

Capítulo 1

Introdução

Entendemos por **informações gráficas** as informações contidas em figuras ou imagens geradas sinteticamente (**imagens sintéticas**) ou capturadas pelos equipamentos sensíveis a sinais eletromagnéticos (p.ex., telescópios, microscópios, tomógrafos) ou acústicos (p.ex., microscópios acústicos) que operam de forma similar à visão humana (**imagens reais**), com o objetivo de transmitir uma idéia (concepção) ou registrar um fato real.

As imagens encontram aplicações em: projetos de engenharia sintetizados em desenhos técnicos/plantas; divulgação de produtos/serviços por cartazes ou filmes; desenhos animados ou filmes de ficção científica; realce e retoque das imagens capturadas pelos satélites (a primeira aplicação foi o retoque das imagens (digitalizadas) de jornais recebidas via cabo entre Londres e Nova Iorque); realce das imagens de transmissão como as imagens médicas para diagnóstico de disfunções; reconhecimento de assinaturas e impressões digitais; e previsão de tempo através da interpretação das imagens de satélites).

Hoje em dia, as imagens são objetos de pesquisa de três domínios de conhecimento: Computação Gráfica, Processamento de Imagens e Visão Computacional. Computação Gráfica foca na geração de imagens sintéticas com uso de computadores. Processamento de Imagens centra no realce e no filtro dos dados contidos nas imagens para facilitar a sua interpretação/compreensão ou simplesmente para melhorar a sua qualidade. Visão Computacional, por sua vez, procura emular os efeitos da visão humana – percepção, interpretação e reconstrução 3D – por meio de máquinas autônomas, ou seja, preocupa-se com a análise de imagens

O objetivo desta disciplina é prover uma visão introdutória destas três áreas correlatas de conhecimento.

1.1 Formação de Imagens

Dois elementos são fundamentais na formação de uma imagem: objeto de interesse e observador. O primeiro elemento é o elemento imageado e o segundo, o elemento que produz a imagem. Usualmente, o segundo elemento é conhecido por **observador**.

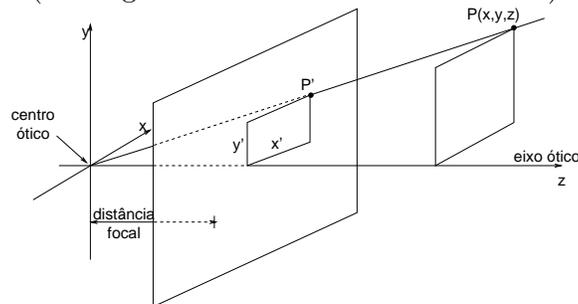
Um exemplo clássico de observador é a visão humana: um sistema complexo que produz imagens intrinsecamente bidimensionais na retina da parte interna de cada olho por meio do cristalino (lente biconvexa flexível), constituindo assim um dos cinco sentidos utilizados pelos homens para perceber e entender o mundo que os rodeia. A retina é uma camada muito fina, sensível à luz, formada pela ramificação do nervo ótico que transmite as sensações luminosas ao cérebro.

(Ver Fig. 2.1 do livro-texto do Gonzalez.)

Observação 1.1 *Distinguem-se na retina dois tipos de células sensíveis à luz: os cones e os bastonetes. A região de máxima sensibilidade à luz só contém cones. Os bastonetes são destinados à visão com pouca luz, sendo insensíveis às diferenças de cor.*

A imagem que se forma na retina é real, invertida e menor do que o objeto. Entretanto, “vemos” os objetos em posição correta graças a uma “educação” do cérebro, pois relacionamos as nossas impressões não com a imagem mas com a extremidade do raio visual, onde se encontra o próprio objeto.

(Ver Fig. 2.3 do livro-texto do Gonzalez.)

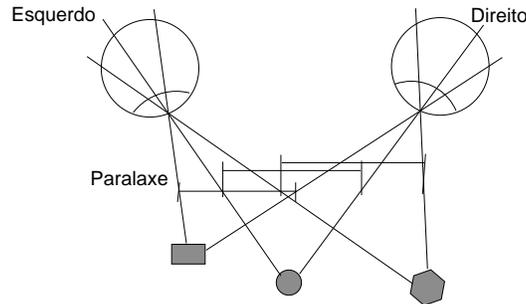


A nossa visão é de fato estéreo. Quando olhamos para um objeto, são formadas duas imagens retinianas, uma em cada olho. Devido ao afastamento entre os dois olhos, as imagens não são idênticas.

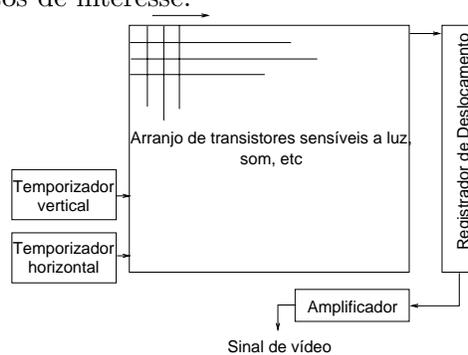
(Ver Fig. 2.21 do livro-texto do Gonzalez.)

A observação simultânea dessas duas imagens ligeiramente diferentes força movimentos musculares dos olhos para a paralaxe na distância en-

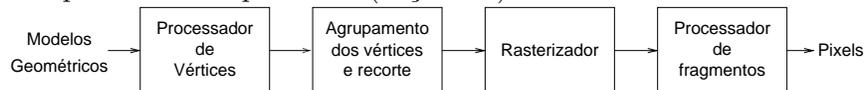
tre dois pontos de diferentes alturas. Estes movimentos permitem que o cérebro diferencie a distância entre tais pontos, dando a sensação de relevo/profundidade.



O modelo de visão humana tem sido utilizado na construção de uma variedade de sistemas de imageamento, como câmeras fotográficas, que procuram reproduzir as sensações percebidas pelo cérebro em outras mídias de armazenamento. O princípio de funcionamento destes dispositivos se baseia no uso de componentes fotosensíveis para distinguir os sinais luminosos oriundos dos objetos de interesse.



Uma outra alternativa para formação de imagens seria com uso de **sistemas gráficos**, que emula o processo de formação de imagens do sistema de visão humano. Neste paradigma, tanto os objetos a serem imageados quanto o observador são sintéticos, como mostra o seguinte fluxo de dados para dispositivos do tipo *raster* (Seção 1.4).



No contexto desta última alternativa, tem-se aprofundado o estudo da habilidade da visão humana para desenvolver tecnologias capazes de produzir imagens sintéticas foto-realistas. Embora esta habilidade seja um fenômeno psicofísico, variando com a sensibilidade de cada indivíduo, constatou-

se que a visão humana apresenta limitações em:

- **acuidade:** é a capacidade para distinguir os detalhes. A visão humana é menos sensível às variações lentas ou bruscas de brilho numa imagem.

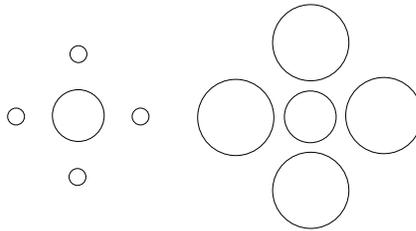
(Ver Fig. 2.7 do livro-texto do Gonzalez.)

Observação 1.2 *Esta característica da percepção humana agrava o fenômeno conhecido como **Bandas de Mach**, quando quantizamos a luminância/brilhância de uma imagem.*

- **percepção de contraste:** é a razão entre a brilhância média do objeto de interesse e a brilhância do fundo. A brilhância percebida depende do contraste.

(Ver Fig. 2.8 do livro-texto do Gonzalez.)

- **precisão:** a disposição geométrica dos contornos pode induzir percepções diferenciadas de uma mesma figura. A ilusão de Ebbinghaus é um exemplo bem conhecido.



- **sensibilidade a cores:** sob condições luminosas normais, o olho humano é mais sensível a cor do que a luminância. Particularmente, o olho humano é mais sensível a cores próximas à faixa espectral correspondente a cor verde-amarelada.

1.2 Modelos de Imagens

Modelos matemáticos são frequentemente utilizados para descrever e processar as imagens em sistemas computacionais. Uma imagem planar (bidimensional) pode ser tratada como uma função f que depende de duas variáveis (coordenadas x e y de um plano) ou de três variáveis (quando ela varia com o tempo, uma terceira variável t é adicionada). O valor da função pode ser um

valor escalar ou um vetor de valores reais, dependendo das informações contidas em cada imagem. O valor de uma **imagem monocromática** pode ser um valor (escalar) da luminância (*luminance*), enquanto numa imagem **colorida** ou **multiespectral** o valor da função pode ser um vetor de n valores reais, cada qual corresponde a uma cor espectral (brilhância – *brightness*).

Considerando uma imagem como uma função, podemos classificá-la em:

imagem contínua: se o domínio e o contra-domínio da função são contínuos,

imagem discreta: se o domínio da função é discreto, e

imagem digital: se o domínio e o contra-domínio da função são discretos.

A discretização do domínio de coordenadas (x, y) é conhecida como **amostragem** e a discretização do contra-domínio, ou seja dos valores ou da amplitude de f , é chamada **quantização**. Particularmente, quando a função f só assume dois valores, dizemos que a imagem é *binária* ou preto-e-branco.

De acordo com a quantidade de elementos no domínio e no contra-domínio, é possível caracterizar o “grau de detalhamento” de uma imagem digital em termos de **resolução**. Distinguem-se quatro tipos de resolução:

resolução espacial (amostragem): define a proximidade entre as amostras de uma imagem discreta/digital (a disposição espacial das amostras no plano forma um reticulado quadrado ou hexagonal).

resolução espectral (quantização): define a quantidade de cores espectrais existentes na imagem.

resolução radiométrica (quantização): define a quantidade de níveis de intensidade (luminância) distinguíveis.

resolução temporal (amostragem): define o intervalo entre duas amostras de imagens. É útil para caracterizar imagens dinâmicas.

1.3 Tipos de Imagens

De acordo com o valor (tipo de informação) contido em cada ponto (x, y) de uma imagem processável por computadores digitais, podemos classificá-la em uma das seguintes imagens:

imagens convencionais ou fotos: a informação em cada ponto da imagem é a sua brilhância (intensidade luminosa captada pelo dispositivo de aquisição), obtida em função da intensidade dos feixes luminosos e da posição e da orientação deste ponto em relação aos feixes incidentes e a posição da câmera/do observador. São também conhecidas como **imagens de intensidade**.

imagens sintéticas: a informação (brilhância) em cada ponto da imagem é determinada a partir dos modelos sintéticos ou das imagens existentes. O uso das imagens existentes tem sido um recurso eficaz para aumentar o realismo das imagens sintéticas.

Imagens de transmissão: a informação em cada ponto da imagem é função de algumas propriedades físicas dos objetos de interesse capazes de alterar de forma diferenciada algum fenômeno físico. Por exemplo, objetos que apresentam propriedades ópticas (difração, reflexão e refração) muito diferenciadas podem modificar de forma perceptível a direção de um feixe luminoso homogêneo que passa por eles (raios X), ou objetos que apresentam comportamentos magnéticos bem distintos podem ter a sua polaridade alterada quando sujeitos a um intenso campo magnético (imagens de ressonância magnética – MRI). Essas diferenças são processadas e codificadas em imagens.

Imagens de profundidade: o valor em cada ponto é a profundidade do ponto em relação ao dispositivo de captura.

Observação 1.3 *Luminância ou brilhância é a razão entre a intensidade luminosa do objeto na direção considerada e a área da projeção do objeto sobre um plano perpendicular a esta direção. Na literatura de sistemas de informações gráficas pode encontrar a distinção no uso dos dois termos: luminância para imagens em níveis de cinza (monocromáticas) e brilhância para imagens coloridas.*

1.4 Exibição de Imagens

Essencialmente, existem duas classes de dispositivos para exibição de imagens:

1. vetoriais (década 60 a década 80), fundamentados no modelo *pen-plotter*. No caso de tecnologia de tubos catódicos (CRT), o controlador de vídeo governa o movimento (aleatório) dos canhões de raios

catódicos e a intensidade destes raios de acordo com as instruções existentes na memória (de desenho). Para evitar cintilamentos na imagem, deve-se assegurar uma taxa de regeneração maior que 30Hz.

(Ver Fig. 1.1 do livro-texto do Foley.)

(Ver Figs. 4.12 – 4.14 do livro-texto do Foley.)

2. digitais (*raster*): fundamentados no modelo *pixel-filling*. No caso de tecnologia de tubos catódicos, o movimento dos canhões é sempre o mesmo (varredura horizontal e retraço) e a função do controlador se limita a monitorar a intensidade dos raios catódicos de acordo com os atributos de cada pixel (*picture element*). Com isso, foi possível desenvolver a tecnologia de monitores **entrelaçados** (as linhas pares e ímpares são regeneradas alternadamente) que duplica a taxa de regeneração em relação aos **não-entrelaçados** (todas as linhas são atualizadas em cada regeneração). Para aumentar a eficiência do sistema armazenam-se os atributos de cada pixel numa memória em separado chamado *frame buffer*.

(Ver Fig. 1.2 do livro-texto do Foley.)

(Ver Figs. 4.18 – 4.22 do livro-texto do Foley.)

Os algoritmos de síntese de imagens são intimamente relacionados com a tecnologia de exibição das imagens. Os dispositivos de saída são hoje predominantemente do tipo *raster*; portanto, nós nos concentraremos nesta disciplina nos métodos de síntese de imagens apropriadas para serem exibidas neles.

1.5 Interação com as Imagens

Os sistemas gráficos, nos quais um usuário pode interagir com os resultados visualizados através de um dos seguintes seis **dispositivos lógicos básicos**, são conhecidos como **sistemas gráficos interativos**:

locator: entra a posição de um ponto da imagem. Exemplos de dispositivos físicos que realizam esta tarefa básica são *mouse* e teclado. Outros dispositivos conhecidos são *joystick*, *trackball* e *tablet*.

stroke: entra uma sequência de pontos. Tipicamente, esta tarefa pode ser realizada com uso de um *mouse*.

pick: seleciona um objeto da imagem. O dispositivo físico que tipicamente realiza esta função é o *mouse*.

valuator: entra um valor numérico. Exemplo típico de um *valuator*] são os *dials* (potenciômetros).

keyboard ou string: entra uma sequência de caracteres alfa-numéricos. O dispositivo físico típico que realiza esta tarefa é o teclado.

choice: entra uma opção dentre um conjunto de alternativas pré-definidas. Exemplo típico de um dispositivo físico que realiza esta tarefa é o *tablet* com funções pré-definidas. É comum, hoje em dia, mapear um conjunto de alternativas às teclas do teclado, tornando-o também um dispositivo lógico de *choice*. Podemos ainda utilizar o *mouse* como um dispositivo lógico de *choice*, se mapearmos cada um dos seus botões a uma das funções pré-definidas ou provermos na interface do aplicativo sub-janelas com as opções pré-definidas.

Pesquisas na área de Interface Homem-Máquina comprovam que a comunicação homem-computador pode ser mais eficiente com auxílio de imagens, o que impulsionou o desenvolvimento de dispositivos de entrada cada vez mais ergonômicos, funcionais e econômicos, desde a idade de *light-pens*.