

# ESTI019 – Codificação Multimídia

Profs. Celso Kurashima e  
Mário Minami

## Introdução à Codificação de Imagem

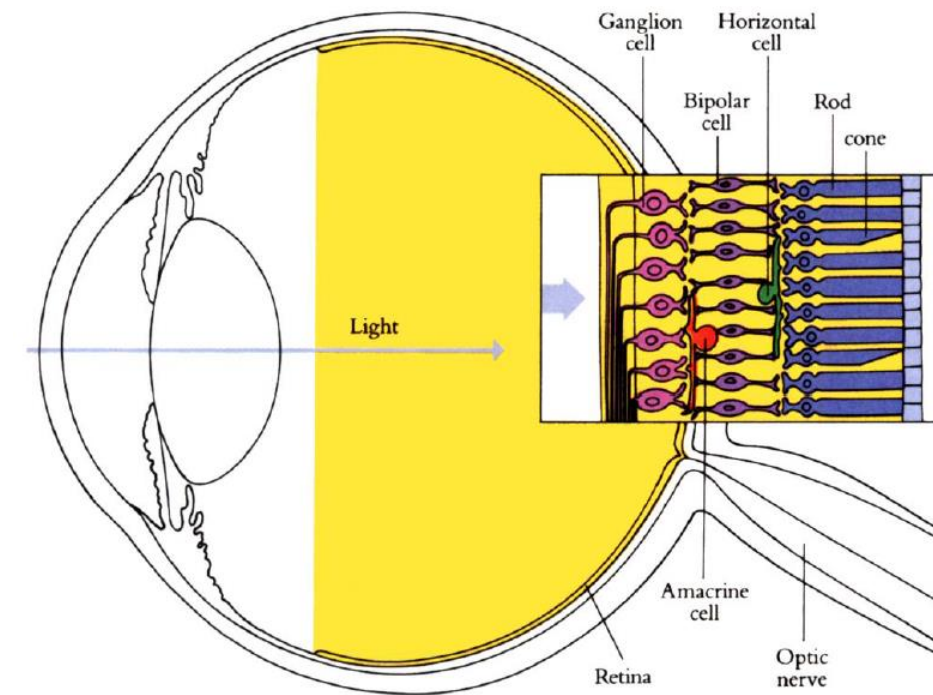
[mario.minami@ufabc.edu.br](mailto:mario.minami@ufabc.edu.br)  
[celso.kurashima@ufabc.edu.br](mailto:celso.kurashima@ufabc.edu.br)

# Tópicos desta aula

- Codificação Visual
- Sub-amostragem no Espaço de Cores
- Codificação por Transformadas
  - Transformada Discreta Cosseno (DCT)
  - Quantização de Coeficientes



Lenna



# Codificação Visual

What are the computational problems of visual coding?

# Redundância de Informações na Imagens

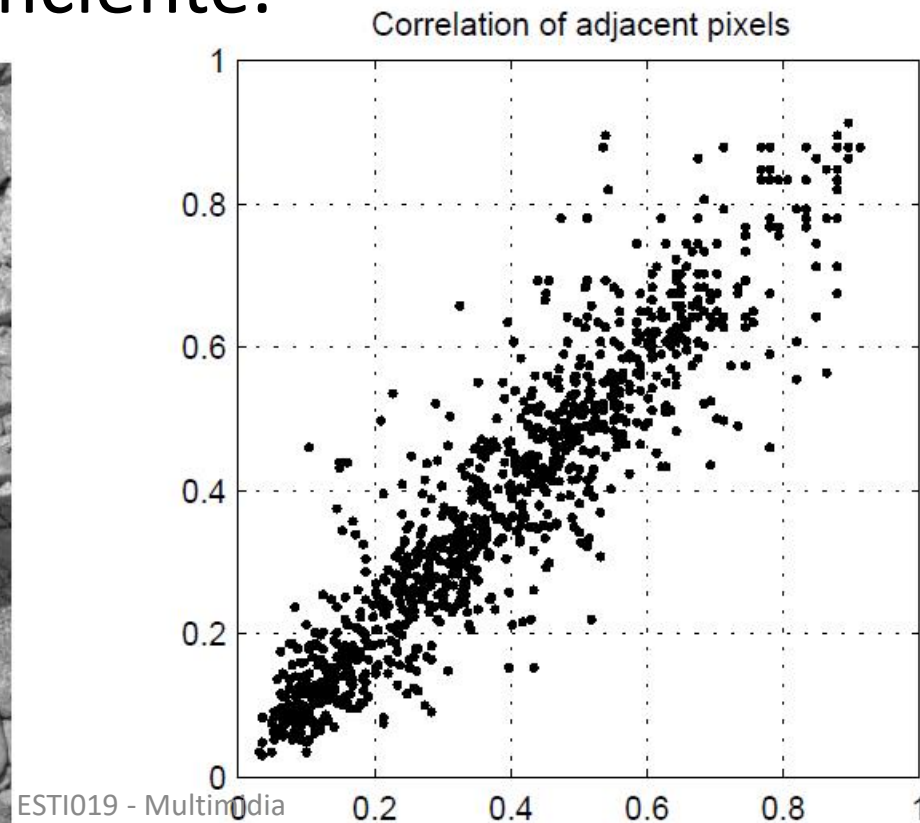
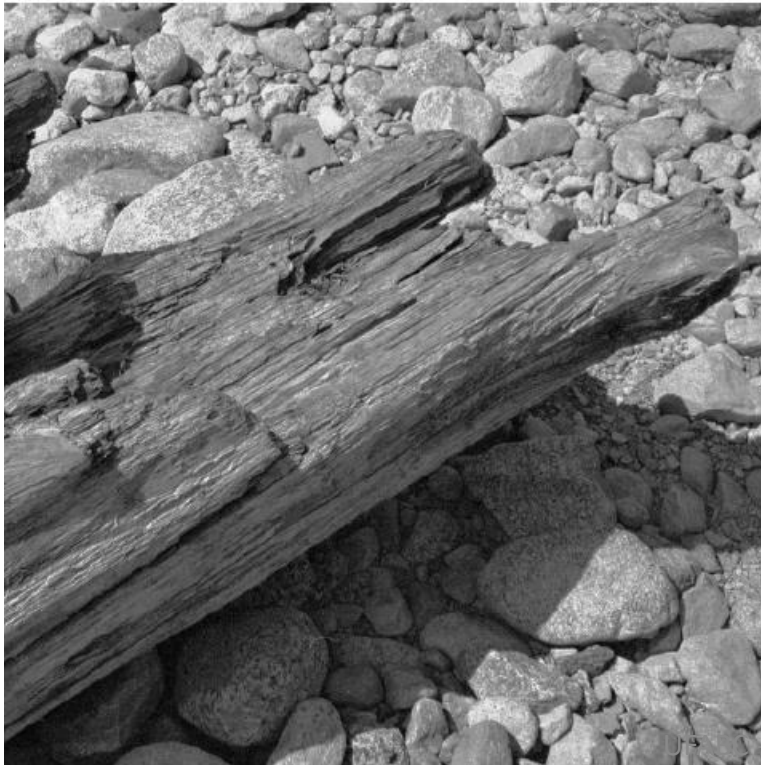
A Fovea já realiza uma grande redução na quantidade de informação visual;

Como podemos transmitir com segurança as informações visuais importantes?



# Uma abordagem geral para a codificação: redução de redundância

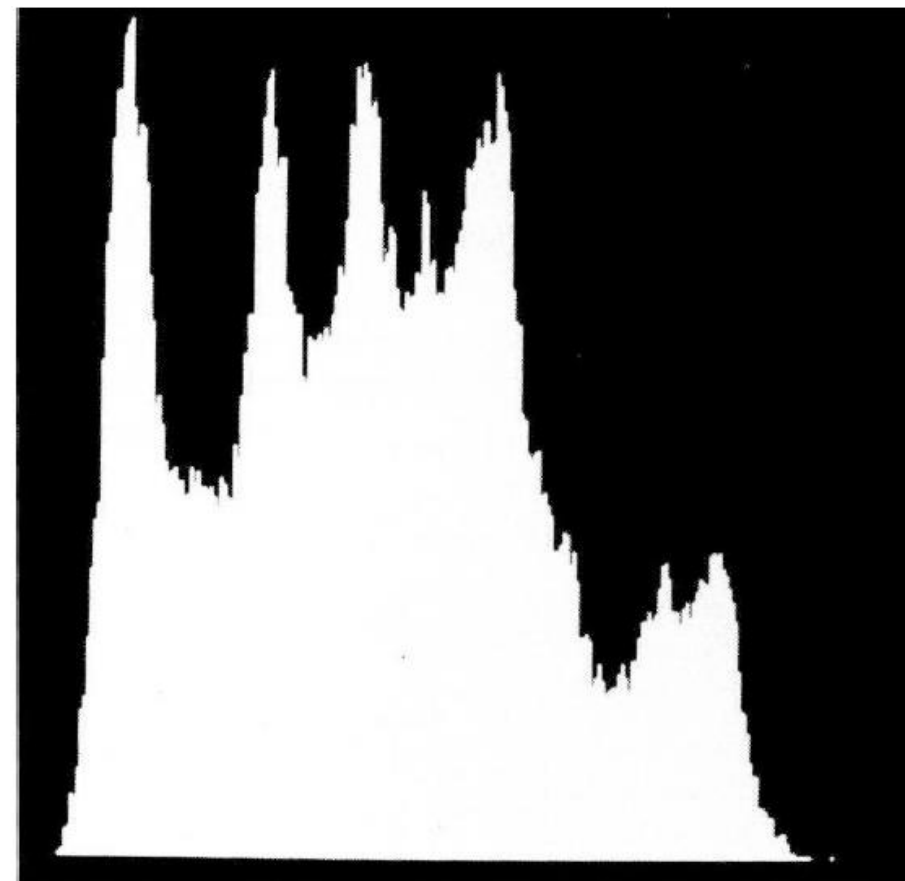
- Por que reduzir redundância?
- Equivalente à codificação eficiente.



# Reduzindo a redundância de pixels



Lena: a standard 8 bit 256x256 gray scale image



histogram of pixel values  
Entropy = 7.57 bits

# Um pouco da Teoria da informação

---

$H(X)$  é uma medida de quanta informação é necessária em média para descrever a variável aleatória  $X$ .

---

Se sabemos  $p(X)$ , podemos calcular a entropia ou otimizar o modelo para os dados,

---

mas o que acontece se não sabemos  $p(X)$  e só podemos aproximá-lo?

---

Quantos bits esta inexactidão nos custa?

---

# Compressão de imagens

- A compressão de imagem é a aplicação de técnicas de compressão de dados em imagens digitais bidimensionais  $I(x, y)$ ,
- para reduzir a redundância dos dados de imagem para armazenamento ou transmissão de uma forma eficiente.
- A compressão de imagens pode ser classificada em duas categorias:
  - Sem perdas ou com perdas



## Compressão Sem Perdas

- A compressão sem perdas, que alcança proporções de compressão menores do que a compressão com perdas, aproveita principalmente o conteúdo da imagem contendo uma distribuição de probabilidade não uniforme para uma representação de comprimento variável dos pixels da imagem.
- Essas imagens incluem desenhos técnicos, ícones ou quadrinhos e conteúdos de alto valor, como imagens médicas ou digitalizações de imagens feitas para fins de arquivamento.

## Compressão Com Perdas

- Os métodos de compressão com perdas, especialmente quando atingem uma proporção de compressão muito elevada, podem introduzir artefactos de compressão.
- No entanto, a compressão com perdas são especialmente adequadas para imagens naturais, tais como fotografias, em aplicações onde uma perda de fidelidade menor (por vezes imperceptível) é aceitável quando é desejável obter uma redução substancial na taxa de bits.
  - Padrões de compressão de imagem do estado-da-arte usam uma combinação de algoritmos com e sem perdas para obter o melhor desempenho.



# Sub-amostragem no Espaço de Cores

O espaço RGB é o formato básico.

Os formatos YIQ, YUV, YCbCr, foram desenvolvidos para sistemas de televisão e vídeo.

# A correção Gamma



Para compensar o processamento não linear das telas, é feita a correção-gamma no dado RGB linear, antes da transmissão:  $R'G'B'$ .



Como resultado, as áreas mais escuras das imagens, possuem um ganho maior do que as áreas mais claras.



**Os benefícios da correção gamma no transmissor são:**

**O custo do receptor de TV é minimizado**

**O processo de transmissão é reduzido em áreas escuras da imagens, resultando na eliminação de artefatos visuais.**



*>>> É possível fazer uma analogia da correção-gamma com o processo de compensação no sinal de voz PCM?*

# A correção Gamma ( $\gamma$ )

- Gamma é definido por  $V_{out} = V_{in}^\gamma$ , onde  $V_{out}$  é o valor da luminância de saída e  $V_{in}$  é o valor de luminância real de entrada.

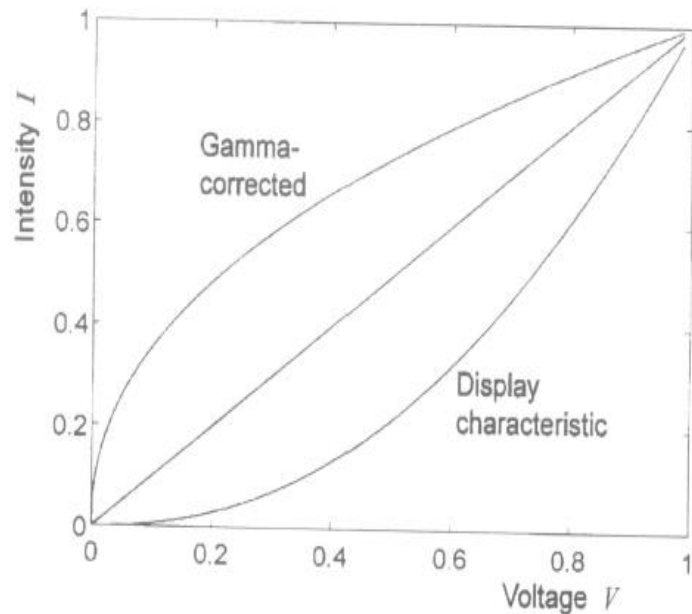
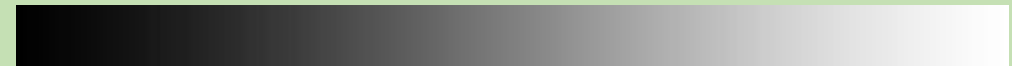


Figure 2.2 The effect of gamma correction with a factor of 2.2.

Faixa de Luminância Original



Codificação Linear



Codificação Gamma



$$I_{display} \propto V_{received}^\gamma$$



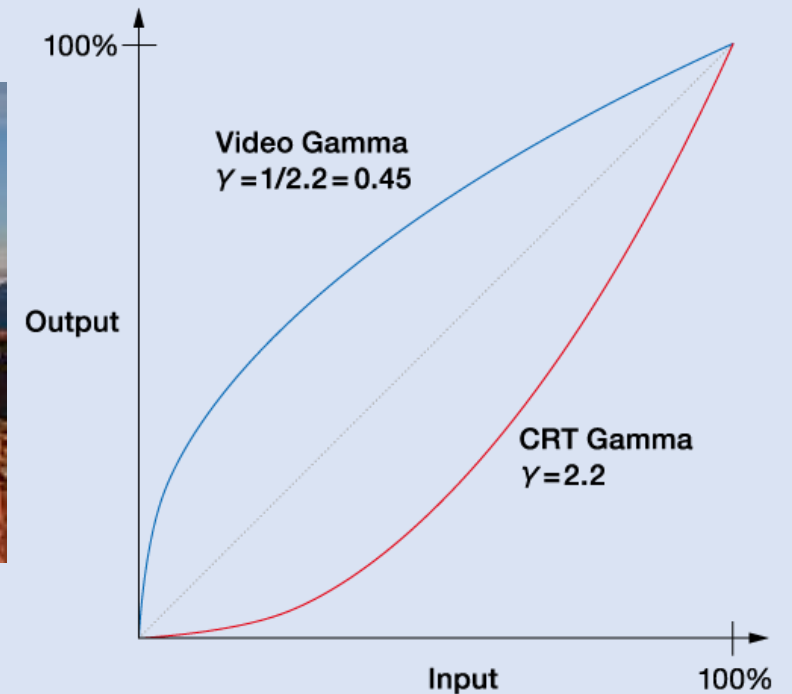
# A correção Gamma 2.2 em imagens



Linear RAW Image  
(image gamma = 1.0)

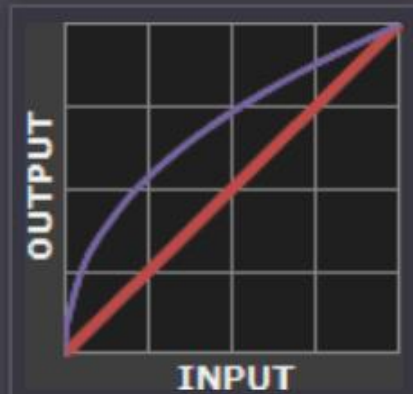


Gamma Encoded Image  
(image gamma = 1/2.2)

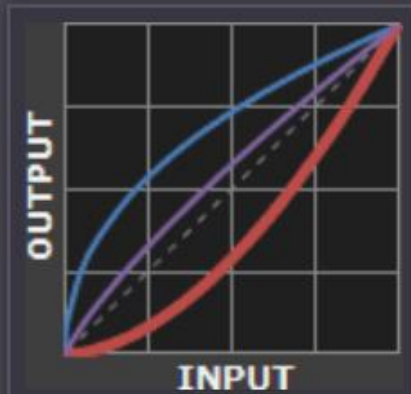


A maioria dos arquivos de imagem usa uma gama de codificação de 1/2.2 (como aqueles que usam sRGB e Adobe RGB 1998), mas a grande exceção é com arquivos RAW, que usam um gamma linear.

# Correção Gamma em Displays e Telas



Display Gamma 1.0



Display Gamma 1.8



Display Gamma 2.2



Display Gamma 4.0



Diagrams assume that your display has been calibrated to a standard gamma of 2.2.

Recall from before that the image file gamma (—) plus the display gamma (—) equals the overall system gamma (—). Also note how higher gamma values cause the colors to tend downward.

# Espaço de Luminância e Crominância

A componente de luminância  $Y$  representa a intensidade da escala de cinza e os componentes de crominância  $CbCr$  são informação de cor.

A razão para fazer essa transformação de cor é que podemos perder mais informações nos componentes de crominância do que no componente de luminância, uma vez que os olhos humanos são menos sensíveis aos conteúdos cromáticos de alta frequência do que aos conteúdos de luminância de alta frequência.

# Espaços “rotacionados”, equações:

**YUV (PAL)**

e

**YIQ (NTSC)**

- $Y' = 0.299R' + 0.587G' + 0.114B'$
- $U = -0.147R' - 0.289G' + 0.436B'$   
 $= 0.492 (B' - Y')$
- $V = 0.615R' - 0.515G' - 0.100B'$   
 $= 0.877(R' - Y')$

- $Y' = 0.299R' + 0.587G' + 0.114B'$
- $I = 0.596R' - 0.275G' - 0.321B'$   
 $= 0.736(R' - Y') - 0.268(B' - Y')$
- $Q = 0.212R' - 0.523G' + 0.311B'$   
 $= 0.478(R' - Y') + 0.413(B' - Y')$

**Nota:**

1.  $Y'$  é muitas vezes referida como “luma”
2. O sinal ' denota as componentes depois da correção gamma

\*K. JACK, Video Demystified.

# Equações YCbCr

- *RGB-YCbCr : SDTV*

- $Y' = 0.299R' + 0.587G' + 0.114B'$
- $Cb = -0.172R' - 0.339G' + 0.511B' + 128$
- $Cr = 0.511R' - 0.428G' - 0.083B' + 128$

- *RGB-YCbCr: HDTV*

- $Y' = 0.213R' + 0.715G' + 0.072B'$
- $Cb = -0.117R' - 0.394G' + 0.511B' + 128$
- $Cr = 0.511R' - 0.464G' - 0.047B' + 128$



RGB

Y'

Cb

Cr

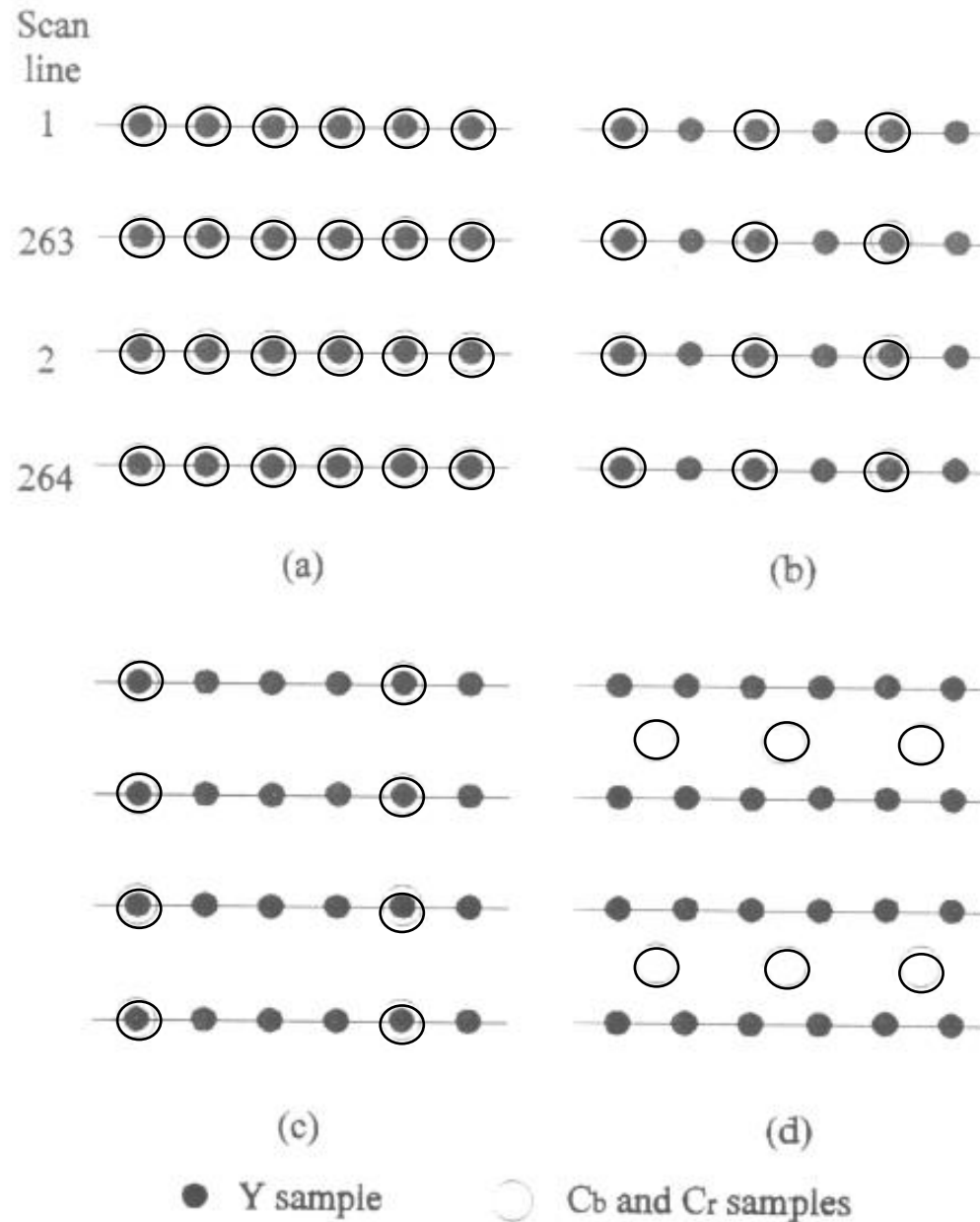
\*K. JACK, Video Demystified.



# Sub-amostragem de Crominância

- Após a transformação, o componente de luminância  $Y$  é deixado em resolução total
  - Quando os componentes de croma  $CbCr$  também são deixados em resolução total o formato é chamado 4: 4: 4,
- Enquanto que os componentes de croma  $CbCr$  são frequentemente reduzidos em 2: 1 ou 4: 1 horizontalmente e em 2: 1 ou 1: 1 (sem alteração) verticalmente.
- No JPEG, essas alternativas são geralmente chamadas de técnicas de subamostragem chroma 4: 1: 1 e 4: 2: 2

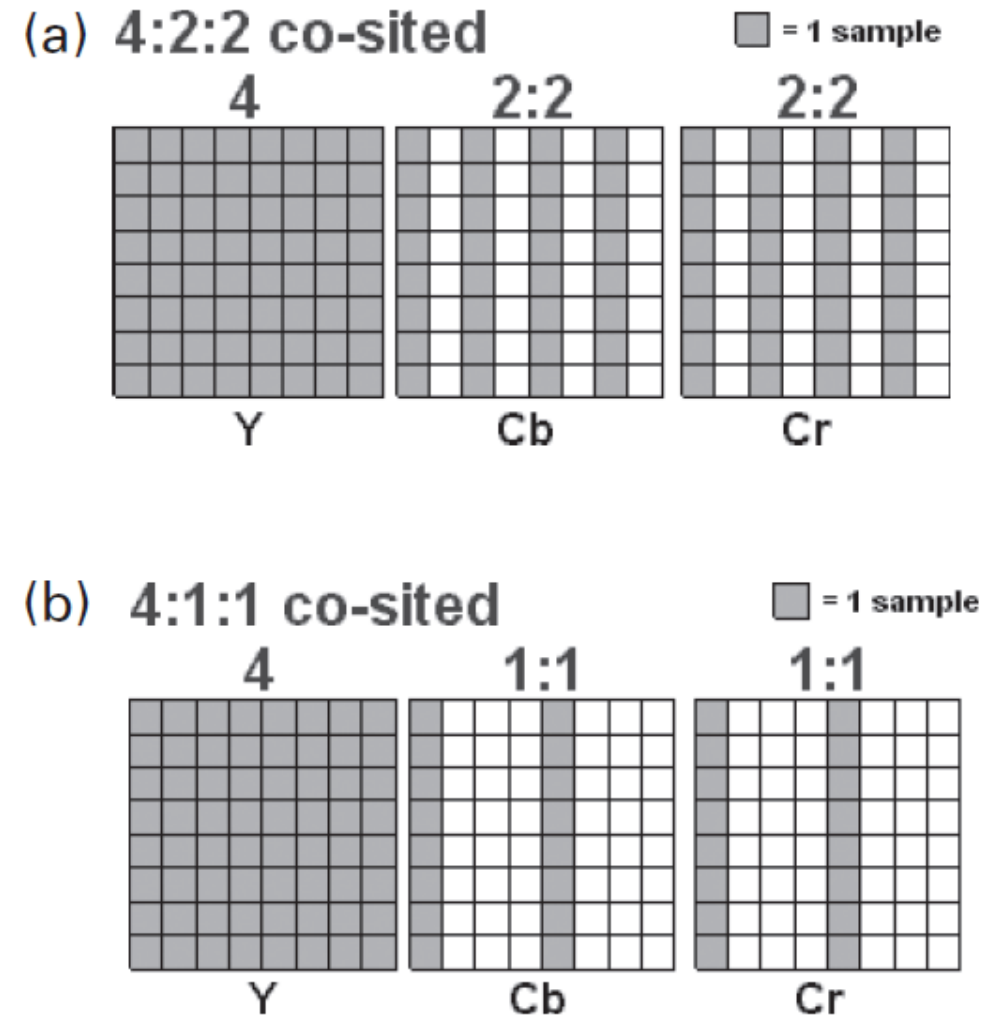
Sub-amostragem  
das  
componentes  
**Cb**, **Cr**



**Figure 2.3** Orthogonal sampling on the scan lines of an interlaced system: (a) 4:4:4; (b) 4:2:2; (c) 4:1:1; and (d) 4:2:0 format.

# JPEG

*Sub-amostragem  
utilizadas  
na compressão JPEG*



# Codificação por Transformadas

Transform Coding

# Codificação por transformadas

- Um sinal é mapeado a partir de um domínio – espacial ou temporal, para o domínio da transformada.
- O sinal pode ser uni-dimensional ou multidimensional.
- Transformadas ortogonais são mais usadas, pois o mapeamento é único e reversível.
- A energia é preservada no domínio da transformada, e o sinal pode ser completamente recuperado pela sua transformada inversa.
- Para compressão de imagens, as principais transformadas discretas são a DCT e a DWT.



# A Vantagem da Ortogonalidade

**C ortogonal:**  $C^T C = I$

Implica em  $C^{-1} = C^T$

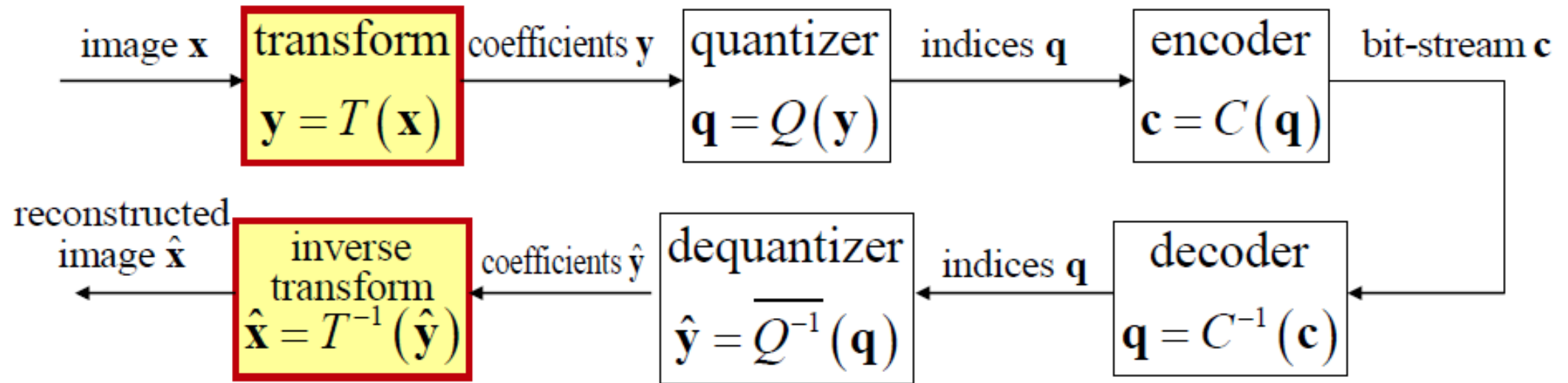
A transformada direta fornece:  $Y = CXC^T$

Resolver  $Y = CXC^T$  para  $X$ :

$$C^T Y = C^T C X C^T = X C^T$$

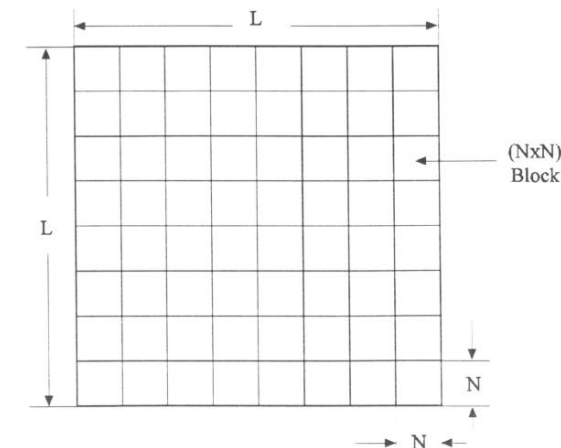
$$C^T Y C = X C^T C = X \quad \text{ou} \quad X = C^T Y C \quad (\text{transformada inversa})$$

# Codificador Estruturado típico



- Transformada  $T(\mathbf{x})$  usualmente inversível
- Quantização  $Q(\mathbf{y})$  não inversível, introduz distorção
- Combinação de encoder  $C(\mathbf{q})$  e decoder  $C^{-1}(\mathbf{c})$  sem perdas

Figure 5.14 Subblock division in transform image coding.



# Codificação por transformada DCT

- Uma imagem ( $L \times L$ ) é dividida em blocos não sobrepostos de tamanho  $N \times N$  cada, as quais são mapeadas num domínio DCT-2D.
- Uma versão simples de codificação de imagens por transformadas é mostrada na figura ao lado.
- Generalizando, a imagem original poderia ser  $P \times L$  e o bloco poderia ser  $M \times N$ .
- Dos coeficientes DCT ( $N \times N$ ), apenas alguns são significantes, e são estes os selecionados para quantização.

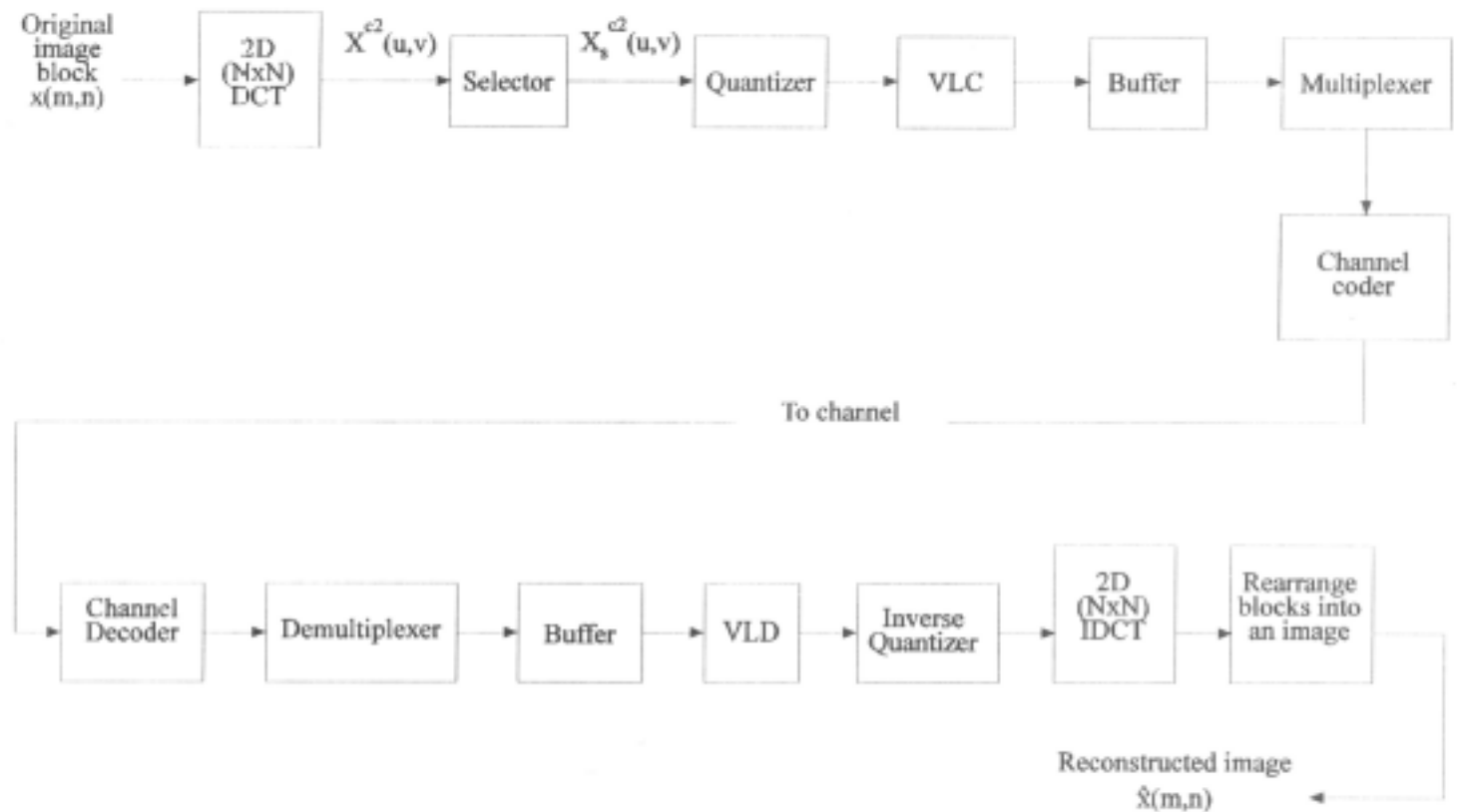


Figure 5.15 Transform image encoder/decoder.

# Transformada Discreta de Cosseno - DCT

- Nesta aula estudaremos a Transformada Discreta de Cosseno, especificamente.
- A DCT foi selecionada por grupos de padronização de compressão de imagem e de vídeo, pois é o fator primário para obtenção de compressão.
- Por si só, a DCT não leva à redução de redundâncias em imagens. Necessita de técnicas complementares como quantização, classificação, codificação de linha (VLC, RLC), etc.; e inclusive a sub-amostragem de cromaticidade.

# Definindo a DCT (1D)

- Discrete cosine transform
  - Projeção de  $f$  sobre as funções de base
  - Todas amostras de  $f$  contribuem para o coeficiente
  - $C(0)$  é a componente de frequência zero – o valor médio!

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos\left(\frac{(2x+1)u\pi}{2N}\right)$$
$$\alpha(k) = \begin{cases} \sqrt{\frac{1}{N}} & \text{if } k = 0 \\ \sqrt{\frac{2}{N}} & \text{otherwise} \end{cases}.$$



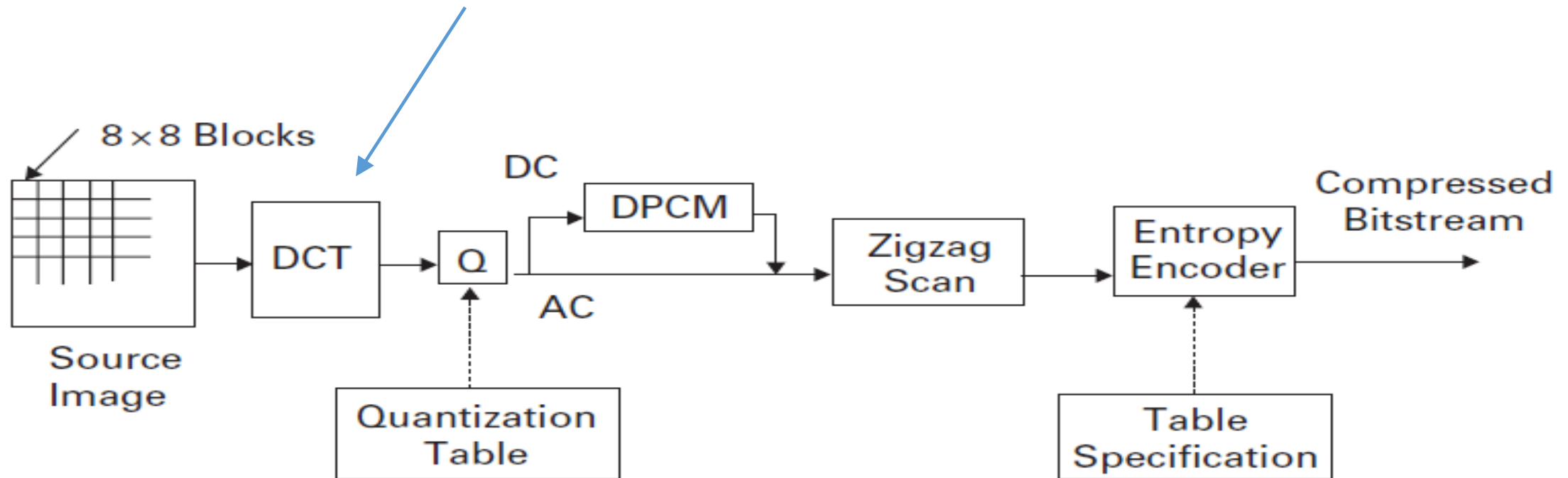
# Transformada Discreta de Cosseno - DCT

A transformada ótima KLT não possui uma transformação fixa, e não possui um algoritmo rápido.

A DCT se aproxima bem da KLT em desempenho, e possui algoritmos rápidos e estruturas recursivas.

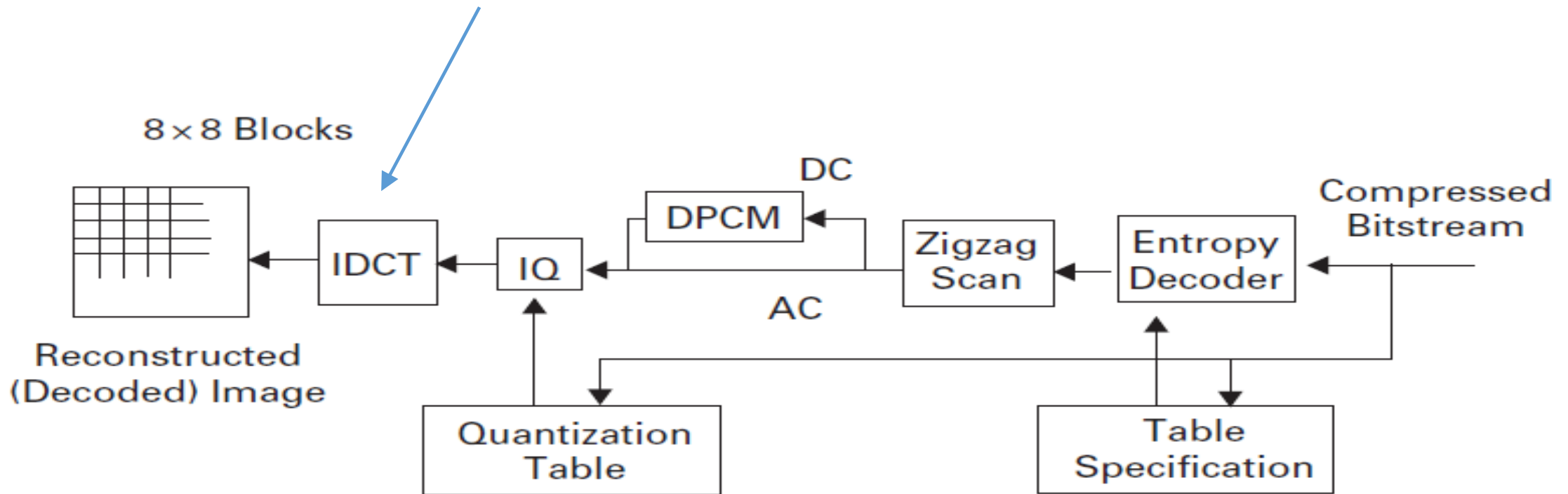
A desvantagem é a estrutura em blocos que domina em taxas baixas de bits.

# Posição da DCT (2D) no JPEG- Encoder



**Figure 4.7** A DCT-based sequential baseline encoder of JPEG [14].

# Posição da IDCT (2D) no JPEG- Decoder



**Figure 4.8** A DCT-based sequential baseline decoder of JPEG [14].

# Equações DCT (2D)

- Para cada componente de imagem, agrupamos os valores dos pixels  $\{f(x,y)\}$  em blocos 8x8 para a DCT

$$F(u, v) = \frac{1}{4} C(u) C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos \left[ \frac{(2x+1)u\pi}{16} \right] \cos \left[ \frac{(2y+1)v\pi}{16} \right],$$

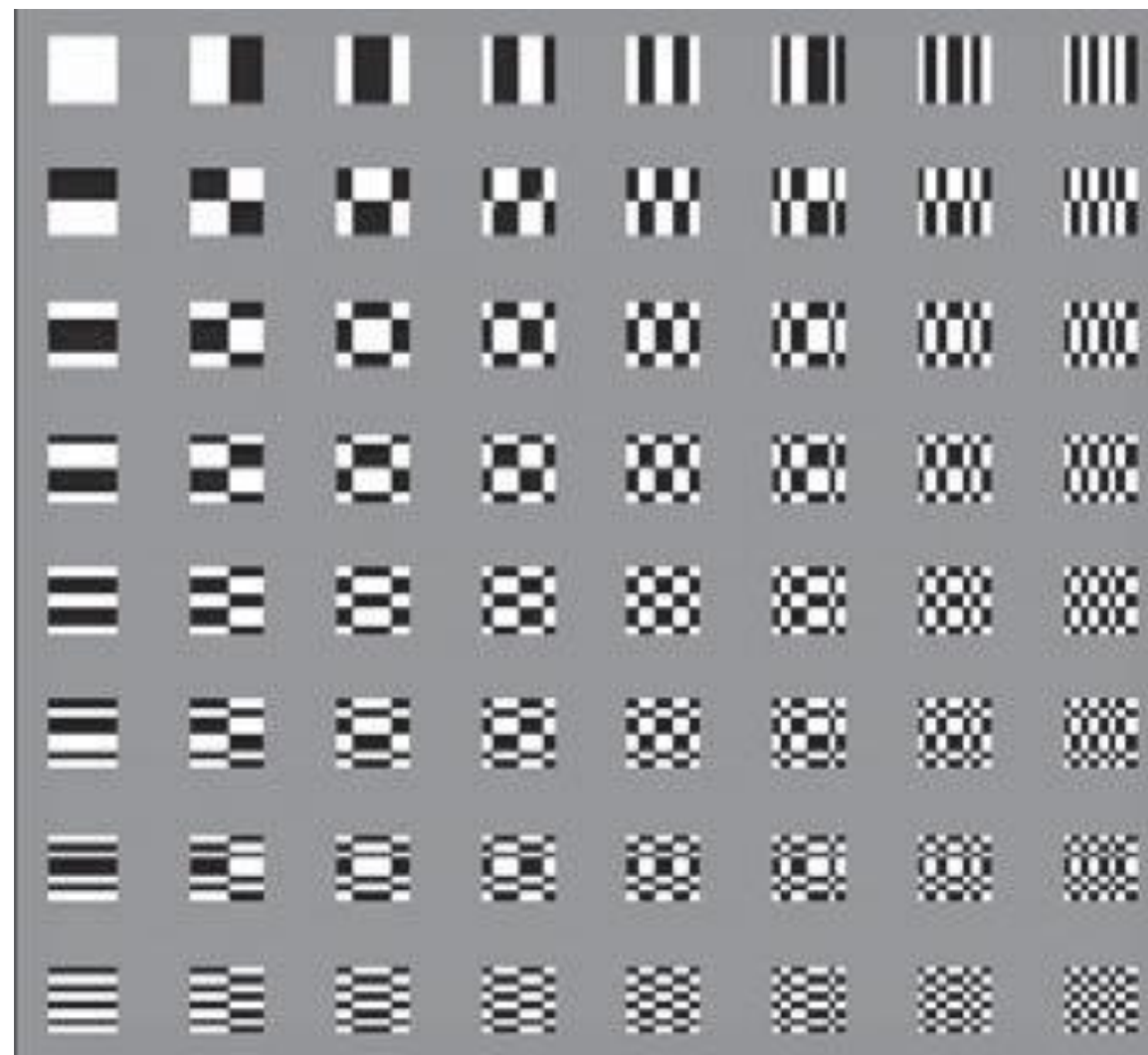
$$f(x, y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u) C(v) F(u, v) \cos \left[ \frac{(2x+1)u\pi}{16} \right] \cos \left[ \frac{(2y+1)v\pi}{16} \right],$$

where

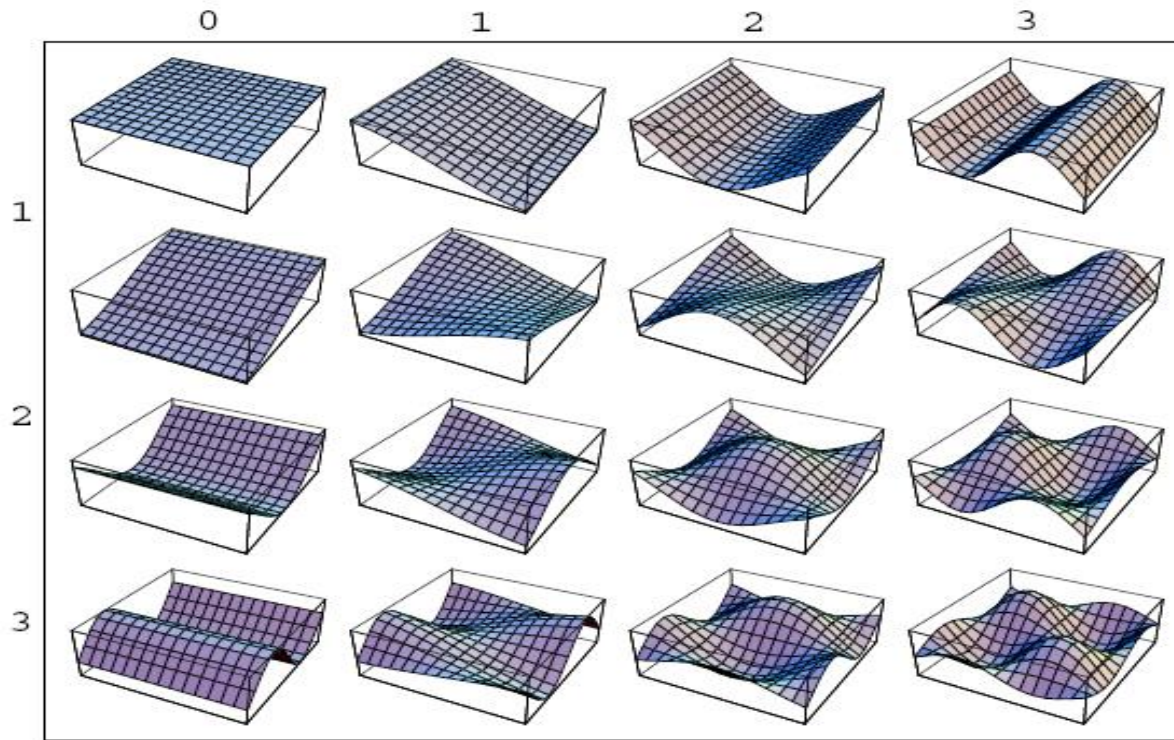
$$C(w) = \begin{cases} \frac{1}{\sqrt{2}}, & w = 0, \\ 1 & \text{otherwise.} \end{cases}$$

# Transformada Discreta de Cossenos

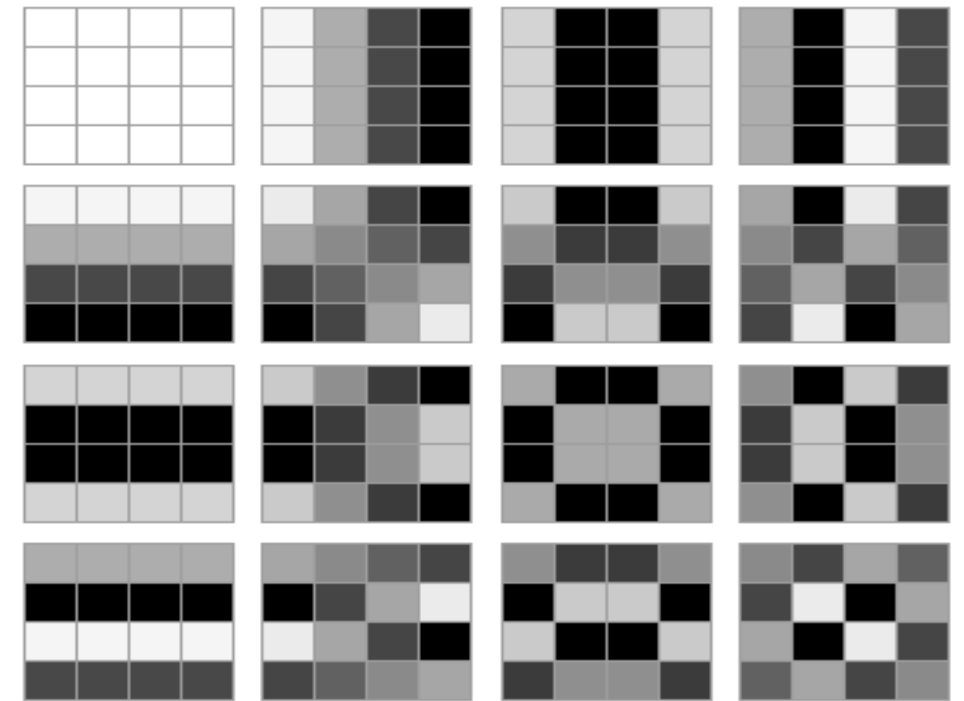
- DCT em duas dimensões utiliza uma base ortogonal de 64 sinais, cada uma
- contendo um das 64 frequências espaciais.



# Funções Base da DCT e os Padrões Base



Basis functions

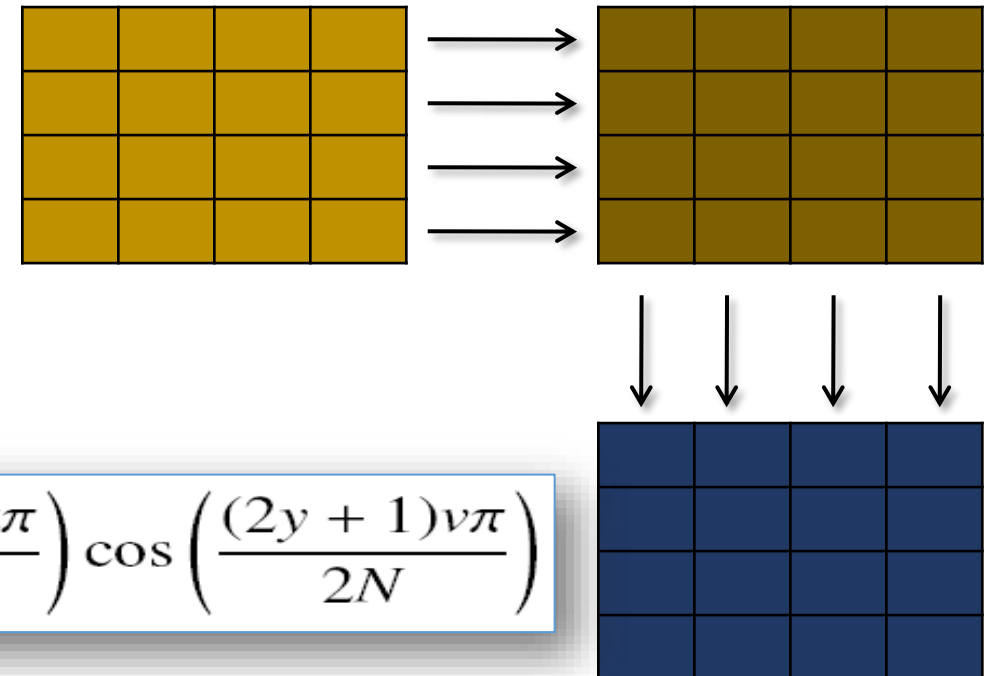


Basis patterns (imaged functions)

Ao discretizar as Funções Base obtem-se os Padrões Base

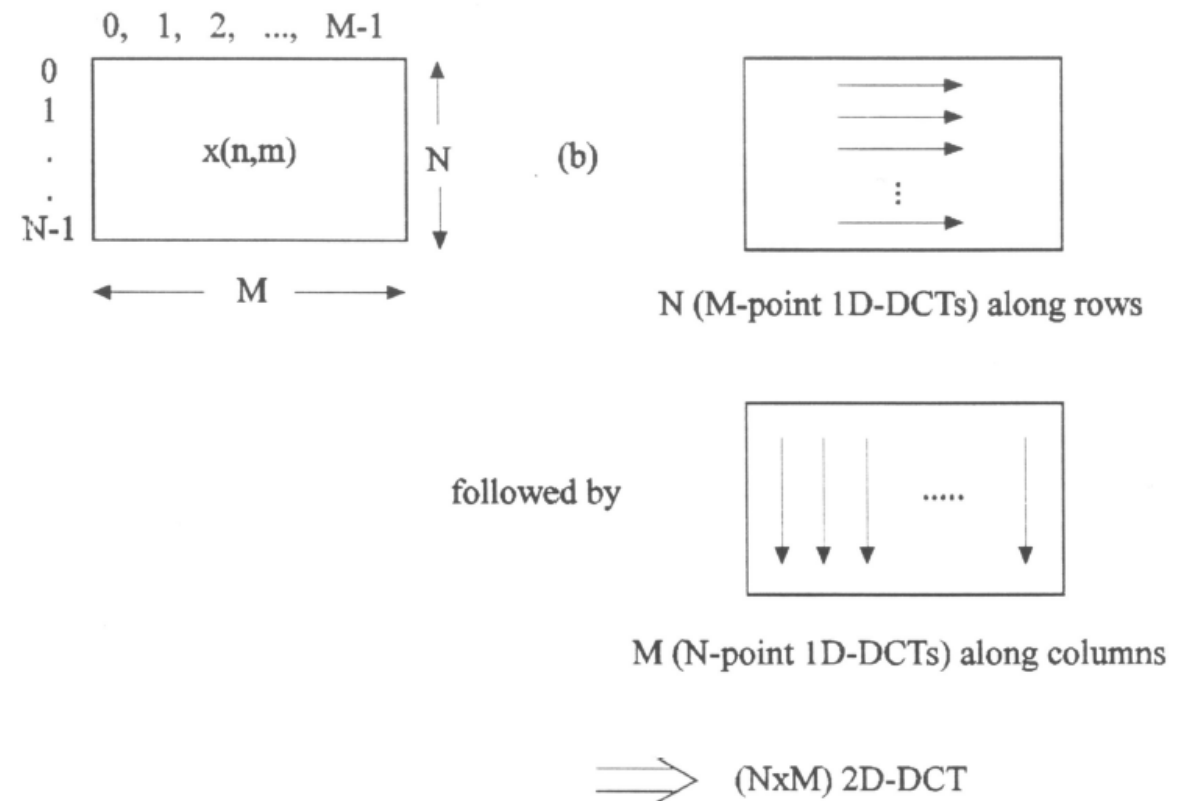
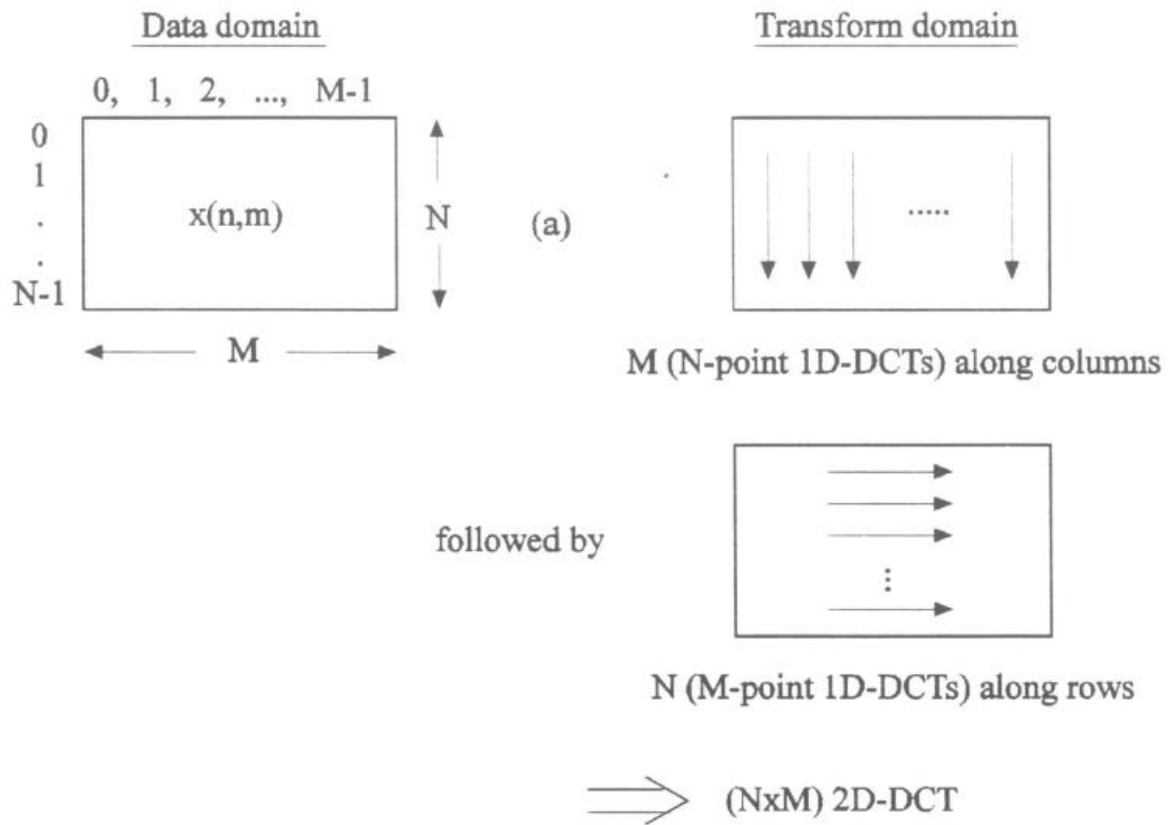
# Separabilidade

- A DCT tem propriedade de separabilidade
  - Os coeficientes podem ser obtidos calculando os coeficientes 1D para cada linha
  - Usando os coeficientes de linha para calcular os coeficientes de cada coluna (usando a transformada direta 1D)



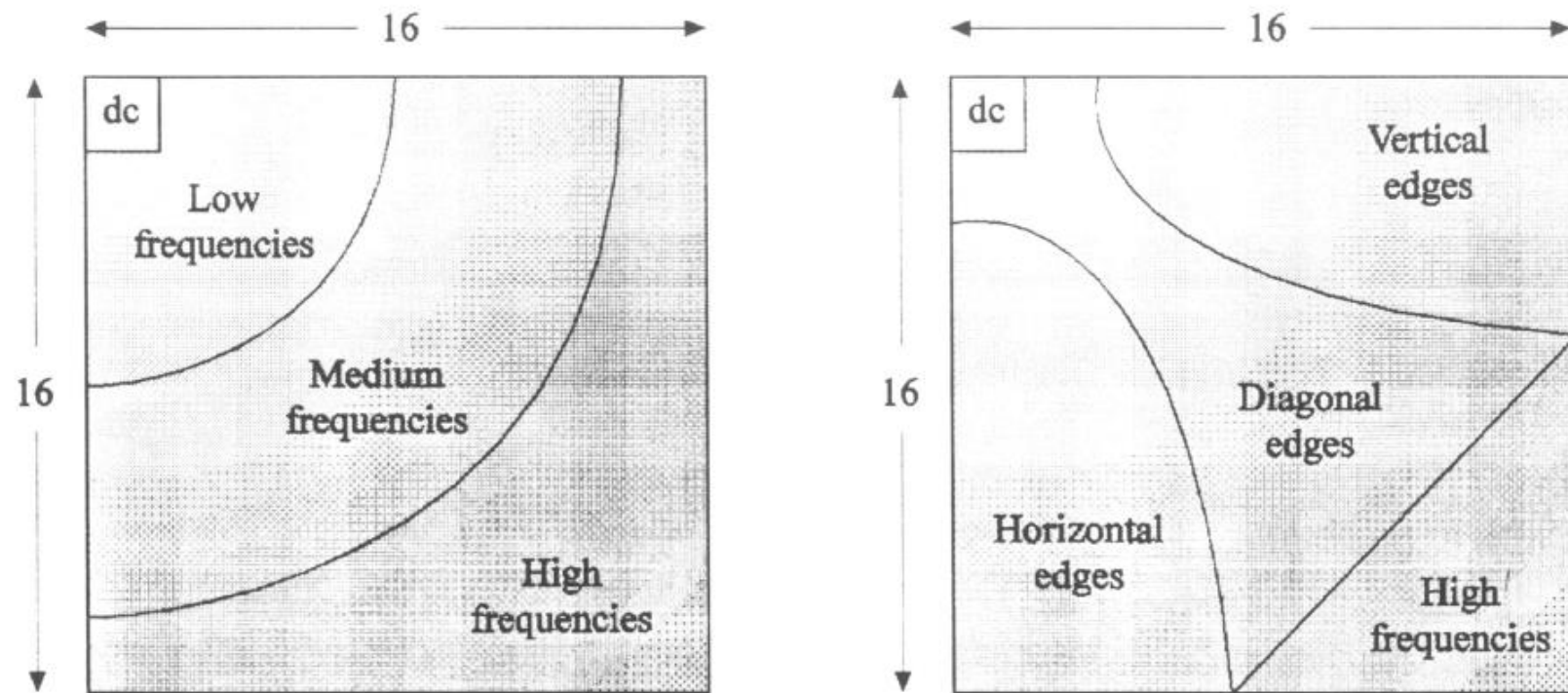
$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{(2x+1)u\pi}{2N}\right) \cos\left(\frac{(2y+1)v\pi}{2N}\right)$$

# Implementação da DCT-2D (NxM), por série de duas DCT-1D





## Características dos coeficientes da DCT



**Figure 5.12** Two-dimensional (16 × 16) DCT: (a) frequency distribution, (b) structural decomposition.

# Resumo sobre a DCT

- A DCT provê compactação de energia
  - Os coeficientes de baixa frequência têm maior magnitude (tipicamente)
  - Os coeficientes de alta frequência têm menor magnitude (tipicamente)
  - A maior parte das informações é compactada nos coeficientes de frequência mais baixa (os coeficientes na parte superior esquerda)
- A compactação pode ser alavancada para compressão
  - Utiliza os coeficientes DCT para armazenar dados de imagem, mas descarta uma certa porcentagem dos coeficientes de alta frequência!
  - JPEG faz isso

# Exercício 3. Coeficientes da DCT-2D

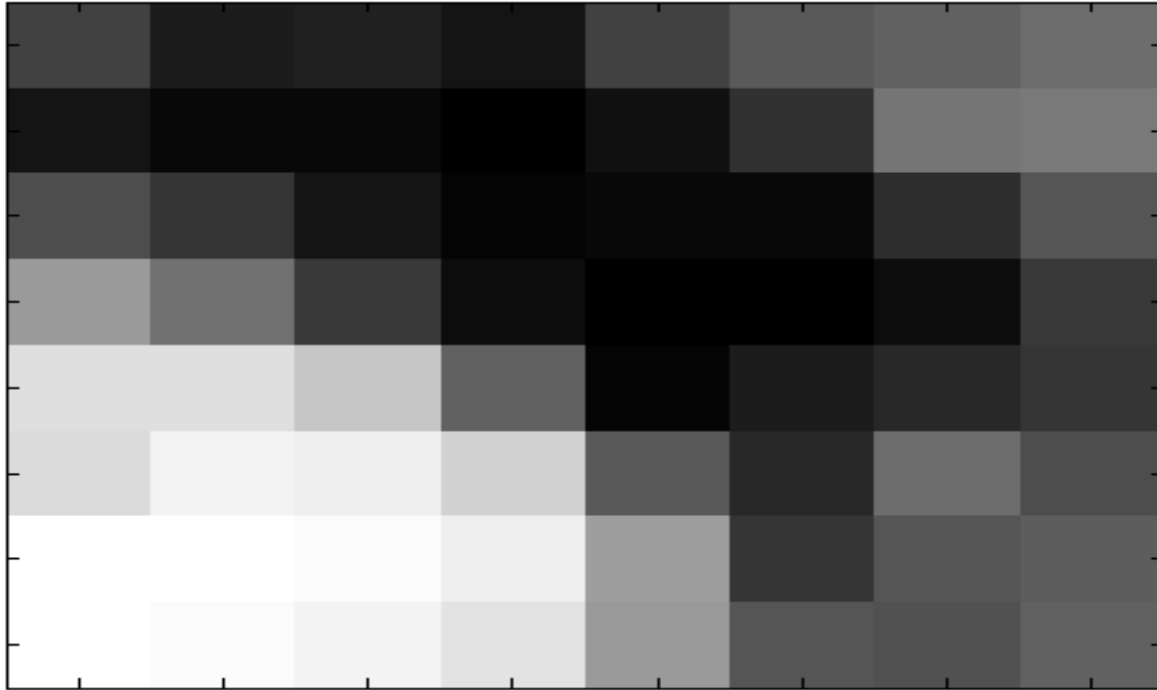
- Calcule os coeficientes  $F(u,v)$  e  $G(u,v)$  da DCT-2D (slide 32), para os blocos de imagem  $f(x,y)$  e  $g(x,y)$ :

$$a. \quad f(x,y) = \begin{bmatrix} 128 & -64 & -64 & 128 & 128 & -64 & -64 & 128 \\ 128 & -64 & -64 & 128 & 128 & -64 & -64 & 128 \\ 128 & -64 & -64 & 128 & 128 & -64 & -64 & 128 \\ 128 & -64 & -64 & 128 & 128 & -64 & -64 & 128 \\ 128 & -64 & -64 & 128 & 128 & -64 & -64 & 128 \\ 128 & -64 & -64 & 128 & 128 & -64 & -64 & 128 \\ 128 & -64 & -64 & 128 & 128 & -64 & -64 & 128 \\ 128 & -64 & -64 & 128 & 128 & -64 & -64 & 128 \end{bmatrix}$$

$$a. \quad g(x,y) = \begin{bmatrix} 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 \\ -32 & -32 & -32 & -32 & -32 & -32 & -32 & -32 \\ 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 \\ -32 & -32 & -32 & -32 & -32 & -32 & -32 & -32 \\ 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 \\ -32 & -32 & -32 & -32 & -32 & -32 & -32 & -32 \\ 32 & 32 & 32 & 32 & 32 & 32 & 32 & 32 \\ -32 & -32 & -32 & -32 & -32 & -32 & -32 & -32 \end{bmatrix}$$

# Compressão de Imagem

8 x 8 Pixels



Imagem

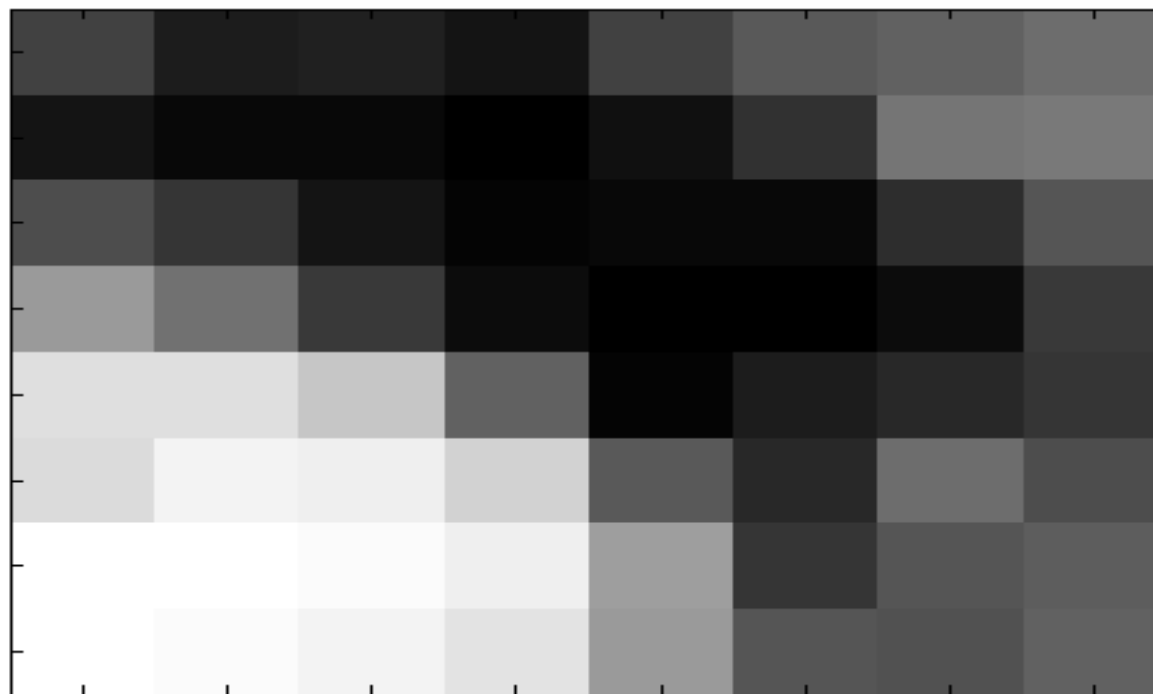


# Compressão de Imagem

- Exemplo Escala de Cinzas (ou Luminância)
- Faixa de Valores: 0 (*black*) --- 255 (*white*)

63	33	36	28	63	81	86	98
27	18	17	11	22	48	104	108
72	52	28	15	17	16	47	77
132	100	56	19	10	9	21	55
187	186	166	88	13	34	43	51
184	203	199	177	82	44	97	73
211	214	208	198	134	52	78	83
211	210	203	191	133	79	74	86

**X**



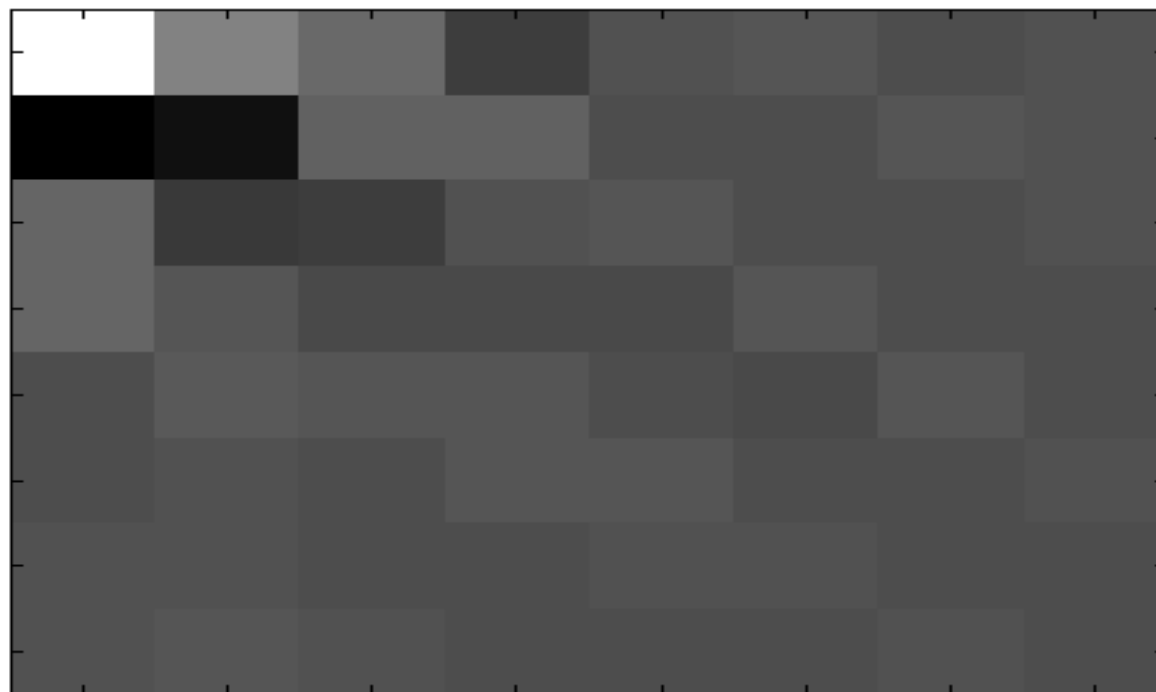
# Aplicando DCT-2D

- DCT-2D da matriz

Estes Números são os  
coeficientes do polinômio

-304	210	104	-69	10	20	-12	7
-327	-260	67	70	-10	-15	21	8
93	-84	-66	16	24	-2	-5	9
89	33	-19	-20	-26	21	-3	0
-9	42	18	27	-7	-17	29	-7
-5	15	-10	17	32	-15	-4	7
10	3	-12	-1	2	3	-2	-3
12	30	0	-3	-3	-6	12	-1

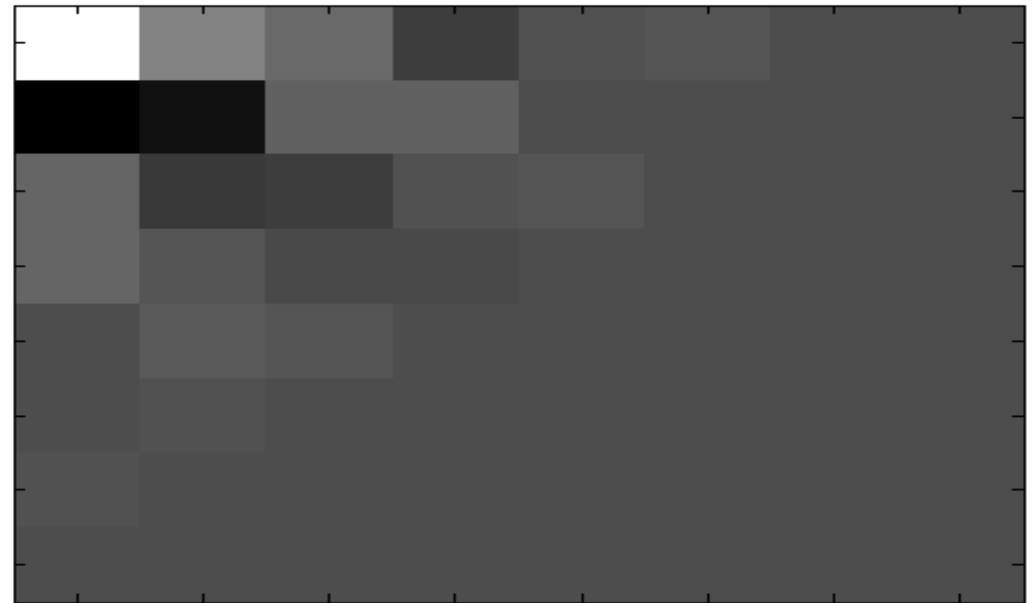
**Y**



# Compressão

- Elimina os componentes menos significativos

-304	210	104	-69	10	20	-12	0
-327	-260	67	70	-10	-15	0	0
93	-84	-66	16	24	0	0	0
89	33	-19	-20	0	0	0	0
-9	42	18	0	0	0	0	0
-5	15	0	0	0	0	0	0
10	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0



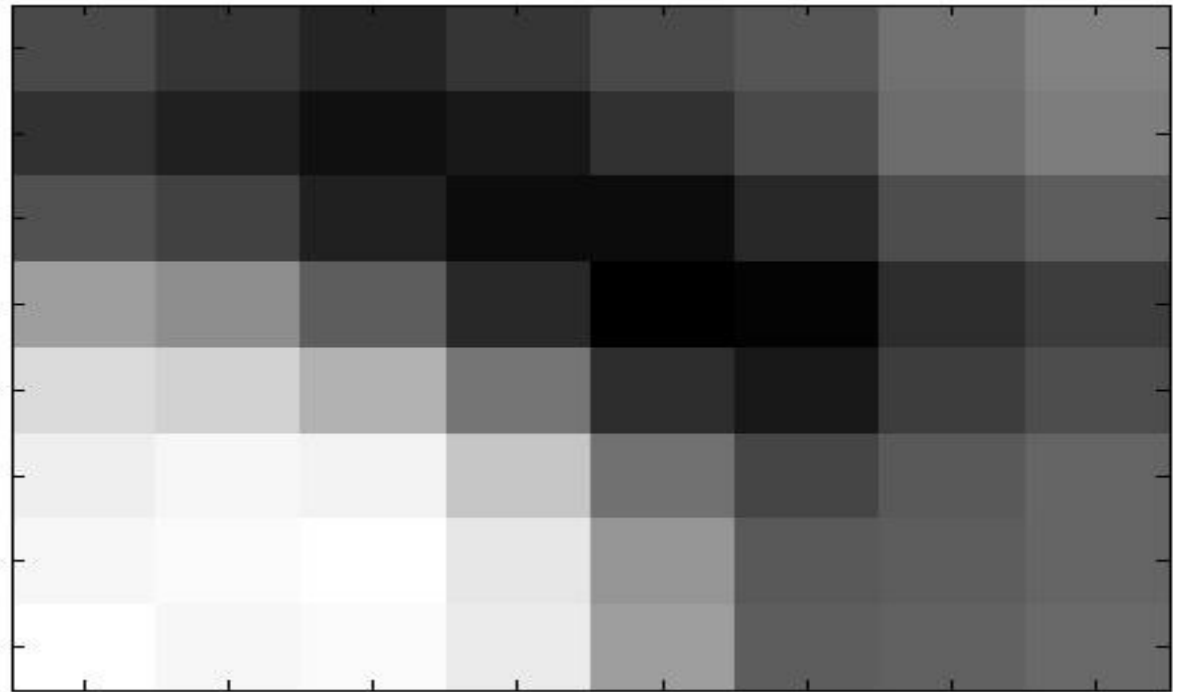
Note que preservamos um pouco mais da metade da memória original.



# Reconstruindo a Imagem

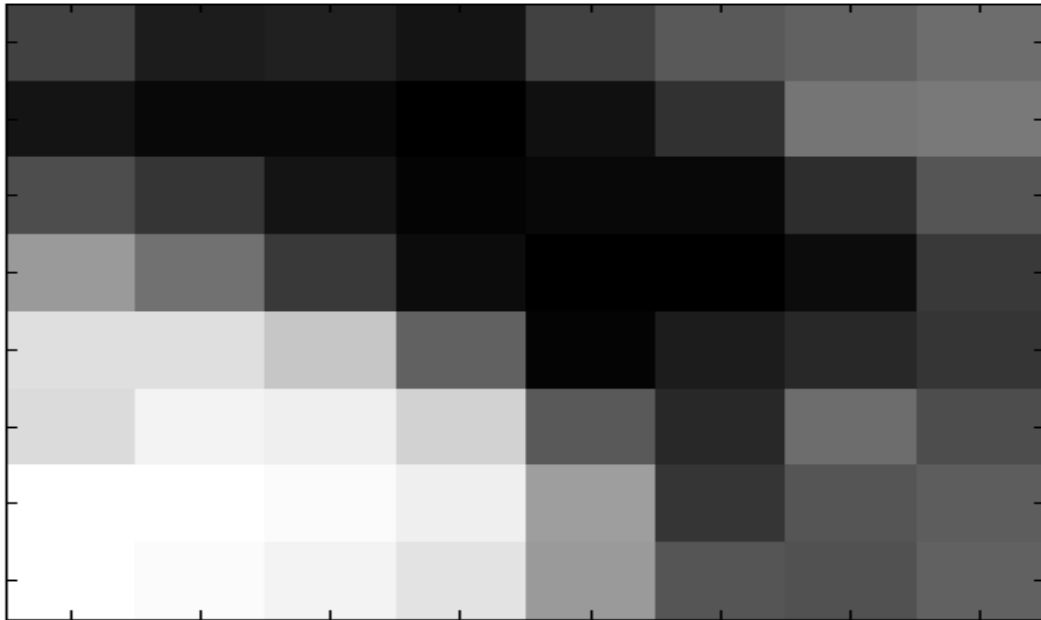
- Nova Matriz e a Imagem Reconstruída (com IDCT-2D)

55	41	27	39	56	69	92	106
35	22	7	16	35	59	88	101
65	49	21	5	6	28	62	73
130	114	75	28	-7	-1	33	46
180	175	148	95	33	16	45	59
200	206	203	165	92	55	71	82
205	207	214	193	121	70	75	83
214	205	209	196	129	75	78	85

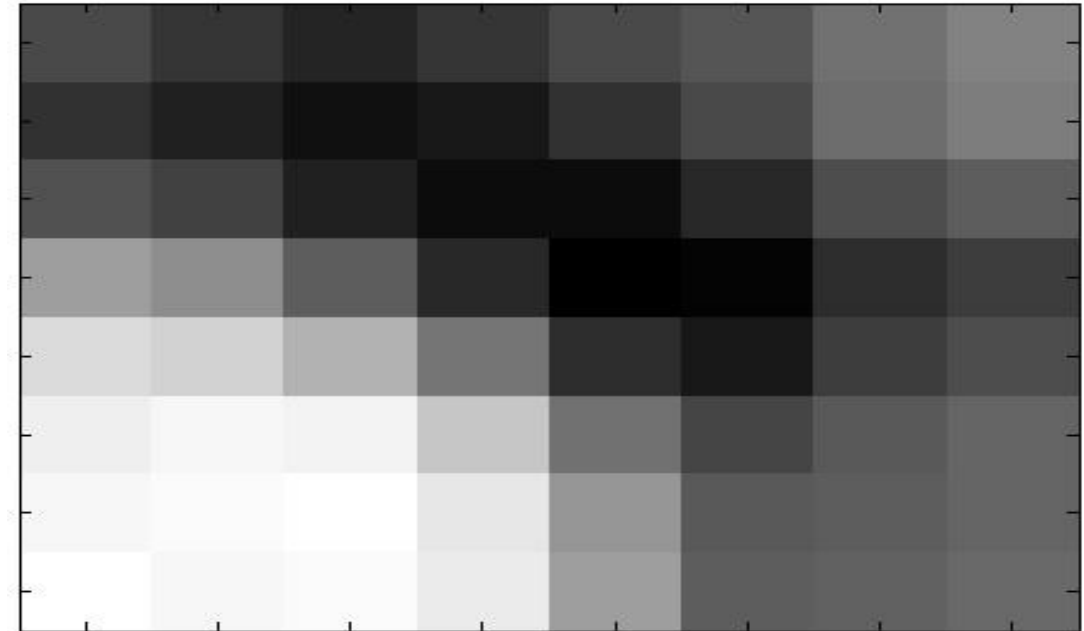


# Qual é a Diferença?

Original



Reconstruída (*após compressão*)



# Voce pode notar a Diferença?

Original

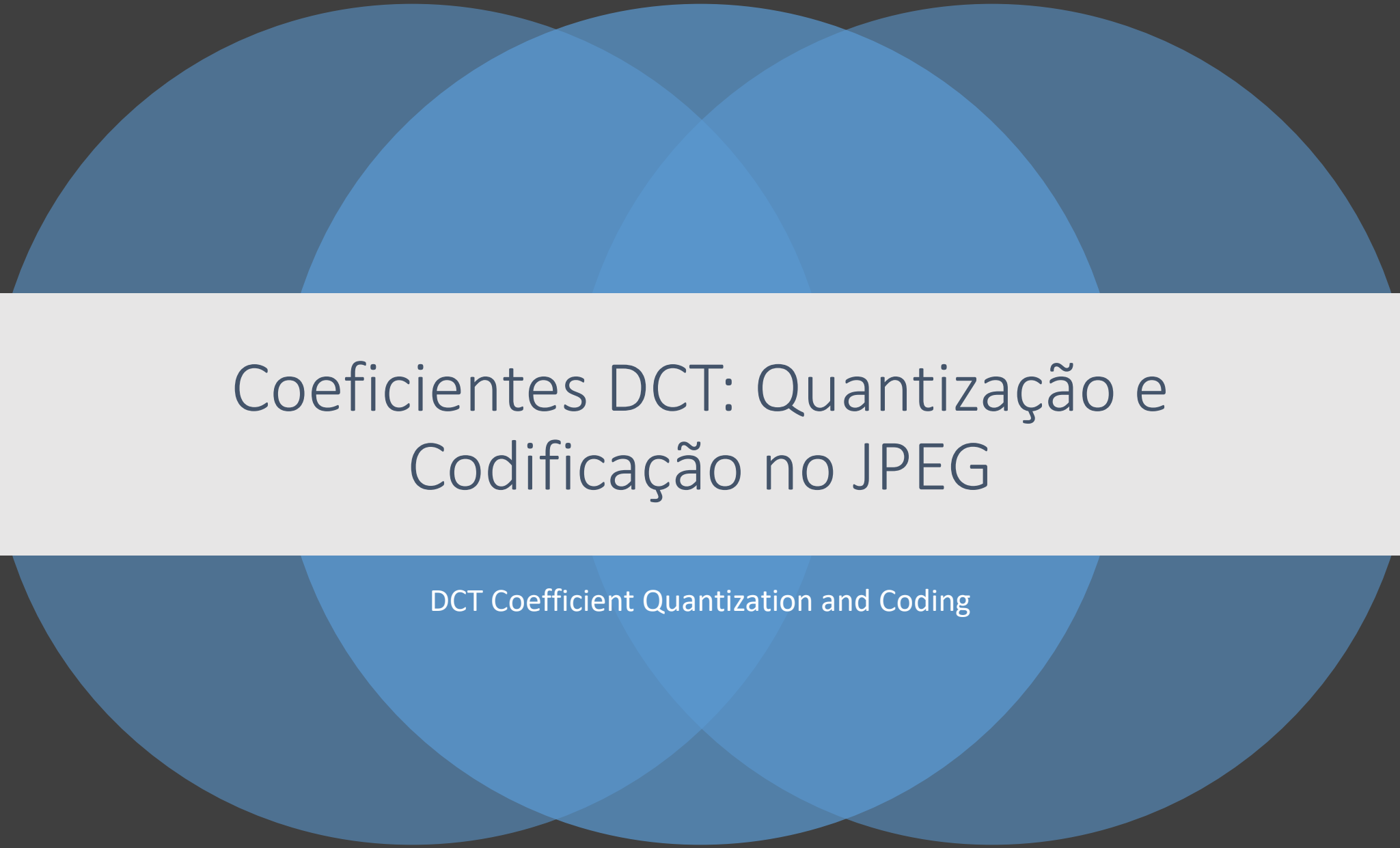


Reconstruída (*após compressão*)



# Resposta: Bronzeado sem Perigo





# Coeficientes DCT: Quantização e Codificação no JPEG

DCT Coefficient Quantization and Coding

# Quantização dos coeficientes da DCT (JPEG)

- Depois da DCT, cada um dos 64 coeficientes de  $F(u,v)$  é uniformemente quantizada utilizando uma tabela pré-definida de quantização  $Q(u,v)$  (também conhecida como matriz de normalização), segundo a expressão

$$F^Q(u, v) = \text{IntegerRound}\left(\frac{F(u, v)}{Q(u, v)}\right)$$

# JPEG: Quantização dos coeficientes da DCT

**Table 4.1** (a) The default quantization table for luminance ( $Y$  component); (b) the default quantization table for chrominance ( $C_b C_r$  components)

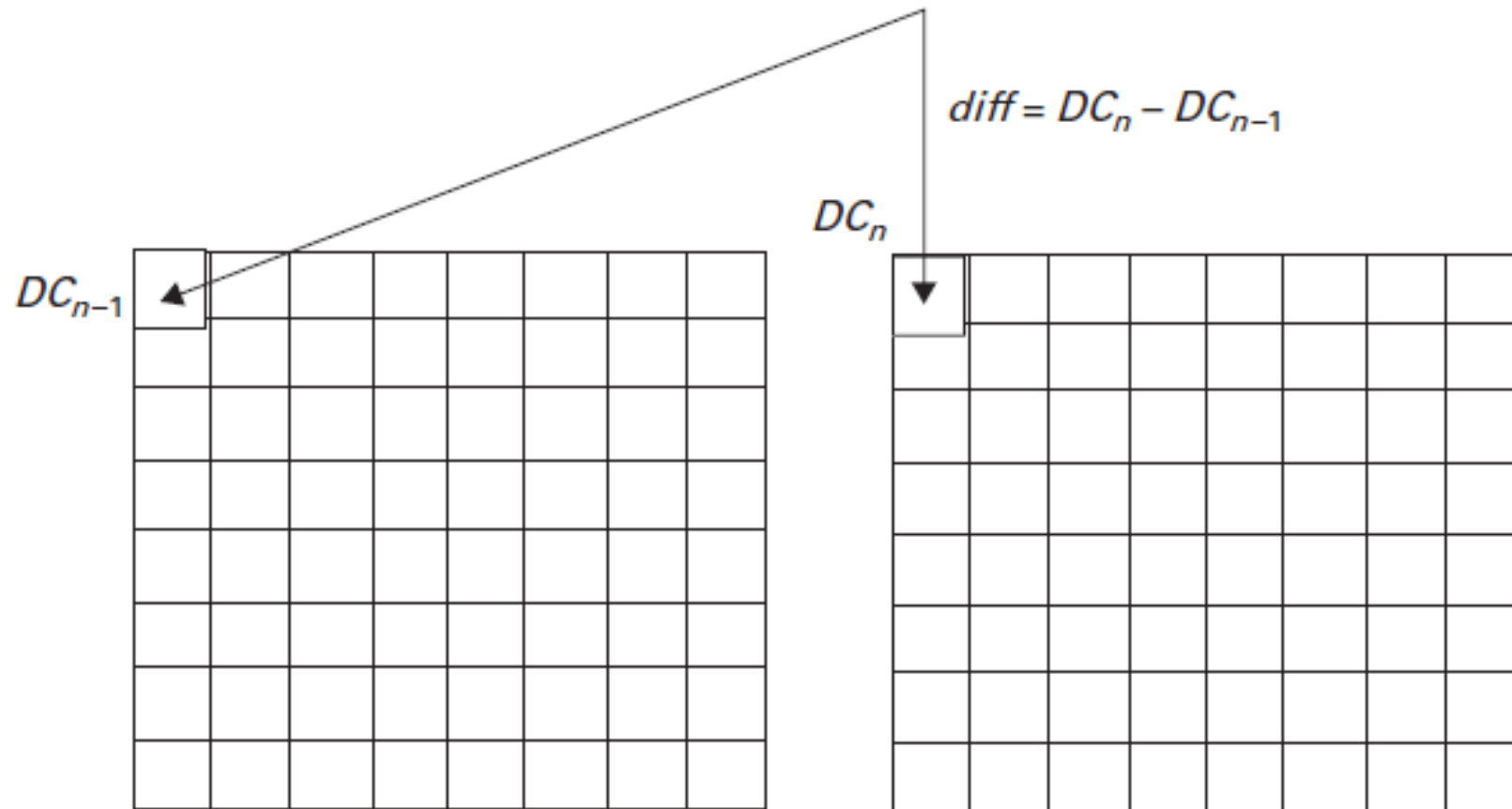
16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99
(a)							
17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
(b)							



# JPEG: Codificação DC e seqüência zigzag

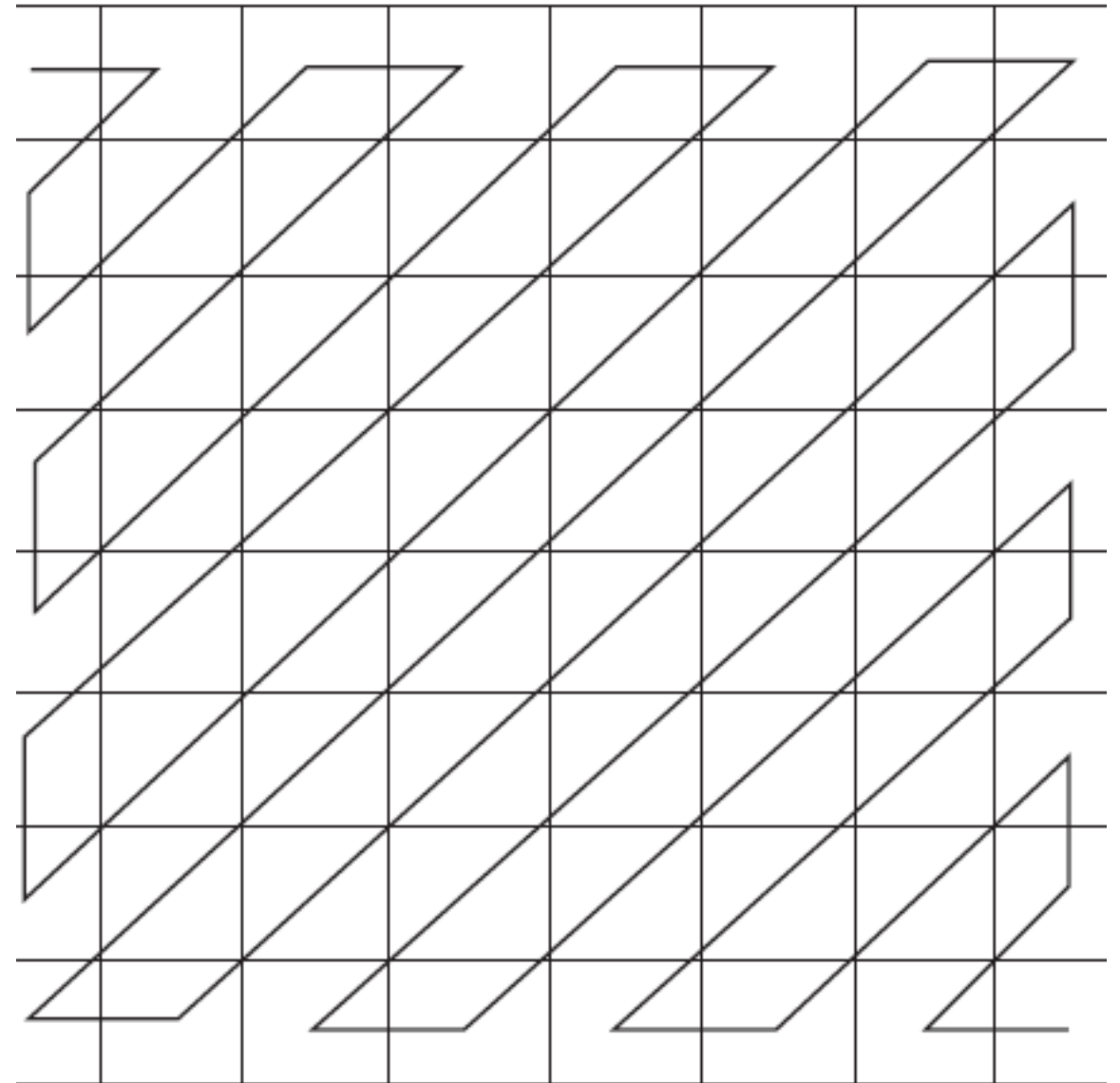
- Por possuir uma maior correlação entre os coeficientes DC dos blocos adjacentes, o coeficiente DC quantizado é codificado como a diferença do termo DC do bloco anterior, seguindo a ordem de codificação.

# JPEG: Codificação DC e seqüência zigzag



# JPEG: seqüência zigzag

- Por fim, todos os coeficientes são ordenados numa seqüência em *zigzag*, para facilitar a codificação por entropia.
- Como muitos das componentes AC são zero, podemos utilizar o esquema de codificação RLC.



# Exemplo JPEG: um bloco 8x8

## Coeficientes DCT calculados

**Table 4.2** (a) The FDCT coefficients  $\{F(u,v)\}$ ; (b) the quantized FDCT coefficients  $\{F^Q(u,v)\}$  derived from Eq. (4.13) [14]

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

(a)

15	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	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

(b)

## Fatores $Q(u,v)$

**Table 4.1** (a) The default quantization table for luminance (Y component); (b) the default quantization table for chrominance ( $C_b, C_r$  components)

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

(a)

$\{3, 0, -2, -1, -1, -1, 0, 0, -1, EOB\}$

Coeficientes DCT Quantizados após zig-zag

## Coeficientes DCT Quantizados

## JPEG: Codificação por entropia

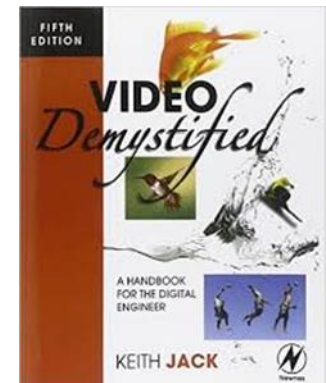
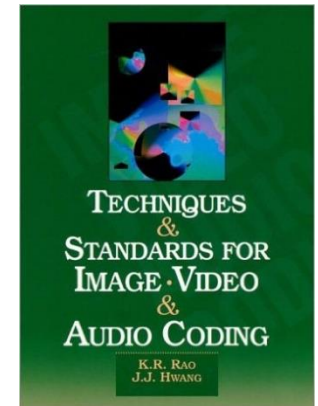
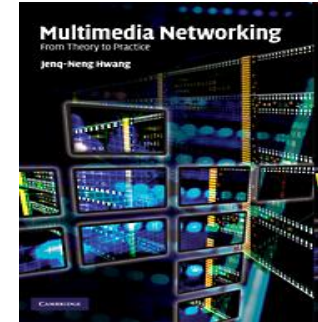
- A codificação por entropia é usada para obter uma compressão sem perdas adicional, pela codificação da seqüência multi-símbolo RLC dos coeficientes do DCT quantizado.

## Comentários sobre o JPEG-2000

- JPEG2000 é um padrão de compressão de imagens baseado em wavelet
- O objetivo é superar o desempenho do padrão JPEG, baseado em DCT
- Em altas taxas de bits: JPEG2000 possui compressão 20% maior que JPEG, em média.
- A baixas taxa de bits, JPEG2000 possui muito mais vantagens sobre o JPEG.
- O ganho de compressão sobre o JPEG é atribuído ao uso da *discrete wavelet transform* (DWT) e esquemas de codificação por entropia mais sofisticados.

# Bibliografia

- HWANG, Jenq-Neng. **Multimedia Networking: From Theory to Practice**, Cambridge University Press, 2009.
- K. R. Rao, J.J. Hwang. **Techniques and Standards for Image, Video, and Audio Coding**, Prentice Hall, 1996.
- K. JACK, **Video Demystified**. 5. ed. Newnes, 2007.



# Referencias

- Cambridge in Color. “UNDERSTANDING GAMMA CORRECTION”.  
<<http://www.cambridgeincolour.com/tutorials/gamma-correction.htm>>.
- Ze-Nian Li, Mark S. Drew, Jiangchuan Liu, **Fundamentals of Multimedia**, 2nd Ed, Springer, USA, 2014.