Processamento e Análise de Imagens

Fundamentos



Percepção Visual

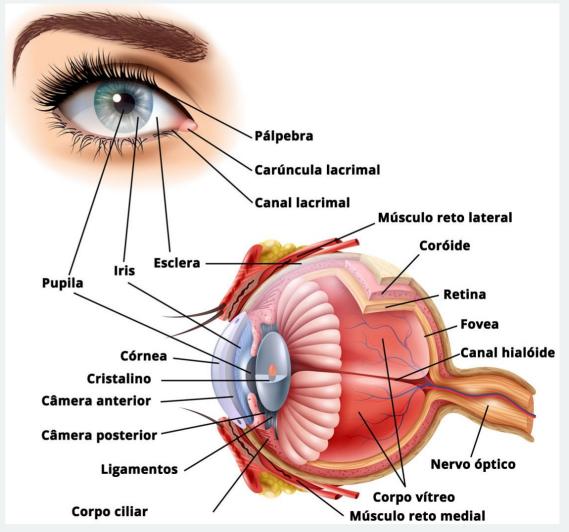
Percepção Visual

- A área de processamento digital de imagens ser fundamentada basicamente em conceitos matemáticos,
 - No entanto, a intuição e a subjetividade humana exercem consideravel influência sobre os resultados produzidos pelos algoritmos;
 - Cada técnica é avaliada, muitas vezes, sobre critérios subjetivos, oriundos da percepção dos analistas, engenheiros e usuários.

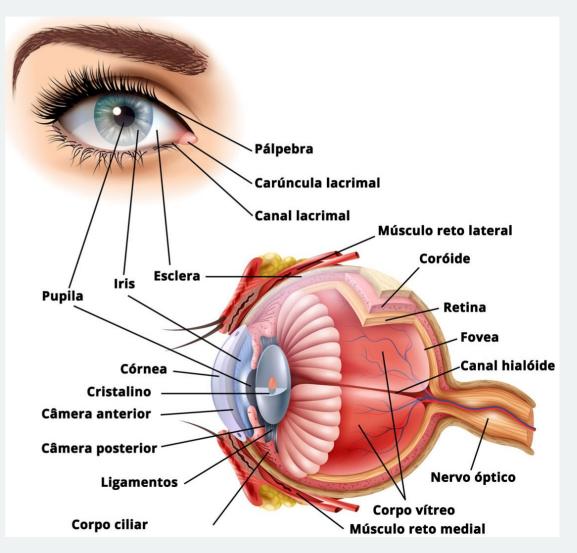
Percepção Visual

- Segundo Gonzalez (2018), a visão é considerado o nosso principal sentido;
- Apesar de sua importância, esta é consideravelmente limitada
 - Seu alcance é limitado, não possibilitando o reconhecimento de objetos muito pequenos ou muito distantes;
 - Além disso, a visão humana está sujeita a "defeitos" fisiológicos (ex.: miopia, hipermetropia, etc.);
 - O processamento das informações também é sujeito a incorreções (ex.: ilusão de ótica).

- Córnea: atribuição de proteger o olho humano contra traumas e contaminações;
- Íris: estrutura mais escura do olho, algumas vezes colorida;
 - Possui o controle de abertura e fechamento da pupila;



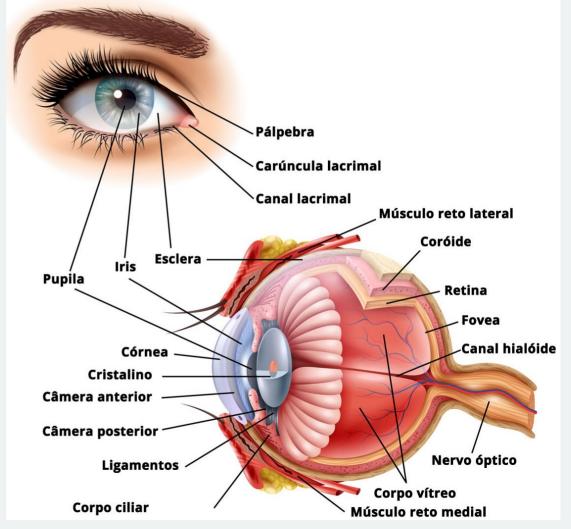
- Pupila: responsável pela permissão de entrada de luz nos olhos, variando conforme a intensidade;
 - Em baixa luminosidade, a pupila dilata, para que entre mais luz;
 - Em alta luminosidade, a pupila se contrai, protegendo o olho;
- Corpo Vítreo: parte gelatinosa que ocupa o centro do olho;



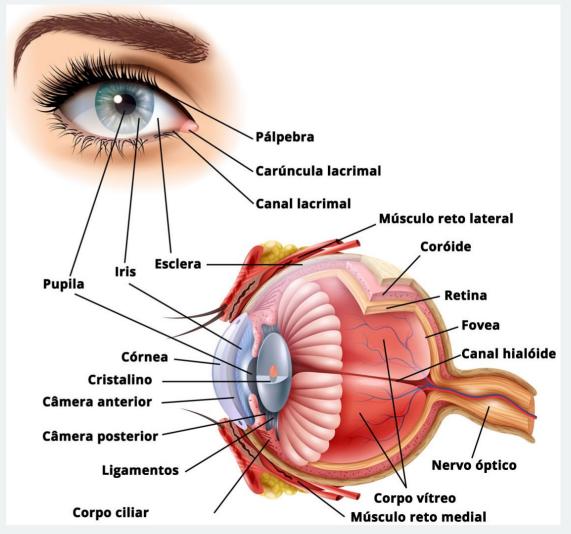
- Cristalino: estrutura convergente que focaliza a luz que entra na pupila e possibilita a formação de imagens na retina
 - Possui uma distância focal variável, conforme a necessidade;
 - Realiza o ajuste fino do foco para as diferentes tarefas;

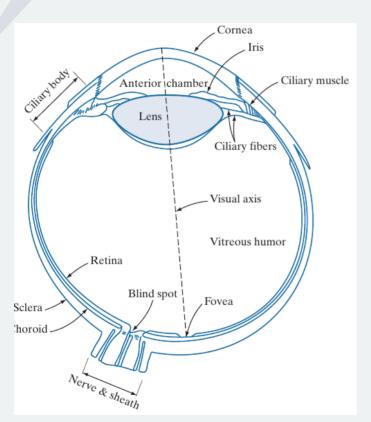
Pálpebra Carúncula lacrimal **Canal lacrimal** Músculo reto lateral Coróide Esclera Pupila Retina **Fovea** Canal hialóide Córnea Cristalino Câmera anterior Câmera posterior Nervo óptico Ligamentos Corpo vítreo Corpo ciliar Músculo reto medial

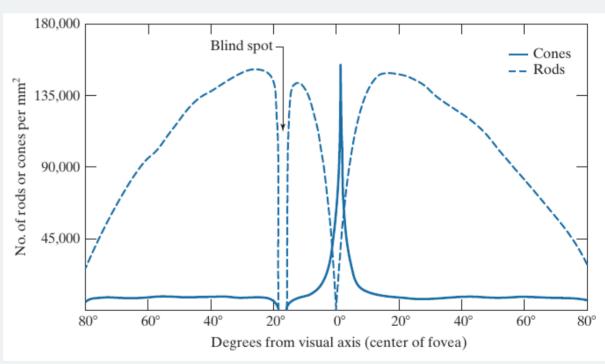
- Retina: parte do globo ocular onde são formadas as imagens
 - Possui células fotossensíveis, os cones e os bastonetes;
 - Cones: percebem a noção de cor;
 - Bastonetes: percebem a noção de intensidade.



- Fóvea: região especializada da retina, que permite perceber detalhes de objetos e das cores;
- Nervo óptico: responsável por enviar os sinais nervosos produzidos pelo olho humano para o cérebro





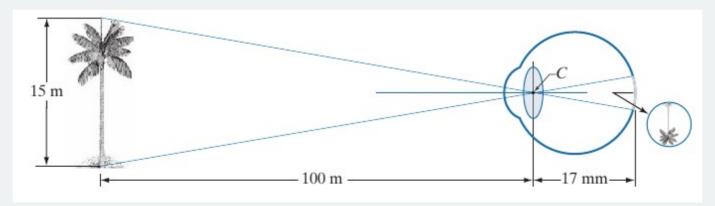


Mudança de Foco

- Para adequação às necessidades humanas, o olho deve possibilitar variação do foco
 - A distância entre o centro da lente e o sensor de imagem (a retina) é fixa, e a distância focal necessária para atingir o foco adequado é obtida variando a forma da lente (cristalino);
 - As fibras do corpo ciliar fazem isso achatando ou engrossando a lente para objetos distantes ou próximos, respectivamente;
 - A gama de distâncias focais é de aproximadamente 14 mm a 17 mm, esta última ocorrendo quando o olho está relaxado e focado a distâncias superiores a cerca de 3m;
 - Câmeras fotográficas, por outro lado, variam a distância entre as lentes e o plano de imagem.

Processamento de Informação Visual

- A percepção então ocorre pela excitação relativa dos receptores de luz, que transformam a energia radiante em impulsos elétricos que, por fim, são decodificados pelo cérebro.
 - A imagem capturada é processada pelo cerébro, encarregado de ajustes necessários para entendimento da imagem.

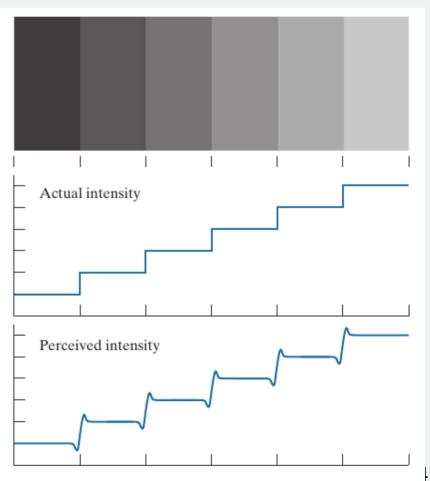


Adaptação à Luz

- Para realização de tarefas, o olho humano deve ser capaz de se adequar a diferentes níveis de luminosidade
 - Existem evidências que a percepção luminosa do olho humano possa ser representada por uma função logarítimica em relação à luz incidente no olho;
 - Um fenômeno conhecido é a capacidade de adaptação à luminosidade, que varia desde a visão escotópica (produzida pelo olho em condição de baixa luminosidade) até o limite de brilho.

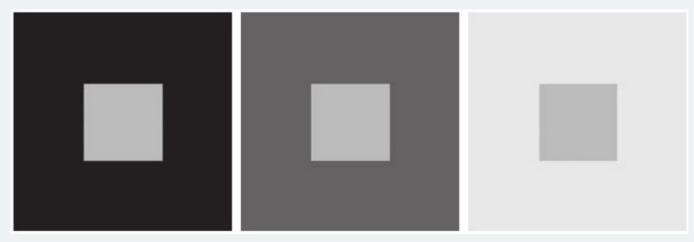
Adaptação à Luz

- No entanto, estudos sugerem que a percepção de luz está limitada a doze níveis máximos de intensidade – a percepção de brilho não pode ser descrita como uma simples função de intensidade
- O olho tende a supervalorizar (overshoot)
 ou subvalorizar (undershoot) as bordas de
 regiões de diferentes intensidades.



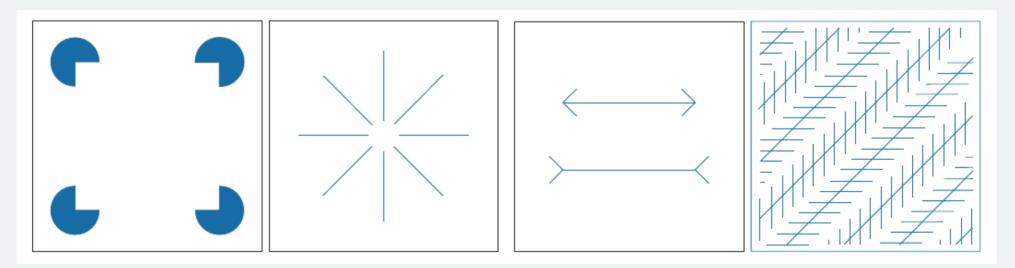
Adaptação à Luz

- A percepção da luz não depende apenas da luminosidade, porém é sujeita à variação do contraste da região subjacente (simultaneous contrast)
 - Os quadrados centrais possuem todos as mesmas cores, porém são percebidos de maneira distinta.



Ilusão de Ótica

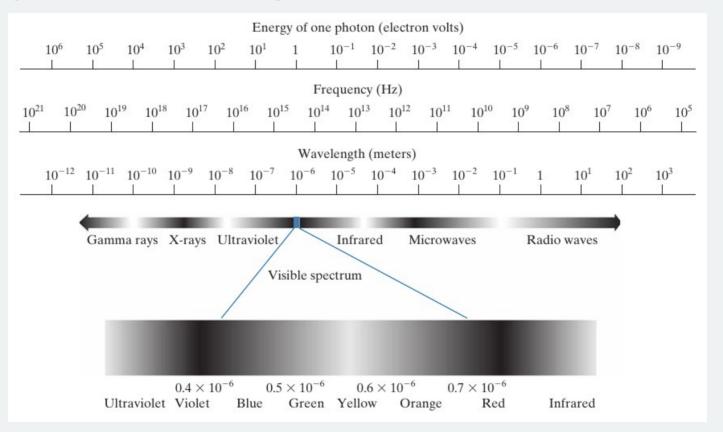
• Ilusão de ótica corresponde a adição de detalhes inexistentes ou a percepção incorreta de propriedades geométricas em imagens.



16

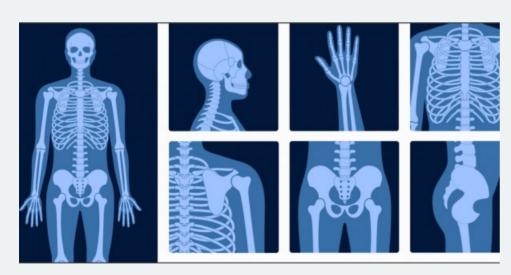
Luz e Espectro Eletromagnético

Luz e Espectro Eletromagnético



Luz e Espectro Eletromagnético

 Apesar de sermos incapazes de ver uma gama grande de espectro luminosos, podemos converter a luz em outros espectros para valores visíveis, com auxílio de equipamentos ou anteparos.



Fonte: DAPI (2023).



Fonte: Brasil Escola (2023).

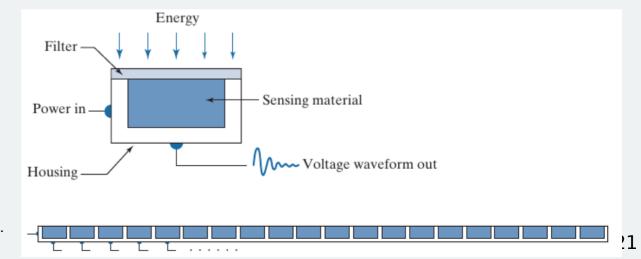
Aquisição de Imagens

Aquisição de Imagens

 A aquisição de imagens de interesse é feita com base em uma fonte de iluminação, seguida da reflexão ou absorção de energia pelos elementos da cena

 Observe que a definição é genérica, sendo adequada desde a cenas naturais até fontes de energia eletromagnética, como radares, infravermelhos e

sistemas de raio x.



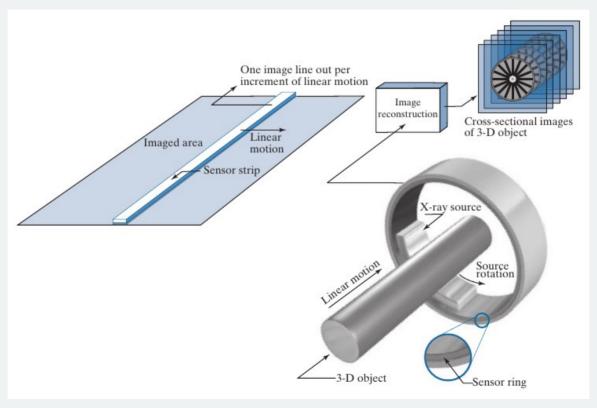
Aquisição de imagens usando faixa de sensores

- Uma faixa de sensores é um mecanismo mais comum que sensores individuais
 - As faixas proveem elementos de imagem em uma direção, enquanto movimentos perpendiculares das faixas proveem imagens em outra direção;
 - Tal mecanismo é usado em sensores de mesa;
 - Esses sensores são comuns em aplicações de sensoreamento aéreo
 - O sistema de imageamento é montado em uma aeronave que voa a uma altitude e velocidade constantes sobre a área geográfica a ser imageada.

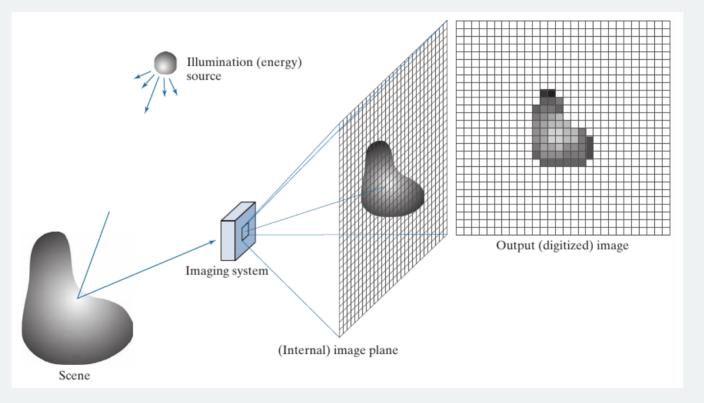
Aquisição de imagens usando faixa de sensores

- Uma faixa de sensores produz uma linha da imagem por vez e o movimento relativo das faixas em outra direção produzem imagens em 2-D.
- Sensores em faixa em uma configuração de anel são usados em setores médicos e industriais, para obter imagens transversais ("fatias") de objetos 3-D
 - Um raio-x rotativo provê iluminação, enquanto sensores de raio-x do lado oposto coletam a energia que passa sobre o sensores;
 - O mecanismo acima descrito é base para tomografias axiais computadorizadas (computerized axial tomography – CAT).

Aquisição de imagens usando faixa de sensores



- Dispositivos eletromagnéticos, ultrasônicos e câmeras digitais frequentemente são dispostos na forma de arrays em 2 dimensões;
 - Câmeras digitais frequentemente utilizam sensores do tipo CCD (dispositivos de carga acoplada), que podem ser fabricados com uma ampla gama de propriedades de detecção, embalados em matrizes 4.000 x 4.000 elementos.
- Para aquisição de imagens usando arrays, não é necessário movimentação dos sensores, uma vez que os mesmos já estão dispostos em grid.



- A resposta de cada sensor é proporcional à da energia da luz projetada em usa superfície do sensor
 - Tal mecanismo, dependendo do tempo de exposição à luz, pode gerar ruídos indesejados;
 - A redução de ruído é obtida ao permitir que o sensor integre o sinal de luz de entrada durante minutos ou mesmo horas.

- Em imagens de telescópios, o tempo de exposição pode atingir horas, permitindo a menor quantidade de ruídos possíveis
 - A imagem ao lado, produzida pelo telescópio espacial James Webb, levou 12,5h para ser tirada.



Fonte: Vox (2022).

Modelos de Imagens

Representação de imagens

• Uma imagem é uma função bidimensional de intensidade de luz f(x, y), onde

$$0 < f(x, y) < \infty$$

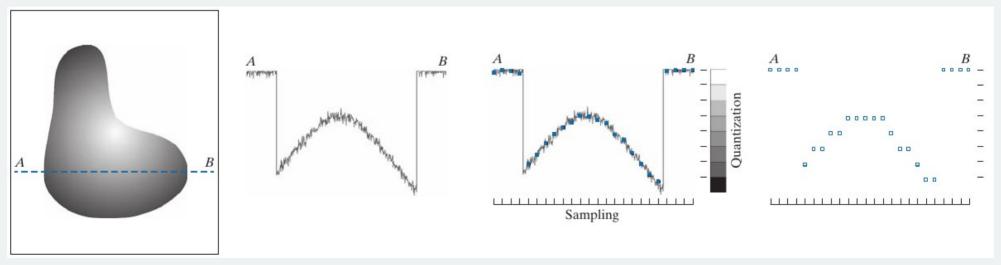
- A natureza dos tons da imagem pode ser caracterizada por duas componentes:
 - Intensidade luminosa, i(x, y), dependende da fonte de energia
 - Reflectância, r(x, y), que depende das propriedades do material.

$$f(x, y) = i(x, y) \times r(x, y)$$

, onde
$$0 < i(x, y) < \infty \ e \ 0 < r(x, y) < 1$$
.

Amostragem e Quantização

- Considere a imagem mais a esquerda, que deve ser convertida para a forma digital
 - Para digitalização, devem ser geradas funções que representem as coordenadas xy e a amplitude.



31

Amostragem e Quantização

- A digitalização das coordenadas é denominada amostragem
 - Deve ser feita em função de linhas e colunas (M_{linhas} × N_{colunas});
- A digitalização das amplitude é denominada quantização
 - No número de níveis utilizados para representar cada pixel, denotado por $G = 2^m$, onde m é a quantidade de bits usada para representação.
 - Com m=8, cada pixel pode representar até 256 tons de cinza, suficientes para distinção pelo olho humano.

Amostragem

- A redução de pontos de amostragem pode causar a perda de informação e detalhes da imagem
 - Como possível "benefício", reduz-se também o espaço necessário para armazenamento da imagem.



Fonte: Wimarshika Thamali (2020).

Quantização

- A redução da quantidade de bits para representação dos pixels, reduz a qualidade da imagem em relação às cores (ou ao tom de cinza)
 - Podem criar falsos contornos, em imagens com 16 tons de cinza ou menos.

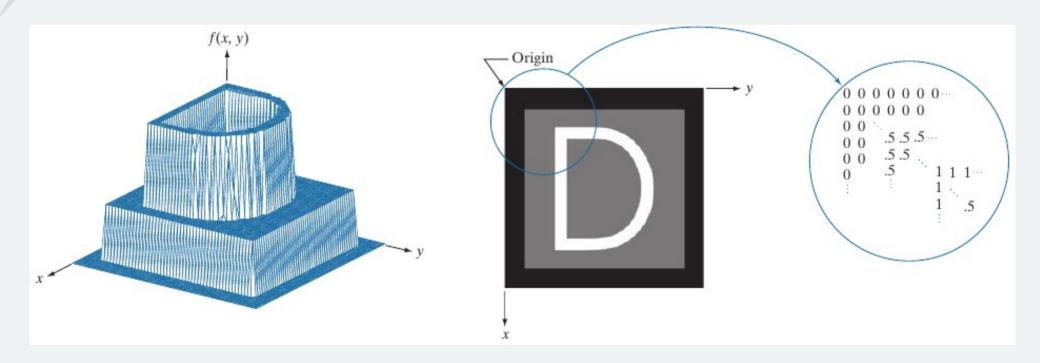


Fonte: Matan Weksler (2020).

Representação de imagens

- Considere uma imagem contínua f(s, t), de duas variáveis s e t;
- A imagem é convertida para uma imagem digital por meio de amostragem e quantização, produzindo f(x, y), com coordenadas discretas
 - A seção do plano real abrangida pelas coordenadas é denominada domínio espacial, com x e y sendo referidos como coordenadas espaciais.
- Uma imagem digital pode ser representada das seguintes formas:
 - Plotagem da imagem como uma superfície;
 - Array com intensidade visual;
 - Array numérico 2-D (matriz).

Representação de imagens



36

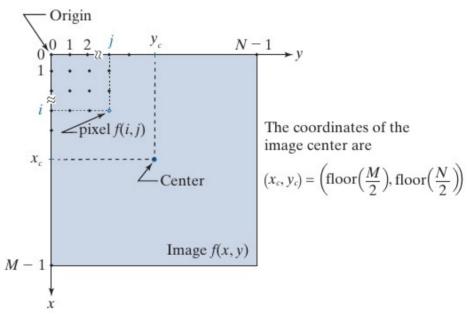
Representação de imagens

- A representação numérica usando matrizes é frequentemente utilizada para processamento computacional;
 - Cada elemento da matriz M × N é um pixel;
 - A origem da matriz é definida no canto esquerdo superior (ordem de varredura para exibição de pixels em monitores e TVs);

$$\mathbf{A} = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & & \vdots \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{bmatrix}$$

Representação de imagens

- Para representação das posiçoes na matriz, Utiliza-se a nomenclatura x e y intercambiavelmente com linhas e colunas.
- Podemos perceber que a imagem tem seu eixo cartesiano rotacionado em 90°, permitindo que a origem esteja no canto superior esquerdo.



Saturação e Ruído

Saturação, Ruído e Contraste

- A faixa de valores abrangidos pela escala de cinza é chamada de faixa dinâmica, um termo usado de maneiras diferentes em diferentes campos;
- Define-se como faixa dinâmica a razão do máximo mensurável de intensidade até o nível mínimo de intensidade detectável no sistema;
 - O limite superior é determinado pela saturação;
 - O limite inferior é determinado pelo ruído (apesar de ruídos estarem presentes em intensidades mais claras);
 - Como as regiões mais escuras são aquelas onde houve menor captura de luminosidade, esta região é aquela onde, frequentemente, há maior quantidade de ruído.

Saturação, Ruído e Contraste

- A imagem ao lado mostra um exemplo de ruído (noise) e saturação (saturation)
 - Pode-se observar que os ruídos ocorrem principalmente em áreas escuras das imagens;
 - A saturação é o limite superior de intensidade luminosa.

Noise Saturation

Saturação, Ruído e Contraste

- A faixa dinâmica estabelece a diferença entre os níveis mais baixos e altos de intensidade que podem ser representados e que a imagem pode ter
 - O contraste da imagem corresponde à diferença de intensidade entre os níveis mais altos e mais baixos;
 - A taxa de contraste é a razão entre ambas as quantidades;
- Quando uma grande quantidade de pixels de uma imagem possui uma alta faixa dinâmica, espera-se que a imagem possua alto contraste
 - Imagens com baixa faixa dinâmica possuem aparência cinza, opaca e desbotada.

 Para armazenamento de imagens monocromáticas, o espaço necessário em termo de bits b, é:

$$b = M \times N \times k$$

- O valor de k corresponde à quantidade de bits necessários para armazenamento da intensidade (tonalidade);
- Quando uma imagem possui 2^k níveis de intensidade, uma prática comum é definila como uma "imagem de k bits" (para 256 cores, a imagem é definida como uma "imagem de 8 bits").

• Ex. 1: Para armazenar uma imagem de 640x480 monocromática de 256 cores, qual a quantidade de bits necessários?

- Ex. 1: Para armazenar uma imagem de 640x480 monocromática de 256 cores, qual a quantidade de bits necessários?
 - Para 256 cores, serão necessários 8 bits (28);
 - Logo, teremos 640x480x8 = 2.457.600 bits
 - Em bits: 2.457.600
 - Em bytes: 307.200
 - Em kilobytes: 300
 - Em megabytes: 0,3

Resolução Espacial e de Intensidade

Resolução Espacial

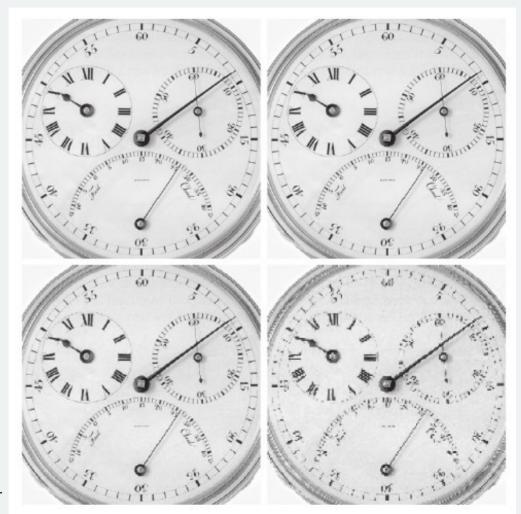
- A Resolução Espacial é a medida do menor detalhe discernível em uma imagem
 - Quantitativamente, pode ser expressa de várias maneiras, sendo pares de linhas por unidade de distância e pixels (dots) por unidade de distância as medidas mais comuns;
 - A unidade pixels por polegadas (DPI dots per inchs) é, talvez, a métrica de resolução mais utilizada para impressão e publicação de informações.
 - Jornais utilizam, em geral, 75 dpi, enquanto revistas utilizam resolução de 133 dpi.

Resolução Espacial

- Resoluções espaciais, para que façam sentido, precisam ter unidades de medidas de espaço a elas relacionadas
 - Uma imagem com resolução de 1024 × 1024 pixels não contém informações suficientes para indicar a qualidade da mesma;
 - Se a imagem for impressa em tamanho A4 e A3, para a mesma imagem, a qualidade será consideravelmente diferente.

Resolução Espacial

- Imagens de mesmo tamanho com resoluções espaciais diferentes possuem qualidade diferente
 - As imagens à direita possuem resoluções de 930, 300, 150 e 72dpi, respectivamente.

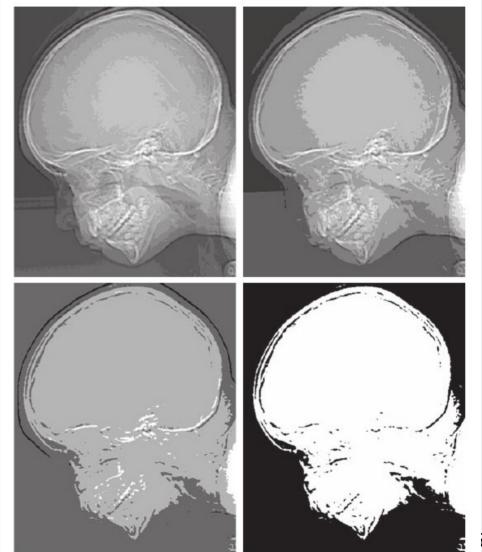


Resolução de Intensidade

- A Resolução de Intensidade se refere à menor mudança perceptível no nível de intensidade.
 - Níveis de intensidade, usualmente definidos em potência de 2 com 8 bits (28), são utilizados para especificar as variações de tonalidade dos pixels;
 - Quantidades superiores de bits para representação de pixels são possíveis, porém pouco usuais;
 - A resolução de intensidade normalmente é definida em função do número de pixels utilizados para quantizar os valores.

Resolução de Intensidade

- Imagens com resoluções de intensidade diferentes, para um valor de resolução espacial, constante possuem qualidades diferentes.
 - As imagens ao lado possuem 16,
 8, 4 e 2 níveis de cores,
 respectivamente.



Interpolação de Imagens

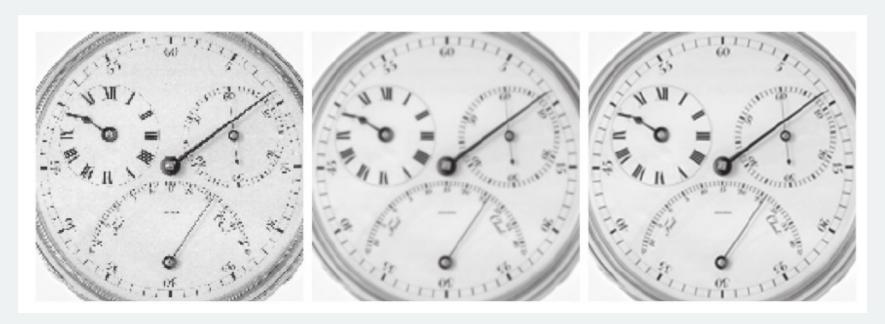
- Interpolação é o processo de estimar pontos desconhecidos com base no pontos existentes
 - No contexto de imagens, a interpolação é utilizada para estimar o valor dos pixels, principalmente em tarefas de ampliação de imagens;
 - Interpolação também pode ser usada em tarefas de redução, rotação e correção de imagens digitais;

Interpolação de Imagens

- Os métodos de interpolação mais comuns são:
 - Nearest neighbor interpolation (interpolação pelo vizinho mais próximo): define para cada nova localização a intensidade do vizinho mais próximo;
 - Possui como desvantagem a produção de efeitos indesejados e distorções;
 - Bilinear interpolation: quatro dos vizinhos mais próximos são utilizados para estimativa da intensidade em um determinado pixel;
 - Bicubic interpolation: envolve a estimativa de intensidade dos 16 vizinhos mais próximos para estimativa de intensidade de um determinado pixel.

Interpolação de Imagens

• As imagens abaixo foram reduzidas para 72 dpi e, em seguida, retornadas para 930 dpi usando interpolação pelo vizinho mais próximo, bilinear e bicubica.



Relações básicas entre pixels

Relações básicas entre pixels

 Um pixel p com coordenadas (x, y) possui vizinhos verticais e horizontais com as seguintes coordenadas

$$(x+1, y), (x-1, y), (x, y+1), (x, y-1)$$

Esse conjunto é denominado 4-conectado (4-neighbors), denotado por N₄(p).

Relações básicas entre pixels

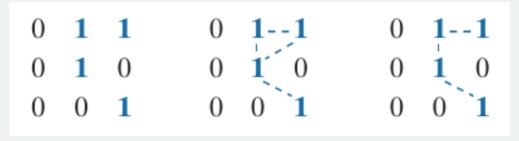
• Além das 4 vizinhanças anteriores, é possível adicionar as vizinhanças nas diagonais (x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)

• Esse conjunto é denominado **8-conectado** (8-*neighbors*), denotado por $N_8(p)$.

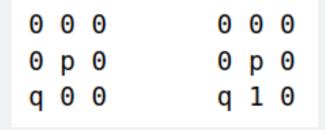
- As relações de adjacência entre pixels são baseadas em um critério de similaridade entre os níveis de cinza semelhantes (definido pela letra "V");
- Para uma imagem binária, V = 1;
- Para uma imagem em tons de cinza, o conjunto é determinado por uma faixa de tons de cinza, V = 16, 17,;

- Tipos de adjacência:
 - **4-adjacência**: 2 pixels, $p \in q$, com valores em V são 4-adjacentes se q está no conjunto $N_4(p)$;
 - 8-adjacência: 2 pixels, $p \in q$, com valores em V são 8-adjacentes se q está no conjunto $N_8(p)$;
 - M-adjacência (adjacência mista): 2 pixels, p e q, com valores em V são madjacentes se:
 - q está no conjunto N₄(p), ou
 - q está em $N_D(p)$ e o conjunto $N_4(p) \cap N_4(q)$ não possuem pixels cujos valores são de V.

- A adjacência mista surgiu como uma modificação da 8-adjacência para eliminar possíveis ambiguidades da última.
 - Entenda por ambiguida, múltiplas conexões entre os vértices;
 - Se um vértice possui 2 relações, a adjacência é definida como ambígua.



- Para entender melhor as ambiguidades possíveis na 8 adjacência, analise o exemplo abaixo:
 - Em ambos os casos, as matrizes são 8-adjacentes;
 - No entanto, somente no primeiro caso a matriz é m-adjacente, se considerarmos que V = {1}.



Fonte: Próprio autor.

Caminho Digital (Curva)

• Um caminho digital (curva) de um pixel p, com coordenadas (x_0 , y_0) para um pixel q, com coordenadas (x_n , y_n) é uma sequência distinta de pixels com coordenadas

$$(x_0, y_0), (x_1, y_1), ..., (x_n, y_n)$$

- onde os pontos (x_i, y_i) e (x_{i-1}, y_{i-1}) são adjacentes para $1 \le i \le n$.
- Nesse caso, n é o tamanho do caminho;
- Se $(x_0, y_0) = (x_n, y_n)$, então o caminho é um caminho fechado;
- É possível definir 4-, 8- ou m-caminhos, dependendo do tipo de adjacência utilizada.

Componentes conexos

- Dois pixels p e q são conexos em S se existe um caminho entre eles consistindo somente de pixels em S;
- Para cada pixel p em S, o conjunto de pixels que são conectados a ele em S são chamados de componentes conexos de S;
- Se existe somente um componente e esse componente é conexo, então S é chamado de conjunto conexo.

Regiões

- Considere que R represente um subconjunto de pixels.
 - Esse subconjunto de pixels é denominado região da imagem se R for um conjunto conexo;
 - Duas regiões R_i e R_j são adjacentes se sua união forma um conjunto conexo;
 - Regiões que não são adjacentes são denominadas disjuntas (em relação ao tipo de adjacência – 4 ou 8 adjacência).

Foreground e Background

- Suponha que uma imagem contenha K regiões disjuntas, R_k , onde k = 1, 2, ..., K;
- Seja R_u, a união de todas as K regiões e (R_u)^c o complemento correspondente
 - Todos os pontos em R_u são denominados foreground ("primeiro plano" ou "frente") e todos os pontos em (R_u)^c pertencem ao background (fundo) da imagem.

0	0	0	0	0
0	1	1	0	0
0	1	1	0	0
0	1	(1)	1	0
0	1	1	1	0
0	0	0	0	0

Bordas

- As bordas (borders, boundaries ou contours) de uma região R são o conjunto de pixels em R que são adjacentes aos pixels do complemento de R
 - Bordas, de forma alternativa, podem ser definidas como o conjunto de pixels de uma região que possuem ao menos um vizinho no background;

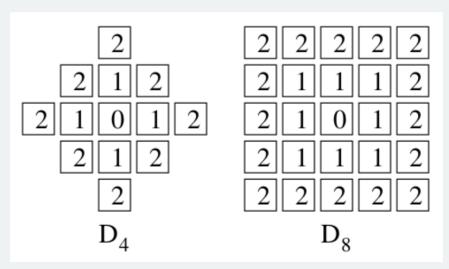
 A definição anterior é classificada como borda interna, enquanto a borda pelo visão do background é definida como borda externa.

Edge

- Edges (limites ou, eventualmente, bordas) é associdado com bordas e regiões
 - Bordas de regiões finitas formam um caminho fechado e são definidas como um conceito global;
- Um edge é um conceito local, baseado na medida do nível de descontinuidade de um ponto
 - Edges são associados a segmentos, de forma que são linkados de uma forma que eles correspondem a bordas (porém nem sempre)
 - Dependendo do tipo de conector e da conectividade usada, o limite (edge) de uma região binária será o mesmo de uma borda de uma região.

Medidas de Distância

- As medidas de distância entre os pixels p(x, y) e q(s, t), podem ser definidas como:
 - Euclidiana: $D_e(p, q) = \sqrt{((x-s)^2 + (y-t)^2)}$
 - D_4 (quarteirão): $D_4(p, q) = ||x-s|| + ||y-t||$
 - D_8 (tabuleiro): $D_8(p, q) = max(||x-s||, ||y-t||)$



Referências

Referências

- Rafael C. Gonzalez, Richard E. Woods. Digital Image Processing 4th Edition.
 2018. Pearson. ISBN: 978-9353062989.
- Agostinho Brito Jr. Processamento digital de imagens Slides de Aula.
 2018.
- DAPI Diagnóstico Avançado por Imagens. Curiosidades sobre o raio X. 2023.
 Disponível em https://dapi.com.br/curiosidades-sobre-o-raio-x/
- Joab Silas da Silva Júnior. **O que é infravermelho?** 2023. Disponível em: https://brasilescola.uol.com.br/o-que-e/fisica/o-que-e-infravermelho.htm.
- Vox. Nasa releases the first Image of James Webb Space Telescope. 2022.
 Disponível em: https://twitter.com/voxdotcom/status/1546866088893329408

Referências

- Wimarshika Thamali. Sampling & Quantization in Digital Image Processing.
 2020. Disponível em: https://wimarshikathamali1995.medium.com/sampling-quantization-in-digital-image-processing-8c4490357039
- Matan Weksler. Deep Dive Into Multi-Bit Weighted Quantization for CNNs.
 2020. Disponível em: https://medium.com/swlh/deep-dive-into-multi-bit-weighted-quantization-for-cnns-d2723afdc5db
- Retina Curitiba. **Quais são as partes que compõem o olho humano?** 2023. Disponível em: https://retinacuritiba.com.br/quais-sao-as-partes-que-compoem-o-olho-humano/