

Capítulo 1

Introdução

Computação Gráfica consiste de um conjunto de métodos e técnicas computacionais que gera imagens sintéticas ou que processa amostras das informações visuais de um ambiente real. Nesta disciplina centramos na transformação de dados em imagens que podem ser visualizadas por dispositivos gráficos do tipo *raster*. Veremos que o processo pode ser dividido em etapas, para cada qual existem diferentes modelos matemáticos que o tornam computacionalmente tratável.

Em termos de áreas de pesquisa, a Computação Gráfica tem uma grande afinidade com várias outras áreas, como Modelagem Geométrica (representação das formas geométricas dos objetos que aparecem nas imagens), Processamento de Imagens (realce, retoque, efeitos especiais nas imagens) e Visão Computacional (renderização com base em imagens adquiridas de uma cena real).

Hoje em dia, as imagens aparecem em uma diversidade de aplicações, tais como: 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 aéreas ou de satélites.

1.1 Um Pouco da Evolução da Computação Gráfica

- 1885: desenvolvimento da tecnologia do tubo de raios catódicos.

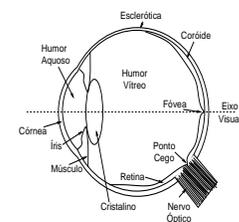
- 1930: é construído o primeiro computador (ENIAC).
- 1947: os Bell Labs inventam o transistor.
- 1950: desenvolvimento do primeiro computador com recursos gráficos pelo MIT (Whirlwind I).
- 1956: o MIT constrói o primeiro computador totalmente transistorizado.
- 1959: surge o termo *Computer Graphics*, criado por Verne Hudson.
- 1962: é divulgada uma das mais importantes publicações da Computação Gráfica: a tese de Ivan Sutherland (*Sketchpad - A Man-Machine Graphical Communication System*)
- 1963: Sutherland apresenta um sistema de desenho interativo de primitivas gráficas 2D baseado em caneta luminosa (*lightpen*). Englebart inventa o dispositivo de interação *mouse*.
- 1965: Roberts cria um algoritmo de remoção de linhas escondidas.
- 1966: surge a primeira empresa de produção de animações e efeitos especiais com uso de computadores.
- 1968: surge a idéia de *ray-tracing*.
- 1969: é criado o grupo de interesse em Gráficos (SIGGRAPH – *Special Interest Group in Graphics*) pela ACM.
- 1971: é publicado o algoritmo de tonalização de *Gouraud*.
- 1974: Catmull desenvolve o algoritmo de *z-buffer*.
- 1975: é publicado o modelo de iluminação de *Phong* e a tonalização de *Phong*.
- 1979: aplicação de *ray-tracing* no tratamento de reflexões especulares e refrações (modelo de iluminação global).
- 1982: Jim Clark e Abbey Silverstone fundam a empresa *Silicon Graphics Inc.*
- 1984: é divulgada a técnica de radiosidade.

- 1985: padronização oficial da especificação GKS-2D para pacotes gráficos 2D.
- 1988: padronização oficial das especificações GKS-3D e PHIGS para pacotes gráficos 3D.
- 1989 (?): almejando uma interface simples e independente dos fabricantes para o desenvolvimento das aplicações gráficas 2D e 3D, SGI propõe a especificação OpenGL API com base nas suas experiências em SGI IRIS GL (*Iris Graphics Language*).
- 1991: a empresa *S3 Graphics Ltd.* desenvolve o primeiro acelerador gráfico 2D para PCs.
- 1992: forma-se o bordo de revisão da arquitetura de OpenGL (ARB – *Architecture Review Board*). OpenGL é uma especificação de API (*application program interface*) para as placas gráficas 2D e 3D.
- 1993: Brian Paul inicia o desenvolvimento da biblioteca gráfica Mesa baseada na especificação de OpenGL.
- 1994: Brian Paul obtém a autorização da SGI para distribuir a biblioteca Mesa através da *Internet*.
- 1999: a empresa nVidia lança a primeira placa gráfica 3D programável (nVidia GeForce 256) para PCs.

Observação 1.1 Em <http://accad.osu.edu/~waynec/istory/timeline.html> encontra-se uma descrição linear da evolução da Computação Gráfica.

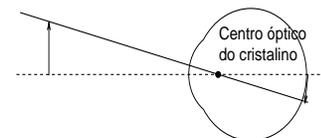
1.2 Imagens

A visão é um dos cinco sentidos (visão, audição, olfato, paladar e tato) mais utilizado pelo homem para perceber e entender o mundo que o rodeia. Esta percepção se dá através da formação de imagens intrinsecamente bidimensionais na retina da parte interna do olho por meio do cristalino (lente biconvexa). A retina é uma camada muito fina, sensível à luz, formada pela ramificação do nervo óptico que transmite as sensações luminosas ao cérebro. Portanto, podemos dizer que a percepção visual decorre da distinção da brilhância (ou luminância) dos objetos de interesse.

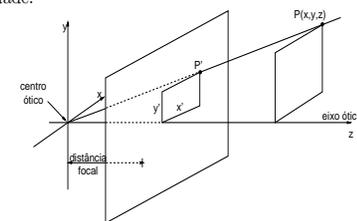


Observação 1.2 Distinguem-se na retina dois tipos de células sensíveis à luz: os cones e os bastonetes. A região de máxima sensibilidade à luz só contém cones. Os bastonetes são destinados à visão com pouca luz, sendo insensíveis às diferenças de cor.

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.

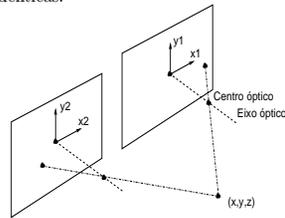


O modelo mais comum para emular a formação de imagens na retina de um olho humano é o modelo de câmera escura (*pin-hole camera*), no qual o plano de imagem é disposto simetricamente em relação ao “plano de retina” para evitar a distinção de coordenadas de sinais diferentes na direção de profundidade.

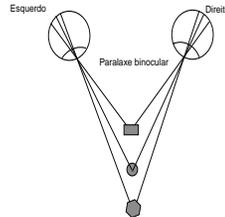


A nossa visão é de fato estéreo. Quando olhamos para um objeto, são formadas duas imagens retinianas, uma em cada olho. Devido ao afasta-

mento entre os dois olhos, as imagens formadas sob dois distintos ângulos de visão não são idênticas.



A observação simultânea dessas duas imagens ligeiramente diferentes força movimentos musculares dos olhos para “ajustar” a paralaxe binocular. Observe que quanto menor for a profundidade do objeto (o retângulo na ilustração), maior será a diferença na posição onde é formada a imagem e maior será o movimento muscular. Estes movimentos permitem que o cérebro diferencie a profundidade entre tais pontos, dando a sensação de relevo/profundidade.



1.3 Tipos de Imagens

De acordo com a maneira como a brilhância em cada ponto da imagem é obtida, distinguem-se quatro classes de imagens:

1. Imagens sintéticas: a 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. Por exemplo, imagens geradas sinteticamente pelo computador.
2. Imagens convencionais: a brilhância em cada ponto da imagem é uma função da intensidade das fontes luminosas de um ambiente real e da

posição e da orientação deste ponto em relação a estas fonte(s). Por exemplo, fotos digitais.

3. Imagens de transmissão: a intensidade e a orientação da fonte luminosa é homogênea para todos os pontos da imagem, como nas imagens de ressonância magnética. A brilhância em cada ponto depende das propriedades radiométricas dos pontos de interesse. Por exemplo, imagens de ressonância magnética.
4. Imagens de profundidade: o valor em cada ponto é, ao invés da brilhância, a profundidade do ponto em relação ao dispositivo de captura. Por exemplo, imagens obtidas com um escaneador de profundidade.

Nesta disciplina estamos interessados na geração de imagens sintéticas com base em modelos geométricos e imagens existentes.

1.4 Percepção Visual de Imagens

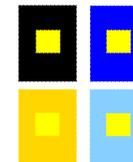
A habilidade da visão humana para distinguir os níveis de brilhância numa imagem é essencial na interpretação e compreensão de uma imagem. Esta habilidade é um fenômeno psicofísico, variando com a sensibilidade de cada indivíduo. Algumas características da percepção humana são:

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

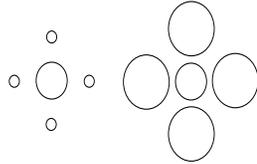
(Ver Fig. 16.17 do livro-texto de Foley.)

Observação 1.3 Esta característica da percepção humana agrava o fenômeno conhecido como **Bandas de Mach**, quando quantizamos a luminância/brilhância de uma imagem.

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



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



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

1.5 Modelos de Imagens

Modelos matemáticos são frequentemente utilizados para descrever e processar as imagens. 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 **a cores** ou **multiespectral** o valor da função pode ser um vetor de n valores reais, cada qual corresponde a uma cor espectral (brilhância – *brightness*).

Observação 1.4 *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 monocromáticas (em tons de cinza) e brilhância para imagens a cores.*

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*.

Nesta disciplina nós nos restringiremos a trabalhar com as imagens digitais, processáveis pelos computadores.

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.

1.6 Exibição de Imagens

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

1. vetoriais (década 60 a década 80): para imagens contínuas cujo conteúdo é representado pelos vetores. 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).

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

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

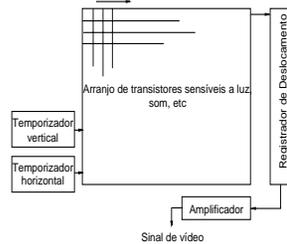
2. digitais (*raster*, a partir da década 70): para imagens digitais. No caso de tecnologia de tubos catódicos, o movimento dos canhões é sempre o mesmo (varredura horizontal e retraço) e a função do controlador se limita a monitorar a intensidade dos raios catódicos de acordo com os atributos de cada pixel (*picture element*). Para aumentar a eficiência do sistema é comum armazenar os atributos de cada pixel numa memória em separado chamado *frame buffer*. Há duas formas de especificar a cor de um pixel: modo indexado através de um *look-up table* ou modo RGB.

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

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

1.7 Aquisição de Imagens

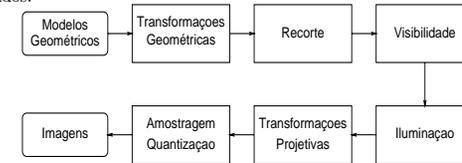
Emulando a visão humana, as câmeras e os escaneadores são os mais conhecidos dispositivos para captura de imagens. O princípio de funcionamento destes dispositivos se baseia na fotosensibilidade de alguns materiais químicos cujas propriedades podem ser alteradas pela luz emitida pelos objetos de interesse. Tais alterações podem ser traduzidas em sinais de vídeo.



Os algoritmos de síntese de imagens são intimamente relacionados com a tecnologia de exibição das imagens. Os dispositivos de saída são hoje predominantemente digitais; portanto, como já mencionamos, nós nos concentraremos nesta disciplina nos métodos de síntese de imagens digitais.

A síntese de imagens consiste, essencialmente, numa sequência de transformações de sai dos modelos geométricos e termina em um conjunto de imagens. Uma possível sequência seria compor uma cena com os modelos geométricos definidos em seus sistemas de referência próprios através das **transformações geométricas**, **recortar** as partes que não estejam dentro do volume de visão, selecionar porções **visíveis**, analisar a interação dos ob-

jetos visíveis com as **fontes luminosas**, **projetar**, **amostrar** e **quantizar** os dados.



Vale ressaltar aqui que a ordem de recorte, análise de visibilidade e iluminação pode mudar nesta sequência, de acordo com o realismo desejado para a imagem final.

1.8 Interação com as Imagens

A crescente demanda pelos sistemas gráficos interativos, nos quais os usuários podem interagir diretamente com o computador através das imagens exibidas, torna imprescindível comentar sobre os dispositivos de entrada em sistemas de informações gráficas. Pesquisas na área de Interface Homem-Máquina comprovam que a comunicação homem-computador pode ser mais eficiente com auxílio de imagens. Os primeiros dispositivos de entrada via imagens – os *light-pens* – foram substituídos nas últimas décadas pelos dispositivos cada vez mais ergonômicos, funcionais e econômicos. Logicamente, esses dispositivos podem ser classificados em seis classes para realizar essencialmente quatro funções (entrada de comandos, entrada de dados, posicionamento e identificação):

1. *locator*: entra a posição de um ponto;

(Ver Figs. 4.34–4.39 do livro-texto de Foley.)

2. *stroke*: entra uma sequência de posições de pontos;

3. *pick*: seleciona um objeto da imagem;

4. *valuator*: entra um valor numérico;

(Ver Fig. 4.40 do livro-texto de Foley.)

5. *keyboard* ou *string*: entra uma sequência de caracteres alfa-numéricos; e

(Ver Fig. 4.41 do livro-texto de Foley.)

6. *choice* : entra uma opção de um menu (usualmente, uma imagem) apresentado.