

Notas de aula de Processamento de imagens e Visão Computacional

Daniel Oliveira Dantas

24 de março de 2020

Sumário

1	Introdução	1
1.1	O que é processamento de imagens	1
1.2	Origens do processamento digital de imagens	1
1.3	Áreas que usam processamento digital de imagens	2
1.4	Passos fundamentais no processamento digital de imagens	2
1.5	Componentes de um sistema de processamento de imagens	2
2	Fundamentos de imagens digitais	4
2.1	Elementos de percepção visual	4
2.2	Luz e o espectro eletromagnético	4
2.3	Aquisição de imagens	5
2.4	Amostragem e quantização	5
2.5	Relações entre pixels	6
2.6	Ferramentas matemáticas	7

Capítulo 1

Introdução

1.1 O que é processamento de imagens

- O que é uma imagem?
 - Uma matriz
 - Uma função
 - Esses conceitos podem ser generalizados para mais dimensões
- Uma definição de processamento de imagens: processo em que tanto a entrada quanto a saída são imagens.
 - Essa definição não engloba, porém, processos como valor médio da imagem, extração de pontos característicos, alguns tipos de reconstrução 3D, aplicações de segurança que detectam atividades suspeitas, reconhecimento de gestos, reconhecimento de caracteres (OCR) e outras aplicações consideradas do campo de visão computacional.
- Uma definição mais abrangente: processos de baixo, médio e alto nível em que tanto a entrada quanto a saída são imagens.
 - Baixo nível: envolve operações primitivas, de redução de ruído, aumento de contraste, aumento de nitidez (*sharpening*), *thresholding* etc. Nos processos de baixo nível, tanto a entrada quanto a saída são imagens.
 - Médio nível: envolve tarefas como segmentação (particionamento), redução dos objetos a uma descrição ou formato apropriado para processamento, classificação (reconhecimento) de objetos. A entrada são imagens, e a saída são atributos extraídos dessas imagens, como bordas, contornos, características ou classe de objetos individuais.
 - Alto nível: envolve atividades cognitivas associadas com a visão. Cognição envolve atividades como memória, compreensão, aprendizado, raciocínio, atenção, resolução de problemas e tomada de decisão.

1.2 Origens do processamento digital de imagens

- 1920: *Bartlane cable picture transmission system*, 5 tons de cinza.
- 1929: idem, 15 tons de cinza.
- 1948: invenção do transistor.
- 1950 a 1960: invenção das linguagens de programação de alto nível, COBOL e Fortran.
- 1958: invenção do circuito integrado.
- 1964: primeira foto da Lua tirada de uma sonda.
- 1968 a 1971: invenção dos primeiros microprocessadores, CADC, TMS1000, 4004.
- 1969: invenção do CCD.
- 1971: primeira tomografia computadorizada

1.3 Áreas que usam processamento digital de imagens

- Medicina, astronomia, meteorologia, indústria, fotografia, editoração, segurança etc.
- Técnicas de obtenção de imagem:
 - Eletromagnética: luz, radiação UV, radiação IR, raios X, raios gama, microondas.
 - Eletrônica: microscopia eletrônica.
 - Mecânica: ondas acústicas, ultrassom.
 - Sintética: computação gráfica, fractais.

1.4 Passos fundamentais no processamento digital de imagens

- Aquisição
- Melhoramento (*image enhancement*)
- Restauração
- Processamento de cores
- *Wavelets* e processamento multirresolução
- Compressão
- Processamento morfológico
- Segmentação
- Representação e descrição
- Reconhecimento de objetos

1.5 Componentes de um sistema de processamento de imagens

- Sensores + *hardware* especializado
- Computador + GPU
- Armazenamento em massa
- *Software* de processamento de imagens
- Monitor de imagem
- Impressora
- Rede

Capítulo 2

Fundamentos de imagens digitais

2.1 Elementos de percepção visual

2.2 Luz e o espectro eletromagnético

- Frequência: f
- Comprimento de onda: λ
- Velocidade da luz: $c = 3 \times 10^8 \text{ m/s}$
- Constante de Planck: $h = 6.6 \times 10^{-34} \text{ Js}$ ($\text{m}^2\text{kg/s}$), $E = hf$.

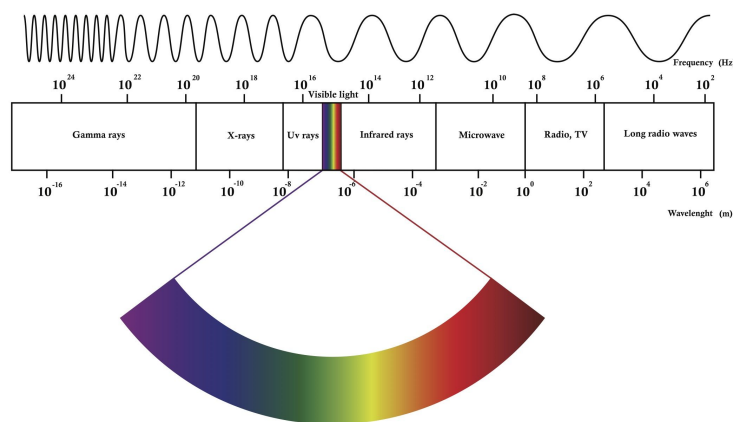


Figura 2.1: Espectro eletromagnético

2.3 Aquisição de imagens

- Normalmente é feita através de componentes sensíveis a alguma faixa específica do espectro eletromagnético
 - Sensores simples: fotodiodo, fototransistor.
 - Sensores lineares: CCD linear.
 - Sensores em matriz: CCD de câmeras fotográficas, mouse ótico.

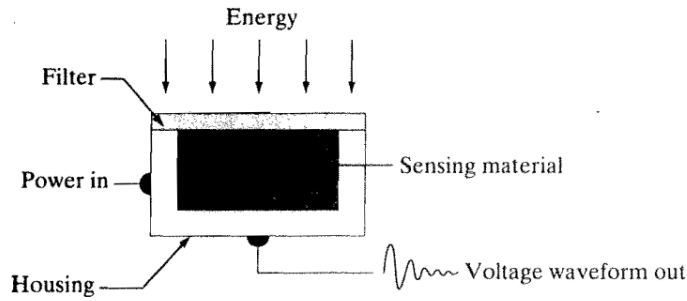


Figura 2.2: Esquema de um sensor

2.4 Amostragem e quantização

- Amostragem: digitalização dos valores das coordenadas, tanto no espaço quanto no tempo.
- Quantização: digitalização dos valores das amplitudes
- Representação:
 - Matriz

$$f(i, j) = \begin{bmatrix} f(0, 0) & f(0, 1) & \dots & f(0, N-1) \\ f(1, 0) & f(1, 1) & \dots & f(1, N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1, 0) & f(M-1, 1) & \dots & f(M-1, N-1) \end{bmatrix}$$

- Função

$$f : \mathbb{Z}^2 \rightarrow \mathbb{Z}$$

$$f : \{0, \dots, M-1\} \times \{0, \dots, N-1\} \rightarrow \{0, \dots, L-1\}$$

onde $L = 2^k$. L é o intervalo dinâmico do sensor e k é o número de bits necessário para representar L .

- Número de bits necessários para armazenar uma imagem $M \times N$:

$$b = MNk$$

- Número de bytes necessários para armazenar uma imagem $M \times N$:

$$B = MNk/8$$

- Redimensionamento de imagem: *zoom* ou *resize*
 - *Nearest neighbor*: ampliar uma imagem replicando cada pixel várias vezes.
 - *Bilinear*: ampliar uma imagem inserindo pixels calculados da interpolação linear entre os pixels mais próximos.

2.5 Relações entre pixels

- 4-adjacência: os vizinhos de (x, y) são $(x, y - 1)$, $(x, y + 1)$, $(x - 1, y)$ e $(x + 1, y)$.
- D-adjacência: os vizinhos de (x, y) são $(x - 1, y - 1)$, $(x - 1, y + 1)$, $(x + 1, y - 1)$ e $(x + 1, y + 1)$.
- 8-adjacência: os vizinhos de (x, y) são a união da 4-adjacência e da D-adjacência.

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

Tabela 2.1: Disposição dos pixels na vizinhança de (x, y) .

- Distância: é uma relação entre dois pixels p e q com as propriedades listadas abaixo.

1. $D(p, q) \geq 0$
2. $D(p, q) = 0 \Leftrightarrow p = q$
3. $D(p, q) = D(q, p)$ simetria
4. $D(p, q) \leq D(p, z) + D(z, q)$ desigualdade triangular

- Distância Euclidiana: $D_e(p, q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}$
- Distância *city-block*: $D_4(p, q) = |p_x - q_x| + |p_y - q_y|$
- Distância *chessboard*: $D_8(p, q) = \max(|p_x - q_x|, |p_y - q_y|)$

2.6 Ferramentas matemáticas

- Operações lineares: seja H uma operação cuja entrada e saída são imagens, sejam f e g duas imagens, e a e b , dois escalares. A operação H é dita linear se segue a relação abaixo

$$H(af + bg) = aH(f) + bH(g)$$

- Operações não-lineares: uma operação é dita não-linear se não é linear.