

Capítulo 3

VISÃO COMPUTACIONAL

Capítulo 3

3.1. Principais Etapas de um Sistema de Visão Computacional

3.2. Visão Humana X Computacional

3.3. A Imagem Digital

3.4. Histograma de imagem digital

3.5. Sistemas de visão binária

3.6. Operações pontuais globais em imagens

3.1. Principais Etapas de um Sistema de Visão Computacional

3.1.1 Aquisição de Imagens

3.1.2 Restauração e Realce

3.1.3. Segmentação

3.1.4. Extração de Atributos ou Características

3.1.5. Classificação e Reconhecimento

3.1.6. Decisão

3.1. Principais Etapas de um Sistema de Visão Computacional

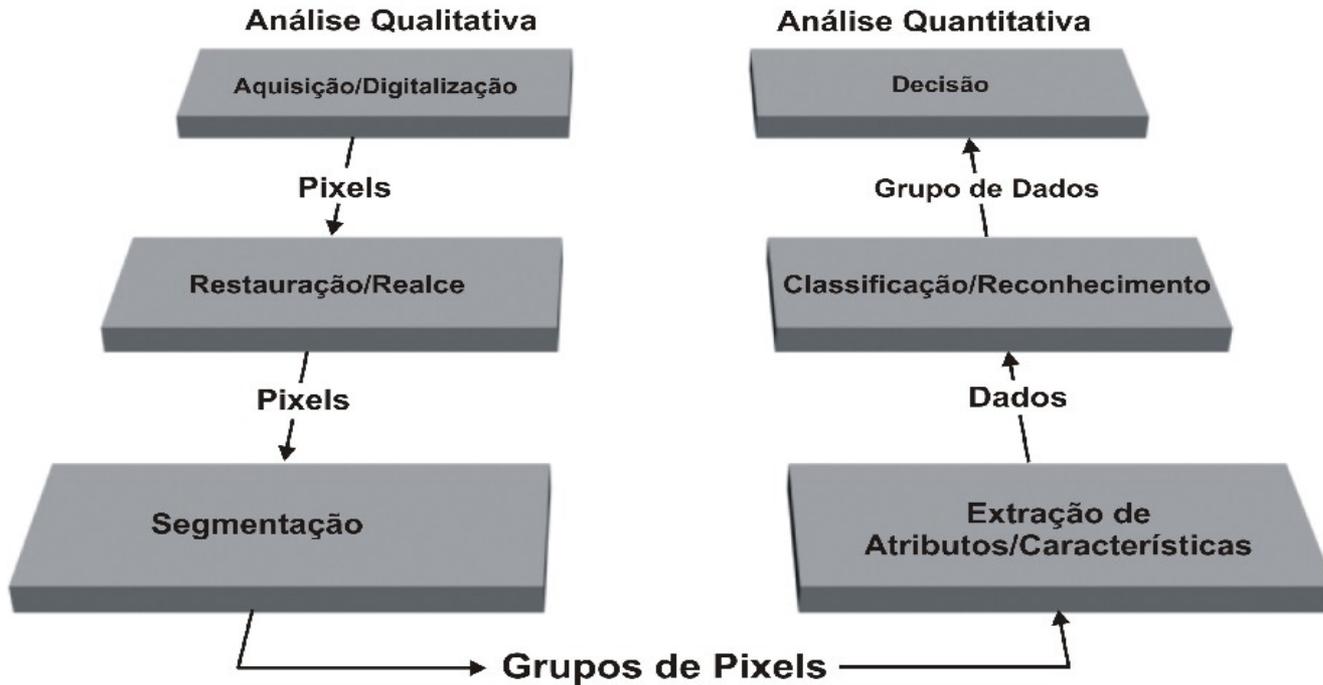
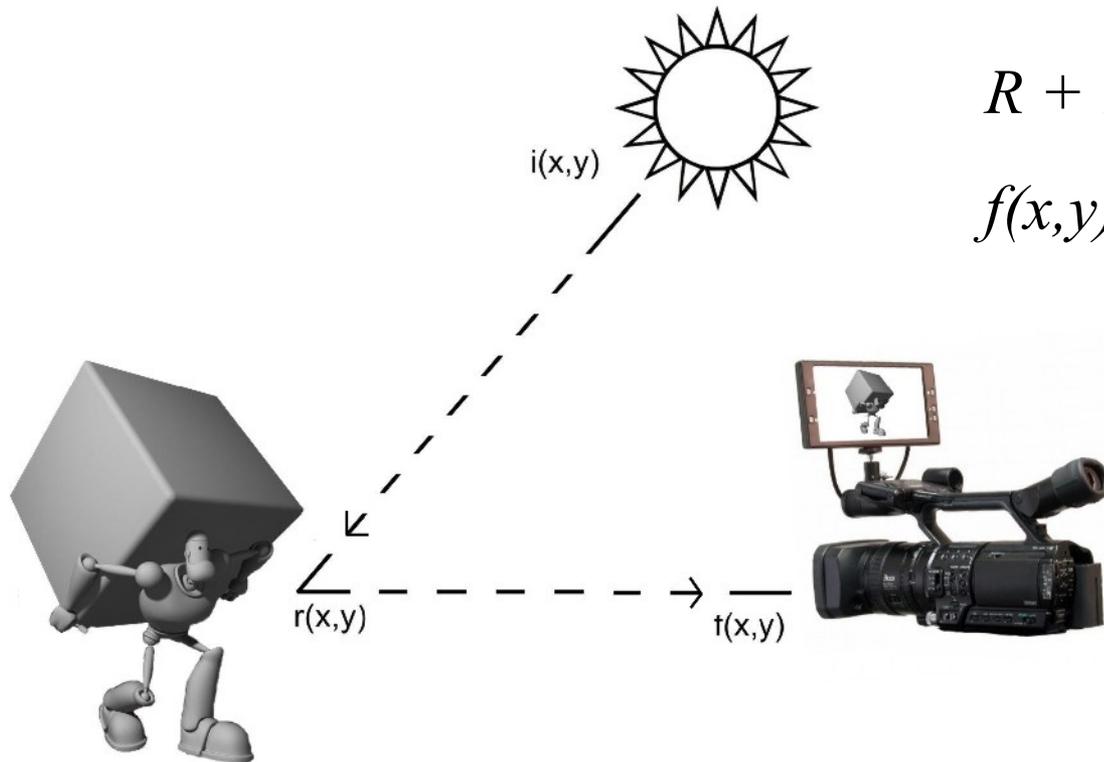


Figura 3.1 – Etapas de um sistema de VC genérico.

3.1.1 Aquisição de Imagens



$$R + T + A = 1 \quad (3.1)$$

$$f(x,y) = i(x,y) \cdot r(x,y) \quad (3.2)$$

Figura 3.2 – Aquisição da Imagem.

Amostragem e Quantização

- Amostragem – refere-se ao número de pontos amostrados de uma imagem digitalizada (resolução).
- Quantização - quantidade de níveis de tons que pode ser atribuído a cada ponto digitalizado.

As imagens reais possuem um número ilimitado de cores ou tons. No processamento de imagens computacional é necessário limitar os níveis de cores ou tons possíveis de serem atribuídos a cada *pixel* da imagem (gradação tonal).

Resolução



256 x 160



128 x 80



64 x 40

Figura 3.3 – Efeito da redução da resolução espacial na qualidade da imagem

Quantização



256 níveis de cinza



16 níveis de cinza

Figura 3.4 – Efeito da variação da gradação tonal ou quantização na qualidade da imagem.

3.1.2 Restauração e Realce

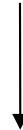
Objetivo



- Restauração - busca compensar deficiências específicas, geradas no momento de aquisição, na transmissão ou em alguma etapa do processamento.
- Realce - destaca detalhes da imagem que são de interesse para análise ou que tenham sofrido alguma deteriorização.

3.1.3. Segmentação

Objetivo



Isolar regiões de pontos da imagem pertencentes a objetos para posterior extração de atributos e cálculo de parâmetros descritivos.

3.1.4. Extração de Atributos ou Características

Objetivo



A partir de imagens já segmentadas (em objeto e fundo) ou binárias busca obter dados relevantes ou atributos, das regiões ou objetos destacados.

Os tipos de atributos ou características mais comuns são: número total de objetos; dimensões; geometria; propriedades de luminosidade e textura.

3.1.5. Classificação e Reconhecimento

Distinguir objetos na imagem agrupando parâmetros de acordo com sua semelhança para cada região de *pixels* encontrada.

Os objetos são:

reconhecidos como pertencentes a um mesmo grupo e então sejam classificados em uma base de imagens.

apresentados para o sistema, que compara suas características com aquelas dos objetos das classes previamente estabelecidas.

3.1.6. Decisão

- O objetivo de um sistema de Visão Computacional é tomar decisões a partir da extração de informações do mundo real através de imagens.
- A tomada de decisão pode ser feita a partir de indagações simples a respeito de parâmetros extraídos dos objetos ou de algoritmos mais complexos de Inteligência Artificial.

3.2. Visão Humana X Computacional

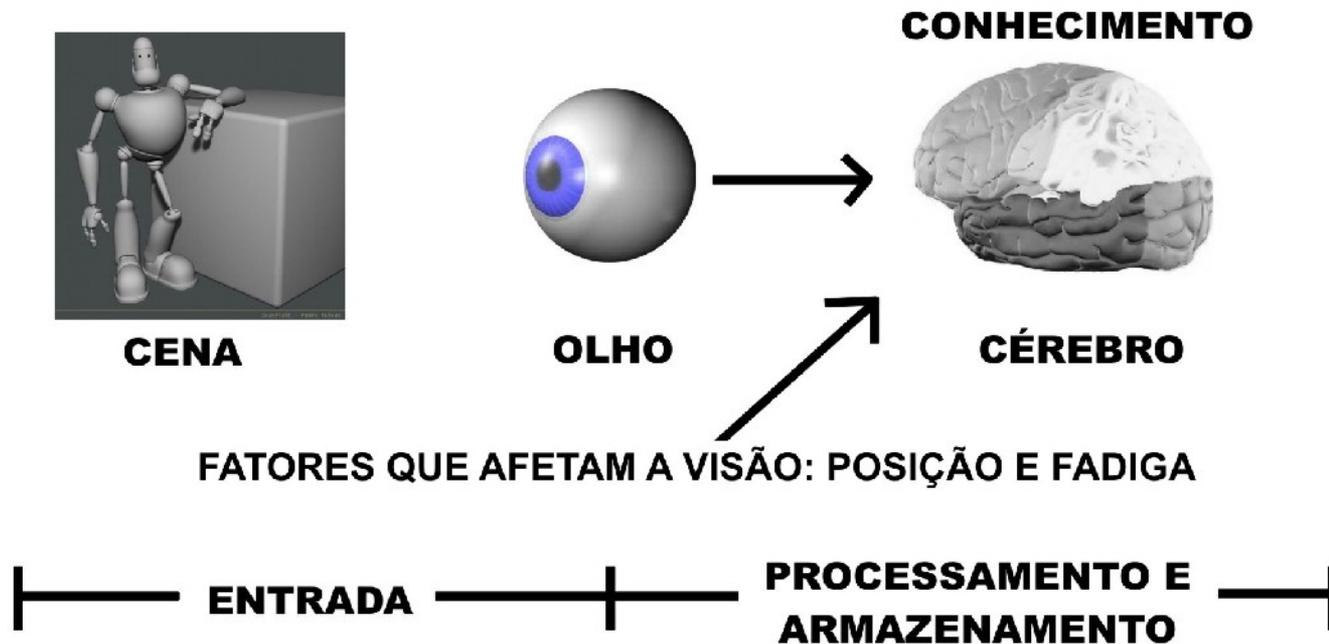


Figura 3.7. Sistema de Visão Humano.

Características

- Adaptabilidade
- Tomada de Decisão
- Qualidade das Medições
- Velocidade de Resposta
- Percepção de Espectros
- Dimensão dos Objetos

Exemplo de sistema de Visão Computacional

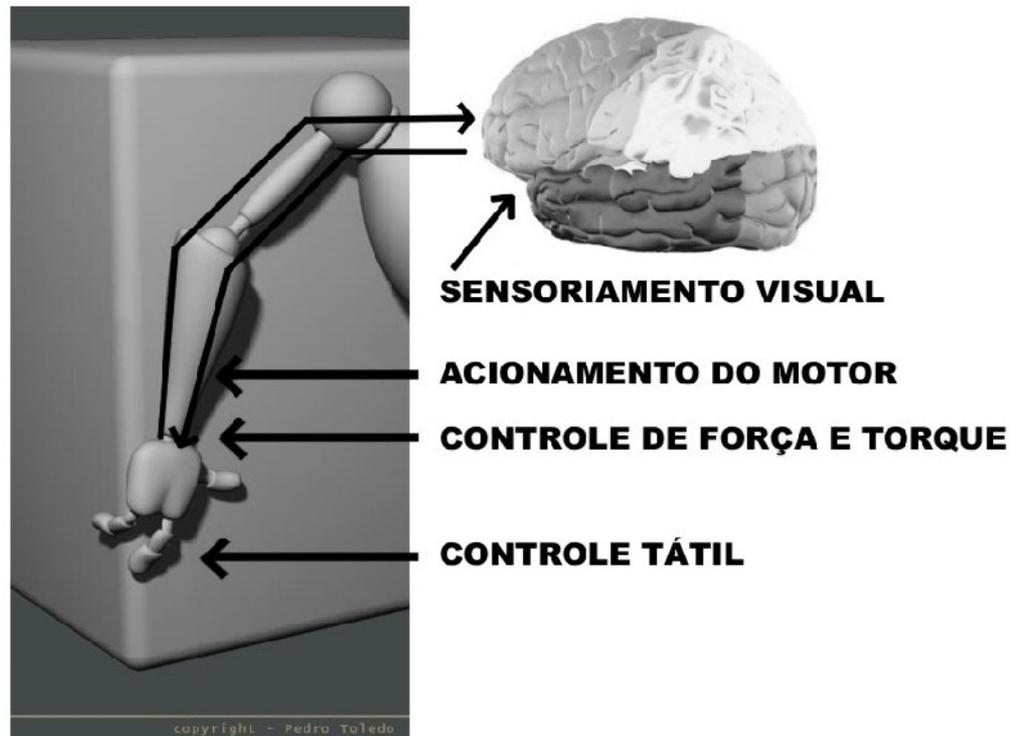


Figura 3.8. Integração da visão com o sistema motor.

Analogia entre a visão humana e a computacional.



Figura 3.9. Analogia do sistema de visão humano e computacional.

3.3. A Imagem Digital

3.3.1. Discretização e Reconstrução

3.3.2. Amostragem e Quantificação

3.3.3. Resolução Espacial

3.3.4. Aliasing

3.3.5. Imagens Monocromáticas

3.3.6. Imagens Coloridas

3.3. A Imagem Digital

- Do latim *imago* - representação visual de um objeto.
- Do grego *eidos* (raiz etimológica do termo *idea* ou *eide*) – considera a *idea* da coisa, a sua imagem, uma projeção na mente.
- Imagem pode ser adquirida ou gerada pelo ser humano.
- Uma imagem consiste em qualquer forma visual de expressão de uma idéia.

Formas de captação de uma imagem

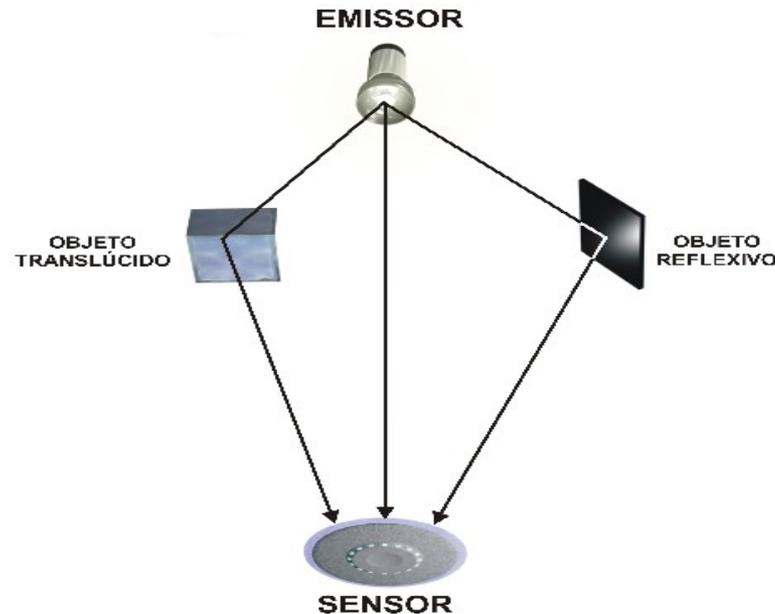


Figura 3.10 - Formas de Captação da Imagem por Radiação

Etapas do processamento de imagem digital

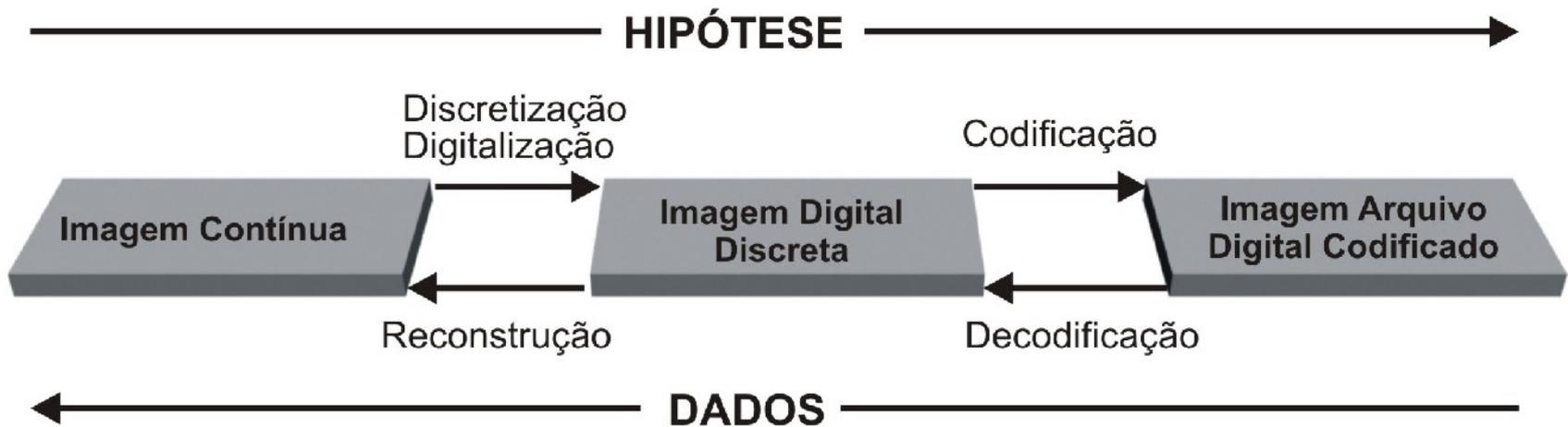


Figura 3.11 - Etapas do Processamento da Imagem

Discretização - conversão da imagem na forma contínua em uma representação discreta.

Reconstrução - processo inverso da discretização.

Codificação - a partir da representação discreta da imagem, gera um conjunto de dados representativos da imagem, dados estes que podem ser transformados no formato de arquivos.

Decodificação - processo oposto à codificação no qual acessam-se informações codificadas na forma de uma representação discreta.

3.3.1. Discretização e Reconstrução

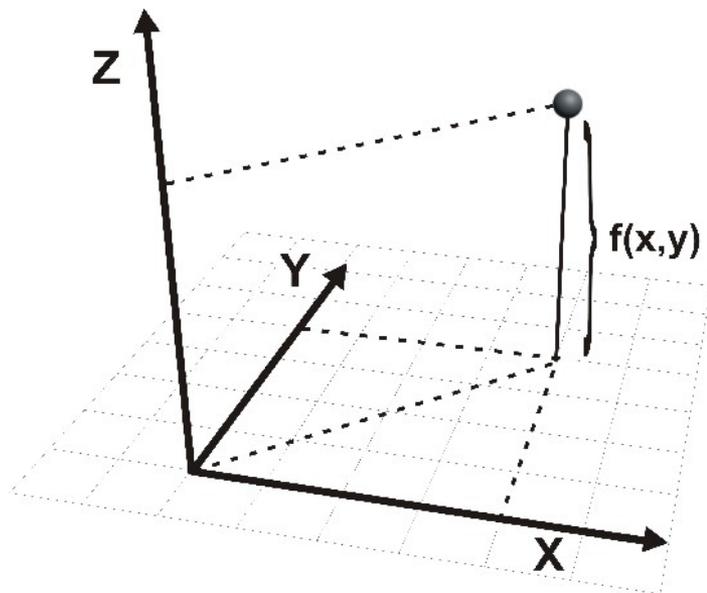
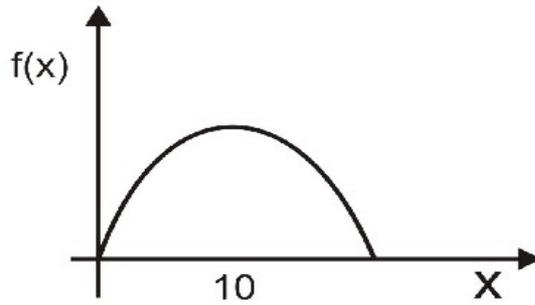


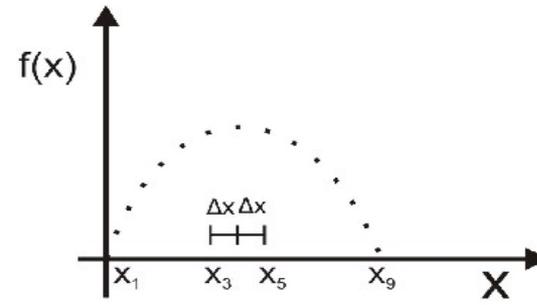
Figura 3.12. Gráfico de uma imagem contínua.

A forma de representar o mundo contínuo ou uma função contínua no computador é discretizando-a.

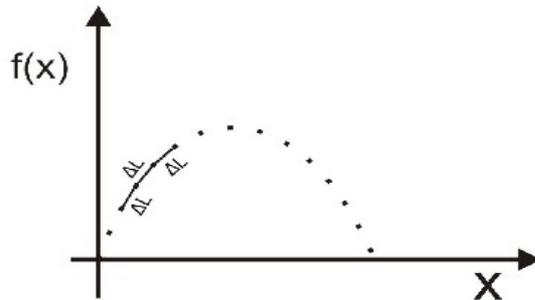
A operação que a partir dos valores discretos retorna uma aproximação da função contínua inicial é chamada de reconstrução.



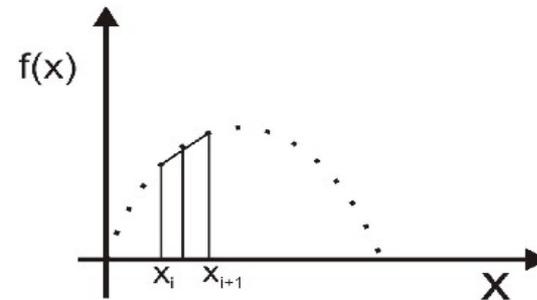
Representação Contínua



Representação Discreta a Intervalos Constantes em x



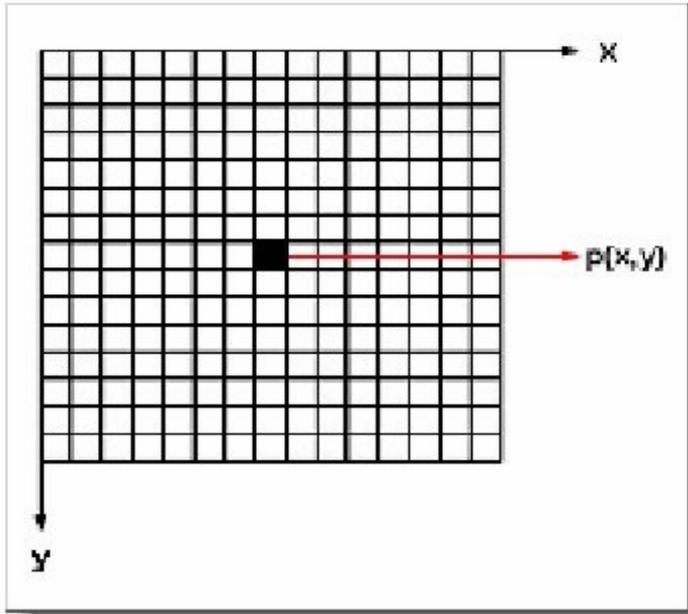
Amostragem a Intervalos Constantes em Relação a Distância entre os pontos $(x, f(x))$



Reamostragem Criando Valores Intermediários por Interpolação Linear por Pares de Valores Anteriores

Figura 3.13 – Formas de Representação de uma função $y = f(x)$.

3.3.2. Amostragem e Quantificação



Uma imagem digital é descrita por uma matriz $N \times M$ de valores de *pixel* ($p(x,y)$) inteiros positivos, que indica a intensidade de cor em cada posição $[x,y]$ da imagem.

Figura 3.15 – Reticulado uniforme da representação matricial da imagem.

Um *pixel* é caracterizado pelo valor de tonalidade de cor e pela sua localização na imagem.



47	52	64	132	153
51	58	121	149	142
49	99	143	144	164
94	135	161	170	199
138	165	180	212	213

Figura 3.16 – Representação matricial de uma região da imagem.

3.3.3. Resolução Espacial

Ao ser digitalizada a imagem assume um tamanho adimensional, em *pixels*.

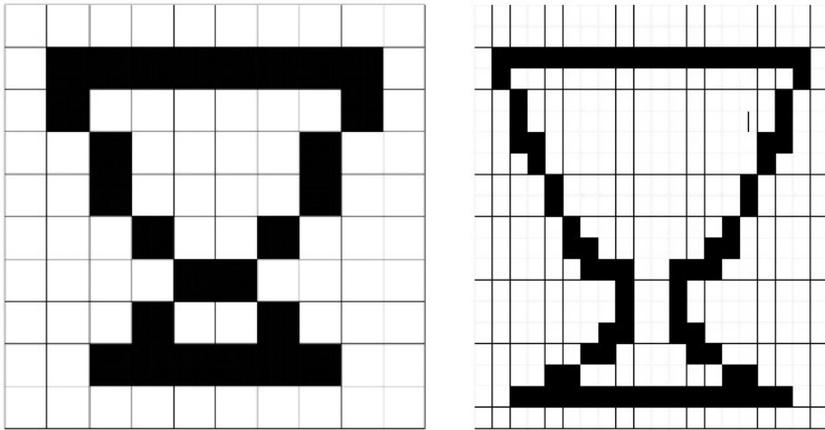


Figura 3.17 – Taça em duas resoluções, mas exibido no seu tamanho original.

Pode-se conhecer o tamanho da amostragem, conhecendo-se a razão entre o número de *pixels* obtido e o tamanho da imagem real no filme fotográfico ou equivalente.

A isso chama-se de resolução espacial, que em geral é medida em pontos por polegada ou dpi (*dots per inch*).

Imagens reais - > Digitais

Para que sejam representadas no meio digital, seu comportamento analógico (contínuo) tem que ser convertido numa série de valores discretos (descontínuos).

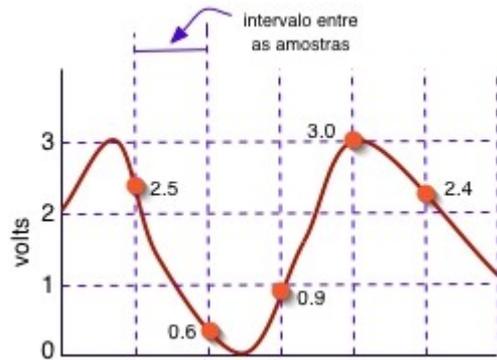
Esses valores são números (dígitos) que representam amostras (samples em inglês)

Amostragem

A conversão do sinal analógico para o digital é realizada por uma sequência de amostras da variação de voltagem do sinal original.

Cada amostra é arredondada para o número mais próximo da escala usada e depois convertida em um número digital binário (formado por "uns" e "zeros") para ser armazenado.

As amostras são medidas em intervalos fixos.



valores das amostras				
2.5	0.6	0.9	3.0	2.4
valores quantizados				
2	0	1	3	2
valores convertidos em dígitos binários				
10	00	01	11	10

O número de vezes em que se realiza a amostragem em uma unidade de tempo é a taxa de amostragem

3.3.4. Aliasing

Ocorre quando a frequência de amostragem é inferior à frequência de Nyquist.

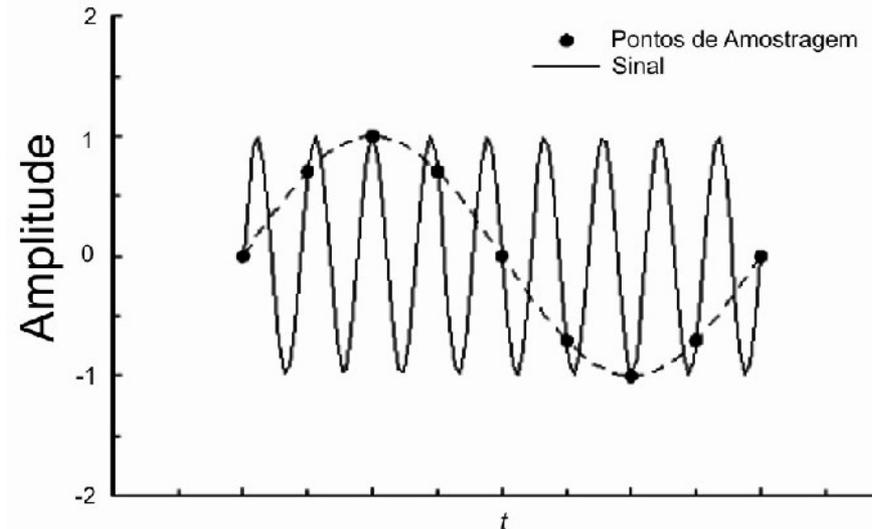


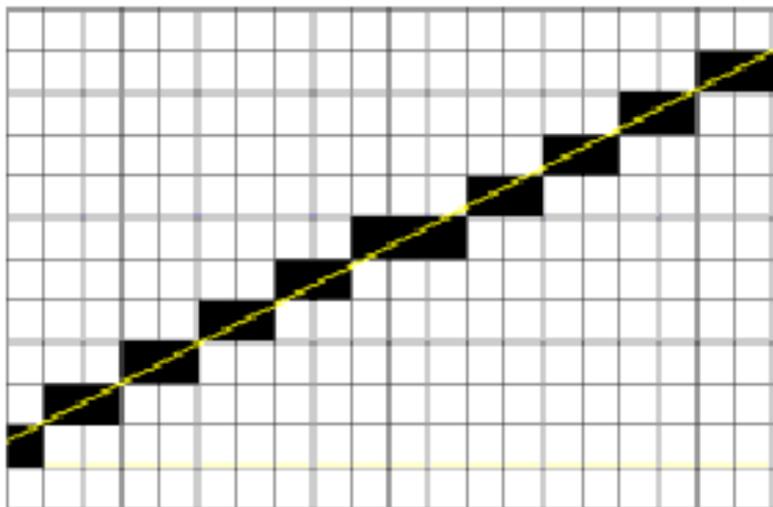
Figura 3.19 – O sinal digitalizado fica completamente diferente do sinal original devido a sua baixa frequência de amostragem.

Teorema de Nyquist

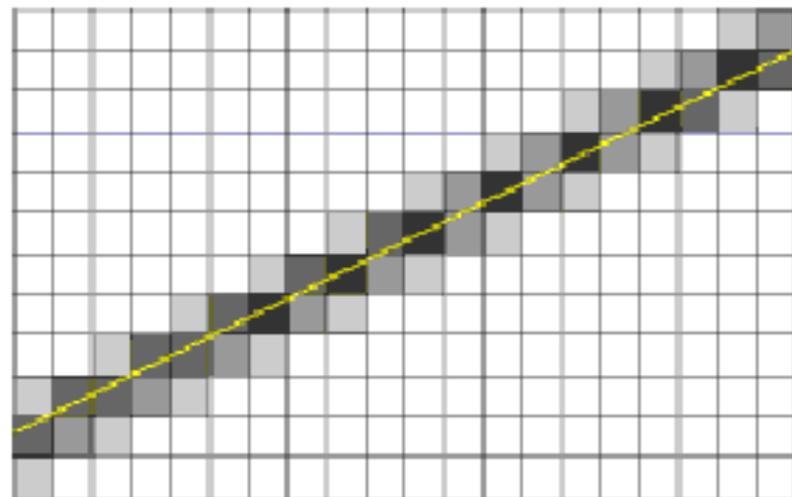
A taxa de amostragem deve ser pelo menos duas vezes a maior frequência que se deseja registrar.

Esse valor é conhecido como frequência de Nyquist.

Ao se tentar reproduzir uma frequência maior do que a frequência de Nyquist ocorre o fenômeno de aliasing (ou foldover)



(a) “Dentes” na representação de retas.



(b) Uma forma de amenizar o problema.

Figura 3.20 - Problema do *aliasing* em monitores de computadores antigos.

Moiré

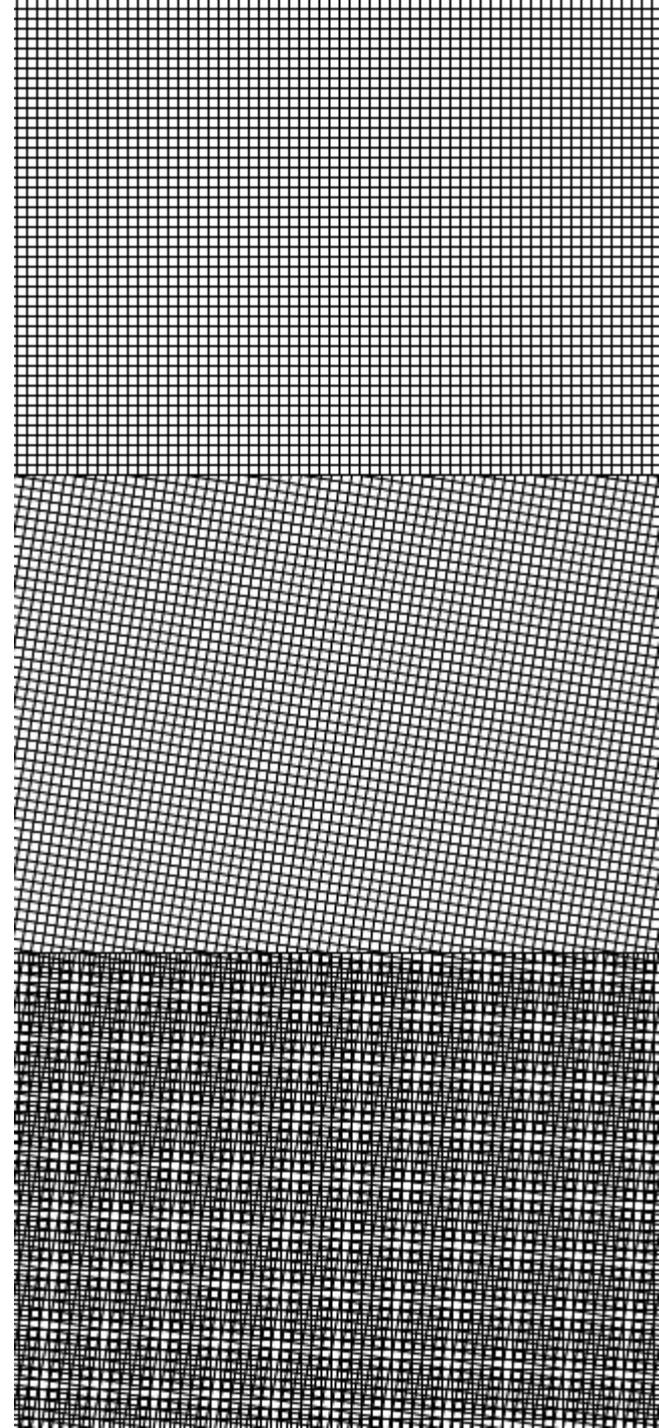
Efeito de segunda ordem:

Duas resoluções diferentes
(escaneado em 200dpi e
impresso em 300 dpi, por
exemplo); ou

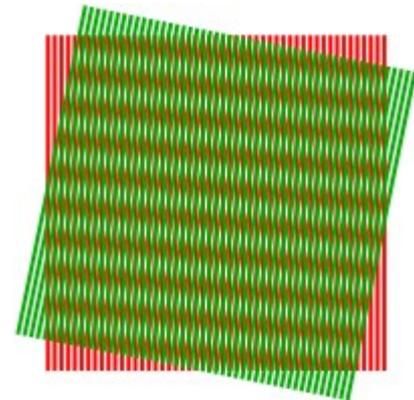
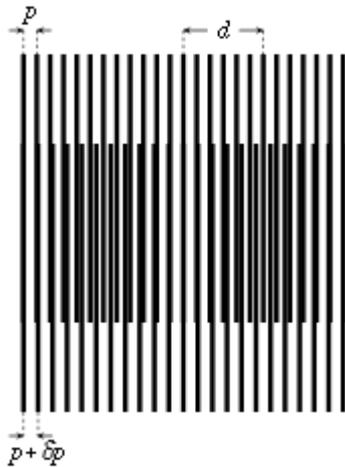
Grid não alinhados (rotacoes
ou aspect ratio)

(site do curso:

<http://www.ic.uff.br/~aconci/moire.html>)



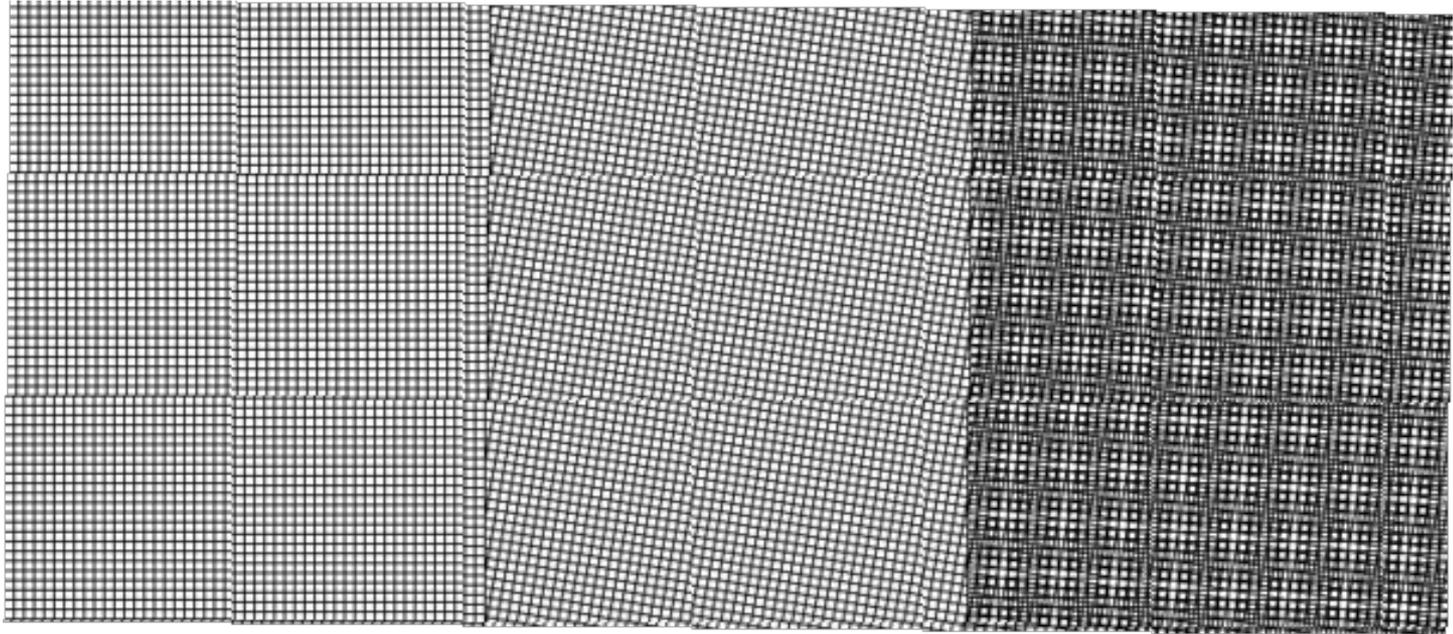
Exemplificando a causa dos padrões



Moiré

é um padrão de interferência criado quando duas grades são sobrepostas em um determinado ângulo, ou quando elas têm pequenas diferenças nos tamanhos da malhas da grade.

padrão moiré



Nem sempre maior resolução é melhor:

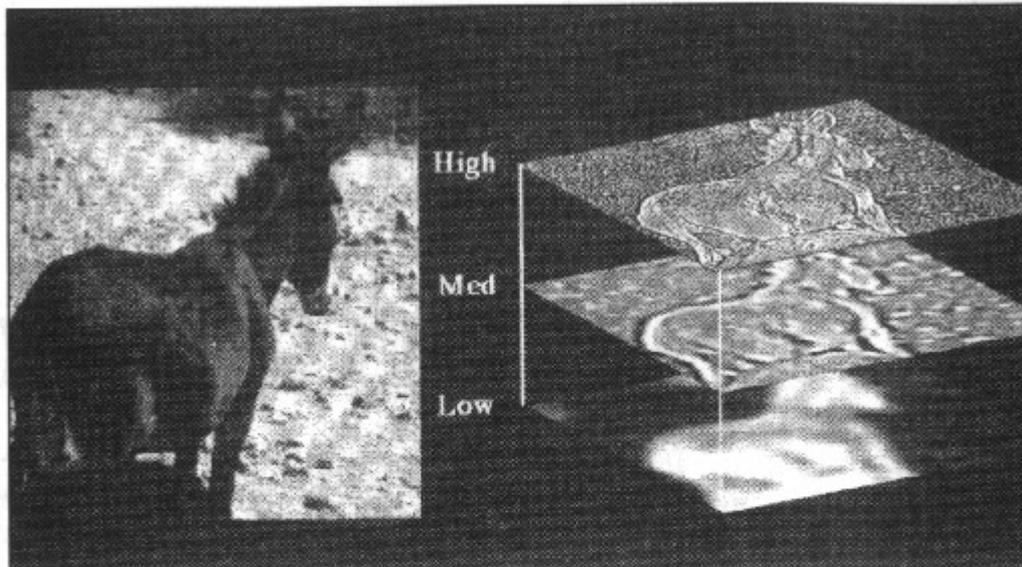


Figure 2. A "natural scene" and its decomposition into high, middle, and low frequencies, using a two-dimensional wavelet transform. The effectiveness of a transform cannot be understood or evaluated in isolation; it depends on the relationship between the transform and the properties of the data to be encoded. Natural scenes have certain statistical properties in common. For example, as shown in the above decomposition, they are redundant: many edges found in low frequencies also exist in middle or high frequencies. Wavelets with a certain narrow bandwidth of frequencies and certain narrow range of orientations can encode natural scenes concisely, with just a few coefficients. (Courtesy of David Field.)

imagem com 140 dp

Observe os padrões
na região escura
da imagem.

Nem sempre maior resolução é melhor:

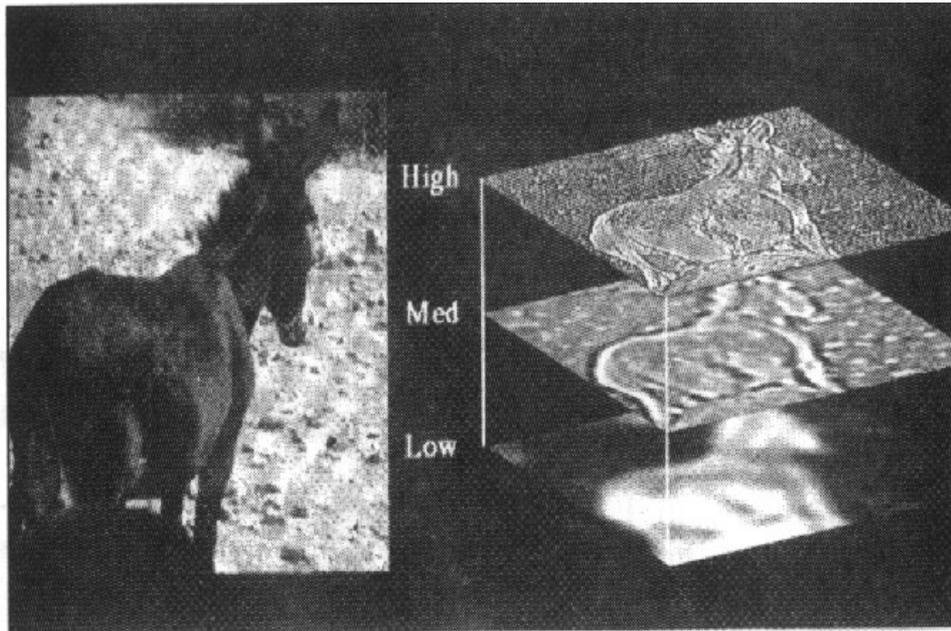


Figure 2. A "natural scene" and its decomposition into high, middle, and low frequencies, using a two-dimensional wavelet transform. The effectiveness of a transform cannot be understood or evaluated in isolation; it depends on the relationship between the transform and the properties of the data to be encoded. Natural scenes have certain statistical properties in common. For example, as shown in the above decomposition, they are redundant: many edges found in low frequencies also exist in middle or high frequencies. Wavelets with a certain narrow bandwidth of frequencies and certain narrow range of orientations can encode natural scenes concisely, with just a few coefficients. (Courtesy of David Field.)

imagem com 180 dp

Observe os padroes

Diminuem bem na
imagem.

Nem sempre maior resolução é melhor:

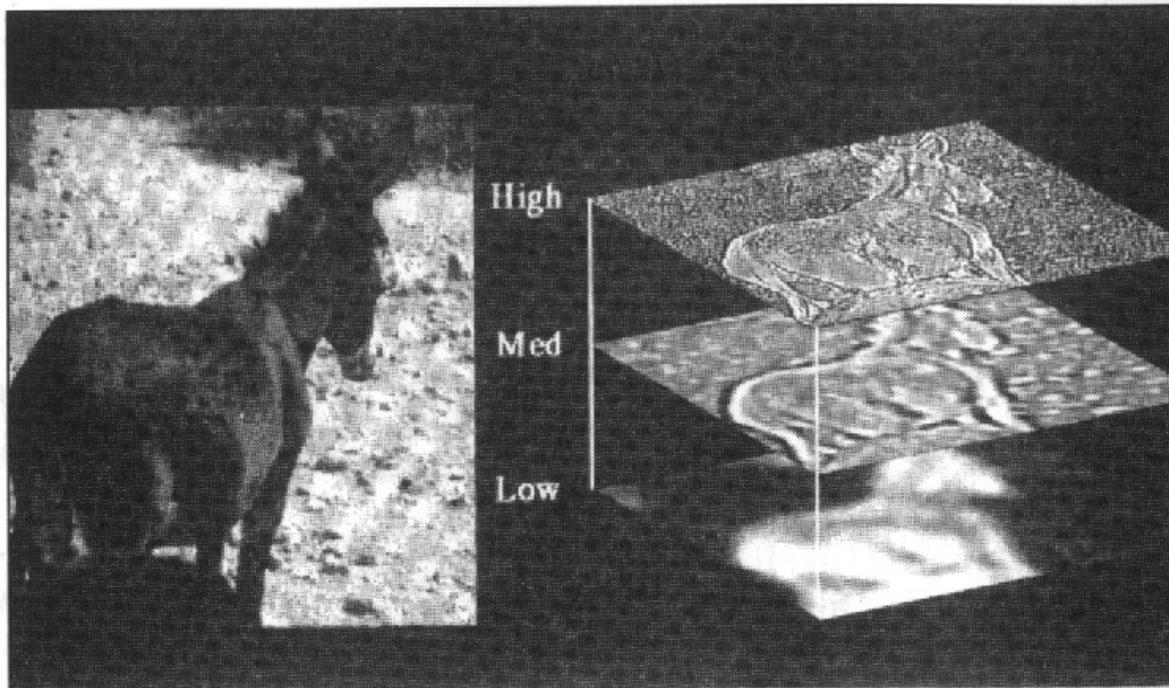


Figure 2. A "natural scene" and its decomposition into high, middle, and low frequencies, using a two-dimensional wavelet transform. The effectiveness of a transform cannot be understood or evaluated in isolation; it depends on the relationship between the transform and the properties of the data to be encoded. Natural scenes have certain statistical properties in common. For example, as shown in the above decomposition, they are redundant: many edges found in low frequencies also exist in middle or high frequencies. Wavelets with a certain narrow bandwidth of frequencies and certain narrow range of orientations can encode natural scenes concisely, with just a few coefficients. (Courtesy of David Field.)

Imagem
200 dp

Padrões
Tanto na
Parte clara
Como na
Escura
da imagem.

3.3.5. Imagens Monocromáticas

Imagens monocromáticas são imagens digitais onde cada *pixel* possui apenas uma banda espectral.



Figura 3.22 – Exemplos de imagens monocromáticas

O número de tons entre os valores limites, branco e preto, que se pode representar em tons, depende de quantos bits são alocados na matriz de imagem para armazenar o tom de cada *pixel*.

Número de elementos na Escala de cinza	Tons de cinza limites	Números de Bits necessários para representação do <i>pixels</i>
2^1 2 valores	0,1	1
2^3 8 valores	0 a 7	3
2^4 16 valores	0 a 15	4
2^8 256 valores	0 a 255	8

Uma imagem monocromática pode ser representada geometricamente também por valores reais quanto à posição dos *pixels* como no gráfico $G(f)$ da função f :

$$G(f) = \{ (x,y,z); (x,y) \in R^2 ; z = f(x,y) \} \quad (3.4)$$

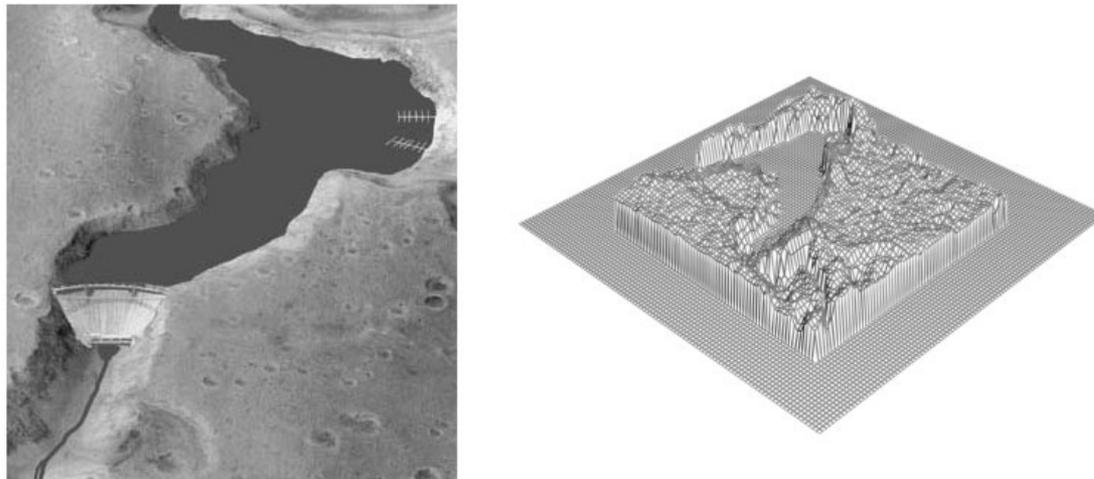


Figura 3.24 – Imagem Monocromática e gráfico 3D de sua função $g(f)$ no R^3 .

3.3.6. Imagens Coloridas

Imagens multibandas são imagens digitais onde cada *pixel* possui n bandas espectrais.

Quando uma imagem é representada pela composição das três bandas visíveis (RGB) tem-se uma imagem colorida aos olhos humanos.

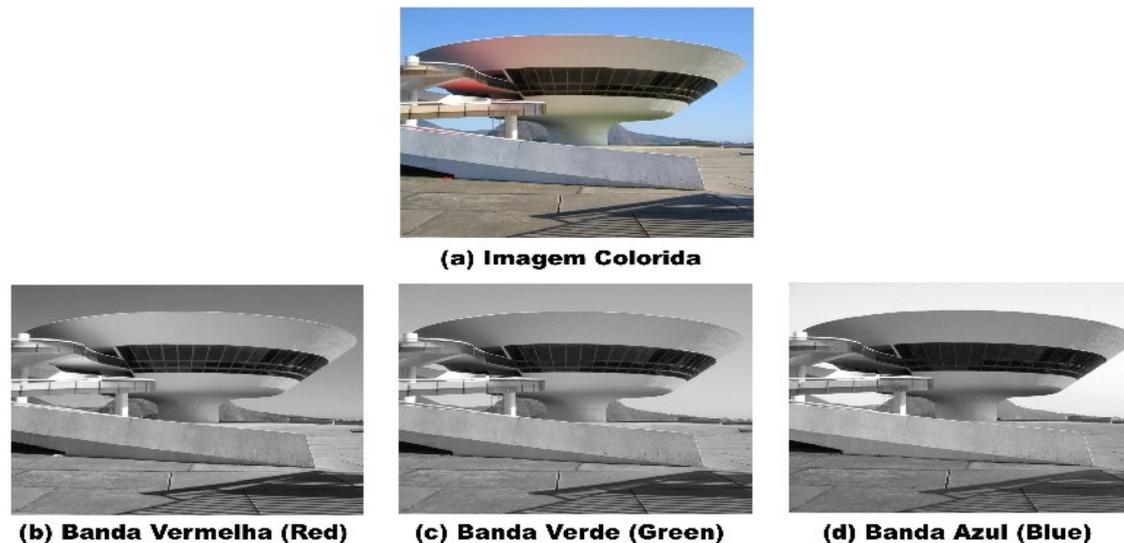


Figura 3.25 – Imagem colorida e cada uma de suas bandas RGB.

3.4. Histograma de imagem digital

O histograma de uma imagem indica o percentual de *pixels* que a imagem tem de determinado nível de cinza ou cor.

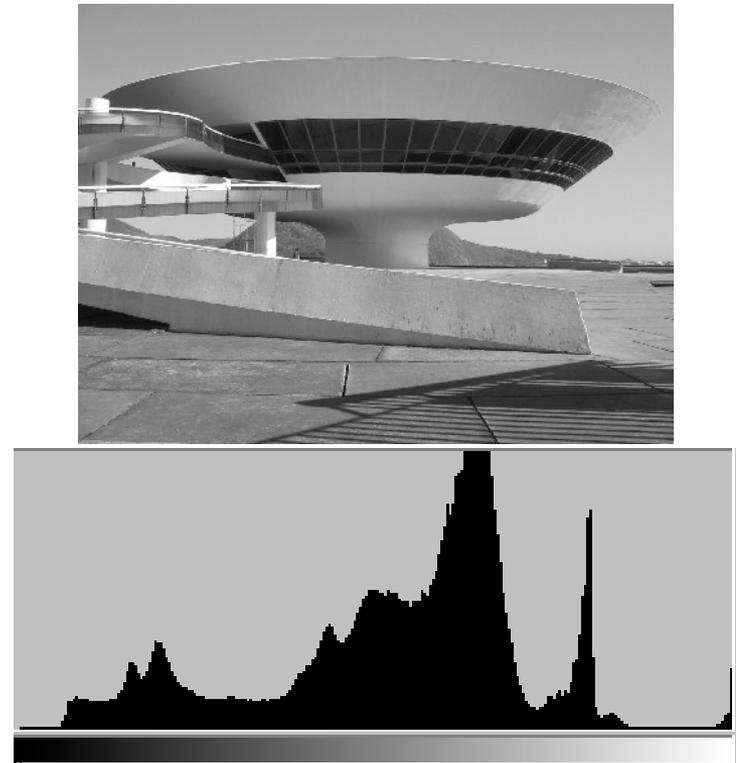


Figura 3.26 – Imagem em tons de cinza e o seu histograma.

O histograma fornece uma indicação da qualidade da imagem quanto ao contraste e intensidade luminosa.

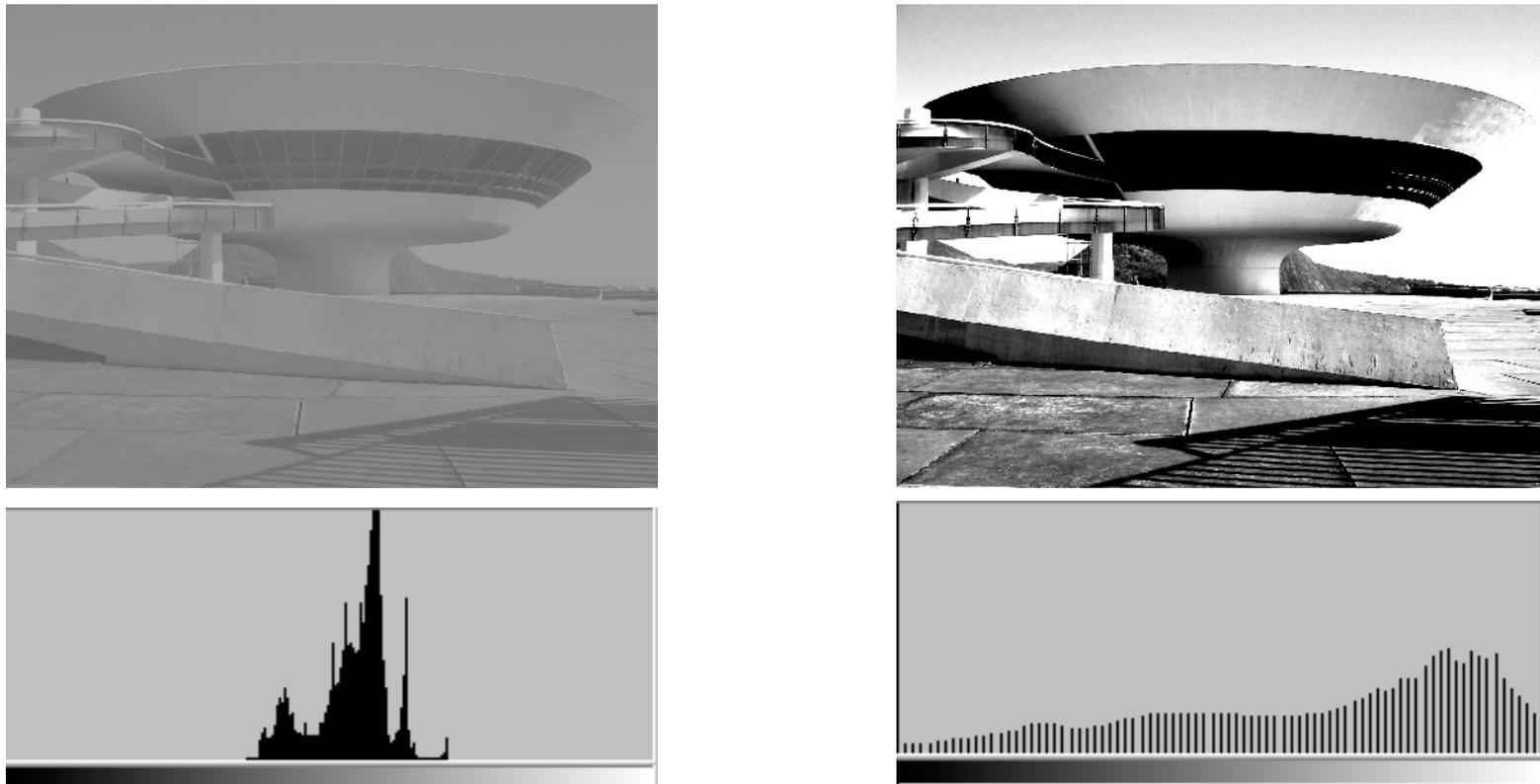


Figura 3.27 – Imagens em tons de cinza e seus respectivos histogramas: com pouco (direita) e muito contraste (esquerda).

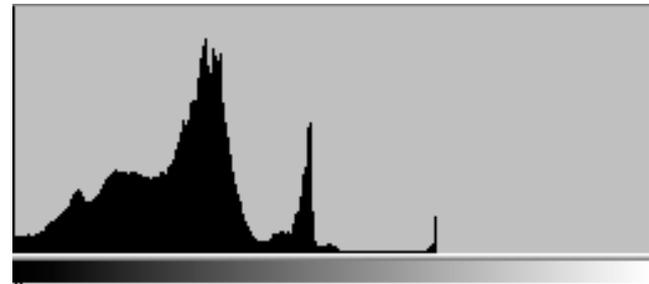
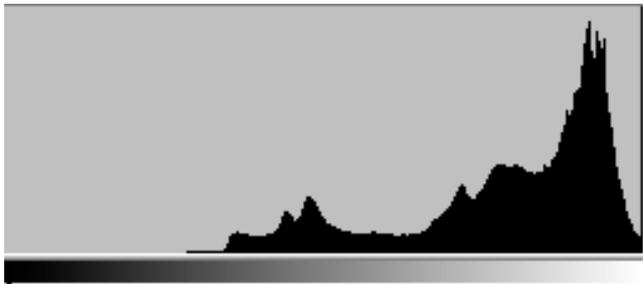


Figura 3.27 (cont.) – Imagens em tons de cinza e seus respectivos histogramas: com alta luminosidade (clara) e com baixa luminosidade (escura).

Histogramas das bandas de uma imagem colorida

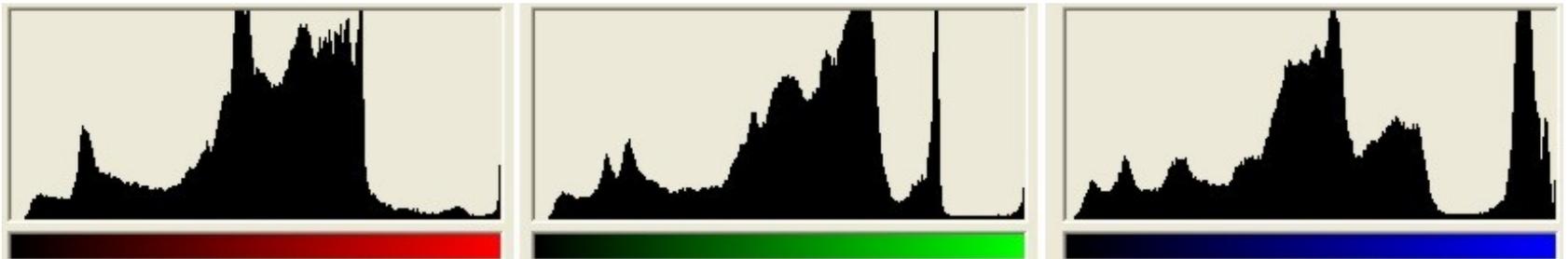
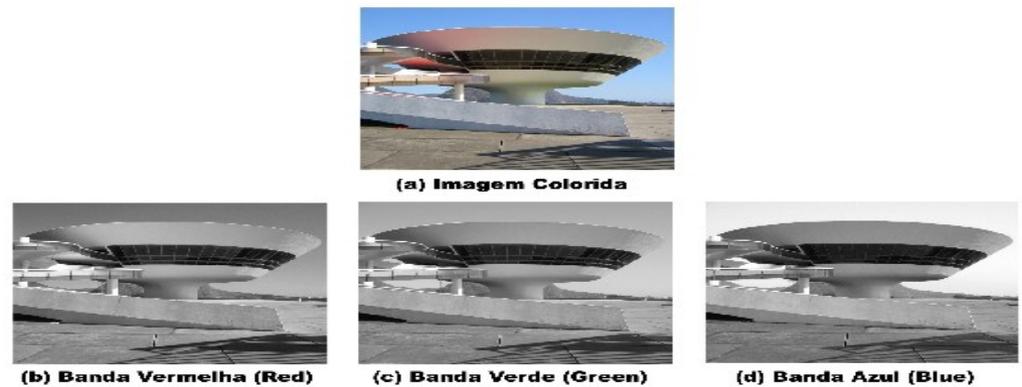


Figura 3.28 - Histogramas da Imagem 3.25 separado por canal R,G B.

3.5. Sistemas de visão binária

3.5.1. Limiarização (Thresholding):

Manual (baseada no histograma)

Automática:

Método de Otsu (*Thresholding* Global)

Método Iterativo

outros (ver site do curso)

3.5.1. Agrupamento por limiar

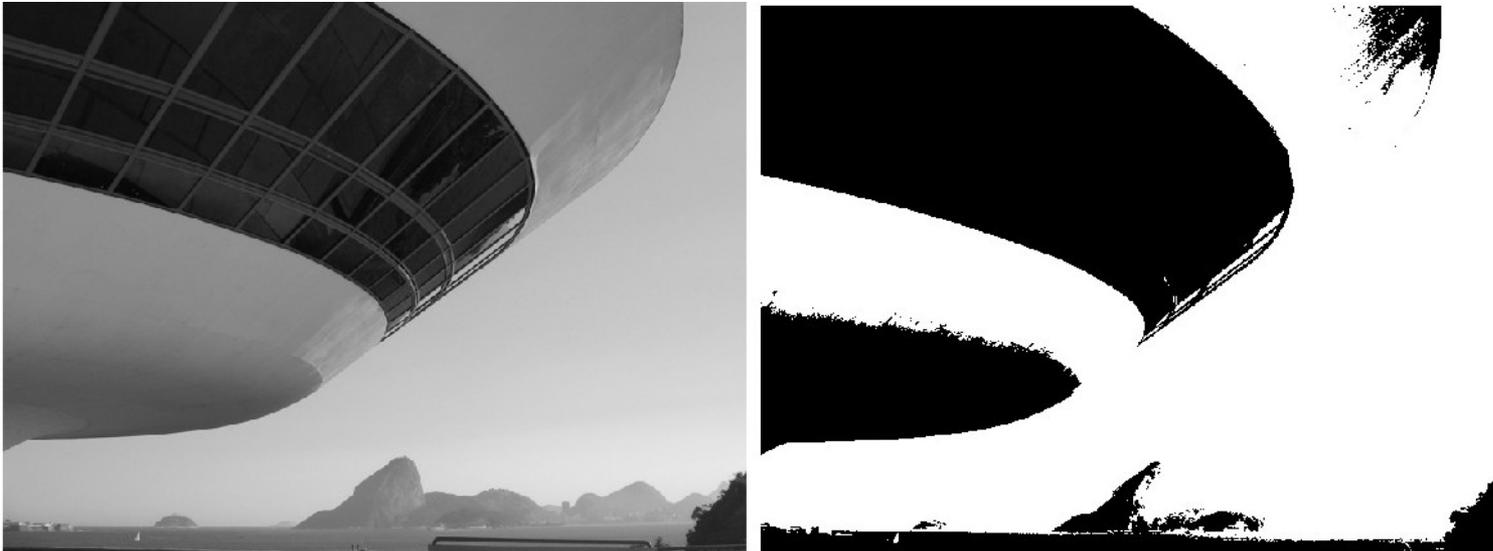


Figura 3.30 - Exemplo de binarização: (a) Imagem em tons de cinza, (b) Imagem binária

3.5.1. Agrupamento por limiar (limiarização)

Uso em imagens em que o objeto a ser segmentado apresenta uma tonalidade bem diferente do fundo da imagem.

Em um histograma bimodal, é possível estabelecer um limiar entre as duas tonalidades.

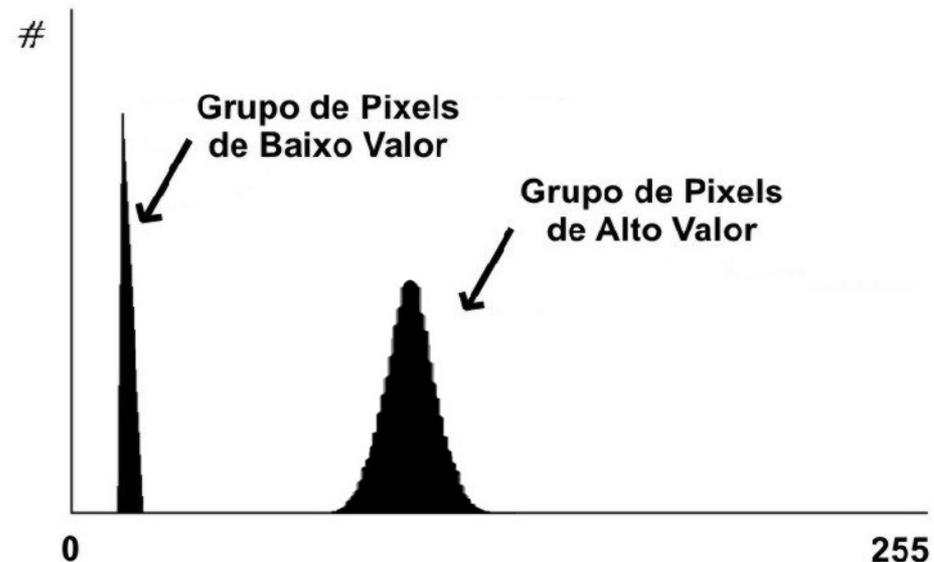


Figura 3.31 - Exemplo de histograma bem-dividido.

Quando a imagem tem mais de dois objetos com cinzas diferentes em um fundo mais escuro, pode ser usada a técnica de limiarização multinível (*multilevel thresholding*).

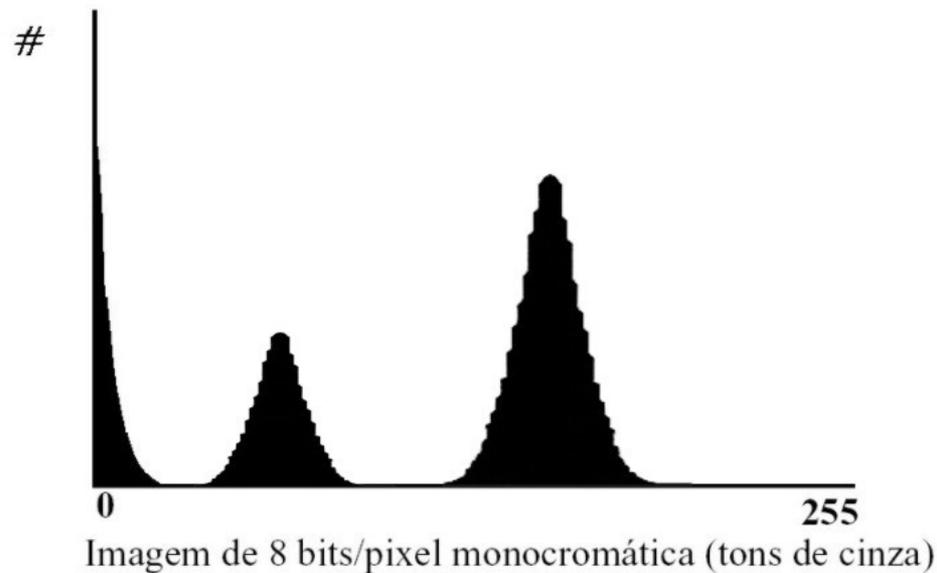


Figura 3.34 - Exemplo de um histograma trimodal.

Há casos em que torna-se necessário o uso de interpolação.

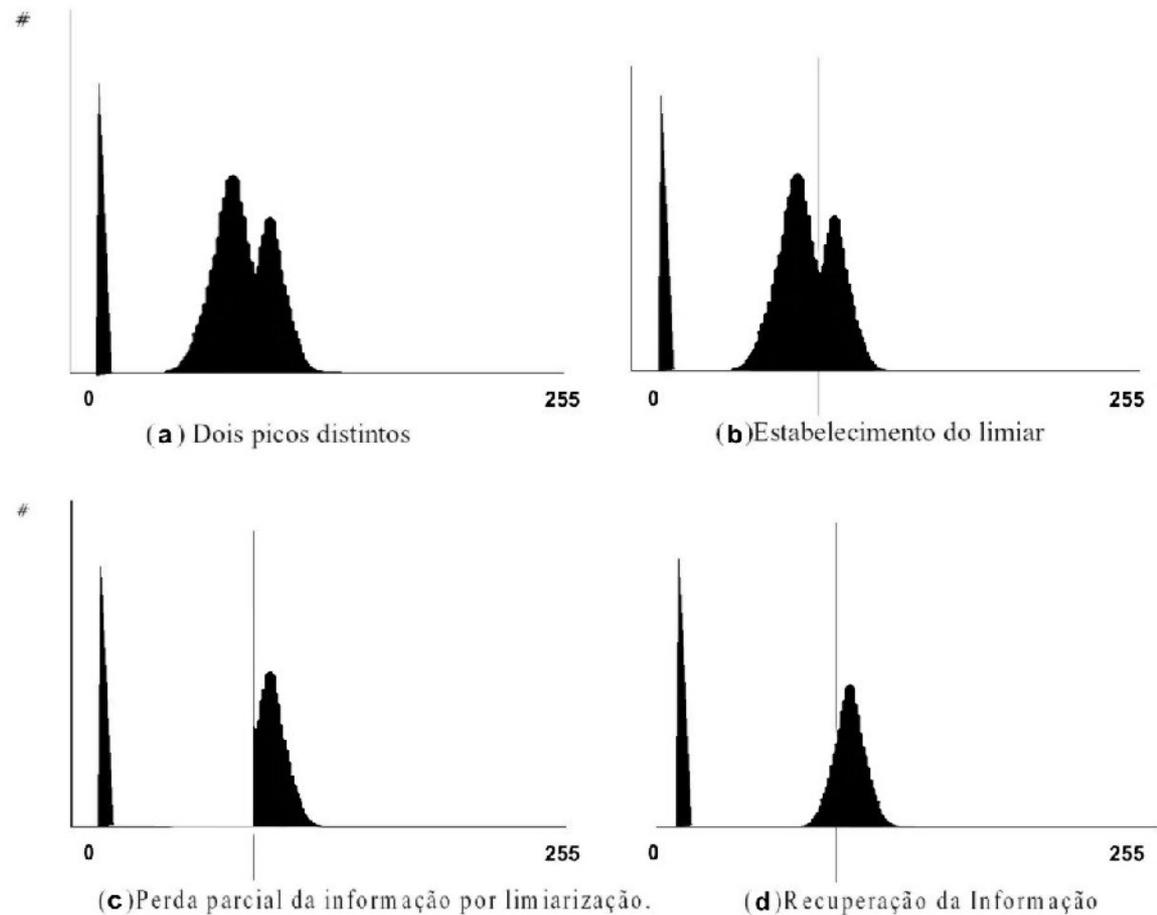


Figura 3.33 - Limiarização por interpolação (Chow e Kaneko, 1972).

A limiarização converte uma imagem de entrada:

$f(x, y)$ de N níveis de cinza

em uma imagem $g(x, y)$, chamada de imagem limiarizada (ou posterizada), com número de níveis de cinza menor do que N .

No limite, $g(x, y)$, terá só dois níveis de cinza, como na equação:

$$g(x, y) = \begin{cases} R_1 & \text{se } f(x, y) \leq T \\ R_2 & \text{se } f(x, y) > T \end{cases} \quad (3.5)$$

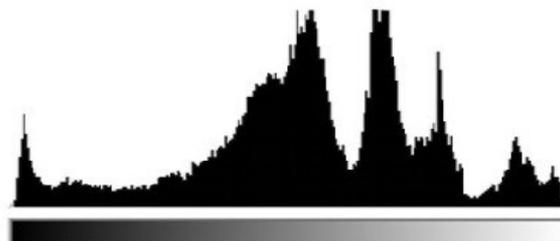
De uma maneira mais genérica, a definição de T é a função de várias variáveis na forma:

$$T = T [x, y, p(x,y), f(x,y)] \quad (3.6)$$

onde $p(x,y)$ é uma propriedade local de x,y .



(a) Imagem com 256 tons de cinza



(b) Histograma



(c) Limiarização com valor 80



(d) Limiarização com valor 150

Figura 3.36 – Influência do valor do limiar sobre a qualidade da limiarização. (As imagens (c) e (d) são posterizadas nesta representação).

3.5.2. Limiarização Automática (*Thresholding*) baseada no histograma

O método mais simples (e mais caro computacionalmente) de calcular automaticamente o *threshold* é o método iterativo.

Este método é descrito por:

1. Selecione um tom T_i . inicial (entre o valor da intensidade mínima e máxima da imagem.)
2. Segmente imagem usando T_i . Isto é, divida-a em dois grupos: (G1) formado com os pixels cujos valores são $> T$ e (G2) com os pixels com tons $< T$.
3. Calcule a intensidade média dos grupos: μ_1 e μ_2
4. Calcule o novo threshold: $T_{i+1} = \frac{1}{2} (\mu_1 + \mu_2)$
5. Repita os passos 2 até 4 até que $T_{i+1} - T_i < \epsilon$ (parâmetro pré-definido)

Método de Otsu

A operação de limiarização é considerada como sendo o particionamento da imagem com L níveis de cinza em duas classes C_0 e C_1 , que representam o objeto e o fundo, ou vice-versa, sendo que esta partição se dará no nível de cinza t , *que maximiza a razão da variância entre as classes e a variância total*

C_0 = pixels da imagem com tons de 0 a t

C_1 = pixels da imagem com tons de t a l

Método de Otsu para escolher o melhor *threshold*: T .

Se baseia no histograma normalizado (da imagem)., que pode ser visto como uma função de **densidade de probabilidade discreta** (**histograma nirmalizado**)

$$p_r(r_q) = \frac{n_q}{n}, q = 0, 1, 2, \dots, L - 1$$

onde n é o número total de píxeis na imagem, n_q é o número de píxeis que tem intensidade r_q

O método de Otsu escolhe o threshod de valor T *que maximiza*:

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2$$

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2$$

onde, $w_0 = \sum_{q=0}^{k-1} p_q(r_q)$

$$w_0 = \sum_{q=k}^{L-1} p_q(r_q)$$

$$m_1 = \sum_{q=k}^{L-1} qp_q(r_q)/\omega_1 \qquad m_0 = \sum_{q=k}^{L-1} p_q(r_q)$$

$$m_T = \sum_{q=0}^{L-1} qp_q(r_q)$$

=

3.6. Operações pontuais em imagens

Global x local

3.6.1. Operações pontuais baseadas na curva de tom

3.6.2. Operações baseadas em histograma

3.6.3. Técnicas Baseadas no Histograma de Imagens Coloridas

3.6. Operações pontuais em imagens

Uma operação pontual **global** m uma imagem digital r é a função $f(r)$ aplicada a **todo** *pixel* da imagem:

$$r: s = f(r) \quad (3.7)$$

3.6.1. Operações pontuais globais baseadas na curva de tom

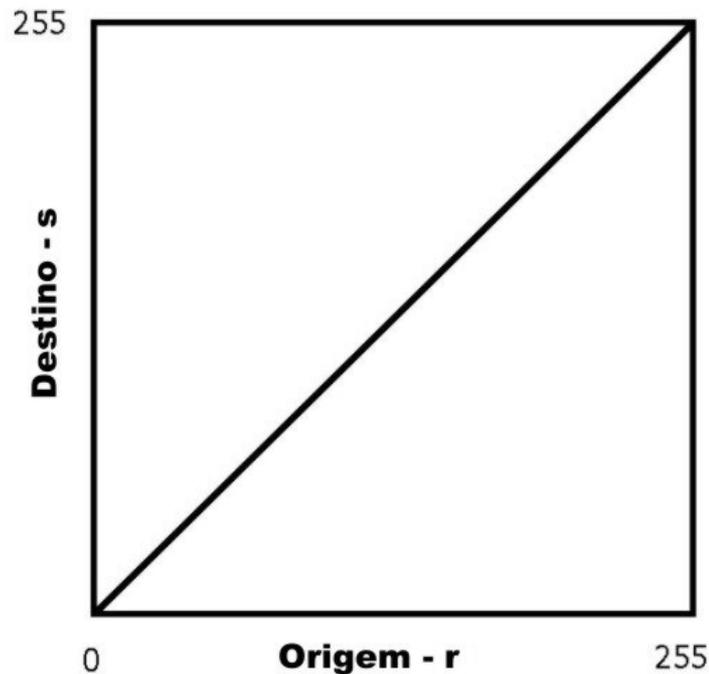


Figura 3.38 - Relação entre imagem origem e destino.

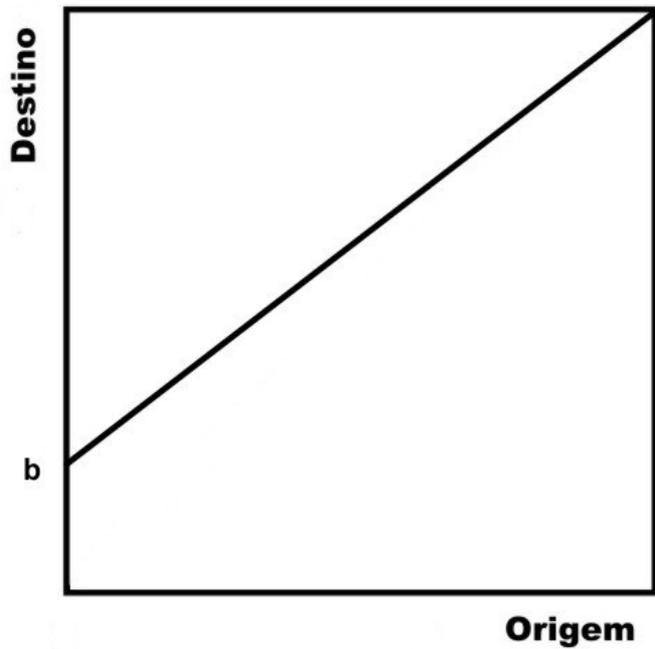


Figura 3.39 - Representação na forma linear

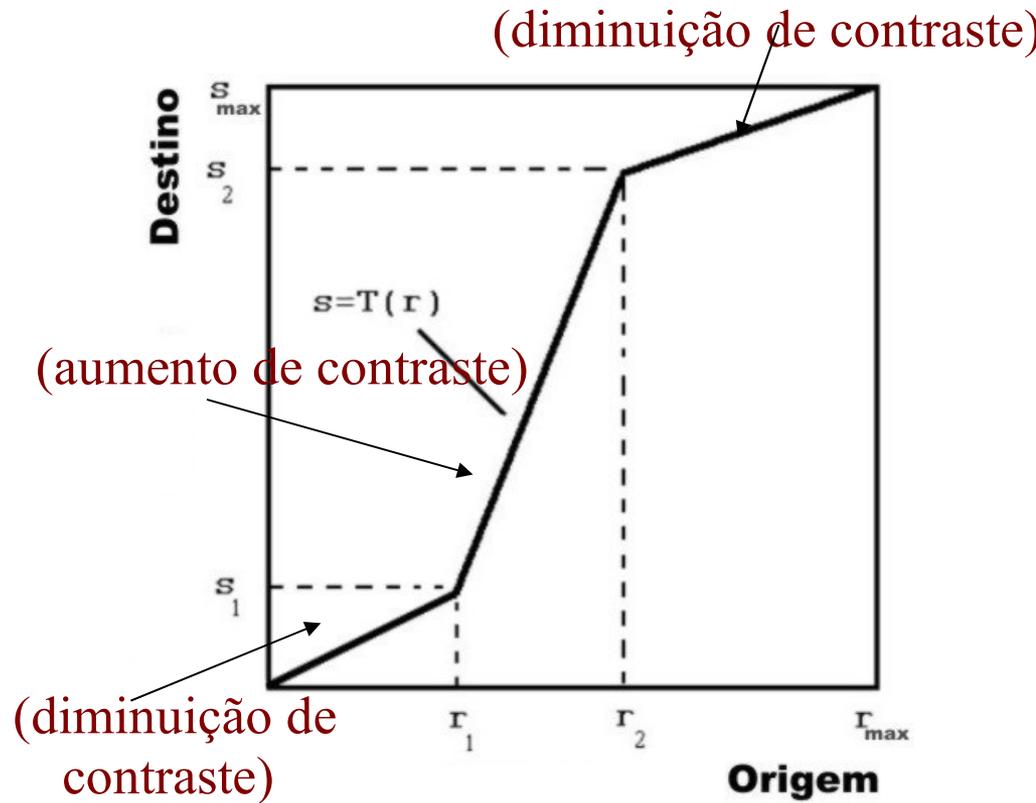


Figura 3.40. Alteração da imagem destino a partir de uma função genérica.

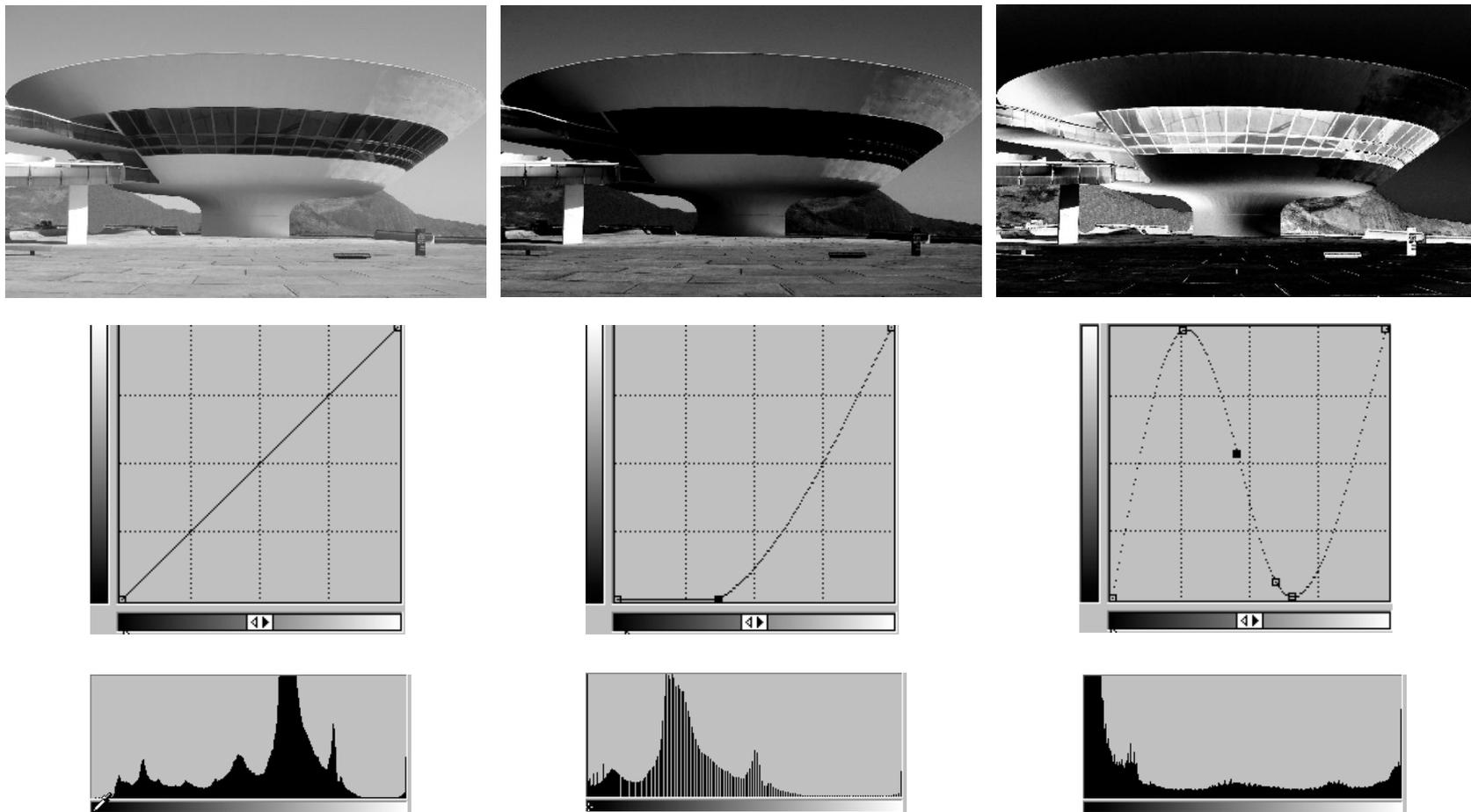


Figura 3.44 Curva de tom original e alteradas, com suas respectivas curvas de tom e histogramas.

Compressão do histograma (diminuição do range)

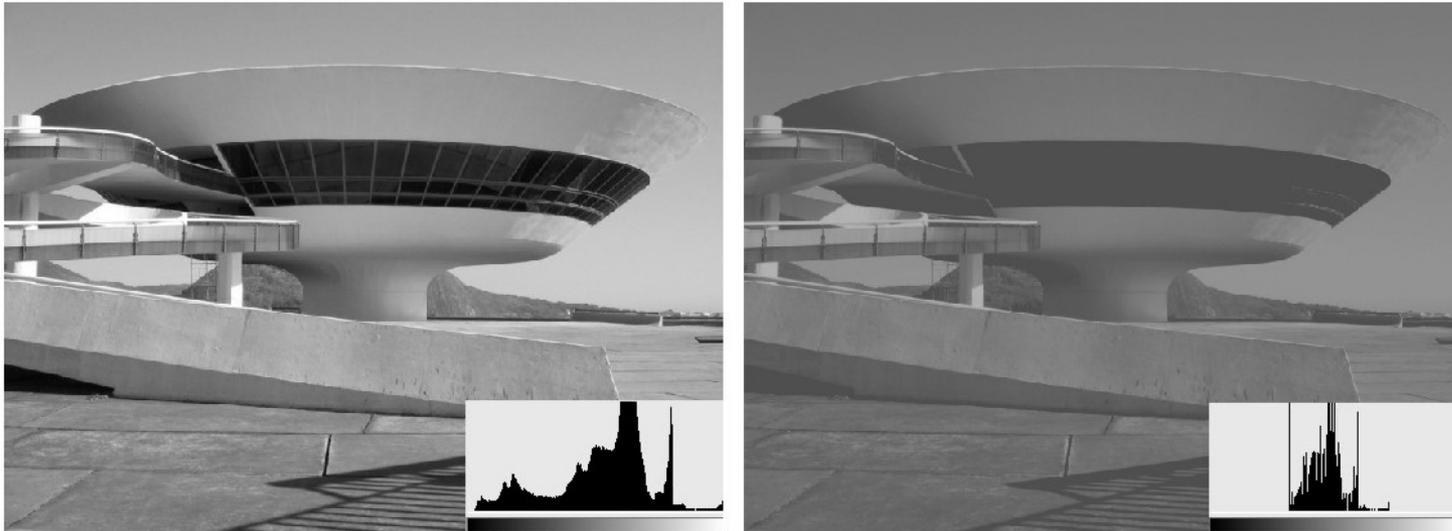


Figura 3.45 - Efeito da Compressão de Histogramas

Expansão de histograma (ampliação do range)



Figura 3.46. Efeito da Expansão de Histogramas.

Imagem negativa (ou com os tons invertidos)

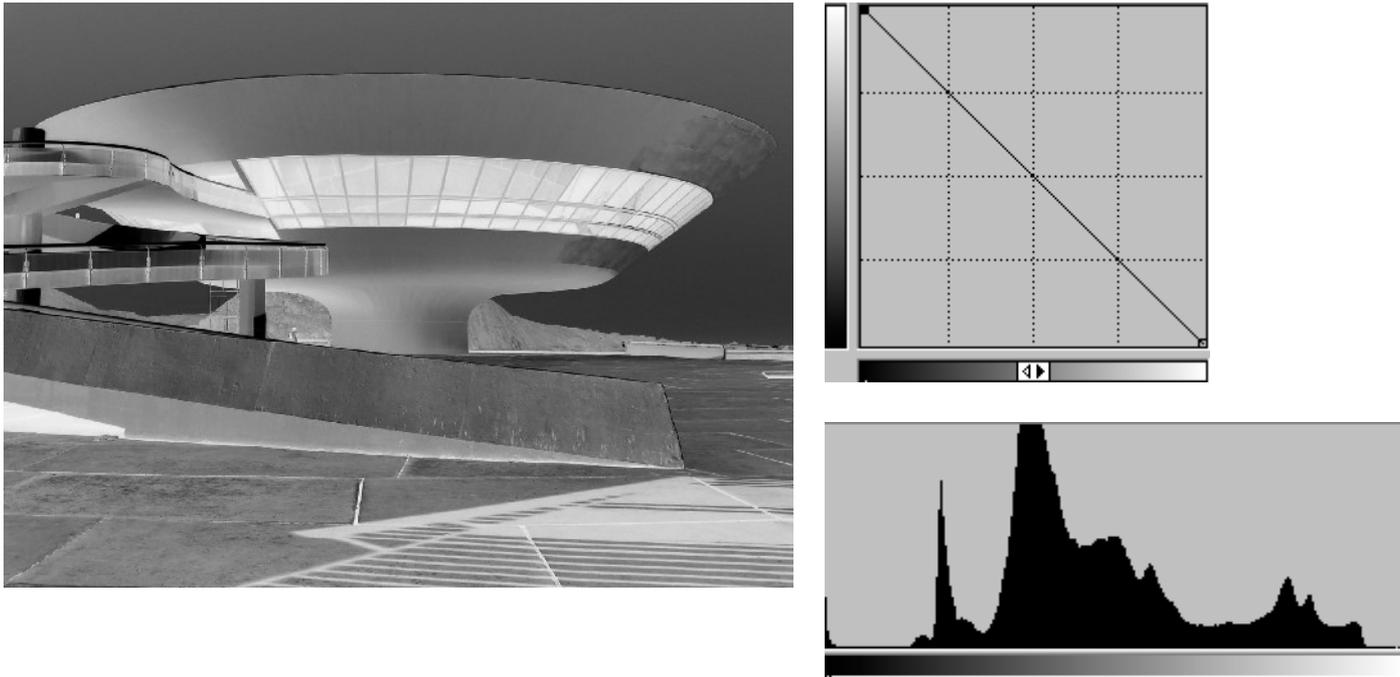


Figura 3.47. Efeito da imagem negativa.

3.6.2. Outras perações globais baseadas em histograma

3.6.2.1. Equalização

O objetivo da equalização de imagens é a melhoria do seu contraste.

Para tanto, o ideal é que os níveis de cinza sejam representados de maneira uniforme e distribuída.

A equalização de histograma consiste na redistribuição dos valores de nível de cinza em uma imagem, de forma que todos os *pixels* tenham a probabilidade de aparecer mais **equalitária possível**.

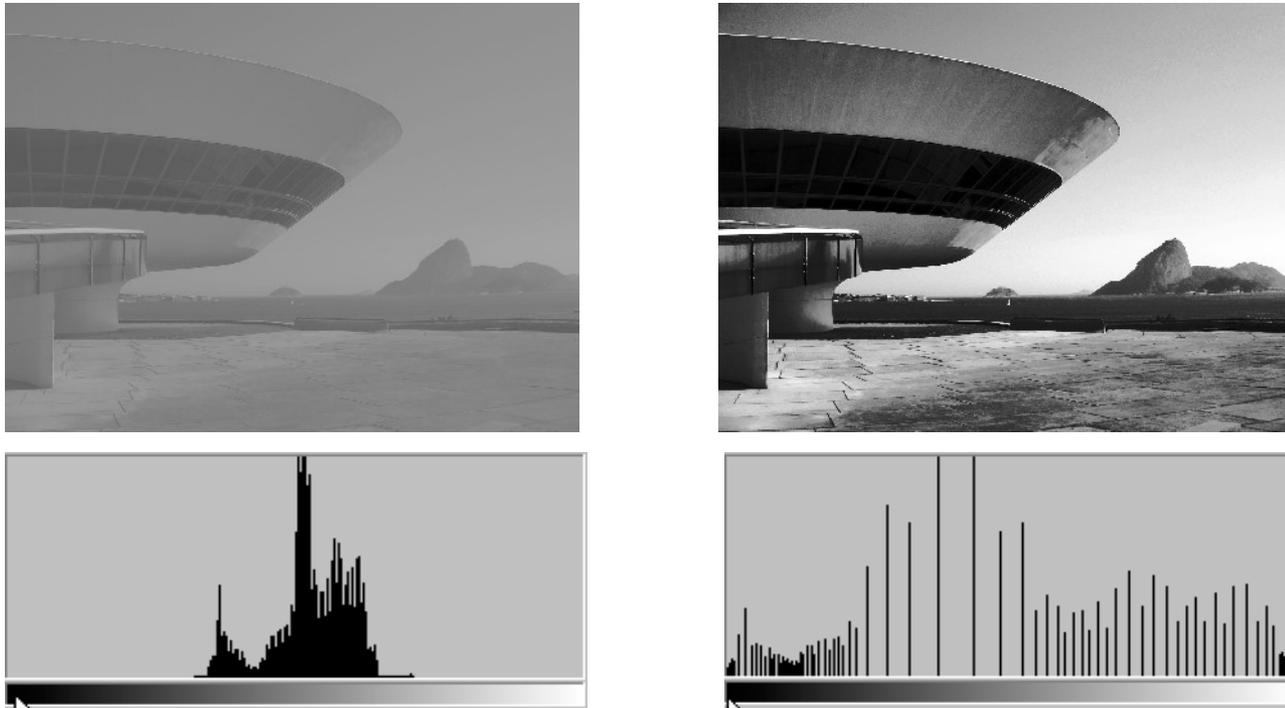


Figura 3.49 – (a) Imagem com baixo contraste e seu histograma.

(b) Imagem após equalização e seu histograma.

3.6.2.1. Especificação Direta

A equalização de histogramas procura distribuir igualmente para todos os níveis de *pixels* da imagem o que não permite estabelecer parâmetros de mudança capazes de ressaltar faixas de níveis de cinza na imagem.

Esta pode ser realizada por outra técnica chamada de especificação direta de histograma.

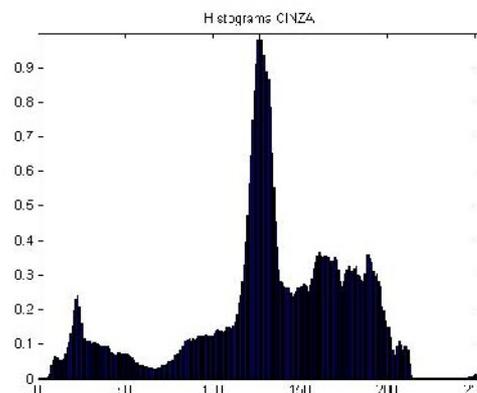


Figura 3.51 – Imagem original e seu histograma normalizado

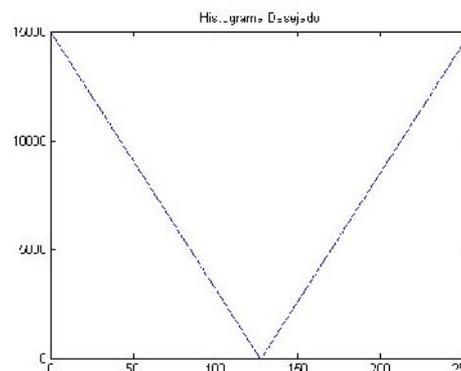


Figura 3.52 – Imagem obtida (a) após a especificação do histograma (b).

3.6.3. Técnicas Baseadas no Histograma de Imagens Coloridas

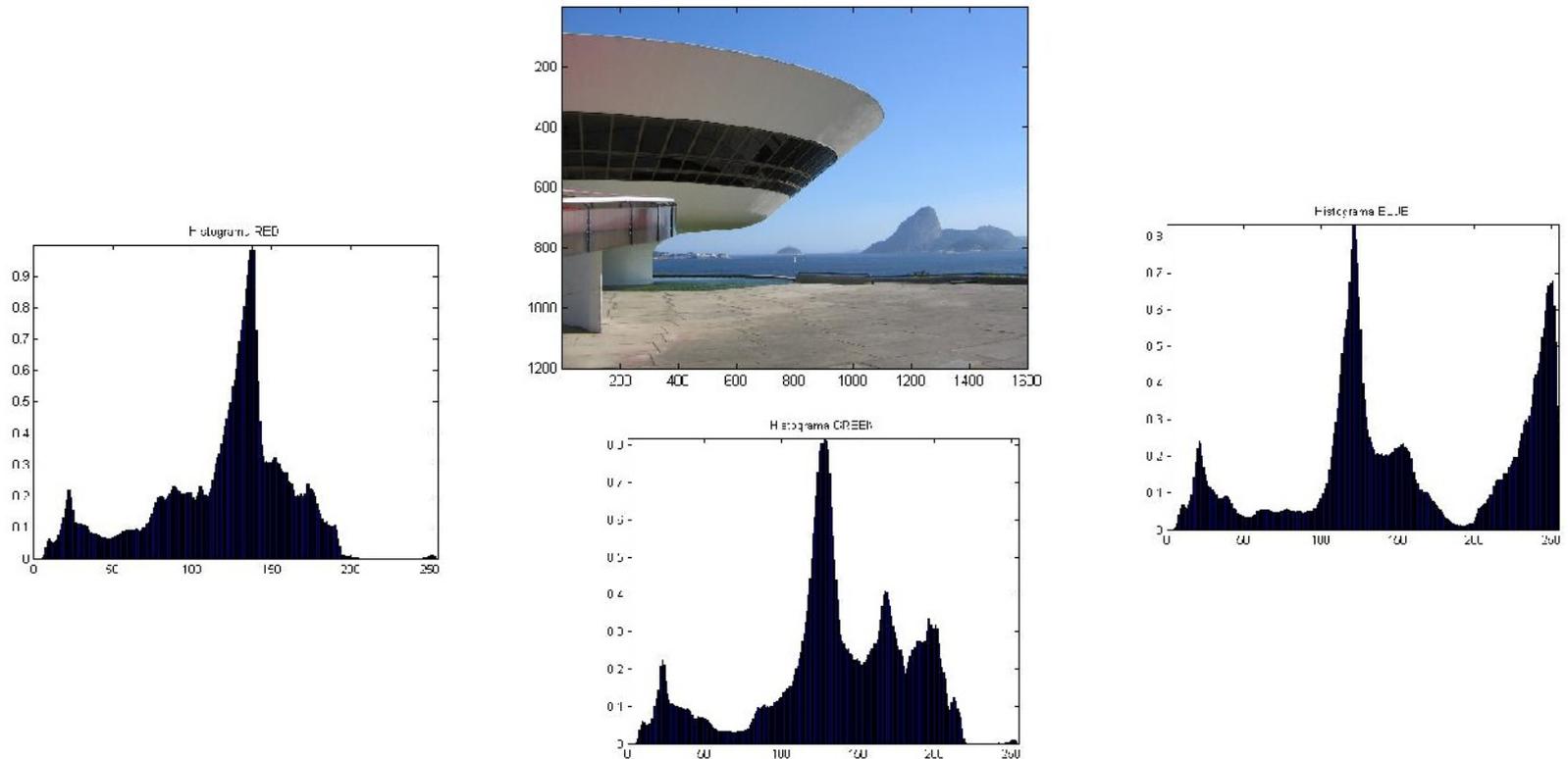


Figura 3.55 – Imagem original e seus histogramas normalizados (Gonçalves, 2006)

Problema:

Quando as técnicas de mudança de histograma para imagens em **tons de cinza** são simplesmente **triplicadas** para cada um dos canais (RGB) independentemente, o resultado será o surgimento de **cores não presentes** na imagem original.

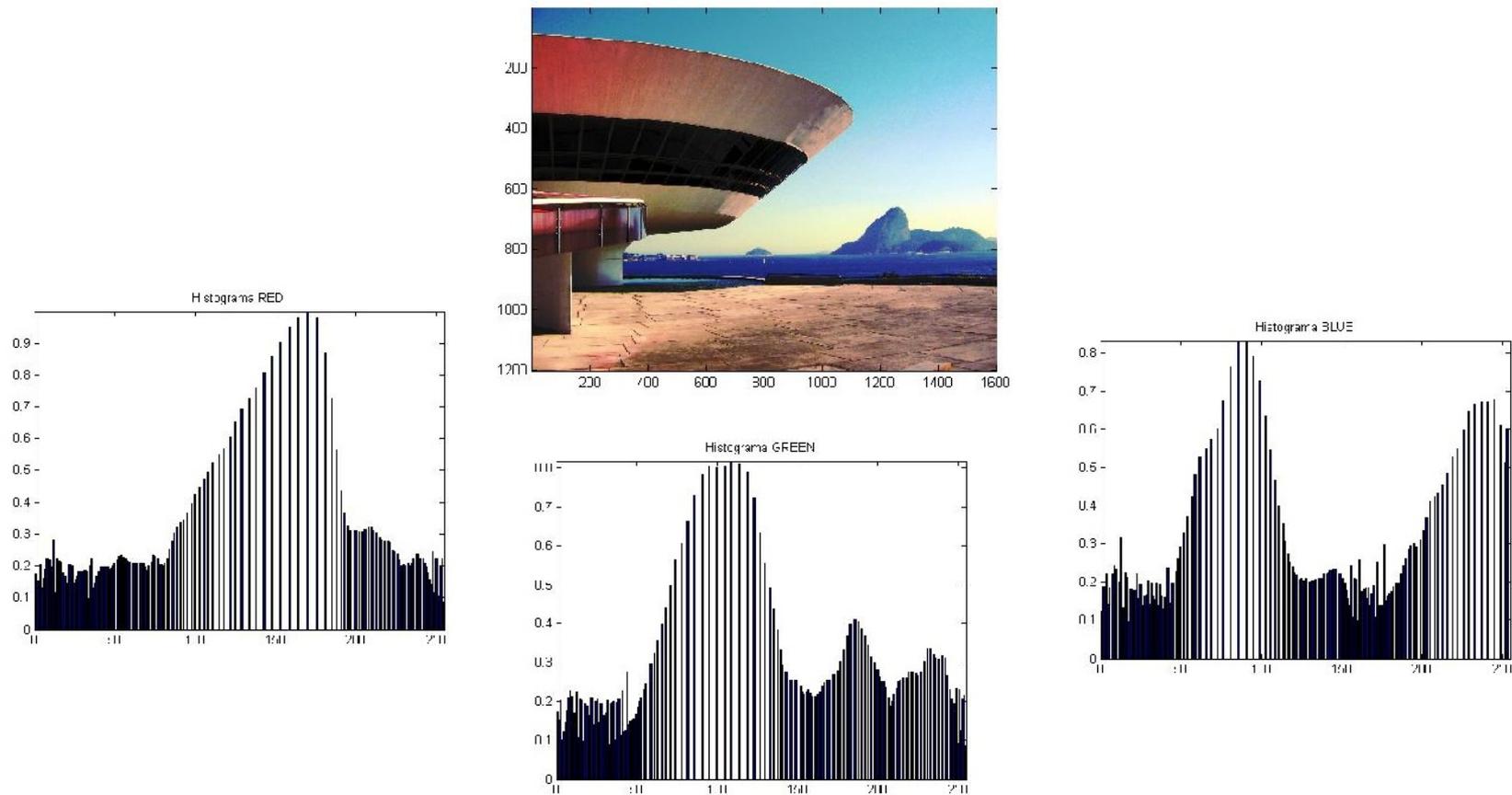


Figura 3.56 – Imagem modificada por equalização e seus histogramas equalizados normalizados .

Soluções:

- 1- Usar melhor canal em melhor espaço de cores

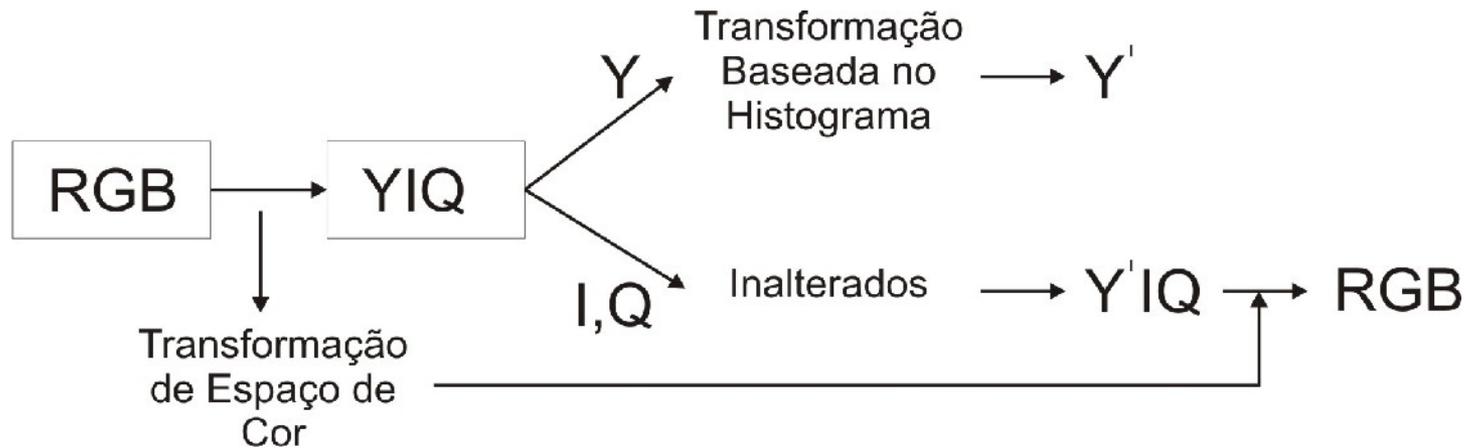


Figura 3.58 – Esquema de modificações baseadas no histograma para imagens coloridas.

Soluções:

2- Usar melhor canal em melhor espaço de cores

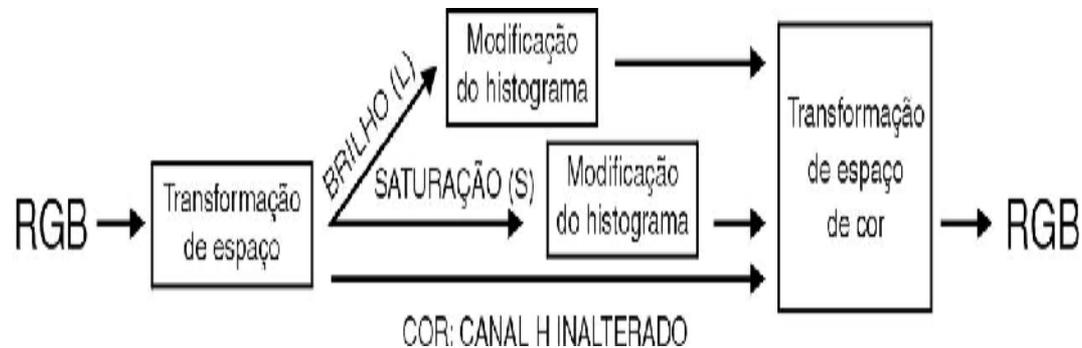


Figura 3.59 – Esquema de modificação para espaços HSV e HLS

Bibliografia Complementar (cap. 3):

<http://www.ic.uff.br/~aconci/OTSUeOutras.pdf>

Limiarização de Pun e Kapur:/[/~aconci/curso/binari~3.htm](http://www.ic.uff.br/~aconci/curso/binari~3.htm)

Chow e Kaneko: <http://www.ic.uff.br/~aconci/limiarizacao.htm>

Otsu : tese Leonardo Hiss Monteiro do IC:2002

Dissertação: Jonh Edson Ribeiro de Carvalho “Uma Abordagem de Segmentação de Placas de Automóveis baseada em Morfologia Matemática- 2006

Flávio L. Seixas, Anita Martins, Arthur R. Stilben, Daniel Madeira, Rafael Assumpção, Saulo Mansur, Silvia M. Victor, Vilson B. Mendes e Aura Conci,
AVALIAÇÃO DOS MÉTODOS PARA A SEGMENTAÇÃO
AUTOMÁTICA DOS TECIDOS DO ENCÉFALO EM RESSONÂNCIA
MAGNÉTICA <http://www.ic.uff.br/~aconci/pub2008.html>

<http://eof.alunos.dcc.fc.up.pt/wiki/index.php/Otsu>