

PADRÃO JPEG DE COMPACTAÇÃO DE IMAGENS

Oswaldo R. T. Hu*
Luís T. M. Raunheite**

Resumo

Este artigo apresenta um estudo sobre o padrão de compressão de imagem JPEG (Joint Photographics Expert Group), que atualmente é um dos mais utilizados formatos de arquivos de imagem. O padrão JPEG divide-se em dois diferentes métodos de compressão: compressão sem perdas (*lossless*) e compressão com perdas (*lossy*). Em particular, estaremos detalhando a compressão com perdas, que utiliza algoritmo baseado na DCT (Discrete Cosine Transform, Transformada Discreta do Coseno, caso particular da Transformada Discreta de Fourier).¹ A aplicação da compressão com perdas resulta em maiores taxas de compressão em imagens estáticas, sendo também utilizada como parte dos algoritmos de compressão de vídeo MPEG e H.261.²

Abstract

This paper provides a brief study of the JPEG (Joint Photographic Expert Group) compression standard that currently is one of the formats of archives more used in multimedia and Internet for storage and transmission of static pictures. The JPEG compression methods are divided in lossless and lossy compression methods. In particular, we will detail the lossless method, which uses an algorithm based on the DCT (Discrete Cosine Transform, a particular case of Discrete Fourier Transform).¹ This method reaches the highest compression rates in static pictures, and is also used as part of video compression algorithm MPEG and H.261.²

* Fundação Armando Álvares Penteado. E-mail: oshu@yahoo.com.br.

** Universidade Presbiteriana Mackenzie. E-mail: raunheite@mackenzie.com.br.

1 INTRODUÇÃO

Este artigo apresenta uma análise sobre a compressão de imagens realizada pelo método JPEG (Joint Photographic Expert Group). Esta também é a denominação do grupo de especialistas que desenvolveu este método de compressão de imagem e posteriormente, o normalizou na ISO (International Organization for Standardization) como um padrão internacional.

Todos os métodos de compressão visam reduzir as redundâncias presentes na informação, e com isso diminuir de forma significativa o tamanho dos arquivos que contêm essas informações, aumentando a velocidade de transmissão e economizando espaço no armazenamento dos mesmos.³

Essa metodologia foi implementada com dois modos básicos: o método de compressão com perdas, que utiliza um algoritmo baseado na DCT, e o método de compressão sem perdas, que se baseia em métodos preditivos. Existem no total 44 variações do método de compressão JPEG, sendo algumas delas utilizadas em aplicações específicas, tais como aplicações médicas e imagens astronômicas, entre outras. Devido a essa quantidade de variações, dificilmente será encontrado um software que processe todas.

Podemos dividir essas variações do método de compressão JPEG em quatro grandes grupos:

- *Baseline/sequencial encoding*: os componentes da imagem são compactados de forma seqüencial da esquerda para a direita e de cima para baixo.
- *Progressive encoding*: a imagem é compactada em um processo de múltiplas linhas de varredura. Esta forma é geralmente utilizada em arquivos que são transmitidos pela Internet, pois possibilita a visualização da imagem inteira, em menor resolução, enquanto o restante da imagem está sendo enviado.
- *Lossless*: neste método não ocorre perda na qualidade da imagem. Existem duas variações: o original, que foi normalizado em 1992, e o novo método JPEG-LS, que deverá deixar obsoleto o formato JPEG *lossless* original.
- *Hierarchical encoding*: outro formato utilizado na Internet. Nele a imagem é comprimida em múltiplas resoluções (vários tamanhos), sendo a de menor resolução (menor tamanho) codificada e transmitida primeiro e, na seqüência, compactadas e transmitidas as de melhor resolução (as maiores)⁴ em ordem crescente de tamanho; por fim, é enviada a imagem completa em seu tamanho original. Os elementos de imagem das resoluções já recebidas são utilizados na próxima resolução, diminuindo dessa forma o tamanho do arquivo.

Com exceção do tipo *lossless*, todas as variantes apresentam perda de qualidade da imagem, sendo essa perda a função da taxa de compressão.*

* A taxa de compressão é definida como a razão entre o tamanho do arquivo de imagem original, sem compressão, e o tamanho do arquivo de imagem comprimido.

Um grupo de especialistas, denominado Independent JPEG Group, desenvolveu um parâmetro chamado Índice de Qualidade (IQ), que varia de “1” (pior qualidade e maior taxa de compressão) a “100”. Estes índices influenciam as tabelas de quantização, que serão vistas a seguir. Com a alteração destas tabelas de quantização, varia-se a taxa de compressão e a qualidade da imagem.⁵

Normalmente com uma taxa de compressão 20:1 já é possível perceber pequenas perdas de qualidade da imagem. Uma taxa de compressão de 100:1 produz uma imagem de baixa qualidade.

É possível visualizar a relação entre o tamanho do arquivo e a qualidade da imagem em função do IQ na Figura 1 e nos exemplos das Figuras de 2 a 9.

Normalmente o algoritmo para a aplicação do padrão JPEG é disponibilizado como um programa gráfico a ser instalado no computador, mas ele pode vir a ser implementado também via hardware, através de uma placa compactadora e descompactadora. Essa placa instalada no computador contém, além do software

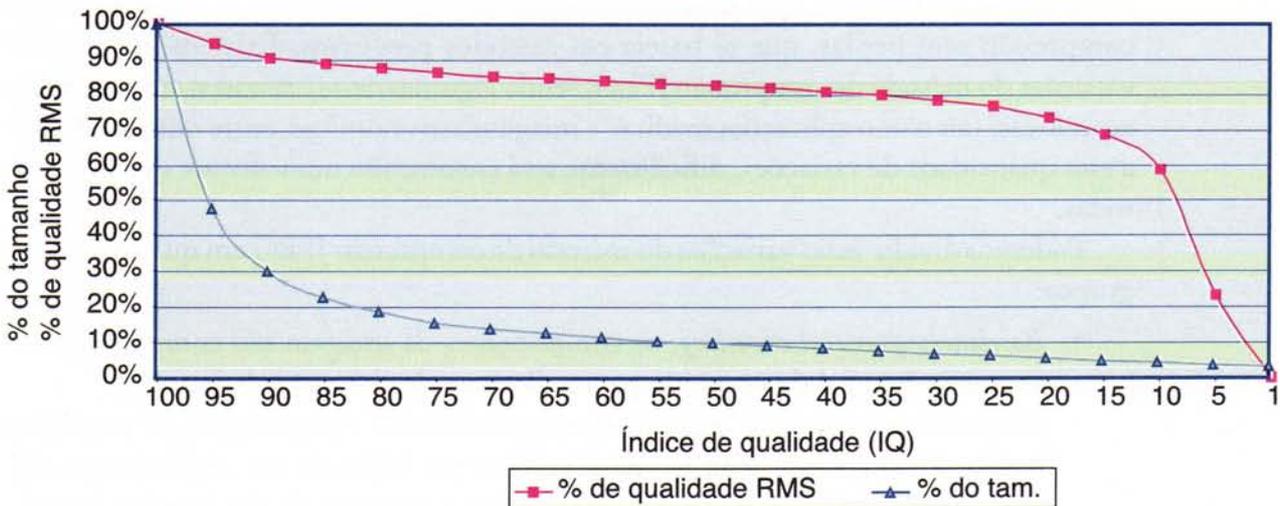


Figura 1 Porcentagem da qualidade RMS e do tamanho do arquivo original x índice de qualidade (IQ)



Figura 2 Arquivo SOL.BMP original. 198.054 bytes



Figura 3 JPEG com IQ = 100; 49.082 bytes; 4,04:1



Figura 4 JPEG com IQ = 75; 9.100 bytes; 21,76:1



Figura 5 JPEG com IQ = 50; 6.040 bytes; 32,79:1



Figura 6 JPEG com IQ = 25; 3.839 bytes; 51,59:1



Figura 7 JPEG com IQ = 10; 2.947 bytes; 67,21:1



Figura 8 JPEG com IQ = 5; 2.293 bytes; 86,37:1

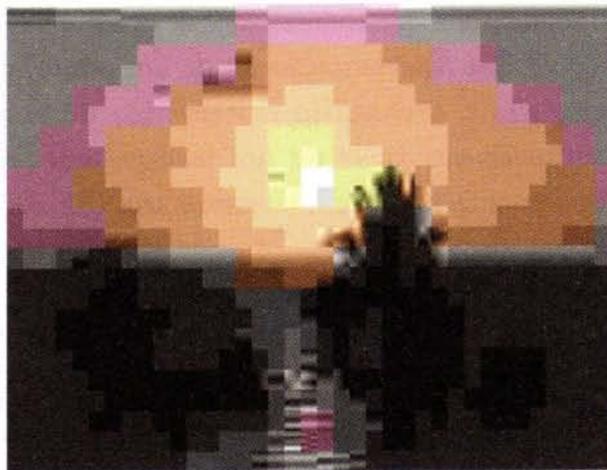


Figura 9 JPEG com IQ = 1; 1.965 bytes; 100,79:1

necessário à compactação, uma CPU (Control Processing Unit) dedicada. Com a utilização dessas placas é possível liberar a CPU do computador para outras tarefas, o que permite uma velocidade global de processamento muitas vezes maior.

Quando a empresa C-Cube Microsystems desenvolveu uma das primeiras placas de compressão JPEG, notou que essa placa conseguia descompactar em 1/30 de segundos, coincidentemente a velocidade necessária para reproduzir vídeos no padrão NTSC (National Television Standards Committee, sistema de TV em cores utilizado nos Estados Unidos e no Japão).⁶ Essa placa compressora possibilitou a realização das primeiras experiências de compressão de vídeo, em que se compactava quadro a quadro, usando padrão JPEG, sem nenhuma preocupação com a redundância temporal. Este foi o ponto de partida para o desenvolvimento de métodos mais elaborados de compressão de vídeo como o MPEG* ou o H.261,** que utilizam o método JPEG como base em sua estrutura de compressão. Esses algoritmos acrescentam novos métodos que eliminam a redundância temporal. Um vídeo comprimido pelo método JPEG é muito maior em tamanho do que um gravado pelo seu predecessor MPEG.

2 AQUISIÇÃO DA IMAGEM

Uma imagem digital pode ser capturada através de scanners, câmeras digitais ou câmeras analógicas ligadas ao computador através de placas digitalizadoras. A câmera digital e o scanner contêm sensores que captam a luz emitida ou refletida dos objetos e a decompõe nos seus componentes fundamentais: vermelho (R), verde (G) e azul (B).

A imagem capturada é bidimensional e pode ser dividida em seus componentes fundamentais, que chamaremos de pixels. Pixel é o menor componente de uma imagem digital. Sua representação é a de um ponto, porém, com um tamanho definido, ou seja, não é o conceito matemático de ponto com tamanho nulo. Esse ponto eventualmente pode ter um formato quadrado e é composto pelos componentes fundamentais: vermelho, verde e azul (RGB).⁷

* MPEG, acrônimo de Moving Picture Expert Group, foi normalizado pela ISO (International Organization for Standardization) em 1991, na versão MPEG-1, para codificações de vídeos. Atualmente, sua versão MPEG-2 é utilizada em DVD e televisão digital.

** O padrão de compressão H.261 foi desenvolvido pela CCITT, também conhecida como ITU-T (International Telecommunication Union), entre 1988 e 1990, e padronizada em 1990. Neste padrão, as imagens são comprimidas a taxas múltiplas de 64 kilobites. É utilizado para transmissão de vídeo pela Internet e para videoconferências.

Estes pixels são enviados para um conversor analógico-digital e quantizados, ou seja, os valores dos componentes fundamentais RGB, que teoricamente poderiam variar de zero ao infinito, são convertidos para poucos valores inteiros, próximos de 256, o que possibilita guardar a representação de cada componente em apenas 8 bits.

Numa segunda etapa, esses componentes RGB são convertidos para componentes de *luminância* (Y) e *crominância* (U e V). A luminância é uma escala de representação numérica do cinza, enquanto a crominância são duas escalas numéricas, que juntas representam as cores.

A escala de luminância, por convenção (ITU-R 601), é quantizada com valores entre “16” e “235”, o valor “16” representando o negro absoluto e o valor “235”, o branco absoluto. As duas escalas de crominância contêm valores que variam entre “16” e “240” (com 128 como valor central). Os valores não utilizados, dos “256” possíveis, servem como códigos de controle.⁸

As matrizes de conversão entre os componentes RGB e os componentes YUV para o sistema PAL (Phase Alternate Line, sistema de TV em cores utilizado na Alemanha, no Brasil e na Argentina) são:

Y	=	0,2990	0,5870	0,1140	R
U	=	-0,1687	-0,3313	0,5000	G
V	=	0,5000	-0,4187	-0,0813	B
R	=	1,00000	0,00000	1,40200	Y
G	=	1,00000	-0,34414	-0,71414	U
B	=	1,00000	1,77200	0,00000	V

Figura 10 Matrizes de conversão YUV x RGB e vice-versa

Estas matrizes de conversão se alteram em função do sistema de aquisição de imagem (NTSC – SECAM, etc.). Para o sistema NTSC, os índices do vetor de luminância (Y) não se alteram, mas os vetores de crominância se alteram e recebem outra denominação (I e Q).

3 PREPARAÇÃO DOS BLOCOS

Após a aquisição e a conversão analógica-digital, as matrizes de crominância são reduzidas à taxa de 2:1, horizontalmente (cada duas linhas convertidas em uma), e

1:1 ou 2:1, verticalmente (cada duas colunas são convertidas em uma). Essas conversões são chamadas de 2h1v sampling e 2h2v sampling. A matriz de luminância não é reduzida, pois o olho humano é mais sensível à luminância (tonalidade de cinza) do que à crominância (tonalidades das cores), o que permite maior taxa de perda de crominância sem que esta perda seja percebida pelo espectador.⁹ Nesta etapa, as informações do quadro são compactadas em duas ou três vezes (a matriz de luminância permanece com o mesmo tamanho e as duas de crominância diminuem para um quarto ou para a metade de seu tamanho original).¹⁰

4 TRANSFORMADA DISCRETA DO CO-SENO

O algoritmo de compressão JPEG divide a matriz de luminância e as duas matrizes de crominância, que são as três matrizes que descrevem a imagem, cada uma com o tamanho de 8×8 elementos. Com isso teremos várias matrizes de 64 elementos, conhecidas como *sample values*. Sobre essas matrizes é aplicado o algoritmo DCT (Discrete Cosine Transform – Transformada Discreta de Coseno), gerando outras matrizes, denominadas coeficientes de DCT, cuja maioria dos elementos tem valor igual a zero.

Esse processo translada a informação do domínio do tempo para o domínio da frequência.¹

A Transformada Discreta de Co-seno converte uma matriz de valores altamente correlacionados, com uma distribuição de probabilidade uniforme, em um conjunto de valores menos correlacionados, e com uma distribuição de probabilidade não uniforme. Em outras palavras, esse algoritmo converte uma matriz numérica em outra, sendo a maior parte dos elementos dessa nova matriz, especificamente os do canto inferior direito, com valor igual a zero.

O valor médio da matriz, que representa a cor fundamental dos 64 pixels, é chamado de componente DC (Direct Current – Corrente Direta) e está localizado no canto superior esquerdo da matriz de coeficientes de DCT. Os outros coeficientes são denominados coeficientes AC (Alternating Current – Corrente Alternada) e representam os valores das pequenas variações de tonalidade e coloração do bloco.

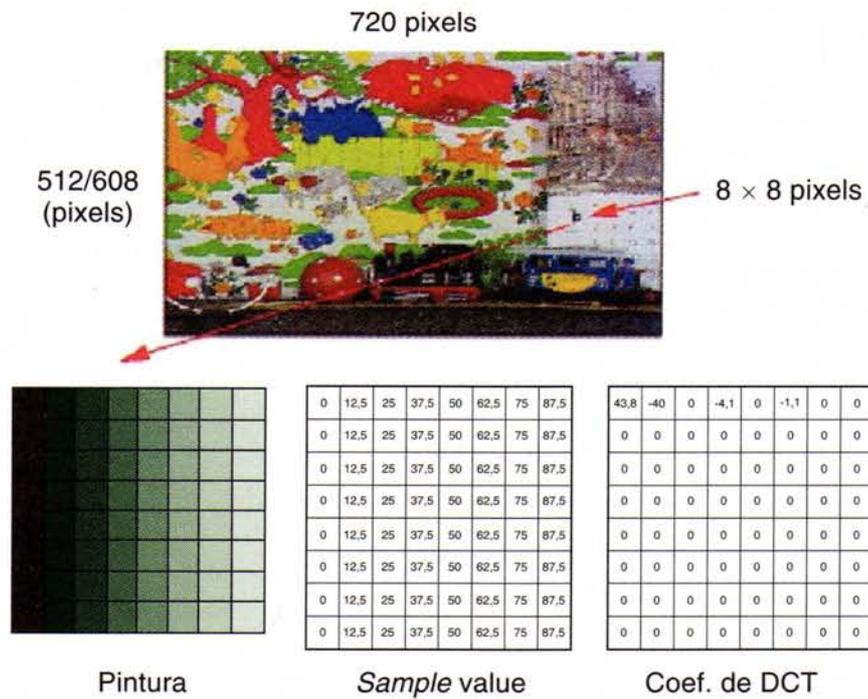


Figura 11 Conversão da pintura para os seus coeficientes e posterior aplicação da DCT¹¹

A DCT, em termos de compactação, gera uma compactação equivalente à da Transformada Discreta de Fourier (FFT). No entanto, no processo inverso, ao se aplicar a antitransformada da DCT consegue-se uma aproximação melhor dos coeficientes originais do que a antitransformada da FFT (ver Figura 12).

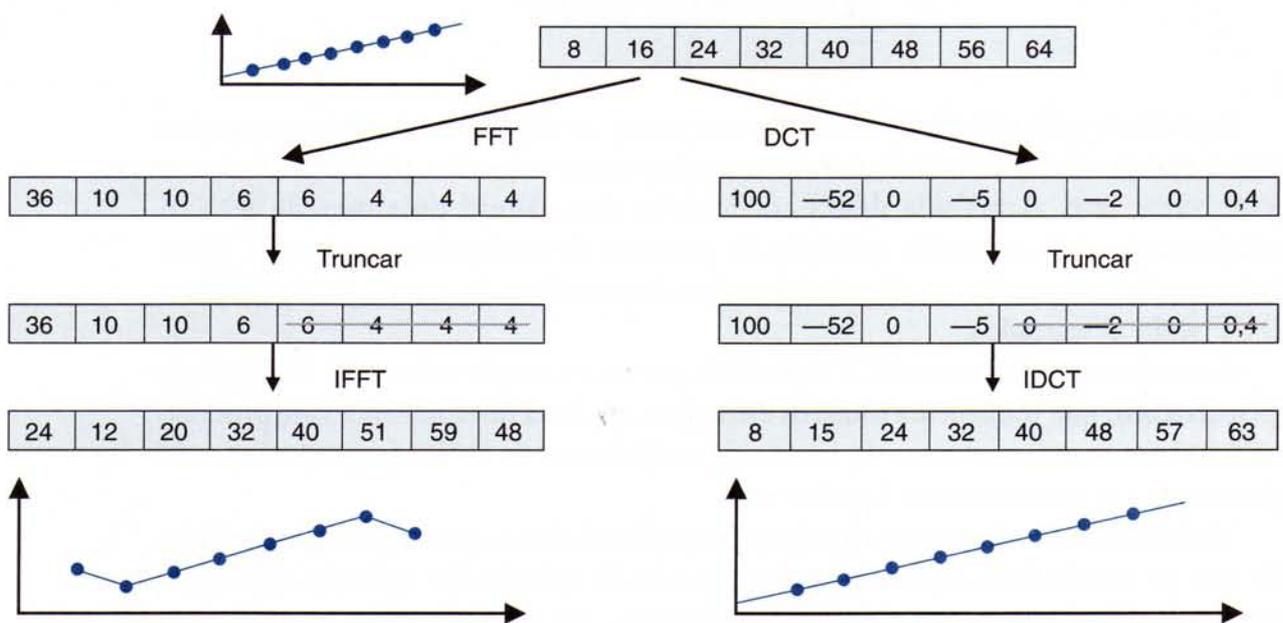


Figura 12 Comparação entre a compactação DCT e a FFT⁹

A Transformada Discreta de Co-seno para matrizes de duas dimensões e tamanho de 8 x 8 pode ser expressa pela seguinte equação:¹

$$F(u,v) = (1/4)C(u)C(v) \sum_{u=0}^7 \sum_{v=0}^7 f(i,j) \cos((2i+1)u/16) \cos((2j+1)v/16)$$

onde:

$F(u, v)$ = coeficiente no domínio da transformada

u = eixo horizontal no domínio da transformada

v = eixo vertical no domínio da transformada

$C(x) = 1/\sqrt{2}$ para $x = 0$
 $C(x) = 1$ para $x \neq 0$ 1

x = u ou v

$f(i, j)$ = amplitude no domínio do tempo

i = eixo horizontal no domínio do tempo

j = eixo vertical no domínio do tempo

5 QUANTIZAÇÃO

Somente a aplicação da DCT resulta em pouca ou quase nenhuma compactação, pois a matriz resultante terá o mesmo tamanho que a original e poucos coeficientes terão valor zero. A maioria destes coeficientes tem valores próximos de zero. A compactação virá a partir da aplicação do processo denominado quantização. Neste processo de compressão, os elementos menos importantes da matriz de coeficientes DCT serão descartados.

A matriz de coeficientes DCT é dividida por outra matriz (chamada de tabela de quantização), que transforma todos os elementos em uma nova matriz.³ Este processo irá acentuar as características da matriz de coeficientes de DCT, forçando todos os elementos a se aproximarem do valor zero.

Todos os valores desta nova matriz serão arredondados, o que implica a introdução de erro na matriz final. Esse erro poderá ser notado quando for aplicado o processo inverso na descompactação e geração da imagem. Na montagem do arquivo serão armazenados somente os valores não zerados.

150	80	40	14	4	2	1	0
92	75	36	10	6	1	0	0
52	38	26	8	7	4	0	0
12	8	6	4	2	1	0	0
4	3	2	0	0	0	0	0
2	2	1	1	0	0	0	0
1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0

1	1	2	4	8	16	32	64
1	1	2	4	8	16	32	64
2	2	2	4	8	16	32	64
4	4	4	4	8	16	32	64
8	8	8	8	8	16	32	64
16	16	16	16	16	16	32	64
32	32	32	32	32	32	32	64
64	64	64	64	64	64	64	64

150	80	20	4	1	0	0	0
92	75	18	3	1	0	0	0
26	19	13	2	1	0	0	0
3	2	2	1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figura 13 Cálculo dos coeficientes DCT quantizados³

Isso possibilita obter uma alta taxa de compressão através da transformação dos coeficientes de DCT em valores com uma precisão não superior à necessária para gerar a imagem na qualidade desejada. Nesta etapa ocorre perda de informação, mas, como a vista humana é mais sensível a baixas frequências do que a altas frequências, e estas perdas são referentes às altas frequências, as perdas serão pouco notadas.

Não existe uma definição na norma JPEG para a tabela de quantização. Cada aplicação pode definir sua própria tabela. A variação na taxa de compressão e conseqüentemente na qualidade da imagem compactada se dá alterando esta tabela. Quanto maiores os coeficientes da tabela de quantização, maior a taxa de compressão e, conseqüentemente, pior a qualidade da imagem gerada.

6 LEITURA EM ZIGUEZAGUE

Esta matriz de valores quantizados é lida em ziguezague, da esquerda para a direita e de cima para baixo, fazendo com que os primeiros elementos lidos sejam valores não zerados. Ignora-se a leitura a partir do momento em que se encontram elementos com valores zerados.¹¹

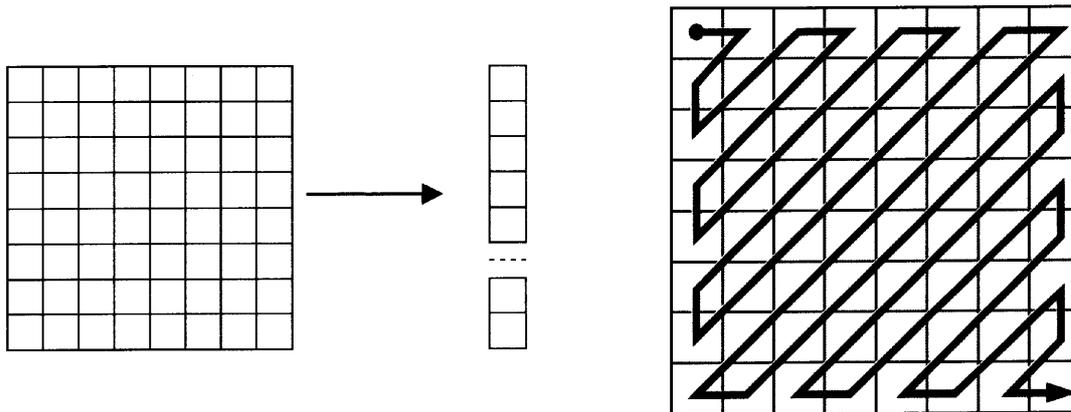


Figura 14 Leitura em ziguezague⁹

No exemplo anterior, esta leitura gerará um vetor de 64 elementos, no qual os 38 últimos serão zeros. Nesse caso, basta substituir esses 38 zeros por um identificador que informe essa quantidade.

7 CODIFICAÇÃO DE ENTROPIA

Após a leitura em ziguezague, os coeficientes serão rearrumados para possibilitar uma nova compressão chamada de codificação de entropia. Essa codificação pode utilizar os algoritmos de Huffman ou a codificação aritmética. Esses algoritmos são do tipo *lossless*, ou seja, não introduzem erros nas informações; logo, essa nova compressão não afetará a qualidade da imagem.³

Essa compressão baseia-se em características estatísticas, ou seja, as informações estatisticamente mais comuns de ocorrerem são codificadas com código de tamanho menor, resultando em menores taxas de transmissão. Esse princípio foi utilizado no telégrafo. No código Morse, as letras mais utilizadas do alfabeto inglês foram codificadas com menores quantidades de símbolos. Para a letra “E” adotou-se “.”, enquanto para a letra “Q” adotou-se “- - . -”.³

8 MONTAGEM DO ARQUIVO

Finalmente são adicionados os *headers e trailers* de controle no arquivo, com todos os parâmetros e tabelas a serem utilizadas na descompressão. Dentre as tabelas que serão incluídas no arquivo estão as de quantização e a de Huffman.

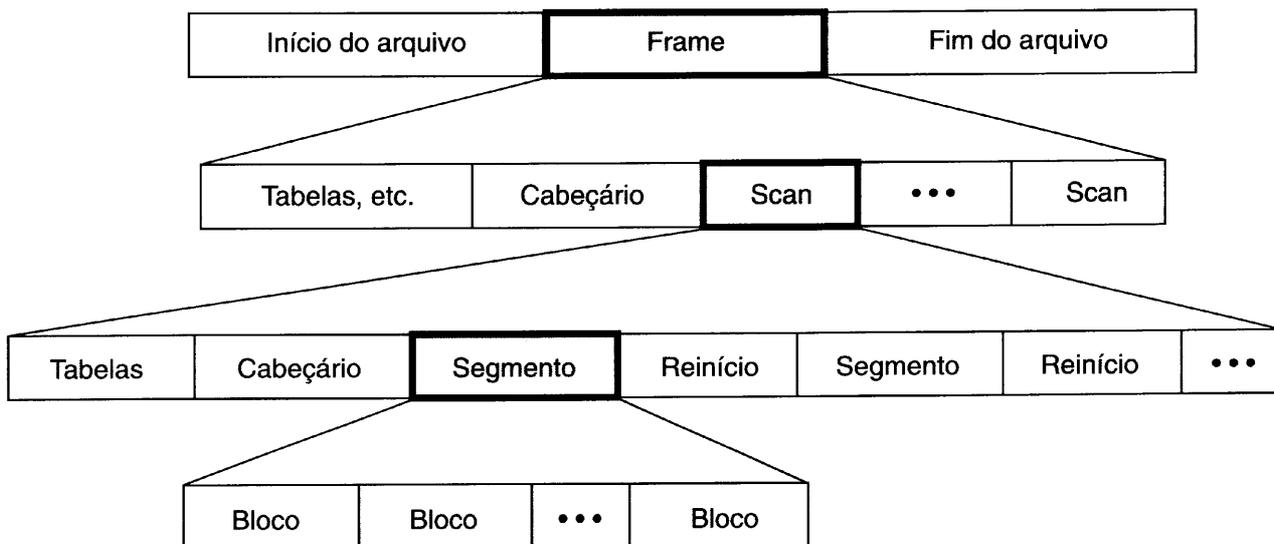


Figura 15 Montagem do arquivo⁹

9 CONCLUSÃO

Esquemáticamente podemos resumir toda a compressão JPEG na Figura 16. Para a descompressão de uma imagem é utilizado o caminho inverso, pois todas as etapas desse método são reversíveis.

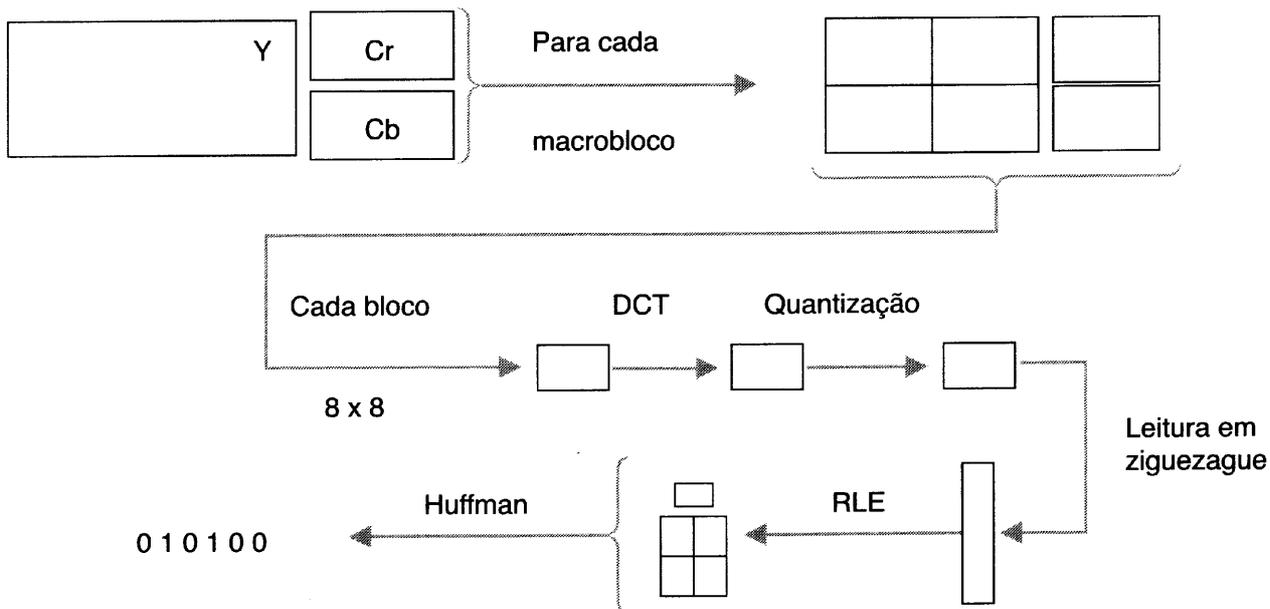


Figura 16 Resumo esquemático⁹

Apesar de o método JPEG ser complicado, tem uma série de qualidades, dentre elas:

- Grande capacidade de compressão.
- Padronizado internacionalmente.
- Pequena perda perceptível. Explora ao máximo as limitações do sistema de visão humana.
- Graças às propriedades do DCT, é possível realizar transformações geométricas (por exemplo, rotação das imagens) diretamente na matriz de transformada.³
- Codificação relativamente simétrica. A codificação demora mais ou menos o mesmo tempo que a decodificação.
- Algoritmos que possibilitam calcular com reduzidas operações aritméticas a DCT e a IDCT (Antitransformada Discreta do Co-seno).⁵
- Ampla aceitação do formato pela comunidade da Internet.

Todas estas qualidades fazem deste método um padrão de fato e de direito.

REFERÊNCIAS BIBLIOGRÁFICAS

1. AHMED, N., Natarajan, T., RAO, K.R. On image processing and a discrete cosine transform. *IEEE Transactions on Computers C-23*, n.1, p.90-93, 1974.
2. HUNG, A.C. *Understanding image compression* [home page]. [citado em 15 jun. 1999] Disponível na Internet: <http://www.igd.fhg.de/icib/telecom/ccitt/rec_H.261-1990/pvrg-descript/chapter2.5.html>.
3. TANENBAUM, A. S. *Redes de computadores*. Rio de Janeiro: Campus, 1977.
4. CONCOLATO, Cyril et al. *What is JPEG* [home page]. [citado em 24 fev. 2000] Disponível na Internet: <<http://www.stud.enst.fr/~gdubois/jpeg/index.htm>>.
5. MIANO, J. *Compressed image file formats – JPEG, PNG, GIF, XBM, BMP*. Reading, Massachusetts: Addison Wesley Longman, 1999.
6. HOLSINGER, E. *Como funciona a multimídia*. São Paulo: Quark do Brasil, 1994.
7. ALBUQUERQUE, M.P. *Processamento de imagens: métodos e análises* [home page]. [citado em 27 mar. 2000]. Disponível na Internet: <<http://mpa.cat.cbpf.br/curso2.htm>>.
8. MELO, Hugo. Compressão digital de vídeo. *Revista de Engenharia de Televisão*, 1999.
9. LI, Ze-Nian, ZHONG, Willian. *CMPT 365 Multimedia systems* [home page]. Simon Fraser University; Canada. [citado em 26 mar. 2000]. Disponível na Internet: <http://www.cs.sfu.ca/CourseCentral/365/li/index_prev.html>.
10. RAUNHEITTE, L.T.M. *Sistemas de comunicação*. São Paulo: Universidade Presbiteriana Mackenzie, 1999, 29p (Apostila de uso do curso).
11. RODRIGUES, A., FIBUSH, D.K., BILOW, S.C. *MPEG-2 Fundamentals for broadcast and post-production engineers* [home page]. Tektronix, 1996. Disponível na Internet: <http://www.tek.com/VND/Support/White_papers/Profile/mpeg2.html>.