

Prof.: Dr. Marcelo Manzato

(mmanzato@icmc.usp.br)

Aula 4 – Percepção, Cor e Imagens.

Instituto de Ciências Matemáticas e de Computação - ICMC Sala 3-160

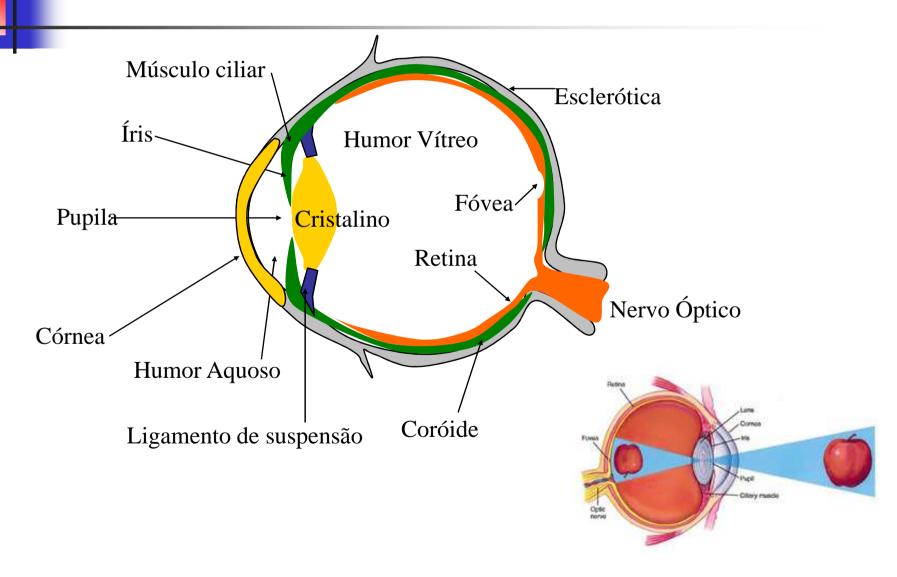
Sumário

- 1 − Introdução
- 2 O Sistema Visual Humano
- 3 Luz e Cor
- 4 Aquisição de Imagens
- 5 − Representação de Imagens
- 6 Compressão de Imagens
- 7 O Padrão JPEG
- 8 JPEG 2000
- 9 Outros Formatos



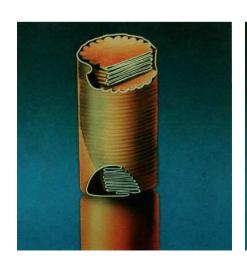
1. Introdução

- 70% das informações que coletamos vêm da visão.
- A visão é o nosso sentido mais importante
 - Relativamente à audição, olfato, tato e paladar.
 - É o mais usado nos sistemas multimídia.
 - É importante estudar o sistema humano de visão.
 - Para usarmos efetivamente a tecnologia multimídia.

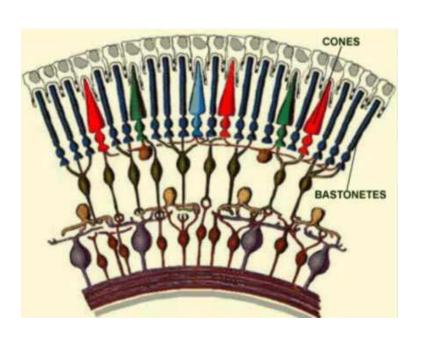




- Cones: fotoreceptores para cor.
 - 6 a 7 milhões.
 - Concentrados na fóvea.
 - Resposta para alta iluminação
- Bastonetes: fotoreceptores para intensidade luminosa.
 - 75 a 150 milhões.
 - Espalhados por toda a retina.
 - Resposta para baixa iluminação









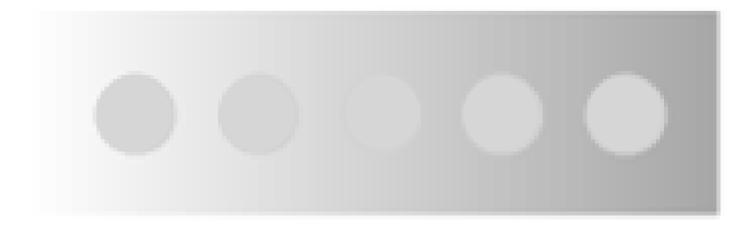
- Adaptação ao brilho (brightness adaptation)
 - Por que é importante?
 - Nível de intensidade de luz:
 - Intervalo de adaptação do olho humano: 10¹º!!!!
 - Não simultâneo!
 - Brilho é subjetivo.
 - Brilho perceptível não é simplesmente uma função de intensidade.
 - Olho humano tende a destacar regiões próximas com diferentes intensidades.



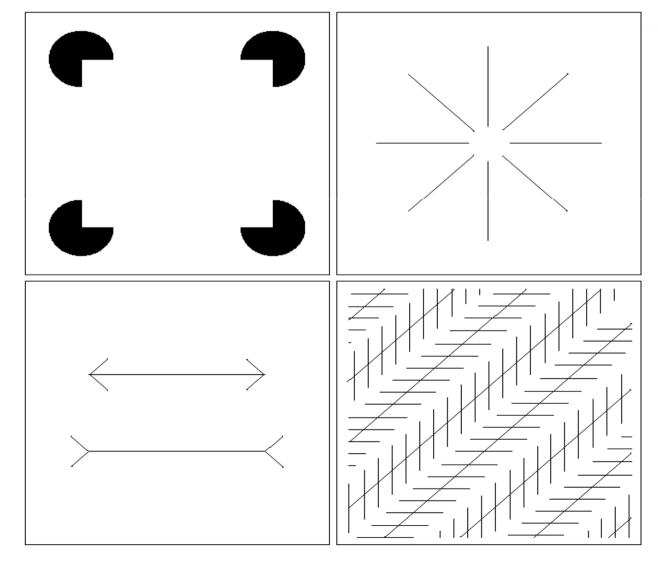
Faixas de Mach:

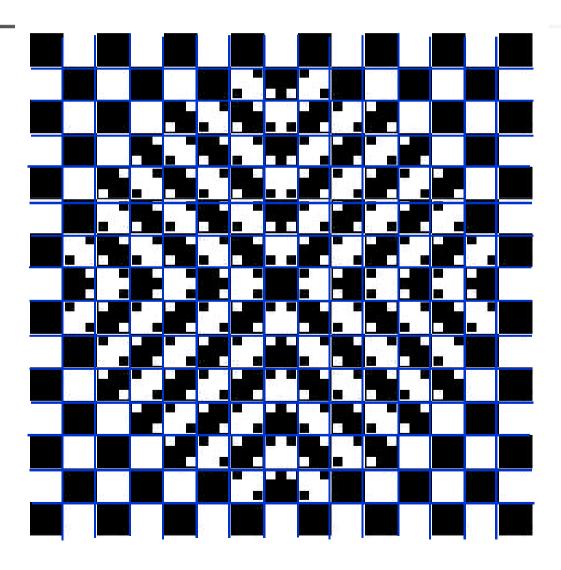


Contraste simultâneo:





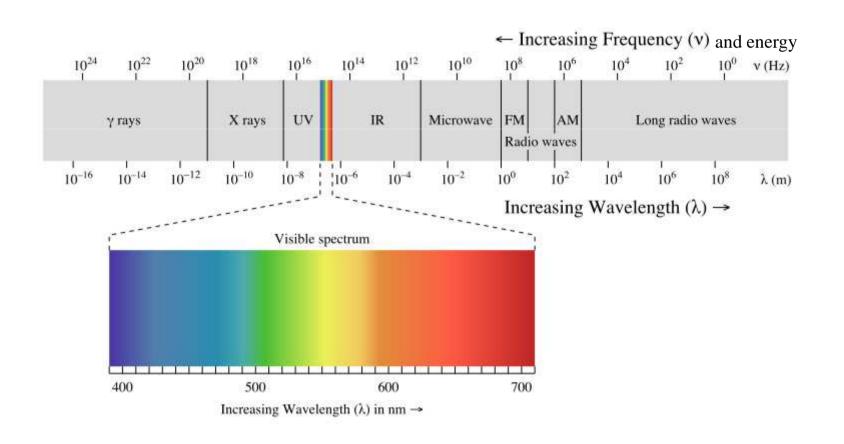






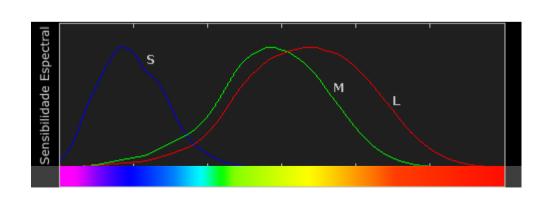
- Cor é importante:
 - Identificação de objetos.
 - Podemos discernir milhares de cores e intensidades comparados a algumas dúzias de cinza.
- Percepção de cores pelo cérebro:
 - É um fenômeno fisiopsicológico.
 - Não é completamente entendido.

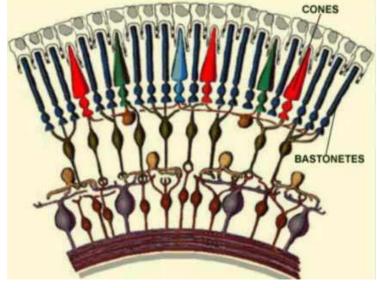
Espectro Eletromagnético



Cones:

- Teoria tristimulus da visão (tristimulus theory of vision)
- Possuem três tipos de fotopigmentos:
 - Azul, Verde e Vermelho.
 - Sensibilidade: 430nm, 530nm e 560nm
 - Espectro visível: ~400nm a 700nm de comprimento de onda.
 - Porcentagem de cones: 4%, 32% e 64%.





3. Luz e Cor

 Além da frequência, três outros aspectos são considerados:

Radiância

 Total de energia que flui da fonte luminosa. Media em watts (W)

Luminosidade

 Quantidade de energia de uma fonte de luz percebida por um observador. Medida em lumens (lm)

Brilho

Noção acromática de intensidade. Subjetivo.

Preto: 0% brilho, 100% cinza

Branco: 100% brilho, 0% cinza

10% cinza = 90% brilho



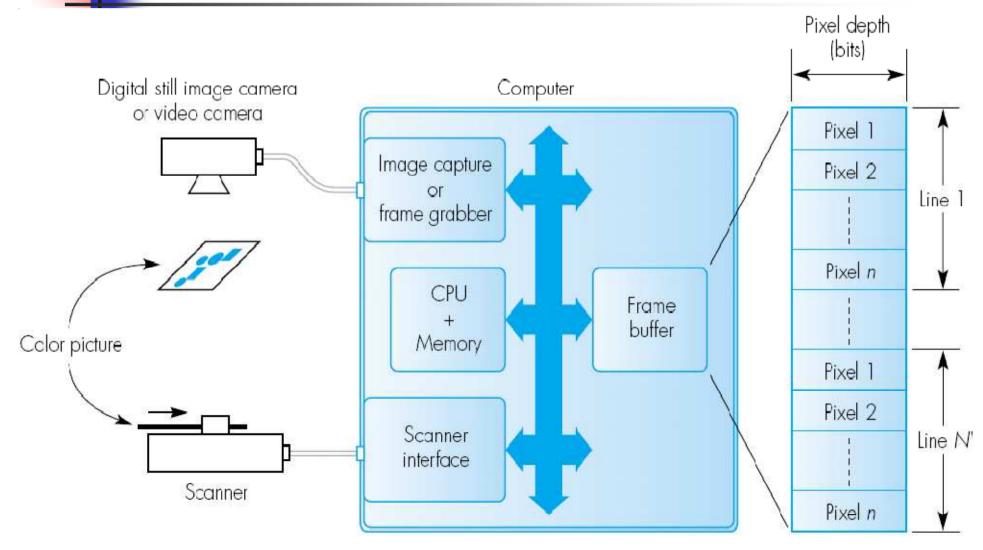
- Como as imagens são capturadas?
 - Tecnologia tenta "imitar" o olho humano.



Olho	Câmera
Cristalino	Lente
Íris	Íris
Cones	CCD
Bastonetes	CCD

CCD = charge-coupled device





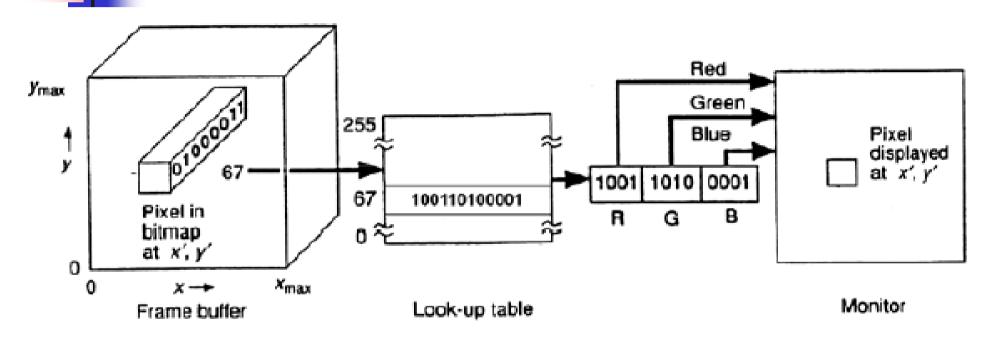


- Imagem é "entendida" como uma matriz de pontos.
 - Pixel ou Pel = Picture element.
- A luz proveniente de cada ponto da imagem é capturada por um sensor (CCD).
 - É composto por uma malha de material fotossensível. Cada célula da malha corresponde a um ponto da imagem.
 - A intensidade da luz incidente em cada célula é convertida em sinal elétrico.
 - Amostragem e quantização.
 - Informação digitalizada é armazenada no Frame Buffer como uma matriz de pontos.



- Imagens coloridas:
 - Utilizam um CCD para cada primária (RGB).
 - Um filtro separa a luz incidente direcionando as componentes para o CCD correto.
 - Cada posição do frame buffer armazena informação dos três componentes.

- Imagem com resolução 640 x 480 com 24 bits por pixel:
 - 921.600 bytes (~1MB)
 - 7.372.800 bits (~7Mbps)
- Frame buffer???



 Look-up table (LUT) ou color look-up table (CLUT)

- Uma imagem na forma digital não apresenta, tecnicamente, uma resolução (dimensão)
 - Exemplo: a que tamanho podemos imprimir uma imagem com 600 x 300 pixels?
- Necessário saber a capacidade do equipamento em reproduzir um número de pixels por unidade de medida
 - Pixels por polegada (ppi) ou Pontos por polegada (dpi)
- Imagem de 600 pixels de largura
 - Impressora de 200dpi: 3" de largura
 - Impressora de 100 dpi: 6" de largura

Outro exemplo

Scanners: 600dpi

Impressoras: 300dpi

Foto 6" x 4" escaneada a 600dpi:

■ Impressora 300dpi: imprime como 12" x 8".

Exercício

A quantos dpi devo realizar a aquisição de uma foto 4" x 3" para exibi-la em um monitor com resolução de 1024 x 768?

Exibição no monitor

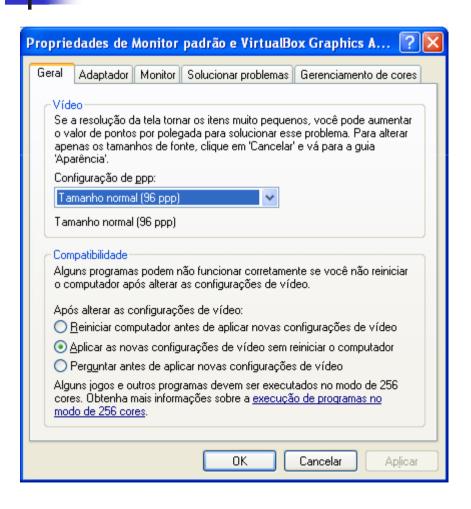


Tela: 640 x 480 pixels Imagem: 400 x 300 pixels



Tela: 800 x 600 pixels

Imagem: 400 x 300 pixels



Além da resolução de tela, pode-se configurar também a quantidade de pixels por polegada (ppi)



5. Representação de Imagens

- Representação de imagens:
 - Na memória do computador:
 - Matriz de pixels armazenada no frame-buffer.

- Em arquivos:
 - Geradas por computador:
 - Gráficos.
 - Digitalizadas:
 - Documentos.
 - Imagens.



5. Representação de Imagens

Imagens

- Adquiridas por scanners ou câmeras.
- Imagens de tom contínuo.
 - Tons de cinza: 8 bits por pixel.
 - Coloridas: de 8, 16, 24 ou 32 bits por pixel.
- Conteúdo do frame-buffer em um arquivo.
 - Normalmente aplica-se compressão.
 - Diversos formatos
 - GIF, PNG, JPEG, ...

6. Compressão de Imagens

- O que é compressão de imagens?
- Necessidade de compressão.
- Princípios de compressão de imagens

6.1 O quê é compressão de imagens?

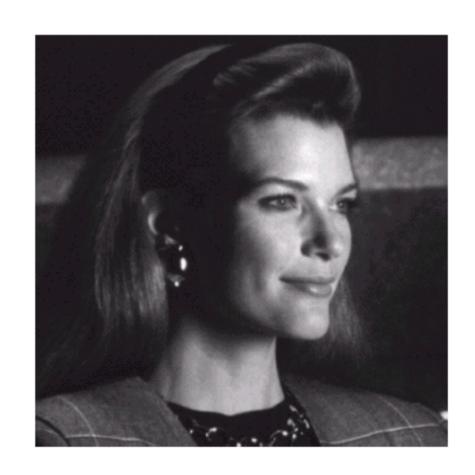
"O termo compressão de imagens refere-se ao processo de reduzir a quantidade de dados necessários para representar uma imagem com uma qualidade subjetiva aceitável."

Dados x informação

- Dados de imagem são altamente redundantes.
 - Remover redundâncias ajuda a alcançar compressão.
 - Redundâncias são matematicamente quantificáveis.
- Redundâncias em imagens:
 - Redundância estatística.
 - Também conhecida como redundância de codificação.
 - Redundância Espacial.
 - Também conhecida como redundância interpixel.
 - Redundância Psicovisual.
 - Utiliza conceitos do HVS.

- Redundância Estatística.
 - Função de Densidade de Probabilidade (pdf).
 - Valores dos pixels em uma imagem tem pdf não uniforme.
 - Métodos de codificação estatística podem ser usados para compressão de imagens.

- Valores dos pixels:
 - Irão variar de modo não uniforme.



- Redundância Estatística.
 - Codificação estatística.
 - Ou Variable Length Coding.
 - Lossless.
 - Códigos menores para símbolos (valores) mais frequentes.
 - Huffman, codificação aritmética.



- Redundância Espacial
 - Refere-se à correlação entre pixels vizinhos em uma imagem.
 - Relação geométrica ou estrutural entre os objetos em uma imagem.

- Redundância Espacial
 - O valor de um pixel pode ser razoavelmente "adivinhado" por meio dos valores de seus vizinhos.
 - Para remover redundância espacial:
 - Matriz de pixels deve ser transformada em um formato mais conveniente.
 - Diferenças entre pixels para representar a imagem.
 - Lossless: Codificação por diferença, codificação runlength.
 - Lossy: codificação preditiva, codificação por transformada.

- Redundância Psicovisual
 - Percepção de brilho.
 - Olho não responde com igual sensibilidade a toda informação visual.
 - Algumas informações tem mais importância relativa que outras.
 - Informação psicovisual redundante.
 - Diferente das outras redundâncias.
 - Está associada com informação visual de fato.
 - Então como é possível eliminá-la?



- Redundância Psicovisual
 - Sua eliminação implica em perda de informação visual quantitativa (real).
 - Daí o nome quantização.
 - É uma operação irreversível.

- Redundância Psicovisual
 - Propriedades do sistema visual humano (HSV).
 - Maior sensibilidade a distorções em áreas suaves (com baixa freqüência espacial).
 - Maior sensibilidade a distorções em áreas escuras de imagens.
 - Em imagens coloridas, maior sensibilidade a mudanças na luminância do que na crominância.

6.3 Princípios de Compressão de Imagens

- Técnicas podem ser combinadas!
 - Estatísticas + Espaciais + propriedades do HSV
 - Vantagem?



7. O Padrão JPEG

- O quê é JPEG?
- Preparação da imagem/bloco.
- Transformada DCT.
- Quantização.
- Codificação por Entropia.
- Construção do Quadro.

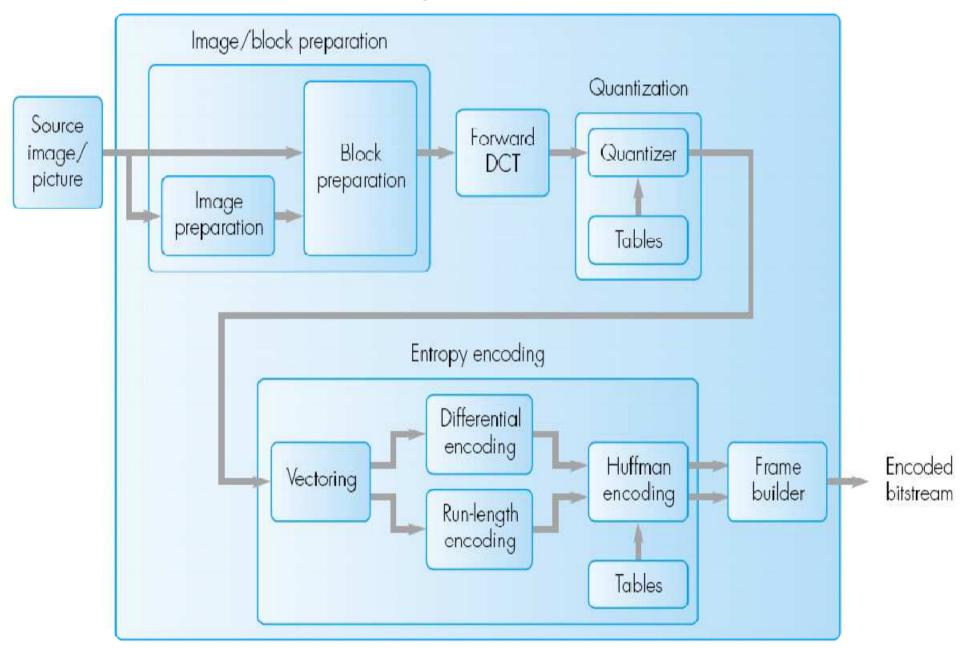
7.1 O quê é JPEG?

- Joint Photographic Experts Group.
 - ISO, CCITT e IEC.
 - Padrão para codificação de imagens estáticas de tons contínuos.
 - Possui 4 modos de operação:
 - Sequencial (baseline mode).
 - Progressivo.
 - Sem perdas.
 - Hierárquico.

7.1 O quê é JPEG?

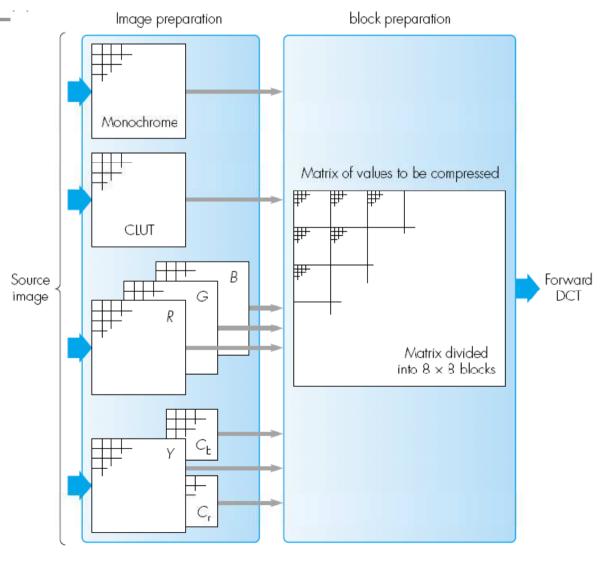
- Modo seqüencial
 - É um método de compressão com perdas.
 - Possui 5 etapas principais:
 - Preparação da imagem/bloco.
 - DCT.
 - Quantização.
 - Codificação.
 - Construção do quadro.

JPEG encoder



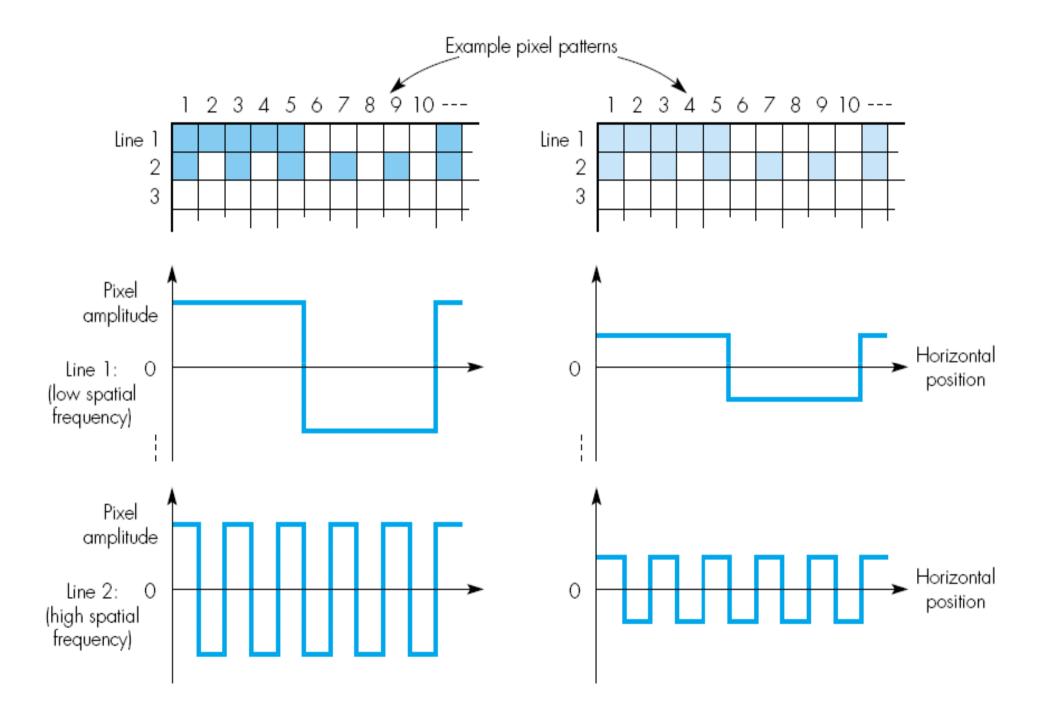
7.2 Preparação da imagem/bloco

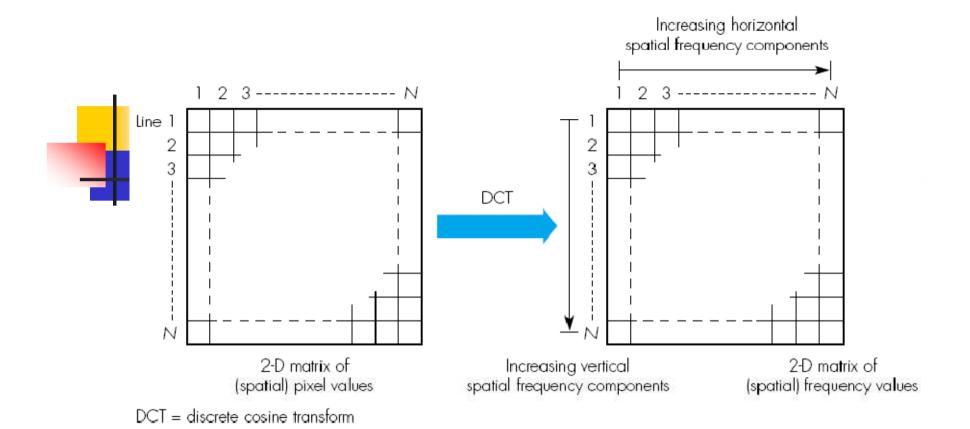
- Imagem é dividida em blocos de 8 x 8 pixels.
- Isso permite aplicação mais eficiente da DCT.





- Transformada Discreta de Cossenos (DCT).
- Transformadas:
 - Transformam a informação de um formato (domínio) para outro.
- Transformada DCT aplicada a imagens:
 - Transforma matriz (imagem) em matriz de frequências espaciais.
 - Não produz perdas.





- Olho humano é menos sensível a distorções em regiões com alta frequência espacial.
- Se a amplitude, nas altas freqüências, está abaixo de um limite, o olho não detecta a informação.
- Matriz transformada ajuda a detectar e eliminar tais informações (redundância psicovisual).



- Transformada Discreta de Cossenos (DCT)
 - Todos os blocos, um a um, são submetidos à DCT.

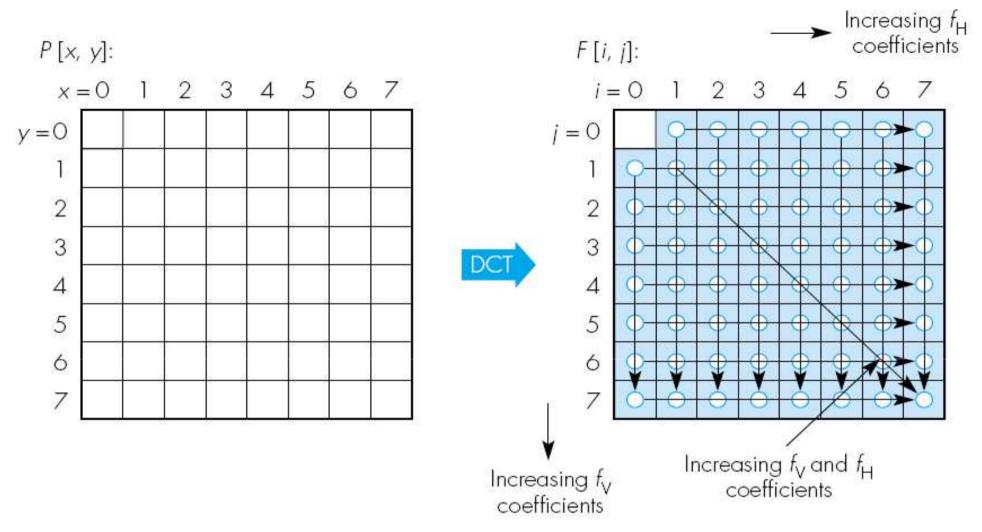
$$F[i,j] = \frac{1}{4}C(i)C(j)\sum_{x=0}^{7}\sum_{y=0}^{7}P[x,y]\cos\frac{(2x+1)i\pi}{16}\cos\frac{(2y+1)j\pi}{16}$$

onde C(i) e $C(j) = 1/\sqrt{2}$ para i,j=0= 1 para todos os outros valores de i e j. x, y, i e j todos variam de 0 a 7.

 Transformada Discreta dos Cossenos Inversa (IDCT)

$$P[x,y] = \frac{1}{4} \sum_{i=0}^{7} \sum_{j=0}^{7} C(i)C(j)F[i,j] \cos \frac{(2x+1)i\pi}{16} \cos \frac{(2y+1)j\pi}{16}$$

onde
$$C(i)$$
 e $C(j) = 1/\sqrt{2}$ para $i,j=0$
= 1 para todos os outros valores de i e j .
 x, y, i e j todos variam de 0 a 7.



 $P[x, y] = 8 \times 8$ matrix of pixel values

 $F[i, j] = 8 \times 8$ matrix of transformed values/spatial frequency coefficients

In F[i, j]: = DC coefficient = AC coefficients

 $f_{\rm H}=$ horizontal spatial frequency coefficient

 $f_{\rm V}$ = vertical spatial frequency coefficient



Após DCT:

- As regiões da imagem que possuem uma única cor geram matrizes com coeficientes DC idênticos (ou próximos) e poucos coeficientes AC.
- As regiões da imagem que possuem transições de cores geram matrizes com coeficientes DC distintos e muitos coeficientes AC.
- Tamanho do bloco na imagem.
- Regiões com pouca/muita transição de cor X coeficientes DC/AC.

7.4 Quantização

- Quantização
 - Olho humano:
 - Boa resposta para coeficientes DC (baixa freq.).
 - Baixa reposta para coeficientes AC (alta freq.).
 - Busca reduzir a quantidade de dados.
 - Limite da amplitude para frequências: divide os valores da matriz transformada pelos valores correspondentes em uma tabela pré-definida.
 - Isso diminui os valores dos coeficientes proporcionalmente à posição dos mesmos na matriz.
 - Ocorre perda. No caso ideal, não perceptível.

DCT coefficients													Quar	ntized	coeffic	cients			
120	60	40	30	4	3	0	0					12	6	3	2	0	0	0	0
70	48	32	3	4	1	0	0					7	3	2	0	0	0	0	О
50	36	4	4	2	0	0	0	Quantizer				3	2	0	0	0	0	0	0
40	4	5	1	ı	0	0	0				2	O	0	0	0	O	0	О	
5	4	0	О	0	0	0	0		Qua	Rudniizer		0	0	0	0	0	0	0	0
3	2	0	О	0	0	0	0					0	0	0	0	0	0	0	О
1	1	0	О	0	0	0	0					0	0	0	0	0	0	0	0
0	0	0	О	0	0	0	0					0	0	0	0	0	0	0	О
						10	10	15	20	25	30	35	40						
						10	1.5	20	25	30	35	40	50						
						15	.20	25	30	35	40	50	60						
						20	25	30	35	40	50	60	70						
						25	30	35	40	50	60	70	80						
						30	35	40	50	60	70	80	90						
						35	40	50	60	70	80	90	100						
						10	50	60	70	80	90	100	110						

Quantization table



7.4 Quantização

- Tabelas de quantização:
 - JPEG define duas tabelas default
 - Uma para luminância.
 - Uma para crominância.
 - JPEG permite a utilização de tabelas personalizadas.



- Explora duas características da matriz quantizada:
 - Coeficiente DC será o maior valor da matriz
 - Muitos dos coeficientes de alta frequência serão zero



- Envolve quatro passos:
 - Vetorização.
 - Codificação por diferença.
 - Codificação por carreira (run-length).
 - Codificação Estatística (método de Huffman).



Vetorização (zig-zag scan)

Quantized coefficients 4 5 6 Linearized vector 63 3 2 AC coefficients in increasing 6 order of frequency DC coefficient 63 5 10 9

4

7.5 Codificação por Entropia

- Codificação por diferença
 - Codificação dos coeficientes DC.
 - DCs possuem alto grau de correlação (redundância espacial).
 - São blocos adjacentes na imagem.
 - Exemplo:
 - Sequência de coeficientes DC de blocos adjacentes: 12, 13, 11, 11, 10, ...
 - Valores codificados: 12, 1, -2, 0, -1, ...

- Codificação run-length
 - Codificação dos coeficientes AC.
 - Vetor de coeficientes possui longas cadeias de zeros.
 - Exemplo:
 - (skip, value): skip indica a quantidade de zeros a ser "pulada"; value é o próximo valor não zero da seq.

(0,6),(0,7),(0,3),(0,3),(0,2), (0,2), (0,2), (0,2), (0,0)



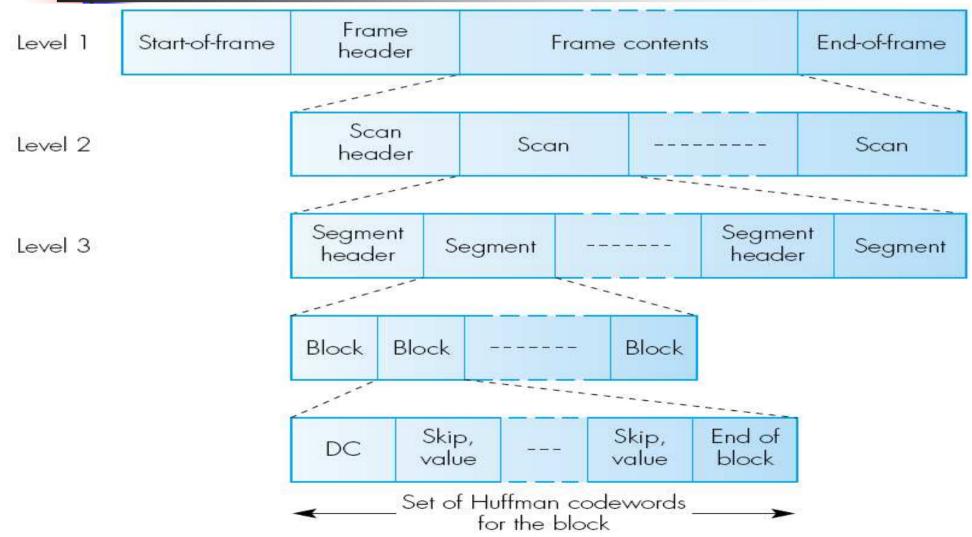
- Codificação estatística
 - Após a codificação Run-Length é aplicada uma codificação estatística.
 - JPEG usa Huffman.
 - A codificação estatística é aplicada no vetor inteiro, o que inclui o resultado das codificações dos DCs e ACs.
 - Vetor possui cadeias de bits apropriado para codificação estatística.
 - JPEG usa tabela de códigos (prefixo).
 - São 256 códigos possíveis.
 - Pré-definida ou enviada junto com o bitstream da imagem.



- Codificação estatística
 - Após a codificação Run-Length é aplicada uma codificação estatística.
 - JPEG usa Huffman.
 - A codificação estatística é aplicada no vetor inteiro, o que inclui o resultado das codificações dos DCs e ACs.
 - Vetor possui cadeias de bits apropriado para codificação estatística.
 - JPEG usa tabela de códigos (prefixo).
 - São 256 códigos possíveis.
 - Pré-definida ou enviada junto com o bitstream da imagem.

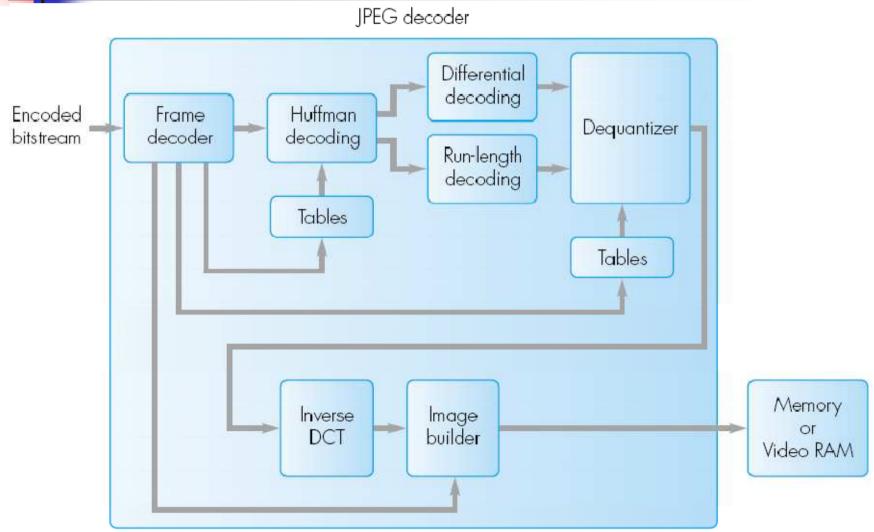


7.6 Construção do Quadro





7.7 Decodificação





Exemplo – bloco original

[140	144	147	140	139	155	179	175]
144	152	140	147	140	148	167	179
152	155	136	167	163	162	152	172
168	145	156	160	152	155	136	160
162	148	156	148	140	136	147	162
147	167	140	155	155	140	136	162
136	156	123	167	162	144	140	147
148	155	136	155	152	147	147	136



Bloco com shifting

[12	16	19	12	11	27	51	47
16	24	12	19	12	20	39	51
24	27	8	39	35	34	24	44
40	17	28	32	24	27	8	32
34	20	28	20	12	8	19	34
19	39	12	27	27	12	8	34
8	28	-5	39	34	16	12	19
20	27	8	27	24	19	19	8]

Após a DCT

[185.88	-17.962	14.943	-9.0778	23.125	-9.0856	-13.901	-19.110
20.365	-34.045	26.557	-9.1747	-11.106	10.935	13.866	6.7143
-10.547	-23.469	-1.6402	5.9121	-18.238	3.3890	-20.329	-1.0530
-8.2518	-5.0009	14.524	-14.729	-8.3648	-2.5596	-3.0050	8.2253
-3.3750	9.5359	8.0480	1.2188	-11.125	18.051	18.450	15.068
3.7574	-2.1876	-18.039	8.4227	8.1706	-3.4929	0.92215	-6.9987
8.8337	0.65168	-2.8289	3.5882	-1.2401	-7.3423	-1.1098	-2.0184
[0.014635]	-7.8035	-2.3794	1.5633	1.1648	4.2876	-6.3987	0.26693



	5	7	9	11	13	15	17
5	7	9	11	13	15	17	19
7	9	11	13	15	17	19	21
9	11	13	15	17	19	21	23
11	13	15	17	19	21	23	25
13	15	17	19	21	23	25	27
15	17	19	21	23	25	27	29
$\lfloor 17$	19	21	23	25	27	29	31

Após quantização

Últimos passos

- Zig-zag sequence
- Intermediate symbol sequence
 - (6)(62), (0,3)(-4), (0,3)(4), (0,2)(-2), (0,3)(-5), (0,2)(2), (0,1)(-1), (0,2)(3), (0,2)(-3), (0,1)(-1), (3,1)(-1), (0,2)(2), (0,1)(-1), (0,1)(-1), (1,1)(1), (0,1)(1), (1,1)(1), (1,1)(1), (0,1)(-1), (0,1)(-1), (0,1)(1), (0,1)(-1), (0,1)(1), (0,1)(1), (0,1)(1), (0,1)(1), (0,1)(1), (1,1)(1), (0,0)
- Encoded bit sequence (total 154 bits)
 - (1110)(111110) (100)(001) (100)(100) (01)(01) (100)(010) (01)(10) (00)(0) (01)(11) (01)(00) (00)(0) (111010)(0) (01)(10) (00)(0) (00)(0) (1100)(1) (00)(1) (1100)(1) (1100)(1) (00)(0) (00)(0) (00)(1) (00)(0) (00)(0) (00)(1) (111010)(0) (1111010)(0) (1100)(0) (111010)(1) (11111011)(1) (1100)(1) (1010)



7.8 Considerações Sobre JPEG

JPEG:

- Padrão abrangente.
- Alcança boas taxas de compressão para imagens de tons contínuos. (até 20:1).
- Desempenho diminui em imagens com muita transição de cores.
- Baseado em particularidades do sistema visual humano:
 - Não é necessário reproduzir cantos com fidelidade.
 - O olho humano não responde bem a transições nas altas frequências espaciais.
 - É adequado para imagens de tom contínuo.

8. JPEG2000

- Uma das maiores limitações do JPEG:
 - Desempenho degrada em baixas taxas de dados (bitrate).
- Emprega transformada Wavelet.
- Melhorias:
 - Codificação estática/dinâmica de alta qualidade de uma região específica.
 - Recuperação de erros.
 - Desempenho: até 50:1 (níveis de cinza), 100:1 (cor).



9. Outros Formatos

- Outros formatos populares são:
 - GIF
 - PNG
 - TIFF, BMP, ...



Para Saber Mais

- Gonzales & Woods. Digital Image Processing.
 2nd ed. Prentice-Hall, 2002. Capítulo 8, seção 8.1.
- Halsall, F. Multimedia Communications: Applications, Networks, Protocols, and Standards, Addison-Wesley Publishing, 2001. ISBN: 0201398184. Capítulo 2, seção 2.4 e capítulo 3, seções 3.2 e 3.4.
- Pennebaker & Mitchell. JPEG Still Image Data Compression Standard. Van Nostrand Reinhold, 1993.