then the precious stone

Of that rich ring were Ormus Ile alone.

The present *Sophie* being sensible of this thorne, which so long had stucke in his predecessours foote, re-
solves vpon a remedy befitting the greatnesse of his
Designe. And euer since his acquaintance with Sir
Robert Sherlie, resolving vpon the conquest of this Iland,
now vndertakes it with the auxiliarie Forces of the
Dutch and English, and that Land Armie which sub-
dued *Rhisme*, now re-inforced with new supplies from
the Continent. The 10. of February 1621. the Persians
landed 3000. of their men, who marched straight vnto
the Towne, which being a place of no great defence,
they tooke it, without any resistance at all; The Portu-
gals & the Moores retiring themselves into the Castle
with their goods and treasure. Vpon this the Chri-
stian Fleete plies the Forte, and the Enemies Fri-
gats which lay vnder the Lee thereof with their Ord-
nance: and the Mahometans Land their whole Ar-
mie, plant their Batteries, and mount their Cannon,
which they had borrowed from the Christians. And
that they might with more expedition and facility,
make themselves Maisters of this peece: they vnder-
mined

Figura 27: Livro século XVII [The Dutch Survey BBD]

Não falleis mais nada.

MUNDO.

Morte, despeja-os, não fique ninguém.

ISA. Oh quem me tivera mais vida alongada
Pera profetar da Virgem sagrada
Cem mil maravilhas que sei muito bem!

MOR. Profetas, nó mais;

Manda o Tempo que logo partais,
Parti-vos comigo, e não mais demoras.

ABR. Ó Morte, quão cruas são tuas esporas!
Quão lastimeiras! MOR. Não vos detenhais;
Andae, que são horas.

MOISES.

Senhor Rei David, não tendes na côrte
Cirurgiães e Fisicos mores,
Astrologos grandes e muitos doctores,
Que vos dem saude e livrem da morte?

MOR. Olhae, não vai nisso;

O mal que se cura não he mal de siso.
Andão deitando remendos á vida;
Mas quanto ao despejo, pois não tens guarida,
Lembra-te, homem, com muito aviso
Que es terra podrida.

BELZEBU.

Ó Morte, ó Morte, sejam bem casada,
Que tão limpa gente nos dás em poder.
Chegae-vos aqui, Senhor Lucifer,

Figura 28: Livro século XVIII [Vicente BBD]

(348)

te leguas de distrito, para que são tantos, quantos vemos? Tantos escriptaens, meirinhos, e alcaides, em cada Cidade, em cada Villa, e Aldea, de que servem; se basta hum para escrevinhar, e meirinhar este mundo, e mais o outro? Este alvitre se deu ao Rey de Castella não ha muitos annos, e não pegou; póde bem ser, que por ser bom para nós. Se esmarmos bem as rendas Reaes das Provincias, e as descutirmos, acharemos que lá ficaõ todas pelas unhas destes galfarros despendidas em salarios, e pitaças. Entremos nas sete Casas desta Corte, mas que seja na Alfandega, e casa da India, acharemos tantos officiaes, e ministros, que não ha quem se possa revolver com elles: e todos tem ordenados: e todos são tão necessarios, q̃ menos póde ser fizessem melhor tudo. A hum Mister de Lisboa ouvi dizer, que bastavaõ na Camera tres Vereadores, e que tinha sete; e que fora melhor poupar quatro mil cruzados para as guerras; e accrescentava: para que são na mesa do Paço oito, ou dez Dezembargadores, se bastaõ quatro, ou cinco? Na casa da Supplicação, para que são vinte, ou trinta, bastando meya duzia? E em todos effes Tribunaes, para que são tantos Conselheiros, que se estorvaõ huns aos outros. Engordaõ particulares com salarios, e em-

magre-

Figura 29: Livro século XIX [Vieira BBD]

São realizadas medições em outros tipos de imagens digitais, como imagens naturais (Lena na figura 52, Baboon na figura 53, Barbara na figura 54), para se obter um patamar de referência de comparação entre o comportamento das imagens de documentos e demais imagens, no tocante à variação da compressão e suas características inerentes. Além disso, para esses testes foram geradas duas imagens sintéticas de referência, a primeira composta por linhas horizontais e a segunda por colunas verticais, ambas de intensidade de brilho variando uniformemente de 0 a 255, conforme podem ser observadas na figura 30. Por ser uma imagem sintética, é possível trabalhar sob uma faixa controlada de valores de pixels, simulando efeitos específicos, como uma mudança constante dos valores de pixels ou avaliar se há diferenças para uma dispersão horizontal ou vertical em uma imagem, do ponto de vista das características e da compressão.

(a) Faixas horizontais variando
de 0 a 255

(b) Listras verticais variando
de 0 a 255

Figura 30: Imagens sintéticas incluída entre as imagens analisadas

Para cada imagem de teste selecionada, são aplicadas as técnicas de compressão de imagens baseadas em DCT e DWT, utilizando respectivamente os programas lib-jpeg [JPG 2010] e Jasper [JasPer 2010], em diferentes níveis de compressão, obtendo no processo diversas imagens comprimidas em JPEG e JPG 2000. Essas imagens comprimidas são, por sua vez, comparadas com a imagem de referência, afim de identificar os impactos da fidelidade visual obtida durante

a variação da taxa de compressão.

Em seguida são realizados os cálculos descritos no capítulo 2 para obter informações sobre as características dessa imagem. Especificamente, calculam-se os valores de média (equação 2.4), variância (equação 2.5), coeficiente de variância (equação 2.6), IAM (equação 2.7) e SAM (equação 2.9). Com essas medições, busca-se encontrar informações relevantes sobre as imagens analisadas e em seguida, verificar se há correlação destas informações com os efeitos visuais da imagem após a compressão.

Anexo à dissertação, encontra-se um CD com os códigos-fonte dos algoritmos empregados e imagens utilizadas. No apêndice A, ao final da dissertação, há uma explicação sobre o funcionamento de cada programa empregado.

4.2 Análise do impacto da taxa de compressão das imagens

Inicialmente, as imagens selecionadas são comprimidas e, no processo, geradas diversas versões de imagens nos formatos JPEG e JPG 2000, cada versão com um grau diferente de compressão, desde uma baixa taxa de compressão, com pequena mudança na fidelidade visual, até níveis altos de taxa de compressão, comprometendo significativamente a imagem final. Em seguida, essas imagens comprimidas são submetidas às equações apresentadas no capítulo 3, sempre em comparação com a imagem de referência, de modo a obter medições objetivas sobre os efeitos da compressão.

A seguir, são apresentados algumas imagens de documentos analisadas e os resultados obtidos na compressão, tanto em JPEG quanto em JPG 2000. As imagens de documento apresentadas são apenas uma pequena amostra do conjunto total analisado, de forma a apresentar os resultados de um exemplar de

cada século e também de algumas imagens manuscritas. Nas figuras 31, 32, 33, 34 e 35 são mostrados apenas fragmentos das imagens, com o intuito de permitir a visualização de detalhes e percepção dos impactos da compressão sobre a fidelidade visual na compressão utilizando a técnica DCT, nas imagens em JPEG e da técnica DWT, nas imagens em JPG 2000, para taxas de compressão de 88% ($CR = 0.88$) e de 98% ($CR = 0.98$). A tabela 10 reapresenta as siglas das métricas utilizadas, junto com sua nomenclatura completa, conforme apresentados no capítulo 3.

(a) Imagem de referência

(b) Imagem em JPEG, com $CR=0.88$ (c) Imagem em JPG 2000, com $CR=0.88$ (d) Imagem em JPEG, com $CR=0.98$ (e) Imagem em JPG 2000, com $CR=0.98$

Figura 31: Compressões da imagem de documento do século XIX [Castro BBD], onde a) representa a imagem de referência (sem compressão) e b) e c) representam as imagens comprimidas a uma taxa de compressão de 0.88 ($CR = 88\%$), nos formatos de imagem JPEG e JPG2000, respectivamente e d) e e) representam as imagens comprimidas à taxa de compressão de 0.98 ($CR = 98\%$)

(a) Imagem de referência

(b) Imagem em JPEG, com CR=0.88 (c) Imagem em JPG 2000, com CR=0.88

(d) Imagem em JPEG, com CR=0.98 (e) Imagem em JPG 2000, com CR=0.98

Figura 32: Compressões da imagem de documento do século XVIII [Garcia BBD]

(a) Imagem de referência

(b) Imagem em JPEG, com CR=0.88 (c) Imagem em JPG 2000, com CR=0.88

(d) Imagem em JPEG, com CR=0.98 (e) Imagem em JPG 2000, com CR=0.98

Figura 33: Compressões da imagem de documento do século XVII [Moreau BBD]

(a) Imagem de referência

(b) Imagem em JPEG, com CR=0.88 (c) Imagem em JPG 2000, com CR=0.88

(d) Imagem em JPEG, com CR=0.98 (e) Imagem em JPG 2000, com CR=0.98

Figura 34: Compressões da imagem de documento do século XVI [Alvarez BBD]

(a) Imagem de referência

(b) Imagem em JPEG, com CR=0.88

(c) Imagem em JPG 2000, com CR=0.88

(d) Imagem em JPEG, com CR=0.98

(e) Imagem em JPG 2000, com CR=0.98

Figura 35: Compressões da imagem de documento manuscrito do século XIX
[Morais BBD]

Tabela 10: Legenda das métricas objetivas utilizadas para averiguar a fidelidade visual após a compressão

Sigla	Nome da Métrica	Equação
CR	Compression Rate	3.14
PSNR	Peak Signal-to-Noise Rate	3.16
AD	Average Difference	3.17
LMSE	Laplacian Mean Square Error	3.18
NK	Normalized Cross-Correlation	3.20
MD	Maximum Difference	3.21
NAE	Normalized Absolute Error	3.22
SC	Structural Content	3.23
SSIM	Structural Similarity Index	3.24

Tabela 11: Compressão da imagem de documento do século XIX, no formato JPEG

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.88	44.94	0.00	0.31	0.99	10.00	0.01	1.00	0.99
0.92	42.55	0.01	0.42	0.99	11.00	0.01	1.00	0.98
0.96	38.08	-0.01	0.72	0.99	24.00	0.02	0.99	0.95
0.97	34.99	-0.01	1.05	0.99	43.00	0.03	0.99	0.92
0.98	31.60	-0.29	1.60	1.00	49.00	0.05	0.99	0.87

Tabela 12: Compressão da imagem de documento do século XIX, no formato JPG 2000

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.85	48.04	-0.07	0.12	1.00	6.00	0.00	0.99	0.99
0.92	42.99	-0.05	0.24	1.00	10.00	0.01	0.99	0.98
0.96	38.21	-0.32	0.55	1.00	20.00	0.02	0.99	0.96
0.98	34.24	-0.45	1.01	1.00	40.00	0.04	0.99	0.93
0.99	30.25	-0.35	1.11	0.99	60.00	0.06	0.99	0.87

Tabela 13: Compressão da imagem de documento do século XVIII, no formato JPEG

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.86	43.18	0.00	0.19	0.99	12.00	0.00	0.99	0.99
0.92	39.12	-0.02	0.35	0.99	20.00	0.01	0.99	0.98
0.94	37.13	-0.01	0.45	0.99	33.00	0.01	0.99	0.96
0.95	36.18	-0.05	0.51	1.00	33.00	0.01	0.99	0.95
0.96	33.15	-0.01	0.75	0.99	42.00	0.02	0.99	0.94
0.97	29.86	0.10	1.14	0.99	61.00	0.03	1.00	0.87

Tabela 14: Compressão da imagem de documento do século XVIII, no formato JPG 2000

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.85	45.47	-0.09	0.06	1.00	7.00	0.00	0.99	0.99
0.92	40.47	-0.16	0.21	1.00	20.00	0.01	0.99	0.97
0.96	34.73	-0.52	0.47	1.00	43.00	0.02	0.99	0.95
0.98	30.10	-0.88	0.85	1.00	80.00	0.03	0.99	0.90
0.99	25.47	-1.23	1.00	1.00	129.00	0.06	0.99	0.82

Tabela 15: Compressão da imagem de documento do século XVII, no formato JPEG

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.87	43.78	0.00	0.21	0.99	11.00	0.00	0.99	0.99
0.92	39.78	0.00	0.37	0.99	19.00	0.01	0.99	0.97
0.94	37.76	-0.01	0.48	0.99	22.00	0.01	0.99	0.96
0.95	36.75	-0.01	0.56	0.99	28.00	0.01	0.99	0.94
0.96	33.62	0.04	0.84	0.99	41.00	0.02	1.00	0.92
0.97	30.25	-0.03	1.29	0.99	65.00	0.03	0.99	0.87

Tabela 16: Compressão da imagem de documento do século XVII, no formato JPG 2000

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.85	45.68	-0.07	0.07	1.00	6.00	0.00	0.99	0.99
0.92	41.00	-0.08	0.23	1.00	16.00	0.01	0.99	0.98
0.96	35.55	-0.42	0.52	1.00	32.00	0.02	0.99	0.95
0.98	31.19	-0.74	0.84	1.00	57.00	0.03	0.99	0.91
0.99	25.95	-0.98	1.04	0.99	120.00	0.05	0.99	0.83

Tabela 17: Compressão da imagem de documento do século XVI, no formato JPEG

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.85	43.79	0.00	0.24	0.99	12.00	0.01	0.99	0.99
0.89	31.42	0.00	0.35	0.99	17.00	0.01	0.99	0.98
0.92	38.42	0.00	0.50	0.99	20.00	0.02	0.99	0.98
0.94	36.42	0.00	0.65	0.99	28.00	0.02	0.99	0.97
0.95	35.13	0.00	0.77	0.99	27.00	0.02	0.99	0.95
0.96	33.18	-0.02	1.00	0.99	36.00	0.03	0.99	0.93
0.97	29.69	0.07	1.60	0.99	65.00	0.05	1.00	0.88

Conforme é verificado nas tabelas apresentadas (a saber, tabelas 11, 12, 13, 14, 15, 16, 17, 18, 19, 20), conforme a taxa de compressão aumenta, valores de

Tabela 18: Compressão da imagem de documento do século XVI, no formato JPG 2000

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.85	45.34	-0.01	0.09	1.00	7.00	0.00	0.99	0.99
0.92	39.80	-0.27	0.32	1.00	21.00	0.01	0.99	0.98
0.96	34.19	-0.34	0.78	1.00	38.00	0.03	0.99	0.96
0.98	29.73	-0.55	0.98	1.00	52.00	0.05	0.99	0.91
0.99	25.54	-0.77	1.19	0.99	107.00	0.08	1.00	0.83

Tabela 19: Compressão da imagem de documento manuscrito do século XIX, no formato JPEG

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.90	47.22	0.00	0.61	0.99	8.00	0.00	1.00	0.99
0.93	44.92	0.00	0.79	0.99	10.00	0.01	1.00	0.98
0.96	41.31	0.00	1.15	0.99	17.00	0.01	0.99	0.96
0.97	38.95	0.01	1.49	0.99	24.00	0.02	1.00	0.95
0.97	36.91	0.02	1.93	0.99	35.00	0.02	1.00	0.93
0.98	33.05	-0.03	3.27	0.99	46.00	0.04	0.99	0.88

Tabela 20: Compressão da imagem de documento manuscrito do século XIX, no formato JPG 2000

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.85	50.03	-0.05	0.29	1.00	5.00	0.00	0.99	0.99
0.92	45.62	-0.15	0.56	1.00	8.00	0.00	0.99	0.98
0.96	42.53	-0.11	0.82	1.00	12.00	0.01	0.99	0.98
0.98	39.08	-0.24	1.07	1.00	28.00	0.02	0.99	0.96
0.99	34.92	-0.63	1.22	1.00	44.00	0.03	0.99	0.93

PSNR, *AD* e *SSIM* diminuem em razão proporcional, enquanto valores de *MD*, *NAE* e *LMSE* aumentam. Por fim, observa-se que há pequena variação nos valores de *NK* e *SC*.

Os gráficos apresentados nas figuras 36 (*PSNR*), 38 (*AD*), 40 (*LMSE*), 42 (*NK*), 44 (*MD*), 46 (*NAE*), 48 (*SC*), 50 (*SSIM*), apresentam a variação dos valores de cada uma dessas medições, conforme altera-se o valor de *CR*, para imagens de documentos nos formatos JPEG. Para as imagens de documentos no formato JPG 2000, são apresentados os gráficos nas figuras 37 (*PSNR*), 39 (*AD*), 41 (*LMSE*), 43 (*NK*), 45 (*MD*), 47 (*NAE*), 49 (*SC*), 51 (*SSIM*). As diversas linhas dispostas

nos gráficos correspondem aos valores coletados de cada imagem do documento ilustrado, para fins de comparações diretas dos valores.

Figura 36: Gráfico de CR por PSNR, para imagens de documento JPEG

Figura 37: Gráfico de CR por PSNR, para imagens de documento JPG 2000

Figura 38: Gráfico de CR por AD, para imagens de documento JPEG

Figura 39: Gráfico de CR por AD, para imagens de documento JPG 2000

Figura 40: Gráfico de CR por LMSE, para imagens de documento JPEG

Figura 41: Gráfico de CR por LMSE, para imagens de documento JPG 2000

Figura 42: Gráfico de CR por NK, para imagens de documento JPEG

Figura 43: Gráfico de CR por NK, para imagens de documento JPG 2000

Figura 44: Gráfico de CR por MD, para imagens de documento JPEG

Figura 45: Gráfico de CR por MD, para imagens de documento JPG 2000

Figura 46: Gráfico de CR por NAE, para imagens de documento JPEG

Figura 47: Gráfico de CR por NAE, para imagens de documento JPG 2000

Figura 48: Gráfico de CR por SC, para imagens de documento JPEG

Figura 49: Gráfico de CR por SC, para imagens de documento JPG 2000

Figura 50: Gráfico de CR por SSIM, para imagens de documento JPEG

Figura 51: Gráfico de CR por SSIM, para imagens de documento JPG 2000

Os experimentos realizados mostram que não há modificação drástica nas correlações estruturais das imagens comprimidas, conforme são apresentados nos gráficos das figuras 42, 43, 46, 47, 48 e 49. Isso se deve à pouca influência da compressão sobre essas métricas, ou seja, a compressão não modifica drasticamente as correlações estruturais das imagens comprimidas, tampouco às correlações cruzadas e o erro absoluto normalizados das imagens. Esses fenômenos se relacionam à forma como são realizadas as compressões nas técnicas baseadas em transformada, onde as perdas e as introduções de artefatos são mais perceptíveis em regiões de bordas, sendo as demais perdas espalhadas uniformemente ao longo da imagem, sem modificação na estrutura geral da imagem.

Valores de *MD* (Maximum Difference) apresenta grande variação quando amostrado para diferentes taxas de compressão. Os resultados para *AD* (Average Difference) apresentam pequena variação, com valores em torno de 0; isso ocorre devido à natureza da compressão por transformadas, onde se mantém o equilíbrio na distribuição dos pixels removidos. Isso. Esses valores, entretanto, demonstram apenas um valor global dos impactos da taxa de compressão, não apresentando diretamente os efeitos visuais causados. Nota-se, por meio desses valores, a quantidade de variação pixel a pixel entre as imagens de referência e comprimidas, de forma a verificar a severidade na introdução de artefatos visuais que podem

impactar na visualização e percepção da imagem.

Já a variação nos valores de *PSNR* (Peak Signal-to-Noise Rate), *LMSE* (Laplacian Mean Square Error) e *SSIM* (Structural Similarity) apresentam correlações com o impacto visual das imagens comprimidas, ao comparar com as respectivas imagens de referência. Conforme há variação na taxa de compressão, esses valores reagem proporcionalmente, permitindo uma leitura da fidelidade visual oferecida, e também a quantidade de influência nas regiões de borda (verificados principalmente pelo *LMSE*) e nas similaridades estruturais de ambas imagens.

Para fins de comparação, apresenta-se também as mesmas medições para imagens naturais, como a Lena, Baboon e Barbara (figuras 52, 53 e 54, respectivamente), e também para a imagem sintética utilizada.

(a) Imagem de referência

(b) Imagem em JPEG, com CR=0.98 (c) Imagem em JPG 2000, com CR=0.98

Figura 52: Compressões da imagem natural Lena

Tabela 21: Compressão da imagem Lena, no formato JPEG

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.77	40.78	0.00	0.35	0.99	21.00	0.01	0.99	0.96
0.85	38.50	0.00	0.50	0.99	29.00	0.01	1.00	0.94
0.92	35.78	0.00	0.70	0.99	42.00	0.02	1.00	0.92
0.93	35.10	-0.03	0.76	0.99	47.00	0.02	0.99	0.90
0.94	34.26	0.00	0.82	0.99	55.00	0.02	1.00	0.89
0.95	32.94	0.01	0.92	0.99	55.00	0.03	1.00	0.87
0.97	30.39	0.01	1.15	0.99	89.00	0.04	1.00	0.82

Tabela 22: Compressão da imagem Lena, no formato JPG 2000

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.75	43.36	0.02	0.11	1.00	9.00	0.01	0.99	0.98
0.87	39.29	-0.17	0.29	1.00	20.00	0.01	0.99	0.95
0.93	36.28	-0.41	0.50	1.00	33.00	0.02	0.99	0.92
0.96	33.14	-0.63	0.72	1.00	51.00	0.03	0.99	0.88
0.98	30.17	-0.93	0.91	1.00	75.00	0.04	0.98	0.83

(a) Imagem de referência

(b) Imagem em JPEG, com CR=0.98 (c) Imagem em JPG 2000, com CR=0.98

Figura 53: Compressões da imagem natural Baboon

Tabela 23: Compressão da imagem Baboon, no formato JPEG

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.57	37.09	0.00	0.08	0.99	20.00	0.02	0.99	0.97
0.76	30.47	0.00	0.34	0.99	49.00	0.04	0.99	0.92
0.82	28.22	0.00	0.51	0.99	74.00	0.05	1.00	0.86
0.88	26.44	0.00	0.66	0.99	92.00	0.06	1.00	0.83
0.91	25.26	0.00	0.76	0.99	97.00	0.07	1.00	0.78
0.95	23.42	-0.02	0.91	0.98	104.00	0.09	1.00	0.68

Tabela 24: Compressão da imagem Baboon, no formato JPG 2000

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.50	43.85	0.13	0.00	0.99	7.00	0.00	1.00	0.99
0.75	34.12	-0.20	0.08	1.00	33.00	0.03	0.99	0.93
0.87	28.59	-0.42	0.28	1.00	62.00	0.05	0.99	0.84
0.93	25.06	-0.80	0.56	0.99	99.00	0.08	0.99	0.72
0.96	22.74	-0.99	0.79	0.99	125.00	0.10	0.99	0.59
0.98	21.27	-1.16	0.91	0.98	126.00	0.12	1.00	0.47

(a) Imagem de referência

(b) Imagem em JPEG, com CR=0.98 (c) Imagem em JPG 2000, com CR=0.98

Figura 54: Compressões da imagem natural Barbara

Tabela 25: Compressão da imagem Barbara, no formato JPEG

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.87	34.00	0.00	0.20	0.99	38.00	0.03	0.99	0.94
0.89	32.88	0.00	0.24	0.99	48.00	0.03	0.99	0.93
0.90	31.76	-0.01	0.30	0.99	59.00	0.03	0.99	0.91
0.92	30.32	0.00	0.40	0.99	79.00	0.04	0.99	0.89
0.96	25.78	-0.02	0.84	0.99	127.00	0.07	1.00	0.76

Tabela 26: Compressão da imagem Barbara, no formato JPG 2000

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.87	36.56	-0.32	0.07	1.00	23.00	0.02	0.99	0.95
0.93	31.41	-0.53	0.21	1.00	45.00	0.04	0.99	0.89
0.96	27.70	-0.75	0.37	1.00	69.00	0.06	0.99	0.81
0.98	25.00	-0.91	0.64	0.99	108.00	0.08	0.99	0.71

(a) Imagem de referência

(b) Imagem em JPEG, com CR=0.98

(c) Imagem em JPG 2000, com CR=0.98

Figura 55: Compressões da imagem sintética Vertical

Tabela 27: Compressão da imagem Vertical, no formato JPEG

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.98	59.90	-0.00	0.70	1.00	1.00	0.00	0.99	0.99
0.99	53.88	0.00	2.30	0.99	1.00	0.00	1.00	0.98
0.99	43.01	0.04	3.43	1.00	4.00	0.01	0.99	0.93
0.99	36.19	0.12	23.91	1.00	9.00	0.02	0.99	0.87

Tabela 28: Compressão da imagem Vertical, no formato JPG 2000

CR	PSNR	AD	LMSE	NK	MD	NAE	SC	SSIM
0.99	60.17	0.06	0.36	0.99	1.00	0.00	1.00	0.99

Verificam-se medições similares entre as imagens de documentos e imagens naturais, quanto ao comportamento das variáveis. Apresentou-se novamente pouca variação nos valores de *SC*, *NAE* e *NK*, (tabelas 21, 22, 23, 24, 25, 26, 27, 28) e variações proporcionais nos demais valores.

Para ilustrar os efeitos visuais da compressão nas imagens analisadas, são apresentados os testes com uma imagem de documento do século XVII, redigida em inglês, conforme a figura 56. Esta imagem corresponde a um fragmento da imagem original, com um zoom de 100% para visualizar os efeitos da compressão em uma área ampliada. As figuras 57 e 58 ilustram uma imagem de documento de referência, e suas versões comprimidas como imagens em JPEG e JPG 2000, ambas às taxas de compressão de 88% ($CR = 0.88$, figura 57) e 98% ($CR = 0.98$, figura 58) ao zoom de 150%, para facilidade de visualização dos artefatos introduzidos e para comparação da fidelidade visual.

Figura 56: Imagem de documento do século XVII [The Dutch Survey BBD], sem compressão

Na tabela 29, os valores correspondentes de PSNR, SSIM e LMSE para essas quatro imagens comprimidas (JPEG a $CR = 88\%$, JPEG a $CR = 98\%$, JPG2000 a $CR = 88\%$ e JPG2000 a $CR = 98\%$) são apresentados. Nas imagens comprimidas a um $CR = 98\%$, a métrica PSNR indica maior fidelidade visual na imagem em JPG 2000 (PSNR=35,441) em relação à imagem em JPEG (PSNR=33,545), quando comparados à imagem de referência, como pode ser visualmente comparada na figura 58.

Tabela 29: Resultados da compressão do documento do século XVII, em JPEG e JPG 2000, para $CR = 0.88$ e $CR = 0.98$

Métrica	JPEG	JPG 2000	JPEG	JPG 2000
CR	0.88	0.88	0.981	0.984
PSNR	43.31	43.67	33.545	35.441
LMSE	0.39	0.23	2.154	1.031
SSIM	0.99	0.99	0.985	0.996

(a) Imagem de referência (b) Imagem em JPEG (c) Imagem em JPG 2000

Figura 57: Imagem de documento do século XVII, onde a) representa a imagem de referência (sem compressão) e b) e c) representam as imagens comprimidas a uma taxa de compressão de 0.88 ($CR = 88\%$), nos formatos de imagem JPEG e JPG2000, respectivamente

(a) Imagem de referência (b) Imagem em JPEG (c) Imagem em JPG 2000

Figura 58: Imagem de documento do século XVII, onde a) representa a imagem de referência (sem compressão) e b) e c) representam as imagens comprimidas a uma taxa de compressão de 0.98 ($CR = 98\%$), nos formatos de imagem JPEG e JPG2000, respectivamente

Em uma imagem de documento, ambas as técnicas de compressão baseadas em DCT e DWT introduzem artefatos visuais nas imagens comprimidas, especialmente nas regiões de bordas, o que afeta diretamente a fidelidade visual à imagem de referência. Para baixas taxas de compressão, os valores de PSNR indicam níveis similares de fidelidade visual entre as duas técnicas de compressão, como pode ser observado na figura 58. Entretanto, para taxas maiores de

compressão (acima de 95% de compressão), as imagens comprimidas em JPEG apresentam menor fidelidade visual quando comparadas às imagens em JPG 2000, o que implica que os artefatos de blocagem introduzidos durante a compressão baseada em DCT causam maiores impactos visuais do que os artefatos de borrões, introduzidos pela compressão de imagens baseada em DWT.

O gráfico da figura 59 ilustra a variações de valores de PSNR quando CR varia. Para valores de CR menores que 94%, a diferença entre os valores de PSNR das imagens em JPEG e JPG 2000 é da ordem de 1.0. Por exemplo, na tabela 29, o valor de PSNR para JPEG é de 43.31, enquanto para a imagem em JPG 2000 é de 43.67. Já para valores superiores, CR acima de 95%, entretanto, uma diferença significativa é observada na mesma tabela, com valores de 33.54 e 35.44 para as imagens em JPEG e JPG 2000, respectivamente. Essa diferença pode ser observada na figura 58, onde a imagem em JPG 2000 apresenta maior fidelidade visual do que a imagem em JPEG.

Figura 59: CR x PSNR

A métrica objetiva LMSE, apresentado no gráfico da figura 60, aponta o fenômeno de introdução de artefatos sobre as regiões de bordas. A partir de valores de $CR > 0.95$, observa-se um crescimento acentuado no valor do LMSE para imagem

em JPEG, enquanto na imagem JPG 2000 o crescimento se manteve constante; isso implica na presença de maiores quantidade de artefatos introduzidos nas imagens em JPEG do que nas imagens JPEG 2000 para altas de compressão, em concordância com os valores apresentados pelo PSNR. Novamente utilizando a figura 58 como base, pode-se perceber que os efeitos de bloco na imagem em JPEG tornam-se bem visíveis principalmente nas regiões do contorno da imagem. Na imagem em JPG 2000, também se observa um acúmulo de artefatos de bordões no contorno da imagem, porém em menor intensidade do que na imagem em JPEG. Essa diferença na introdução de artefatos pode ser medido pela diferença dos valores de *LMSE* apresentados. A saber, na figura 58, para $CR = 98\%$, os valores de *LMSE* para imagens em JPEG e JPG 2000, respectivamente, são 2.15 e 1.03.

Por fim, o índice de SSIM indica pequena diferença entre ambas as imagens e valores altos, o que simboliza efeitos semelhantes na fidelidade visual quanto à similaridade estrutural em relação à imagem de referência. Para taxas elevadas de compressão (CR acima de 95%), observa-se uma disparidade entre os valores das imagens de JPEG e JPEG 2000 (0.985 e 0.996, respectivamente), acompanhando os resultados obtidos pelas demais métricas, com as imagens em JPEG 2000 apresentando maior similaridade estrutural do que as imagens em JPEG a altas taxas de compressão, como pode ser observado no gráfico da figura 61.

Figura 60: CR x LMSE

Figura 61: CR x SSIM

É interessante notar a diferença dos valores apresentados entre ambos tipos de documentos. Os gráficos nas figuras 62, 64 e 66 apresentam a dispersão dos valores das imagens de documentos frente às imagens naturais, onde todas as imagens estão no formato JPEG. As figuras 63, 65 e 67 ilustram os mesmos experimentos para imagens no formato JPG 2000.

Figura 62: Gráfico de CR por PSNR, para diversas imagens JPEG

Figura 63: Gráfico de CR por PSNR, para diversas imagens JPG 2000

Figura 64: Gráfico de CR por LMSE, para diversas imagens JPEG

Figura 65: Gráfico de CR por LMSE, para diversas imagens JPG 2000

Figura 66: Gráfico de CR por SSIM, para diversas imagens JPEG

Figura 67: Gráfico de CR por SSIM, para diversas imagens JPG 2000

Pode-se observar pelas figuras 62 e 63 que, para uma mesma taxa de compressão, observam-se valores superiores de *PSNR* nas imagens de documentos,

em relação às imagens naturais. Isso implica em maior fidelidade visual, quando comparados a imagem comprimida e a respectiva imagem de referência. Também apresentam-se variações compatíveis de *SSIM* e *LMSE*, apontando comportamentos similares ao *PSNR*, entre os tipos de imagens.

Na seção seguinte, são analisadas as características das imagens analisadas que correlacionem com essa diferença de comportamento durante a compressão.

4.3 Análise das características das imagens

Submetendo as mesmas imagens utilizadas na seção anterior às equações apresentadas no capítulo 2, busca-se extrair as características que podem distinguir o comportamento das imagens de documento frente às imagens naturais ou as imagens sintéticas. Para essa experimentação, utilizam-se apenas as imagens de referência, sem compressão.

Utilizando a mesma imagem de documento do século XVII da seção anterior, porém sem zoom, visualizando a imagem inteira, ilustrado pela figura 68, como exemplo dos cálculos realizados para obtenção das características da imagem de documento. A tabela 30 apresenta os valores obtidos sobre as características desta imagem.

Figura 68: Imagem de documento do século XVII [The Dutch Survey BBD], sem zoom

Tabela 30: Características do documento do século XVII

Característica	Valor	Característica	Valor
Dimensões	1584 x 2293	Tamanho arquivo	3745 KB
Profundidade de cores	8 bits	Coef. de Variação (η)	0.345
Média (m)	137.135	Variância (σ)	37.177
IAM	4.119	SAM	6364.709

Em seguida, foram aplicados os mesmos cálculos para as demais imagens analisadas. A seguir, apresenta-se a tabela 31, contendo os valores obtidos tanto para diversas imagens de documentos quanto para imagens naturais, além das duas imagens sintéticas, para fins de comparação.

Dentre os valores obtidos na tabela 31, pode-se perceber que os resultados de média, variância e coeficiente de variação são retratos globais da imagem, verificando o comportamento global quanto à dispersão de pixels e a complexidade da imagem. Esses resultados auxiliam indiretamente a caracterização de uma imagem de documento, quando analisados junto a outros parâmetros. No caso,

Tabela 31: Características de algumas imagens analisadas

Imagem	m	σ	η	SAM	IAM
Imagens de Documentos					
Sec XIX (1)	166.81	35.88	0.21	1962	6.87
Sec XIX (2)	135.64	30.75	0.22	2240	5.73
Sec XVIII (1)	145.11	31.03	0.21	5414	4.91
Sec XVIII (2)	164.4	38.33	0.22	2285	6.27
Sec XVII (1)	164.08	37.87	0.23	1604	6.84
Sec XVII (2)	126.1	30.9	0.24	2976	5.44
Sec XVI (1)	155.15	47.38	0.30	2349	6.45
Sec XVI (2)	116.45	33.01	0.28	3169	5.62
Manusc. (1)	185.40	23.99	0.13	14110	3.17
Manusc. (2)	100.54	37.11	0.37	7268	3.81
Imagens Naturais					
Lena	124.04	47.85	0.38	904	6.15
Baboon	129.14	42.30	0.32	99.7	17.31
Barbara	112.44	47.19	0.41	458.2	11.54
Peppers	104.2	57.40	0.55	570	6.83
Vertical	126.94	73.99	0.58	Inf	2.48
Horizontal	126.94	73.99	0.58	Inf	2.48

esses valores foram empregados na métrica objetiva de compressão $SSIM$.

Os valores de IAM e SAM apontaram características que podem distinguir o padrão de imagens de documentos de imagens naturais, por exemplo. Observa-se que, em geral, imagens de documentos e também as imagens sintéticas, apresentam valores muito superiores de SAM , no domínio espectral e valores inferiores de IAM , no domínio espacial, em relação às imagens naturais. Isso ocorre devido à alta redundância espacial das imagens de documento [Jayant e Noll 1984] e sua menor entropia [Saha e Vemuri 2000], efeitos representados nas regiões de valores de brilho bem delimitados, como as partes do fundo do papel e a disparidade com os caracteres que compõem os textos. Valores altos de IAM e valores baixos de SAM , analogamente, representam baixa redundância espacial, ou seja, mais componentes de alta frequência [Saha e Vemuri 2000].

A figura 69 apresenta algumas imagens de documento analisadas, como um manuscrito ($IAM = 3.1$ e $SAM = 14.110$), texto do século XVII ($IAM = 4.9$ e

Figura 69: IAM e SAM de imagens de documentos

$SAM = 2776$), texto do século XVIII ($IAM = 4.1$ e $SAM = 6364$) e texto do século XIX ($IAM = 2.1$ e $SAM = 23771$).

Os mesmos experimentos foram realizados com imagens naturais. A figura 70 apresenta os resultados obtidos com as imagens Lena ($IAM = 6.2$ e $SAM = 904$), Baboon ($IAM = 17.3$ e $SAM = 99.7$), Peppers ($IAM = 6.8$ e $SAM = 570$) e Goldhill ($IAM = 11.5$ e $SAM = 458$). Observa-se que as imagens naturais apresentam em geral valores superiores de IAM e inferiores de SAM, quando comparados com imagens de documentos. Isso ocorre devido à quantidade de detalhes decorrentes da representação de imagens complexas da natureza, com diversas variações nos valores de brilhos, desde suaves representando um degradê das tonalidades da pele, até bruscos na representação de cada detalhe dos pelos

na imagem Baboon ou dos cabelos na Lena.

(a) Lena

$IAM = 6.2$

$SAM = 904$

(b) Baboon

$IAM = 17.3$

$SAM = 99.7$

(c) Peppers

$IAM = 6.8$

$SAM = 570$

(d) Barbara

$IAM = 11.5$

$SAM = 458$

Figura 70: IAM e SAM de imagens naturais

Por fim, esses cálculos foram aplicados às duas imagens sintéticas da figura 30. Ambas imagens apresentaram valores idênticos, tanto para a caracterização quanto nos resultados da compressão, demonstrando que a orientação espacial da imagem não afeta essas propriedades nesses processos. Estas imagens foram utilizadas por representarem imagens com valor baixo de IAM ($IAM = 2.4$) e valor muito elevado de SAM ($SAM = inf$), para avaliação dos resultados obtidos.

4.4 Correlações entre compressão e características das imagens

Os resultados experimentais obtidos apontam correlações entre as características IAM e SAM das imagens de referência com o impacto das técnicas de compressão com perdas, em relação às taxas de compressão e também da fidelidade visual. De acordo com os resultados comparativos entre as técnicas de compressão, observou-se que as imagens em JPEG 2000 apresentam qualidade visual superior às imagens em JPEG para altas taxas de compressão (CR acima de 95%), de acordo com as três métricas objetivas empregadas. Porém, correlacionando com as características intrínsecas das imagens digitais, pode-se perceber que os níveis de qualidade visual apresentados pelas compressões a altas taxas de compressão são condicionados aos valores de IAM e SAM [Cui et al. 2008]. Valores altos de SAM e valores baixos de IAM implicam em valores superiores de PSNR e SSIM e valores inferiores de LMSE, ou seja, com maior fidelidade visual à imagem de referência, do que valores baixos de SAM e valores elevados de IAM, para uma mesma taxa de compressão.

As tabelas 33 e 34 apresentam a comparação dos valores de IAM, SAM e os valores obtidos de PSNR, LMSE e SSIM para as técnicas de compressão baseadas em DCT e DWT, respectivamente. Nessas tabelas encontram-se os resultados obtidos com diversas imagens analisadas, tanto de imagens de documentos quanto imagens naturais. Na tabela 32, apresenta-se uma legenda de orientação de cada imagem de documento analisada nas tabelas 33 e 34. No apêndice B da dissertação, as imagens utilizadas de cada obra são descritas com mais detalhes e podem ser encontradas no CD anexo ao trabalho.

Tabela 32: Tabela de referência das imagens de documentos analisadas

Imagem	Referência
Manuscrito (1)	[Sá BBD]
Manuscrito (2)	[Morais BBD]
Séc. XVI (1)	[Alvarez BBD]
Séc. XVI (2)	[Freyre BBD]
Séc. XVI (3)	[Rouen BBD]
Séc. XVI (4)	[Thevet BBD]
Séc. XVII (1)	[Moreau BBD]
Séc. XVII (2)	[Montoya BBD]
Séc. XVII (3)	[Vieira BBD]
Séc. XVII (4)	[Banesio BBD]
Séc. XVIII (1)	[Eca BBD]
Séc. XVIII (2)	[Duraõ BBD]
Séc. XVIII (3)	[Garcia BBD]
Séc. XVIII (4)	[Vandelli BBD]
Séc. XIX (1)	[Castro BBD]
Séc. XIX (2)	[Roussin BBD]
Séc. XIX (3)	[Rousseau BBD]
Séc. XIX (4)	[Vicente BBD]

Tabela 33: Tabela comparativa para imagens em JPG

Imagem	SAM	IAM	CR	PSNR	LMSE	SSIM
Natural						
Baboon	99.7	17.3	0.95	23.43	0.92	0.69
Barbara	458.2	11.5	0.97	25.78	0.84	0.76
Lena	904.7	6.16	0.97	30.4	1.16	0.82
Peppers	570.5	6.83	0.97	30.18	0.95	0.79
Vertical	Inf	2.48	0.99	47.66	1.08	0.92
Horizontal	Inf	2.48	0.99	60.17	0.36	0.95
Documentos						
Manuscrito (1)	14110	3.2	0.98	32.94	1.42	0.90
Manuscrito (2)	7368	3.8	0.98	32.14	3.27	0.88
séc. XVI (1)	5743	3.1	0.98	32.6	2.15	0.93
séc. XVI (2)	7645	4.2	0.97	32.08	1.71	0.90
séc. XVI (3)	3471	4.76	0.98	31.6	1.53	0.88
séc. XVI (4)	2194	6.89	0.97	32.3	0.78	0.91
séc. XVII (1)	2776	4.9	0.97	30.55	0.97	0.89
séc. XVII (2)	18312	3.3	0.97	37.44	1.25	0.95
séc. XVII (3)	2919	6.87	0.97	33.6	0.84	0.92
séc. XVII (4)	2976	5.4	0.97	34.2	1.03	0.92
séc. XVIII (1)	6759	2.6	0.98	33.18	1.05	0.93
séc. XVIII (2)	5512	3.1	0.98	36.15	0.59	0.95
séc. XVIII (3)	5414	4.91	0.97	35.44	1.01	0.94
séc. XVIII (4)	11683	2.99	0.98	37.87	1.03	0.94
séc. XIX (1)	4375	2.7	0.98	37.39	1.24	0.92
séc. XIX (2)	4325	6.4	0.97	31.87	1.0	0.77
séc. XIX (3)	5252	4.76	0.97	35.16	0.93	0.87
séc. XIX (4)	4685	3.55	0.98	36.24	0.88	0.93

Tabela 34: Tabela comparativa para imagens em JPG 2000

Imagem	SAM	IAM	CR	PSNR	LMSE	SSIM
Natural						
Baboon	99.7	17.3	0.94	25.1	0.56	0.72
Barbara	458.2	11.5	0.97	27.7	0.38	0.81
Lena	904.7	6.16	0.98	30.7	0.92	0.83
Peppers	570.5	6.83	0.97	32.97	0.60	0.81
Vertical	Inf	2.48	0.99	60.17	0.36	0.95
Horizontal	Inf	2.48	0.99	60.17	0.36	0.95
Documentos						
Manuscrito (1)	14110	3.2	0.98	34.96	0.85	0.94
Manuscrito (2)	7368	3.8	0.98	39.08	1.07	0.96
séc. XVI (1)	5743	3.1	0.98	35.4	1.03	0.96
séc. XVI (2)	7645	4.2	0.97	36.4	0.74	0.90
séc. XVI (3)	3471	4.76	0.98	34.5	0.86	0.94
séc. XVI (4)	2194	6.89	0.97	34.8	0.49	0.94
séc. XVII (1)	2776	4.9	0.97	33.55	0.44	0.93
séc. XVII (2)	18312	3.3	0.97	40.12	0.78	0.98
séc. XVII (3)	2919	6.87	0.97	35.6	0.52	0.95
séc. XVII (4)	2976	5.4	0.98	33.0	0.84	0.93
séc. XVIII (1)	6759	2.6	0.98	37.0	0.45	0.96
séc. XVIII (2)	5512	3.1	0.98	39.95	0.2	0.97
séc. XVIII (3)	5414	4.91	0.98	35.23	0.90	0.95
séc. XVIII (4)	11683	2.99	0.98	41.16	0.48	0.96
séc. XIX (1)	4375	2.7	0.98	40.08	0.70	0.95
séc. XIX (2)	4325	6.4	0.97	32.94	0.64	0.84
séc. XIX (3)	5252	4.76	0.98	34.70	0.81	0.87
séc. XIX (4)	4685	3.55	0.98	37.96	0.56	0.95

5 CONSIDERAÇÕES FINAIS

Este trabalho inicialmente analisou os impactos da compressão de imagens de documentos antigos sobre a fidelidade visual das imagens resultantes para diferentes taxas de compressão, através do uso de métricas objetivas de qualidade visual.

Os resultados obtidos apontam correlações das características espaciais e espectrais das imagens de documento em relação à fidelidade visual das imagens de documentos, quando submetidas às variações da taxa de compressão e de diferentes métodos de compressão com perdas. Especificamente as métricas objetivas PSNR, LMSE e SSIM apontaram correlações com os valores de IAM e SAM, identificando uma relação entre o tipo de imagem e o impacto da compressão sobre a mesma.

Na comparação entre as técnicas de compressão com perdas, para taxas de compressão abaixo de 90% ($CR < 0.90$), os resultados para *PSNR*, *LMSE* e *SSIM* entre as imagens comprimidas com as técnicas baseadas em DCT e em DWT apresentam valores similares, ou seja, não há diferenças visuais entre uma imagem comprimida no formato JPEG e uma no formato JPG 2000. Ao aumentar a taxa de compressão ($CR > 0.95$), entretanto, é possível perceber uma diferença perceptível da fidelidade visual entre esses dois formatos de imagem. Os valores de *PSNR* e *SSIM* tendem a cair com menor intensidade no formato JPG 2000 do que no formato JPEG, indicando que há maior preservação de fidelidade visual

no primeiro formato, para altas taxas de compressão.

Para a avaliação da fidelidade visual entre as imagens comprimidas e a respectiva imagem de referência, optou-se pelo emprego de métricas objetivas. Estas apresentam a vantagem de poderem ser validadas em outros ambientes de experimentação, além de permitirem a análise automatizada das respostas obtidas. As métricas subjetivas de avaliação da fidelidade visual ainda apresentam resultados superiores quanto à correlação com o sistema visual humano (HVS), porém as mesmas, demandam mais tempo por exigirem a presença e avaliação de especialistas na temática de processamento de imagens. Desta forma, métricas subjetivas de avaliação muitas vezes se mostram inviáveis para aplicação em um conjunto diversificado de amostras, como é o caso de uma biblioteca digital.

Ainda no campo de avaliação da fidelidade visual, as métricas existentes podem ser classificadas em três categorias: referência completa (Full Reference – FR), referência reduzida (Reduced Reference – RR) e sem referência (No Reference – NR). Neste trabalho foram adotadas apenas métricas de referência completa, uma vez que todo o processo de compressão é realizado sobre as imagens originais digitalizadas (master), que por sua vez também são armazenadas para efeitos de preservação digital. Portanto, considerou-se que as imagens de documentos comprimidas sempre poderiam ser comparadas à respectiva imagem de referência original.

Foram utilizadas as seguintes métricas objetivas para avaliação da fidelidade visual das imagens comprimidas em relação à imagem de referência:

- *PSNR* – Peak Signal-to-Noise Rate;
- *AD* – Average Difference;
- *LMSE* – Laplacian Mean Square Error;

- *NK* – Normalized Cross-Correlation;
- *MD* – Maximum Difference;
- *NAE* – Normalized Absolute Error;
- *SC* – Structural Content;
- *SSIM* – Structural Similarity Index;

Essas métricas foram analisadas frente à taxa de compressão (CR – Compression Rate) da imagem comprimida. Verificou-se que, ao variar a taxa de compressão, as métricas *PSNR*, *LMSE* e *SSIM* se correlacionam com a fidelidade visual obtida, apresentando uma variação proporcional. Com isso, o restante dos experimentos quanto à análise da compressão se basearam nessas três métricas, deixando as demais de lado para alcançar os resultados do trabalho.

Paralelamente aos testes de compressão, as imagens foram analisadas na extração de características que correlacionem com os resultados da compressão. Foram coletadas as seguintes características das imagens selecionadas:

- *m* – Média;
- σ – Variância;
- η – Coeficiente de Variância;
- *SAM* – Spectral Activity Measure;
- *IAM* – Image Activity Measure;

Dentre essas características, a variação nos valores de IAM e SAM demonstraram correlacionar com os resultados da compressão, além de auxiliarem a compreender o comportamento apresentado pelas imagens de documentos, frente

a outras categorias de imagem. As métricas de m (média), σ (variância) e η (coeficiente de variância) não apresentam correlações diretas com a compressão, por apresentarem caracterizações globais da dispersão de valores de brilho das imagens. Entretanto, essas métricas são utilizadas no cálculo da métrica de compressão *SSIM*, utilizando tanto os valores obtidos na imagem de referência quanto na imagem comprimida, portanto são características importantes a serem medidas para se obter resultados consistentes.

Ao comparar imagens de documento com imagens naturais, verificou-se que as imagens de documentos apresentam valores maiores de IAM e valores baixos de SAM. Isso se deve à alta redundância na distribuição espacial dos pixels e menor entropia espectral. Em geral, imagens de documentos são compostas de grandes áreas monocromáticas, representando o fundo da página, com contornos constantes – os caracteres – apresentando variações bruscas de valores de brilho, porém com um comportamento previsível e estável. Já imagens naturais são compostas por diversos valores distintos de brilho, dispostos de forma imprevisível e variável, de forma a representar a riqueza de detalhes presentes nesse quadro.

Por fim, traçou-se as correlações entre as características das imagens de referência com os impactos da compressão com perdas dessas imagens. Observou-se que, dada uma taxa de compressão fixa para as técnicas baseadas em DCT e DWT, valores mais altos de IAM e valores baixos de SAM na respectiva imagem de referência resultam em menor fidelidade visual nas imagens comprimidas, segundo os resultados de *PSNR*, *LMSE* e *SSIM*.

Portanto, para compressão de imagens de documentos em que há a necessidade ou exigência de altas taxas de compressão, sugere-se analisar inicialmente os valores de *IAM* e *SAM* da imagem a ser comprimida. Para valores de *IAM* < 5.0 e valores de *SAM* > 1000 , pode-se empregar técnicas baseadas em DWT para a compressão, com taxa de compressão de até 95%, de forma a obter uma imagem

de tamanho reduzido e visualmente fiel em relação à imagem de referência.

5.1 Trabalhos futuros

Este trabalho abordou questões sobre a correlação das características das imagens com os impactos visuais ao variar a taxa de compressão sobre essas imagens. Foram utilizadas duas técnicas de compressão com perdas baseadas em transformadas, uma utilizando transformadas discretas de cosseno (DCT) e a outra com transformadas discretas de wavelets (DWT). Um desdobramento da pesquisa é replicar a abordagem deste trabalho, utilizando outra técnica de compressão com perdas bastante pesquisada [Xu, Chen e Dai 2008], [Selim et al. 2008], [Jeng, Tseng e Hsieh 2009], a compressão por fractais. A complexidade dos algoritmos empregados na compressão dos fractais, entretanto, ainda apresenta-se como um desafio adicional, principalmente na aplicação em um largo conjunto de imagens, como se apresentam as imagens de documentos em uma biblioteca digital.

Na área de métricas de qualidade da imagens, abordou-se apenas métricas objetivas de referência completa, ou seja, comparou-se as imagens comprimidas com a respectiva imagem de referência. Para complementar os resultados obtidos, pode-se aplicar métricas objetivas sem referência, analisando apenas os artefatos introduzidos durante a compressão. Esta abordagem é interessante em uma análise da qualidade visual de imagens de documentos em um biblioteca digital quando torna-se custoso, tanto em termos de tempo quanto de recursos, a comparação com a imagem de referência, que está armazenada para fins de preservação digital.

Por fim, há a proposta de empregar a tecnologia de reconhecimento óptico de caracteres (OCR), para fazer a avaliação subjetiva da fidelidade visual das ima-

gens comprimidas. Comparando-se a taxa de acertos do OCR sobre uma imagem de documento de referência e sobre uma imagem comprimida, o fato da última apresentar resultados inferiores pode ser um indicador do valor limite de compressão que pode ser aplicado sobre a imagem, pois a legibilidade e compreensão do conteúdo, e consequentemente a qualidade visual, foram afetadas durante a compressão.

REFERÊNCIAS

- [Acharya 2005]ACHARYA, T. JPEG 2000 standard for image compression: concepts, algorithms and VLSI. [S.l.]: John Wiley & Sons, 2005.
- [Agosti, Ferro e Orio 2005]AGOSTI, M.; FERRO, N.; ORIO, N. Annotating illuminated manuscripts: an effective tool for research and education. In: JCDL '05: Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries. New York, NY, USA: ACM, 2005. p. 121–130.
- [Alencar BBD]ALENCAR, J. de. As azas de um anjo. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/00175600>>. Acesso em 20 de julho de 2011, BBD.
- [Alvarez BBD]ALVAREZ, F. Historia de las cosas de ethiopia. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/00463100>>. Acesso em 20 de julho de 2011, BBD.
- [Aung, Ng e Poh 2009]AUNG, A.; NG, B.; POH, S. C. T. A new transform for document image compression. In: Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on, p. 1 –5, Dec. 2009.
- [Bananere BBD]BANANERE, J. La divina increnca. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01072400>>. Acesso em 20 de julho de 2011, BBD.
- [Banesio BBD]BANESIO, F. N. De saluberrima potione chahve. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/02245500>>. Acesso em 20 de julho de 2011, BBD.
- [BBD 2010]BBD. Biblioteca brasileira digital. Disponível em: <<http://www.brasiliana.usp.br>>. Acesso em: 20 de agosto de 2010, 2010.
- [BDTD 2010]BDTD. Biblioteca digital brasileira de teses e dissertações. Disponível em: <<http://bdtb.ibict.br/>>. Acesso em: 10 de maio de 2010, 2010.
- [Bloom 1996]BLOOM, C. Lzp: a new data compression algorithm. In: Data Compression Conference, IEEE Computer Society, Los Alamitos, CA, USA, p. 425, 1996.
- [Bourgeois et al. 2004]BOURGEOIS, F. L. et al. Document images analysis solutions for digital libraries. Document Image Analysis for Libraries, 2004. Proceedings. First International Workshop on, p. 2 – 24, 2004.

- [Bradley e Brislawn 1994]BRADLEY, J.; BRISLAWN, C. The wavelet/scalar quantization compression standard for digital fingerprint images. In: Circuits and Systems, 1994. ISCAS '94., 1994 IEEE International Symposium on, v. 3, p. 205 –208 vol.3, May 1994.
- [Brislawn 2002]BRISLAWN, C. M. The fbi fingerprint image compression standard. Disponível em: <http://www.c3.lanl.gov/brislawn/FBI/FBI.html>. Acesso em 05 de abril de 2010., 2002. Disponível em: <http://www.c3.lanl.gov/brislawn/FBI/FBI.html>.
- [Castro BBD]CASTRO, J. M. F. de. Almanaque do portugal luzo-brazileiro para 1818. Disponível em: <http://www.brasiliana.usp.br/bbd/handle/00407100>. Acesso em 20 de julho de 2011, BBD.
- [Catecismo da diocese BBD]CATECISMO da diocese. Disponível em: <http://www.brasiliana.usp.br/bbd/handle/02903600>. Acesso em 20 de julho de 2011, BBD.
- [Chandler e Hemami 2005]CHANDLER, D.; HEMAMI, S. Dynamic contrast-based quantization for lossy wavelet image compression. Image Processing, IEEE Transactions on, v. 14, n. 4, p. 397 –410, April 2005.
- [Conci, E. e R. 2008]CONCI, A.; E., A.; R., L. F. Computação Gráfica: Teoria e Prática. [S.l.]: Elsevier, 2008.
- [Cui et al. 2008]CUI, Z. et al. What statistics is psnr related using jpeg2000. In: Cybernetics and Intelligent Systems, 2008 IEEE Conference on, p. 984 –989, September 2008.
- [Cunha BBD]CUNHA, R. da. Relação da entrada que fez o exmo. bispo - isidoro da fonseca. Disponível em: <http://www.brasiliana.usp.br/bbd/handle/03908100>. Acesso em 20 de julho de 2011, BBD.
- [Ding et al. 2004]DING, X. et al. Document digitization technology and its application for digital library in china. Document Image Analysis for Libraries, 2004. In: Proceedings. First International Workshop on, p. 46 – 53, 2004.
- [DLF 1998]DLF. A working definition of digital library. Disponível em: <http://www.diglib.org/about/dldefinition.htm>. Acesso em 04 abril 2010., 1998. Disponível em: <http://www.diglib.org/about/dldefinition.htm>.
- [Drira 2006]DRIRA, F. Towards restoring historic documents degraded over time. Document Image Analysis for Libraries, 2006. DIAL '06. Second International Conference on, p. 8 pp. –357, April 2006.
- [Dung et al. 1998]DUNG, L. P. et al. A measure for image quality. In: SAC '98: Proceedings of the 1998 ACM symposium on Applied Computing. New York, NY, USA: ACM, 1998. p. 513–519.

- [Duraó BBD]DURAO, J. de S. R. Caramurú : poema epico do descobrimento da bahia. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/00880400>>. Acesso em 20 de julho de 2011, BBD.
- [Eca BBD]ECA, M. A. R. da Silva de. Reflexões sobre a vaidade dos homens ou discursos moraes sobre os efeitos da vaidade. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/00456800>>. Acesso em 20 de julho de 2011, BBD.
- [Esakkirajan et al. 2006]ESAKKIRAJAN, S. et al. Fingerprint compression using contourlet transform and multistage vector quantization. International Journal of Biological and Life Sciences, p. 140–147, 2006. Disponível em: <<http://www.waset.org/journals/ijbls/v1/v1-2-19.pdf>>.
- [Eskicioglu e Fisher 1995]ESKICIOGLU, A.; FISHER, P. Image quality measures and their performance. Communications, IEEE Transactions on, v. 43, n. 12, p. 2959–2965, dec 1995. ISSN 0090-6778.
- [Ferreira 2002]FERREIRA, A. J. Aplicação da Análise em Componentes Independentes na Compressão de Imagem. Dissertação (Mestrado) — Centro de Cálculo do Instituto Superior de Engenharia de Lisboa, 2002.
- [Ficher BBD]FICHER, J. Indice alphabetico do dictionario bibliographico brasileiro de s. blake. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/00697500>>. Acesso em 20 de julho de 2011, BBD.
- [Figuerola-Villanueva, Ratha e Bolle 2003]FIGUEROA-VILLANUEVA, M. A.; RATHA, N. K.; BOLLE, R. M. A comparative performance analysis of JPEG 2000 vs. WSQ for fingerprint image compression. Springer Berlin / Heidelberg, Springer Berlin / Heidelberg, v. 2688, p. 1–8, 2003. Disponível em: <<http://www.springerlink.com/content/bqlh4w53dxpnp5vr/>>.
- [Freyre BBD]FREYRE, B. Nova lusitania , historia da guerra brasilica. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/00727010>>. Acesso em 20 de julho de 2011, BBD.
- [Gallica 2010]GALLICA. Bibliothèque numérique gallica. Disponível em: <<http://gallica.bnf.fr/>>. Acesso em: 24 de agosto de 2010, 2010.
- [Garcia BBD]GARCIA, G. Origen de los indios de el nuevo mundo, e indias occidentales. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/00897000>>. Acesso em 20 de julho de 2011, BBD.
- [Gardner 2006]GARDNER, P. Digital libraries come of age. Computing in Science & Engineering, v. 8, n. 5, p. 6–10, Sept.-Oct. 2006.
- [Gazan 2005]GAZAN, R. Digital library evaluation: A longer view. DLib Magazine, p. 1–2, 2005.

- [Gonzales e Woods 2000]GONZALES, R. C.; WOODS, R. E. Processamento de Imagens Digitais. [S.l.]: Editora Edgard Blucher, 2000.
- [Google 2010]GOOGLE. Google books. Disponível em: <<http://books.google.com.br/>>. Acesso em: 24 de agosto de 2010, 2010.
- [Grgic, Grgic e Mrak 2004]GRGIC, S.; GRGIC, M.; MRAK, M. Reliability of objective picture quality measures. Journal of Electrical Engineering, v. 55, p. 3–10, 2004. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.105.3785>>.
- [Gutenberg 2010]GUTENBERG. Free ebooks by project gutenberg. Disponível em: <<http://www.gutenberg.org/>>. Acesso em 10 de maio de 2010., 2010.
- [Haddad et al. 2010]HADDAD, Z. et al. Image quality assessment based on wave atoms transform. In: Image Processing (ICIP), 2010 17th IEEE International Conference on. [S.l.: s.n.], 2010. p. 305 –308.
- [Heminger e Robertson 1998]HEMINGER, A.; ROBERTSON, S. Digital rosetta stone: a conceptual model for maintaining long-term access to digital documents. System Sciences, 1998., In: Proceedings of the Thirty-First Hawaii International Conference on, v. 2, p. 158 –167 vol.2, jan 1998.
- [IJG 2010]IJG. Independent jpeg group. Disponível em: <<http://www.iwg.org/>>. Acesso em: 24 de maio de 2010, 2010.
- [IM 2010]IM. Imagemagick. Disponível em: <<http://www.imagemagick.org>>. Acesso em: 14 de fevereiro de 2010, 2010.
- [Jackson e Hannah 1993]JACKSON, D.; HANNAH, S. Comparative analysis of image compression techniques. System Theory, 1993. In: Proceedings SSST '93., Twenty-Fifth Southeastern Symposium on, p. 513 –517, March 1993.
- [JasPer 2010]JASPER. Jasper project. Disponível em: <<http://www.ece.uvic.ca/~mdadams/jasper>>. Acesso em: 2 de fevereiro de 2010, 2010.
- [Jayant e Noll 1984]JAYANT, N.; NOLL, P. Digital Coding of Waveforms, Principles and Applications to Speech and Video. [S.l.]: Prentice Hall, 1984.
- [Jeng, Tseng e Hsieh 2009]JENG, J.-H.; TSENG, C.-C.; HSIEH, J.-G. Study on huber fractal image compression. In: Image Processing, IEEE Transactions on, v. 18, n. 5, p. 995–1003, May 2009.
- [Jianhai e McDonough 2009]JIANHAI, R.; MCDONOUGH, J. Preserving born-digital cultural heritage in virtual world. In: IT in Medicine Education, 2009. ITIME '09. IEEE International Symposium on, v. 1, p. 745 –748, aug. 2009.
- [JPG 2010]JPG. Lib-jpeg. Disponível em: <<http://www.iwg.org/files>>. Acesso em: 2 de novembro de 2011, 2010.

- [Kavallieratou e Stathis 2006]KAVALLIERATOU, E.; STATHIS, S. Adaptive binarization of historical document images. In: Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, v. 3, p. 742 –745, 0-0 2006.
- [Khalifa e Dlay 1998]KHALIFA, O.; DLAY, S. Wavelets image data compression. Industrial Electronics, 1998. In Proceedings. ISIE '98. IEEE International Symposium on, v. 2, p. 407 –410 vol.2, July 1998.
- [Kornblum 2008]KORNBLUM, J. Using jpeg quantization tables to identify imagery processed by software. Digital Investigation, v. 5, p. S21–S25, set. 2008. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1742287608000285>>.
- [Lindenberg BBD]LINDENBERG, V. G. e A. Norte do brasil através do amazonas, do pará e do maranhão. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/00921600>>. Acesso em 20 de julho de 2011, BBD.
- [Lins et al. 2009]LINS, R. et al. Image classification to improve printing quality of mixed-type documents. In: Document Analysis and Recognition, 2009. ICDAR '09. 10th International Conference on, p. 1106 –1110, July 2009.
- [Lu, Zhang e Tan 2004]LU, Y.; ZHANG, L.; TAN, C. Retrieving imaged documents in digital libraries based on word image coding. Document Image Analysis for Libraries, 2004. In: Proceedings. First International Workshop on, p. 174 –187, 2004.
- [Maio et al. 2002]MAIO, D. et al. Fvc2000: fingerprint verification competition. Pattern Analysis and Machine Intelligence, IEEE Transactions on, v. 24, n. 3, p. 402 –412, March 2002.
- [Marcellin et al. 2000]MARCELLIN, M. et al. An overview of JPEG-2000. Data Compression Conference, Citeseer, p. 523–544, 2000. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.7046>>.
- [Marinai, Marino e Soda 2006]MARINAI, S.; MARINO, E.; SODA, G. Font adaptive word indexing of modern printed documents. In: Pattern Analysis and Machine Intelligence, IEEE Transactions on, v. 28, n. 8, p. 1187 –1199, August 2006.
- [MATLAB 2010]MATLAB. The mathworks: Matlab and simulink for technical computing. Disponível em: <<http://www.mathworks.com/>>. Acesso em: 19 de janeiro de 2010, 2010.
- [MIS 2010]MIS. Medical image samples. Disponível em: <<http://barre.nom.fr/medical/samples>>. Acesso em: 17 de agosto de 2010, 2010.
- [Monteiro 1996]MONTEIRO, H. Avaliação da compressão com perdas em imagens cardíacas de medicina nuclear. Dissertação (Mestrado) — Universidade Estadual de Campinas, 1996.

- [Montoya BBD]MONTTOYA, A. R. de. Conquista espiritual. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01184200>>. Acesso em 20 de julho de 2011, BBD.
- [Morais BBD]MORAIS, A. J. de M. Physiologia das paixões e afecções, precedida de uma noção philosophica geral e por um estudo aprofundado e descrições anatômicas do homem e da mulher. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01193110>>. Acesso em 20 de julho de 2011, BBD.
- [Moreau BBD]MOREAU, P. Klare en waarachtige beschryving van de leste. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01138100>>. Acesso em 20 de julho de 2011, BBD.
- [Mrak, Grgic e Grgic 2003]MRAK, M.; GRGIC, S.; GRGIC, M. Picture quality measures in image compression systems. In: EUROCON 2003. Computer as a Tool. The IEEE Region 8, v. 1, p. 233–236 vol.1, September 2003.
- [NDLTD 2010]NDLTD. Networked digital library of theses and dissertations. Disponível em: <<http://www.ndltd.org/>>. Acesso em: 13 de maio de 2010, 2010.
- [Nirmala, Dandapat e Bora 2007]NIRMALA, S. R.; DANDAPAT, S.; BORA, P. K. Image quality assessment in retinal image compression systems. In: Information and Communication Technology in Electrical Sciences (ICTES 2007), 2007. ICTES. IET-UK International Conference on, p. 737–742, December 2007.
- [Oddi BBD]ODDI, L. D. Della vita del ven. servo di dio p. giuseppe anchieta. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01354400>>. Acesso em 20 de julho de 2011, BBD.
- [Phillips 2000]PHILLIPS, D. Image processing in C. [S.l.]: R & D Publications, 2000.
- [Rani, Bansal e Bansal 2009]RANI, B.; BANSAL, R.; BANSAL, S. Comparison of jpeg and spiht image compression algorithms using objective quality measures. In: Multimedia, Signal Processing and Communication Technologies, 2009. IMPACT '09. International, p. 90–93, March 2009.
- [Reavy e Boncelet 2001]REAVY, M.; BONCELET, C. An algorithm for compression of bilevel images. In: Image Processing, IEEE Transactions on, v. 10, n. 5, p. 669–676, May 2001.
- [Rehman e Wang 2010]REHMAN, A.; WANG, Z. Reduced-reference ssim estimation. In: Image Processing (ICIP), 2010 17th IEEE International Conference on. [S.l.: s.n.], 2010. p. 289–292.
- [Relacion de la victoria que los portugueses de Pernambuco BBD]RELACION de la victoria que los portugueses de Pernambuco. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01590400>>. Acesso em 20 de julho de 2011, BBD.

- [Rouen BBD]ROUEN. C est la dedution du sumptueux ordre... Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/00565110>>. Acesso em 20 de julho de 2011, BBD.
- [Rousseau BBD]ROUSSEAU. Cartas botanicas. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/03893000>>. Acesso em 20 de julho de 2011, BBD.
- [Roussin BBD]ROUSSIN, A. R. Le pilote du bresil - 1a ed. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01629900>>. Acesso em 20 de julho de 2011, BBD.
- [Sá BBD]SÁ, I. C. de. Irmandade da virgem sra. do roário dos pretos. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/02441100>>. Acesso em 20 de julho de 2011, BBD.
- [Saha 2000]SAHA, S. Image compression—from dct to wavelets: a review. Crossroads, ACM, New York, NY, USA, v. 6, n. 3, p. 12–21, 2000.
- [Saha e Vemuri 2000]SAHA, S.; VEMURI, R. How do image statistics impact lossy coding performance? In: Information Technology: Coding and Computing, 2000. Proceedings. International Conference on, p. 42 –47, 2000.
- [Selim et al. 2008]SELIM, A. et al. A simplified fractal image compression algorithm. In: Computer Engineering Systems, 2008. ICCES 2008. International Conference on, p. 53 –58, November 2008.
- [Sheikh e Bovik 2006]SHEIKH, H.; BOVIK, A. Image information and visual quality. In: Image Processing, IEEE Transactions on, v. 15, n. 2, p. 430 –444, fev. 2006.
- [Skodras, Christopoulos e Ebrahimi 2001]SKODRAS, A.; CHRISTOPOULOS, C.; EBRAHIMI, T. The jpeg 2000 still image compression standard. IEEE Signal processing Magazine, v. 18, p. 36–58, 2001.
- [Solomon 2004]SOLOMON, D. Data Compression: The Complete Reference. [S.l.]: Springer-Verlag New York, 2004.
- [Staden BBD]STADEN, H. Hans staden: suas viagens e captiveiro entre os selvagens. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/02441100>>. Acesso em 20 de julho de 2011, BBD.
- [Systems 1992]SYSTEMS, A. Tiff specification. Disponível em: <<http://partners.adobe.com/public/developer/tiff/index.html>>. Acesso em 29 de abril de 2010, 1992.
- [The Dutch Surva BBD]THE Dutch Surva. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/02112500>>. Acesso em 20 de julho de 2011, BBD.

- [Thevet BBD]THEVET, A. Historia dell'india americana detta altramente francia antartica. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01918410>>. Acesso em 20 de julho de 2011, BBD.
- [Tong et al. 2004]TONG, H. et al. No-reference quality assessment for jpeg2000 compressed images. In: Image Processing, 2004. ICIP '04. 2004 International Conference on. [S.l.: s.n.], 2004. v. 5, p. 3539 – 3542 Vol. 5.
- [Turaga, Chen e Caviedes 2002]TURAGA, D.; CHEN, Y.; CAVIEDES, J. No reference psnr estimation for compressed pictures. In: Image Processing. 2002. Proceedings. 2002 International Conference on. [S.l.: s.n.], 2002. v. 3, p. III–61 – III–64 vol.3.
- [UNESCO 2010]UNESCO. United nations educational, scientific and cultural organization. Disponível em: <www.unesco.org/>. Acesso em: 2 de fevereiro de 2010, 2010.
- [USP 2010]USP, T. Biblioteca de teses e dissertações da usp. Disponível em: <<http://www.teses.usp.br/>>. Acesso em 10 de maio de 2010, 2010.
- [Vandelli BBD]VANDELLI, D. Diccionario dos termos technicos de historio natural. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/06002700>>. Acesso em 20 de julho de 2011, BBD.
- [Vemuri e Saha 2000]VEMURI, R.; SAHA, S. An analysis on the effect of image features on lossy coding performance. In: Signal Processing Letters, IEEE, v. 7, n. 5, p. 104 –107, May 2000.
- [Vicente BBD]VICENTE, G. Auto da barca do inferno. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01853710>>. Acesso em 20 de julho de 2011, BBD.
- [Vieira BBD]VIEIRA, A. Arte de furtar. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01950700>>. Acesso em 20 de julho de 2011, BBD.
- [Vieira BBD]VIEIRA, A. Sermoens do p. antonio vieira (volume 05). Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/01952550>>. Acesso em 20 de julho de 2011, BBD.
- [W3C 1996]W3C. Press release: W3c issues first recommendation for png. Disponível em: <<http://www.w3.org/Press/PNG-PR.en.html>>. Acesso em: 29 abril 2010, 1996.
- [Wallace 1992]WALLACE, G. The jpeg still picture compression standard. Consumer Electronics, IEEE Transactions on, v. 38, n. 1, p. xviii –xxxiv, February 1992.
- [Wang et al. 2004]WANG, Z. et al. Image quality assessment: from error visibility to structural similarity. In: Image Processing, IEEE Transactions on, v. 13, n. 4, p. 600 –612, april 2004.

- [WDL 2010]WDL. World digital library. Disponível em: <<http://www.wdl.org>>. Acesso em: 15 de maio de 2010, 2010.
- [Welch 1984]WELCH, T. A technique for high-performance data compression. *Computer*, v. 17, n. 6, p. 8 –19, June 1984.
- [WFCAG 2010]WFCAG. The waterloo fractal coding and analysis group. Disponível em: <<http://links.uwaterloo.ca/Repository.html>>. Acesso em: 21 de março de 2010, 2010.
- [Witten et al. 1994]WITTEN, I. H. et al. Textual image compression: Two-stage lossy/lossless encoding of textual images. In: *Proceedings of the IEEE*, p. 878–888, 1994.
- [Witten et al. 2000]WITTEN, I. H. et al. Greenstone: A comprehensive open-source digital library software system. In: *Proceedings of the Fifth ACM International Conference on Digital Libraries*, 2000. Disponível em: <<http://citeseer.ist.psu.edu/witten99greenstone.html>>.
- [Xu, Chen e Dai 2008]XU, X.-Y.; CHEN, P.; DAI, J. Hybrid encoding analysis of fractal image compression method based on wavelet transform. In: *Machine Learning and Cybernetics, 2008 International Conference on*, v. 5, p. 2886 –2889, July 2008.
- [Zhang e Le 2010]ZHANG, J.; LE, T. A new no-reference quality metric for jpeg2000 images. In: *Consumer Electronics, IEEE Transactions on*, v. 56, n. 2, p. 743 –750, maio 2010.
- [Zheng, Li e Doermann 2004]ZHENG, Y.; LI, H.; DOERMANN, D. Machine printed text and handwriting identification in noisy document images. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 26, n. 3, p. 337 –353, March 2004.
- [Ziv e Lempel 1977]ZIV, J.; LEMPEL, A. A universal algorithm for sequential data compression. In: *Information Theory, IEEE Transactions on*, v. 23, n. 3, p. 337 – 343, May 1977.
- [Ziv e Lempel 1978]ZIV, J.; LEMPEL, A. Compression of individual sequences via variable-rate coding. In: *Information Theory, IEEE Transactions on*, v. 24, n. 5, p. 530 – 536, September 1978.

APÊNDICE A – CÓDIGOS-FONTES DOS ALGORITMOS UTILIZADOS

No CD anexo à dissertação, encontram-se os códigos-fontes dos algoritmos utilizados.

Estão dispostos em dois diretórios. O primeiro é chamado "Códigos MatLab" e o segundo "Códigos Shell Script". A seguir, são apresentadas as listagens dos arquivos contidos em cada diretório, junto com uma breve descrição de sua função:

Para a reproduzir os resultados do trabalho, basta executar o programa "executaTudo.sh".

Códigos MatLab

- IAM.m – Cálculo da Medição da Atividade da Imagem (IAM – Image Activity Measure), utilizando uma imagem de referência;
- SAM.m – Cálculo da Medição da Atividade Espectral (SAM – Spectral Activity Measure), utilizando uma imagem de referência;
- Caracteristicas.m – Utilizando uma imagem de referência, realiza os cálculos de:
 - m – Média,
 - σ – Variância;
 - η – Coeficiente de Variação;

- `compressionMetrics.m` – Utilizando uma imagem comprimida e sua respectiva imagem de referência, realiza os cálculos de:
 - Erro Médio Quadrático (MSE – Mean Square Error);
 - Pico da Razão do Sinal-Ruído (PSNR – Peak Signal-to-Noise Rate);
 - Erro Médio Quadrático Laplaciano (LMSE – Laplacian Mean Square Error);
 - Diferença Média (AD – Average Difference);
 - Diferença Máxima (MD – Maximum Difference);
 - Correlação-Cruzada Normalizada (NK – Normalized Cross-Correlation);
 - Erro Absoluto Normalizado (NAE – Normalized Absolute Error);
 - Conteúdo Estrutural (SC – Structural Content);
- `SSIM.m` – Cálculo do Índice de Similaridade Estrutural (SSIM – Structural Similarity), utilizando uma imagem comprimida e sua respectiva imagem de referência. Código-fonte extraído de [Wang et al. 2004];

Códigos Shell Script

- `converteArq.sh` – Gera os arquivos nos formatos JPEG e JPG 2000;
- `DCT.sh` – Executa os cálculos de fidelidade visual (`compressionMetrics.m` e `SSIM.m`) para as imagens em JPEG, em relação à imagem de referência;
- `DWT.sh` – Executa os cálculos de fidelidade visual (`compressionMetrics.m` e `SSIM.m`) para as imagens em JPG 2000, em relação à imagem de referência;
- `geraSAMeIAM.sh` – Realiza os cálculos de SAM e IAM (`SAM.m` e `IAM.m`) para as imagens de referência;

- `executaCaracteristicas.sh` – Realiza os cálculos das características da imagem de referência (`Caracteristicas.m`): Média (m), Variância (σ) e Coeficiente de Variação (η);
- `executaTudo.sh` – Prepara todo o ambiente e executa sequencialmente todos os programas em Shell Script;

APÊNDICE B – IMAGENS UTILIZADAS NO TRABALHO

No CD anexo à dissertação, encontram-se as 76 imagens utilizadas. As imagens estão divididas em três diretórios: imagens de documentos, imagens naturais e imagens sintéticas.

Todas as imagens utilizadas estão no formato BMP, em grayscale, com 8 bits de tonalidade de cinza. As imagens geradas em JPEG e JPG 2000 não foram disponibilizadas no CD por serem geradas automaticamente durante a execução dos códigos-fontes apresentados no apêndice A.

A seguir, são listados as imagens utilizadas, na seguinte organização:

Referência; Nome do Diretório – Arquivos de imagens no formato bmp.

Imagens de documentos

- [Lindenberg BBD]; Lindenberg – 000035.bmp, 000061.bmp, 000083.bmp, 000253.bmp;
- [Garcia BBD]; Garcia – 000061.bmp, 000161.bmp, 000361.bmp;
- [Duraó BBD]; Duraó – 000061.bmp, 000151.bmp, 000161.bmp;
- [Eca BBD]; Eca – 000061.bmp, 000161.bmp, 000361.bmp;
- [Freyre BBD]; Freyre – 000321.bmp, 000341.bmp, 000421.bmp;

- [Rouen BBD]; Rouen – 000043.bmp, 000091.bmp, 000131.bmp;
- [Alvarez BBD]; Alvarez – 000021.bmp, 000051.bmp, 000101.bmp;
- [Castro BBD]; Castro – 000063.bmp, 000123.bmp, 000163.bmp;
- [Bananere BBD]; Bananere – 000017.bmp, 000037.bmp, 000057.bmp;
- [Montoya BBD]; Montoya – 000041.bmp, 000099.bmp, 000199.bmp;
- [Moreau BBD]; Moreau – 000021.bmp, 000029.bmp, 000099.bmp;
- [Thevet BBD]; Thevet – 000021.bmp, 000121.bmp, 000141.bmp;
- [Roussin BBD]; Roussin – 000023.bmp, 000123.bmp, 000223.bmp;
- [Vandelli BBD]; Vandelli – 000061.bmp, 000161.bmp, 000261.bmp;
- [Banesio BBD]; Banesio – 000021.bmp, 000041.bmp, 000051.bmp;
- [Catecismo da diocese BBD]; Catecismo – 000023.bmp, 000123.bmp, 000223.bmp;
- [Rousseau BBD]; Rousseau – 000061.bmp, 000063.bmp, 000123.bmp;
- [Sá BBD]; Sa – Manuscrito.bmp;
- [Morais BBD]; Morais – Philosophia.bmp;

Imagens naturais

- Lena – Lena.bmp;
- Baboon – Baboon.bmp;
- Barbara – Barbara.bmp;
- Peppers – Peppers.bmp;
- Bird – Bird.bmp;

Imagens sintéticas

- Horizontal – Horizontal.bmp;
- Vertical – Vertical.bmp;