



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
FACULDADE DE ENGENHARIA DA COMPUTAÇÃO E
TELECOMUNICAÇÕES

Por definir.

Luan Assis Gonçalves

BELEM - PARÁ

2015



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
FACULDADE DE ENGENHARIA DA COMPUTAÇÃO E
TELECOMUNICAÇÕES

Luan Assis Gonçalves

Por definir.

Trabalho de Conclusão de Curso apresentado para
obtenção do grau de Engenheiro em Engenha-
ria da Computação, do Instituto de Tecnologia,
da Faculdade de Engenharia da Computação e
Telecomunicações.

BELÉM - PARÁ

2015

Por definir.

Este trabalho foi julgado adequado em xx/xx/2015 para a obtenção do Grau de Engenheiro da Computação, aprovado em sua forma final pela banca examinadora que atribui o conceito _____.

Prof. Dr. Ronaldo de Freitas Zampolo
ORIENTADOR

Prof. Dr. Evaldo Gonçalves Pelaes
MEMBRO DA BANCA EXAMINADORA

Prof. Dr. Francisco Carlos Bentes Frey Müller
MEMBRO DA BANCA EXAMINADORA

Prof. Dr. Francisco Carlos Bentes Frey Müller
DIRETOR DA FACULDADE DE ENGENHARIA
DA COMPUTAÇÃO E TELECOMUNICAÇÕES

Aos meus genitores, tutores, redentores e embaixadores

Helena e Antônio.

Agradecimentos

*"Qualquer um que nunca tenha cometido um erro nunca
tentou algo novo."
Albert Einstein*

Resumo

Palavras-chave: Compreensão, Quantização, Codificação perceptual.

Abstract

Keywords: Compression, Quantization, Perceptual Coding.

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

DCT Discrete Cosine Transform

HVS Human Visual System

JPEG Joint Photographic Experts Group

MPEG Moving Picture Experts Group

Sumário

1	Introdução	14
1.1	Motivação	14
1.2	Visão geral do trabalho	14
1.3	Organização do trabalho	15
2	Fundamentos de compressão de imagens e vídeos	16
2.1	Introdução	16
2.2	Redundância	16
2.3	Teoria da Informação	17
2.3.1	Primeiro teorema de Shannon	17
2.3.2	Entropia	18
2.4	Alguns métodos básicos de compressão	18
2.4.1	Métodos	18
2.5	Os padrões JPEG e MPEG	19
3	Qualidade visual	20
3.1	Introdução	20
4	Proposta	21
4.1	Introdução	21
5	Resultados	22
5.1	Introdução	22
6	Considerações finais	23
6.1	Conclusões	23
6.2	Trabalhos futuros	23

Referências Bibliográficas

Capítulo 1

Introdução

1.1 Motivação

Com a expansão de tecnologias como a TV digital, streaming de vídeo e DVD, a compressão de vídeo tornou-se elemento indispensável para as transmissões broadcast e mídias de entretenimento. Neste contexto, pesquisas relacionadas à área foram potencializadas bem como o surgimento de novos padrões de compressão de imagens e vídeos, a fim de otimizar a utilização da banda de transmissão.

Embora o poder de processamento computacional, armazenamento e a largura de banda disponível tenham evoluído, ainda assim, a necessidade de métodos que possam prover taxas consideráveis de compressão e baixas taxas de distorção são um desafio para tais recursos. Esta realidade é evidenciada em trabalhos como [1], [2], [3] e [4].

A situação apresentada anteriormente potencializa-se com a expansão dos dispositivos móveis. Em [5] é feita uma projeção de que nos próximos 5 anos o planeta terá mais dispositivos móveis conectados à internet do que habitantes, gerando um fluxo de cerca de 10 imagens e 1 vídeo por habitante ao dia.

1.2 Visão geral do trabalho

Na literatura há trabalhos que abordam de maneira profunda as bases teóricas, enquanto outros tem como foco principal os aspectos práticos e de implementação. A nível de graduação observa-se que os discentes muitas vezes são expostos a conceitos teóricos de teoria de informação e processamento de sinais e poucas vezes tem contato com suas aplicações.

Portanto, este trabalho tem como objetivo abordar teoria e prática, fundamentando os conceitos de compreensão de imagens e vídeos através de implementações dos padrões *JPEG* e *MPEG*, bem como abordar os aspectos de qualidade visual.

1.3 Organização do trabalho

O trabalho apresentado está distribuído da seguinte forma:

- Capítulo 2: são apresentados conceitos básicos relacionados a compressão de imagens e vídeos e suas áreas a fins, como teoria de informação e processamento de sinais. Por fim, uma visão geral dos padrões *JPEG* e *MPEG* é apresentada;
- Capítulo 3: neste capítulo, o foco é a avaliação da qualidade visual de imagens e vídeos depois dos processo de descompressão;
- Capítulo 4: um método de gerar tabelas de quantização baseada no sistema visual humano é apresentado, a fim de melhorar a qualidade visual dos vídeos depois do processo de compressão.
- Capítulo 5: após a teorização dos tópicos precedentes, compete a este capítulo a análise dos resultados gerados;
- Capítulo 6: por fim, serão apresentadas as conclusões tomadas a partir de vantagens e desvantagens encontradas na proposta apresentada.

Capítulo 2

Fundamentos de compressão de imagens e vídeos

2.1 Introdução

Neste capítulo abordam-se as bases teóricas deste trabalho, apresentando os conceitos básicos do processo de compressão de imagens e vídeos. Primeiramente, o conceito de redundância, presente em arrays $2D$, será abordado, fazendo-se a relação com a teoria da informação. E por fim, é feita a diferenciação dos tipos de codificação (*com e sem perda*).

2.2 Redundância

O processo de compressão de dados consiste em reduzir a quantidade de dados necessária para representar uma dada informação. Neste contexto, os conceitos de dado e informação são diferentes, em que os dados são os meios pelos quais as informações são transmitidas [6].

Seguindo esta linha de raciocínio, uma informação pode ser representada de infinitas maneiras. Desta forma, o questionamento a ser respondido quando se objetiva a compressão de dados é: qual representação forneceria o menor volume de dados sem que haja perda de informação?

A compressão é obtida através da eliminação dos dados relevantes presentes na representação de uma informação. Em se tratando de arrays bidimensionais, os principais tipos de relevância são:

- *Redundância de codificação*: surge quando a quantidade de bits utilizada para representar

as intensidades presentes em um array $2D$ é superior a quantidade necessária.

- *Redundância espacial e temporal*: devido grande parte dos pixels presentes em um array $2D$ estarem espacialmente correlacionados, surge a redundância espacial. Já as sequências de vídeo estão sujeitas a mais outro tipo de redundância, redundância temporal, em que os pixels de quadros vizinhos encontram-se correlacionados, devido a maioria dos quadros possuírem muitas regiões similares.
- *Redundância psicovisual*: é originada a partir das características do HVS. Sua resposta aos estímulos visuais é uma função não linear de grandezas físicas, como intensidade luminosa e cores. Para o sistema visual humano (HVS), as informações não são igualmente percebidas, tornando algumas informações mais importantes do que outras.

A quantificação prévia do volume de dados redundantes presente em uma representação de imagem é necessária, a fim de estabelecer um limite teórico para eliminação de dados. Sendo assim, assumindo que b e b' são, respectivamente, o volume de dados presentes na representação real de uma imagem e o volume de dados presentes em uma representação comprimida da mesma. A *redundância relativa* R é dada pela seguinte equação,

$$R = 1 - \frac{1}{c} \quad (2.1)$$

em que c é a *taxa de compressão*, definida pela função (2.2).

$$c = \frac{b}{b'} \quad (2.2)$$

2.3 Teoria da Informação

2.3.1 Primeiro teorema de Shannon

Segundo a equação 2.3, pode-se provar, através do teorema da codificação sem perda [7], que é possível representar a saída de um sistema sem memória com uma média H de unidades de informação por píxel,

$$\lim_{n \rightarrow \infty} \left[\frac{L_{avg,n}}{n} \right] = H \quad (2.3)$$

em que, $L_{avg,n}$ é o número médio de símbolos código necessários para representar todos os grupos de n símbolos e H é a entropia.

2.3.2 Entropia

A entropia, equação 2.4, é uma métrica utilizada em diversas áreas do conhecimento, como na química e física. Em se tratando de informações, a entropia representa o grau de incerteza de uma fonte.

Sendo (a_1, a_2, \dots, a_J) o conjunto de saídas de uma fonte, a entropia é dada pela informação média de suas saídas.

$$H = - \sum_{j=1}^J P(a_j) \log P(a_j) \quad (2.4)$$

No caso de imagens digitais, em que a unidade de representação é o *bit*, a base dois será utilizada no logaritmo da equação 2.4.

2.4 Alguns métodos básicos de compressão

De maneira geral, existem dois tipos de compressão, com e sem perda, com as seguintes características:

- Sem perda: objetiva comprimir uma determinada informação sem que a mesma seja afetada. Para isso códigos diferentes do código natural são atribuídos aos símbolos.
- Com perda: objetiva alcançar um maior nível de compressão através da eliminação de elementos sem que a informação seja fortemente afetada, de forma que a mesma possa ser entendida.

2.4.1 Métodos

Existem vários métodos de codificação com e sem perda. Dentre eles, o mais utilizados em compressão de imagens são:

1. Huffman: a codificação de Huffman é um método de codificação sem perda que consiste na atribuição de palavras código menores para símbolos mais frequentes e maiores para os símbolos menos frequentes, [8].
2. Codificação aritmética: é um método de codificação sem perda que trabalha de maneira bem diferente em relação a codificação de Huffman. Sua abordagem atribui faixas de valores para cada símbolo e por fim atribui palavras código aos intervalos, [9].

3. Codificação baseada em transformada de blocos: é uma técnica de compressão que consiste em dividir uma imagem em blobs não sobrepostos de tamanhos iguais (8x8). Uma transformada linear reversível, como a transformada de Fourier e a transformada cosseno, é utilizada para mapear estes blocos em um conjunto de coeficientes que por fim serão submetidos a um processo de quantização, [6].

2.5 Os padrões JPEG e MPEG

Informações visuais exercem uma grande influência sob a percepção humana: cerca de 80 – 90% dos neurônios estão relacionados com o processamento de informações visuais [?].

Capítulo 3

Qualidade visual

3.1 Introdução

Capítulo 4

Proposta

4.1 Introdução

Capítulo 5

Resultados

5.1 Introdução

Capítulo 6

Considerações finais

6.1 Conclusões

6.2 Trabalhos futuros

Referências Bibliográficas

- [1] A. Redondi, M. Cesana, and M. Tagliasacchi, “Low bitrate coding schemes for local image descriptors,” 2012, pp. 124–129, cited By 6. [Online]. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84870605750&partnerID=40&md5=36a3c310d83eb80c36005d94ec72ffaa>
- [2] C. Singh, N. Singh, and R. Tripathi, “Optimization of standards for video compression tools over wireless networks,” 2012, pp. 114–118, cited By 0. [Online]. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84862111611&partnerID=40&md5=c04186a826548dc0706b41c12308327f>
- [3] H. Song and C.-C. Jay Kuo, “Rate control for low-bit-rate video via variable-encoding frame rates,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 4, pp. 512–521, 2001, cited By 88. [Online]. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0035307517&partnerID=40&md5=b2d06ea8d0e2cac993c70f68745c1039>
- [4] B. Vizzotto, B. Zatt, M. Shafique, S. Bampi, and J. Henkel, “A model predictive controller for frame-level rate control in multiview video coding,” 2012, pp. 485–490, cited By 1. [Online]. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84868110949&partnerID=40&md5=1ef2b60451ea8a334c78401dca4ddff0>
- [5] “Em 5 anos, o planeta terá mais dispositivos móveis ligados à Internet do que habitantes.” [Online]. Available: <http://startupi.com.br/2013/02/em-5-anos-o-planeta-tera-mais-dispositivos-moveis-ligados-a-internet-do-que-pessoas/>
- [6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.

- [7] C. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.
- [8] D. A. Huffman, “A method for the construction of minimum-redundancy codes,” *Proceedings of the Institute of Radio Engineers*, vol. 40, no. 9, pp. 1098–1101, September 1952.
- [9] V. Bhaskaran and K. Konstantinides, *Image and Video Compression Standards: Algorithms and Architectures*, 2nd ed. Norwell, MA, USA: Kluwer Academic Publishers, 1997.