

Capítulo 1

Introdução

1.1 Processamento de Imagens: breve histórico e exemplos de aplicações

A área de processamento de imagens vem sendo objeto de crescente interesse por permitir viabilizar grande número de aplicações em duas categorias bem distintas: (1) o aprimoramento de informações pictóricas para interpretação humana; e (2) a análise automática por computador de informações extraídas de uma cena. Ao longo deste livro, reservaremos a expressão 'processamento de imagens' para designar a primeira categoria, adotando os termos 'análise de imagens', 'visão por computador' (ou 'visão computacional') e 'reconhecimento de padrões' para a segunda.

Uma das primeiras aplicações na primeira categoria remonta ao começo deste século, onde buscavam-se formas de aprimorar a qualidade de impressão de imagens digitalizadas transmitidas através do sistema Bartlane de transmissão de imagens por cabo submarino entre Londres e Nova Iorque. Os primeiros sistemas Bartlane, no início da década de 20, codificavam uma imagem em cinco níveis de intensidade distintos. Esta capacidade seria expandida, já em 1929, para 15 níveis, ao mesmo tempo em que era desenvolvido um método aprimorado de revelação de filmes através de feixes de luz modulados por uma fita que continha informações codificadas sobre a imagem.

Mas o grande impulso para a área de Processamento de Imagens viria cerca de três décadas mais tarde, com o advento dos primeiros computadores digitais de grande porte e o início do programa espacial norte-americano. O uso de técnicas computacionais de aprimoramento de imagens teve início no *Jet Propulsion Laboratory* (Pasadena, Califórnia - EUA)¹ em 1964, quando imagens da lua transmitidas por uma sonda Ranger² eram processadas por computador para corrigir vários tipos de distorção inerentes à câmera de TV acoplada à sonda. Estas técnicas serviram de base para métodos aprimorados de realce e restauração de imagens de outros programas espaciais posteriores, como as expedições tripuladas da série Apollo, por exemplo.

De 1964 aos dias atuais, a área de processamento de imagens vem apresentando crescimento expressivo e suas aplicações permeiam quase todos os ramos da atividade humana. Em Medicina, o uso de imagens no diagnóstico médico tornou-se rotineiro e os avanços em processamento de imagens vêm permitindo tanto o desenvolvimento de novos equipamentos quanto a maior facilidade de interpretação de imagens produzidas por equipamentos mais antigos, como por exemplo o de raio X. Em Biologia, a capacidade de processar automaticamente imagens obtidas de microscópios, por exemplo contando o número de células de um certo tipo presentes em uma imagem, facilita sobremaneira a execução de tarefas laboratoriais com alto grau de precisão e repetibilidade. O processamento e a interpretação automática de imagens captadas por satélites auxiliam os trabalhos nas áreas de Geografia, Sensoriamento Remoto, Geoprocessamento e Meteorologia, dentre outras. Técnicas de restauração de imagens auxiliam arqueologistas a recuperar fotos borradas de artefatos raros, já destruídos. O uso de robôs dotados de visão artificial em tarefas tais como controle de qualidade em linhas de produção aumenta a cada ano, num cenário de crescente automação industrial. Inúmeras outras áreas tão distintas como Astronomia, Segurança, Publicidade e Direito – para citar apenas algumas – vêm sendo beneficiadas com os avanços nas áreas de processamento de imagens e visão por computador.

¹ "<http://www.jpl.nasa.gov>"

² "<http://www.jpl.nasa.gov/missions/ranger/>"

Leitura complementar

Mascarenhas [Mascarenhas 1990] apresenta um abrangente texto introdutório aos assuntos abordados neste livro.

1.2 Um sistema de processamento de imagens e seus componentes

Os elementos de um sistema de processamento de imagens de uso genérico são mostrados na figura 1. Este diagrama permite representar desde sistemas de baixo custo até sofisticadas estações de trabalho utilizadas em aplicações que envolvem intenso uso de imagens. Ele abrange as principais operações que se pode efetuar sobre uma imagem, a saber: aquisição, armazenamento, processamento e exibição. Além disso, uma imagem pode ser transmitida à distância utilizando meios de comunicação disponíveis. Todas estas operações são descritas a seguir.

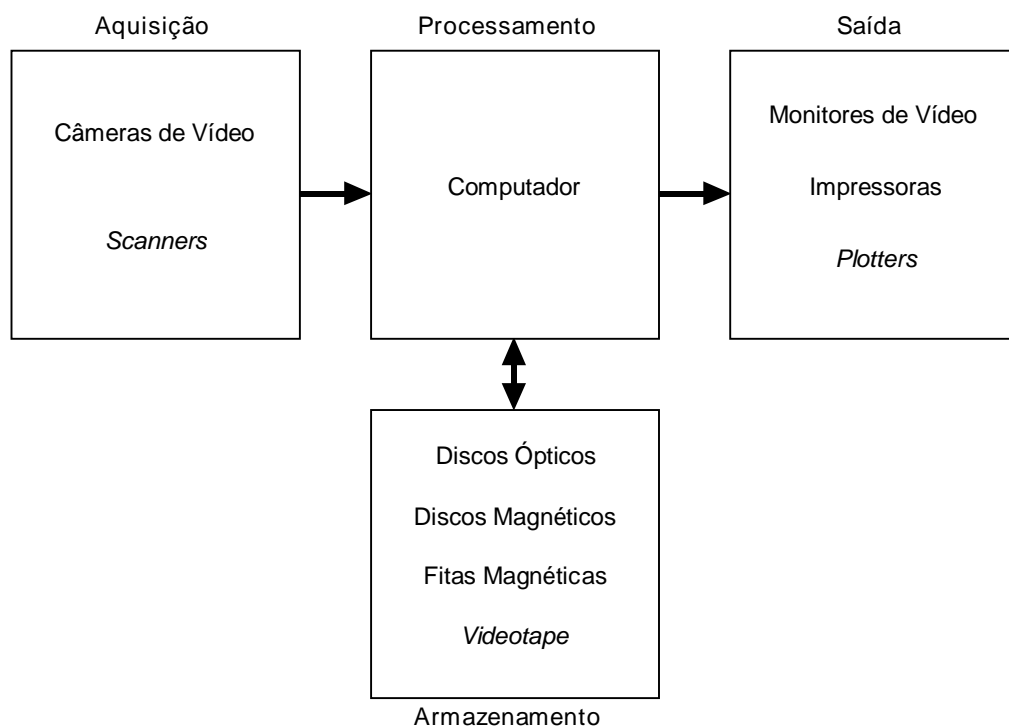


Figura 1 - Elementos de um sistema de processamento de imagens.

1.2.1 Aquisição

A etapa de aquisição tem como função converter uma imagem em uma representação numérica adequada para o processamento digital subsequente. Este bloco compreende dois elementos principais. O primeiro é um dispositivo físico sensível a uma faixa de energia no espectro eletromagnético (como raio X, ultravioleta, espectro visível ou raios infravermelhos), que produz na saída um sinal elétrico proporcional ao nível de energia detetado. O segundo – o digitalizador propriamente dito – converte o sinal elétrico analógico em informação digital, isto é, que pode ser representada através de bits 0s e 1s. Um módulo de aquisição de imagens é normalmente conhecido pelo nome de *frame grabber*. Os capítulos 2 e 7 deste livro trazem mais detalhes sobre os aspectos envolvidos na aquisição de imagens digitais.

1.2.2 Armazenamento

O armazenamento de imagens digitais é um dos maiores desafios no projeto de sistemas de processamento de imagens, em razão da grande quantidade de bytes necessários para tanto. Este armazenamento pode ser dividido em três categorias: (1) armazenamento de curta duração de uma imagem, enquanto ela é utilizada nas várias etapas do processamento, (2) armazenamento

de massa para operações de recuperação de imagens relativamente rápidas, e (3) arquivamento de imagens, para recuperação futura quando isto se fizer necessário. O espaço de armazenamento requerido é normalmente especificado em bytes (8 bits) e seus múltiplos: KB (kilobyte \approx 1000 bytes), MB (megabyte \approx 1 milhão de bytes), GB (gigabyte \approx 1 bilhão de bytes) e TB (terabyte \approx 1 trilhão de bytes). O capítulo 7 discute aspectos de armazenamento de imagens em maior profundidade.

Para o armazenamento de curta duração, a alternativa mais simples é utilizar parte da memória RAM do computador principal. Outra opção consiste no uso de placas especializadas, chamadas *frame buffers*, que armazenam uma ou mais imagens completas e podem ser acessadas a uma alta velocidade, tipicamente 30 imagens completas por segundo. O uso de *frame buffers* permite também que operações de *zoom* (ampliação ou redução para fins de visualização), *scroll* (rolagem na vertical) e *pan* (rolagem na horizontal) sejam executadas de forma praticamente instantânea. Placas *frame buffers* disponíveis no mercado atualmente apresentam capacidade de armazenamento na faixa de alguns MB de memória.

A segunda categoria de armazenamento normalmente requer o uso de discos magnéticos de no mínimo algumas centenas de MB e recentemente passou a utilizar também discos magneto-ópticos, por vezes agrupados em *jukeboxes* contendo de 30 a 100 discos. Nesta categoria o fator 'tempo de acesso' é tão ou mais importante que a capacidade (em bytes) do meio de armazenamento. Através de cálculos simples (n° de pixels na horizontal \times n° de pixels na vertical \times n° de bits necessários para a escala de cinza / 8), pode-se estimar a quantidade de bytes necessários para armazenar uma imagem monocromática em disco. Este cálculo entretanto considera uma imagem representada como uma matriz, cujos elementos são os valores de tons de cinza dos respectivos pixels.³ Na prática, informações adicionais (tamanho da imagem e número de cores ou tons de cinza, no mínimo) são necessárias. Estas informações costumam ser colocadas em um cabeçalho (*header*) no início do arquivo. Infelizmente, não existe um único cabeçalho ou formato de armazenamento de imagens padronizados. Alguns dos formatos mais comuns são o BMP, PCX, TIFF, JPEG e GIF. Estes formatos de arquivos de imagem, além de muitos outros, são apresentados no Apêndice A.

Finalmente, o arquivamento de imagens é caracterizado por quantidades gigantescas de bytes contendo imagens cuja recuperação é esporádica. Nesta categoria, as fitas magnéticas estão dando lugar aos discos ópticos WORM (*Write-Once-Read-Many*), com capacidade que pode chegar a mais de 10 GB por disco, e que também podem ser agrupados em *jukeboxes*, com capacidade total de armazenamento superior a 1 TB.

1.2.3 Processamento

O processamento de imagens digitais envolve procedimentos normalmente expressos sob forma algorítmica. Em função disto, com exceção das etapas de aquisição e exibição, a maioria das funções de processamento de imagens pode ser implementada via software. O uso de hardware especializado para processamento de imagens somente será necessário em situações nas quais certas limitações do computador principal (por exemplo, velocidade de transferência dos dados através do barramento) forem intoleráveis.

A tendência atual do mercado de hardware para processamento de imagens é a comercialização de placas genéricas compatíveis com os padrões de barramento consagrados pelas arquiteturas mais populares de microcomputadores e estações de trabalho. O software de controle destas placas é que determinará sua aplicação específica a cada situação. As vantagens mais imediatas são: redução de custo, modularidade, reutilização de componentes de software em outra aplicação rodando sobre o mesmo hardware e independência de fornecedor. Convém notar, entretanto, que sistemas dedicados continuam sendo produzidos e comercializados para atender a tarefas específicas, tais como processamento de imagens transmitidas por satélites.

³ Para imagens coloridas, a situação é um pouco mais complexa. Normalmente estes números serão índices (endereços) de uma tabela de cores, denominada palheta ou *palette*. Para maiores detalhes, veja a seção 4.6 e o Resumo da Teoria da Prática de Laboratório nº 1.

1.2.4 Transmissão

Imagens digitalizadas podem ser transmitidas à distância utilizando redes de computadores e protocolos de comunicação já existentes. O grande desafio da transmissão de imagens à distância é a grande quantidade de bytes que se necessita transferir de uma localidade a outra, muitas vezes através de canais de comunicação de baixa velocidade e banda passante estreita. Este problema é ainda mais sério quando se deseja transmitir seqüências de vídeo (imagens em movimento com áudio associado) em tempo real, onde outros fatores, como por exemplo sincronização, devem ser considerados. Nestes casos, o uso de técnicas de compressão e descompressão de imagens, como as descritas no capítulo 6, é mandatório.

1.2.5 Exibição

O monitor de vídeo é um elemento fundamental de um sistema de processamento de imagens. Os monitores em uso atualmente são capazes de exibir imagens com resolução de pelo menos 640 x 480 pixels com 256 cores distintas. A tecnologia mais usual ainda é o TRC (Tubo de Raios Catódicos).

Um TRC para um sistema de processamento de imagens normalmente segue um padrão de vídeo. O padrão de vídeo mais comum para sistemas monocromáticos é o RS-170. Ele prevê 480 linhas horizontais entrelaçadas, isto é, a varredura de um quadro é feita em duas etapas, abrangendo primeiramente as linhas ímpares e posteriormente as linhas pares. Cada uma destas etapas é denominada campo. O tempo necessário para percorrer um campo é 1/60 s; conseqüentemente, o tempo total de um quadro é 1/30 s. As características de persistência visual do olho humano fazem com que, nesta velocidade, a varredura individual de cada campo não seja perceptível, bem como dão a impressão de que a seqüência de quadros explorados é perfeitamente contínua.

O padrão RS-170 especifica resolução vertical de 480 pixels, sendo a resolução horizontal determinada pelos circuitos eletrônicos do monitor e pelo tamanho dos pontos de fósforo na tela. Este número costuma ser 512, proporcionando imagens de 512 x 480 pixels.

A resolução espacial dos monitores é normalmente especificada em pontos por polegada (*dots per inch - dpi*). Um valor típico de resolução é 72 dpi, suficiente para exibir uma imagem de 1024 x 1024 pixels em um monitor de 19 polegadas ou uma imagem de 640 x 400 pontos em uma tela cuja diagonal meça 12 polegadas. A título de comparação, uma tela de TV tem resolução na faixa de 40 dpi.

Um TRC colorido difere radicalmente de seu antecessor monocromático, por apresentar três feixes eletrônicos, cada um correspondente a uma das três cores primárias (vermelho, verde e azul). A superfície interna da tela é constituída por três tipos de fósforo, dispostos de forma triangular, cada qual sensível a uma das cores primárias e excitado pelo respectivo canhão eletrônico. Isto significa dizer que, do ponto de vista construtivo, cada pixel é na verdade uma combinação de três pequenos pixels, um para cada cor primária.

A indústria de dispositivos de exibição vem apresentando sistematicamente novas tecnologias de fabricação de monitores de vídeo, dentre eles os monitores de cristal líquido (LCD), cada vez mais populares graças à disseminação dos computadores portáteis (*notebooks*).

Existem diversas formas de reprodução de imagens em papel. A melhor, e mais cara, é a reprodução fotográfica, onde o número de gradações de cinza é função da densidade dos grânulos de prata no papel. Outra possibilidade é o uso de papel sensível a temperatura, cuja composição química faz com que ele apresente coloração mais escura à medida que a temperatura aumenta. Este tipo de impressão ainda é o mais difundido em equipamentos de fax. Uma de suas desvantagens é o desvanecimento das imagens com o tempo. Nos últimos anos aumentou consideravelmente a oferta de impressoras térmicas coloridas no mercado. Estas impressoras baseiam-se na deposição de cera colorida sobre um papel especial para produzir a impressão. O capítulo 7 apresenta maiores informações sobre dispositivos de exibição e impressão de imagens.

Dispositivos periféricos de saída especializados na produção de cópias da imagem em forma de fotografias, slides ou transparências também estão se tornando cada vez mais usuais.

Uma alternativa às técnicas fotográficas consiste no uso de técnicas de *halftoning*. É o método usado por jornais e por impressoras convencionais (laser, matriciais ou a jato de tinta) para a impressão de imagens. Esta técnica consiste basicamente em imprimir pontos escuros de diferentes tamanhos, espaçados de tal maneira a reproduzir a ilusão de tons de cinza. À medida que a distância entre o observador e a imagem impressa aumentam, os detalhes finos vão desaparecendo e a imagem parece cada vez mais uma imagem contínua monocromática.

No jargão computacional, dá-se o nome de *dithering* ao processo de produção do efeito de *halftoning*, bem como a todas as técnicas de conversão de uma imagem para adaptá-la a resoluções menores, tanto para efeito de exibição como para impressão. Existem vários algoritmos de *dithering*, sendo o mais comum o de Floyd-Steinberg, que consiste de um processo adaptativo no qual o padrão de *dither* a ser atribuído a um pixel depende de seu tom de cinza e de seus vizinhos. Uma discussão mais detalhada destes algoritmos foge ao escopo deste livro.

Leitura complementar

O capítulo 7 de [Lindley 1991] e o capítulo 11 de [Rimmer 1993] trazem explicações teóricas e código-fonte em C para impressão de imagens monocromáticas e/ou coloridas utilizando *dithering*.

O capítulo 10 de [Dougherty 1994] é uma excelente referência para um estudo mais aprofundado dos conceitos e técnicas de *halftoning*.

1.3 O sistema visual humano

A figura 2 mostra um corte horizontal do olho humano. O globo ocular tem formato aproximadamente esférico e um diâmetro de cerca de 20 mm. Ele é envolvido por três membranas: a camada externa formada pela córnea e pela esclerótica, a coróide e a retina, que é a camada interna. A córnea é uma película transparente que cobre a parte anterior do olho. Dando continuidade à córnea, a esclerótica é uma membrana opaca que reveste o globo ocular.

A coróide está situada abaixo da esclerótica. Essa membrana contém uma rede de vasos sanguíneos que servem como a principal fonte de nutrição do olho. O revestimento da coróide é fortemente pigmentado, o que ajuda a reduzir a quantidade de luz que entra no olho. Ela é dividida em corpo ciliar e diafragma da íris, sendo este último responsável por controlar a quantidade de luz que deve penetrar no olho. O diâmetro da abertura central da íris (pupila) varia entre 2 mm e 8 mm. A parte frontal da íris contém o pigmento visível do olho, enquanto sua porção posterior possui um pigmento negro.

A membrana mais interna do olho é a retina, situada na sua parede posterior. Quando o olho focaliza uma cena, a imagem correspondente é projetada sobre a retina, na qual estão distribuídos dois tipos de receptores de luz discretos: os cones e os bastonetes. Os cones são em número de 6 a 7 milhões em cada olho e estão localizados na porção central da retina, chamada de fóvea. Eles são altamente sensíveis a cor e cada qual está conectado a uma terminação nervosa dedicada. O número de bastonetes é bastante maior, cerca de 75 a 150 milhões, distribuídos em toda a superfície da retina. Os bastonetes servem para dar uma visão geral da imagem captada no campo de visão. Eles não distinguem cores, mas são sensíveis a baixos níveis de iluminação.

A fóvea é uma reentrância circular na retina com aproximadamente 1,5 mm de diâmetro. De forma aproximada, podemos considerá-la um sensor de área quadrada de 1,5 mm por 1,5 mm. A densidade de cones nesta área da retina é de aproximadamente 150.000 elementos por mm^2 . Baseando-se nessas aproximações, o número de cones na região de maior acuidade do olho é de aproximadamente 337.000 elementos. Para efeito comparativo, esta resolução é facilmente alcançada pela tecnologia atual, usando CCD de área não superior a 7 mm x 7 mm.

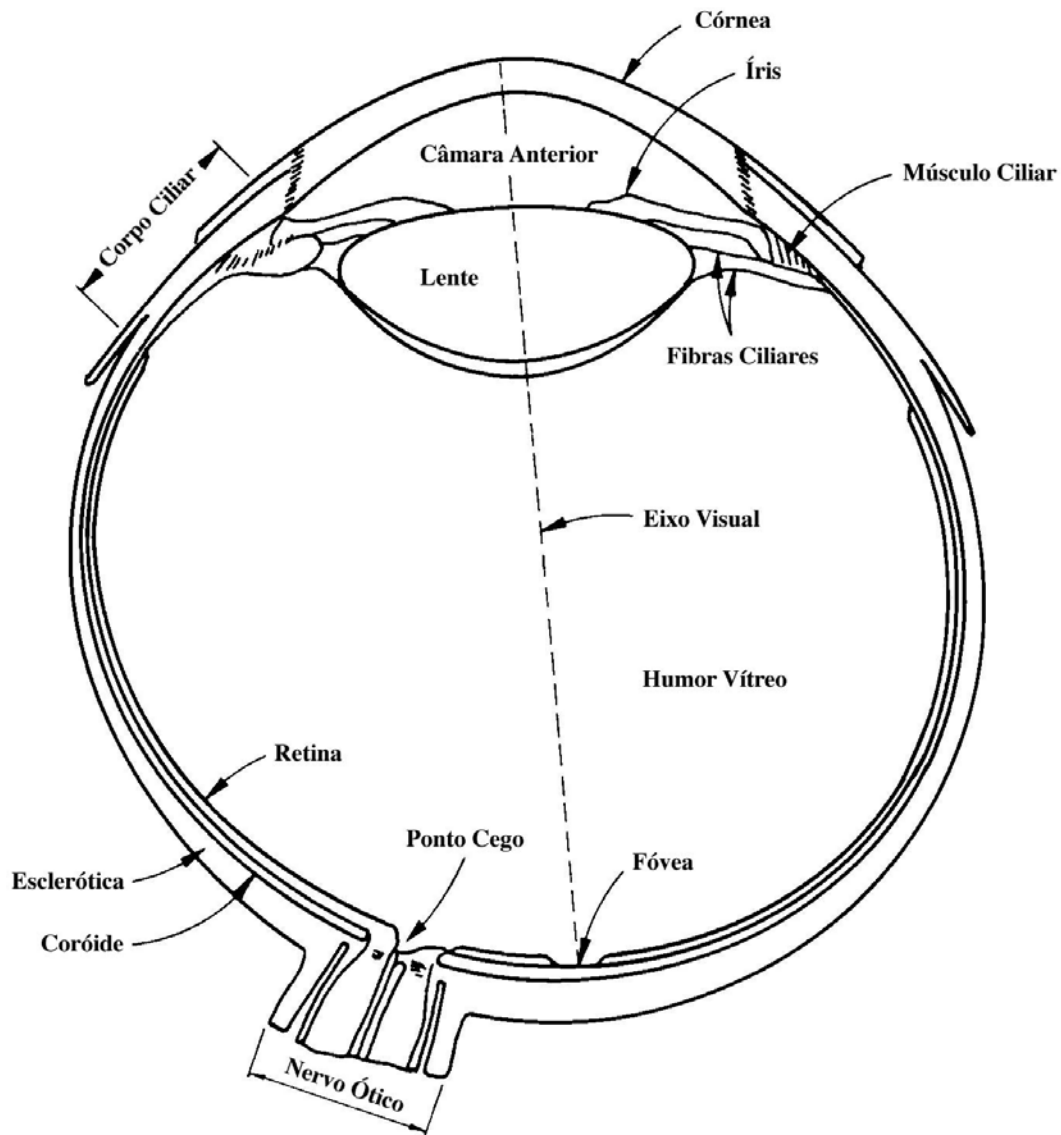


Figura 2 - Vista em corte do olho humano (adaptado de Gonzalez, R.C., Woods, R.E., *Digital Image Processing*, Addison-Wesley, 1992).

Dentre as características do sistema visual humano relevantes para o projeto de sistemas de visão artificial, podemos destacar a enorme faixa de valores de intensidade luminosa (da ordem de 10^{10}) aos quais o olho pode se adaptar e a relação aproximadamente logarítmica entre os níveis de intensidade luminosa presentes na cena e a sensação subjetiva de brilho, ilustrada na figura 3. Convém notar, entretanto, que o olho humano não pode operar sobre toda esta faixa simultaneamente. Ao invés disso, ele excursiona ao longo de toda esta faixa através de adaptações em sua sensibilidade global, um fenômeno conhecido como adaptação de brilho. A faixa total de valores de intensidade que o olho pode discriminar simultaneamente é comparativamente pequena em relação à faixa total de adaptação. Para um certo conjunto de condições, o nível de sensibilidade atual do sistema visual é chamado nível de adaptação de brilho, como por exemplo o valor B_a na figura 3. A pequena porção de curva que intercepta a curva principal representa a faixa de brilho subjetivo que o olho pode perceber quando adaptado a este nível. Esta faixa é bastante restrita, existindo um nível de brilho B_b abaixo do qual todos os estímulos são indistinguíveis. O trecho tracejado da curva não é, na verdade, restrito, mas se estendido além de um limite perde seu sentido, porque neste caso o olho humano buscaria outro nível de adaptação maior que B_a .

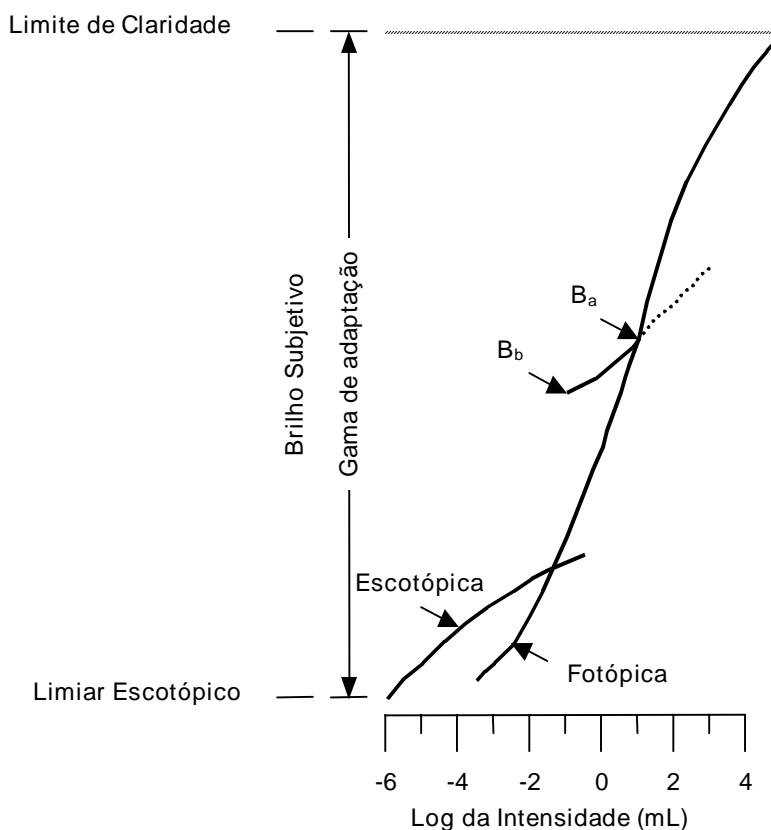


Figura 3 - Curva que relaciona a sensação subjetiva de brilho com o logaritmo da intensidade luminosa incidente sobre o olho humano.

A tabela 1 apresenta uma visão comparativa entre o sistema visual humano e um sistema de visão artificial.

Leitura complementar

Para maiores informações sobre aspectos do processo de percepção visual humana e suas implicações no projeto de sistemas de visão artificial, uma referência obrigatória é [Marr 1982].

A seção 2.1 de [Gonzalez e Woods 1992] apresenta interessantes fenômenos de percepção visual tais como contraste simultâneo, discriminação de brilho usando a razão de Weber e o padrão de faixas de Mach.

1.4 Sistemas de Visão Artificial: fundamentos e desafios

Nas primeiras páginas de seu livro *Digital Image Processing and Computer Vision*, Robert J. Schalkoff [Schalkoff 1989] formula uma sugestiva pergunta quando questiona: "O que estamos tentando fazer e por que isto é tão difícil?"

Uma resposta simplista à primeira parte da pergunta acima é: "Estamos tentando ensinar robôs a enxergar." Ao relacionarmos as dificuldades inerentes ao processo de dotar o computador de uma capacidade visual semelhante à dos seres humanos, deparamo-nos com três admiráveis características do processo de percepção visual humano, que são:

- uma base de dados muito rica;
- altíssima velocidade de processamento; e
- a capacidade de trabalhar sob condições muito variadas.

Os avanços na tecnologia de dispositivos de armazenamento de massa e o surgimento de novas CPUs e arquiteturas computacionais cada vez mais rápidas, com alto grau de paralelismo, nos induzem a crer que dispomos de condições cada vez melhores de modelar as duas primeiras características relacionadas acima. O grande desafio permanece sendo o de fazer com que os sistemas de visão artificial trabalhem em diferentes condições de luminosidade, contraste,

posicionamento relativo dos objetos em uma cena sem perder a capacidade de interpretar a cena, de forma análoga a nossa capacidade de reconhecer um amigo ou parente com relativa facilidade, independentemente de ele estar usando óculos ou não, ter deixado crescer a barba ou estar no carro ao lado do nosso em uma esquina num final de tarde, onde não dispomos de outra imagem senão a vista de perfil e onde as condições de luminosidade são bastante inferiores às que obteríamos ao meio-dia.

Tabela 1 - Comparação entre o sistema visual humano e um sistema de visão artificial.

	Sistema visual humano	Sistema de visão artificial
Espectro	Limitado à faixa de luz visível (300 nm a 700 nm) do espectro de ondas eletromagnéticas.	Pode operar em praticamente todo o espectro de radiações eletromagnéticas, dos raios X ao infravermelho.
Flexibilidade	Extremamente flexível, capaz de se adaptar a diferentes tarefas e condições de trabalho.	Normalmente inflexível, apresenta bom desempenho somente na tarefa para a qual foi projetado.
Habilidade	Pode estabelecer estimativas relativamente precisas em assuntos subjetivos.	Pode efetuar medições exatas, baseadas em contagem de pixels e, portanto, dependentes da resolução da imagem digitalizada.
Cor	Possui capacidade de interpretação subjetiva de cores.	Mede objetivamente os valores das componentes R, G e B para determinação de cor.
Sensibilidade	Capaz de se adaptar a diferentes condições de luminosidade, características físicas da superfície do objeto e distância ao objeto. Limitado na distinção de muitos níveis diferentes de cinza, simultaneamente.	Sensível ao nível e padrão de iluminação, bem como à distância em relação ao objeto e suas características físicas. Pode trabalhar com centenas de tons de cinza, conforme projeto do digitalizador.
Tempo de resposta	Elevado, da ordem de 0,1 s.	Dependente de aspectos de hardware, podendo ser tão baixo quanto 0,001 s.
2-D e 3-D	Pode executar tarefas 3-D e com múltiplos comprimentos de onda (dentro do espectro de luz visível) facilmente.	Executa tarefas 2-D com relativa facilidade, mas é lento e limitado em tarefas 3-D.
Percepção	Percebe variações de brilho em escala logarítmica. A interpretação subjetiva de brilho depende da área ao redor do objeto considerado.	Pode perceber brilho em escala linear ou logarítmica.

1.4.1 Estrutura de um Sistema de Visão Artificial

Definiremos um Sistema de Visão Artificial (SVA) como um sistema computadorizado capaz de adquirir, processar e interpretar imagens correspondentes a cenas reais. A figura 4 mostra esquematicamente um diagrama de blocos de um SVA. Suas principais etapas são explicadas a seguir, partindo da premissa de que um problema prático, por exemplo a leitura do Código de Endereçamento Postal (CEP) de um lote de envelopes, deve ser solucionado.

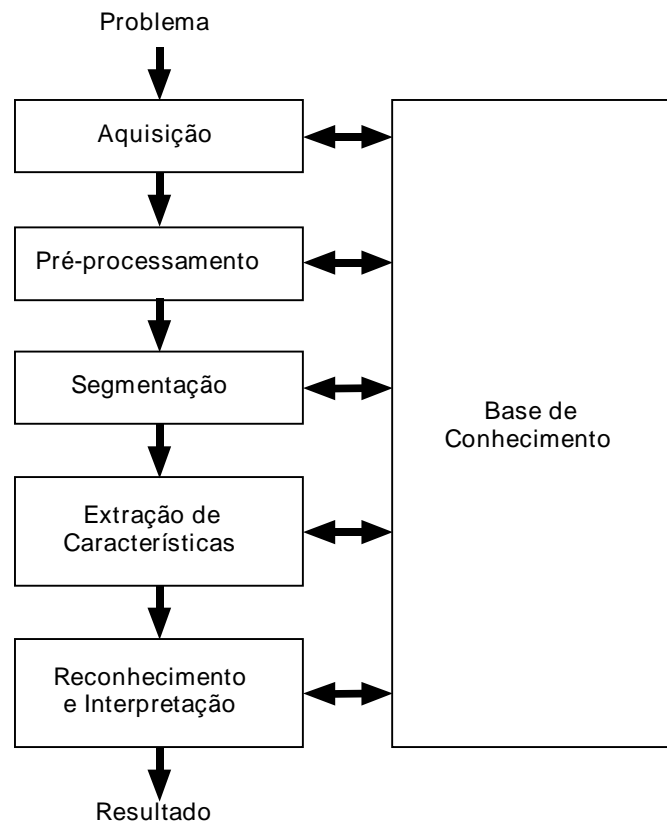


Figura 4 - Um Sistema de Visão Artificial (SVA) e suas principais etapas.

1.4.2 Domínio do problema e resultado

O domínio do problema, neste caso, consiste no lote de envelopes e o objetivo do SVA é ler o CEP presente em cada um deles. Logo, o resultado esperado é uma seqüência de dígitos correspondentes ao CEP lido.

1.4.3 Aquisição da imagem

O primeiro passo no processo é a aquisição de imagens dos envelopes. Para tanto são necessários um sensor e um digitalizador. O sensor converterá a informação óptica em sinal elétrico e o digitalizador transformará a imagem analógica em imagem digital.

Dentre os aspectos de projeto envolvidos nesta etapa, pode-se mencionar: a escolha do tipo de sensor, o conjunto de lentes a utilizar, as condições de iluminação da cena, os requisitos de velocidade de aquisição (assumindo que os envelopes passariam pela frente da câmera a uma certa velocidade), a resolução e o número de níveis de cinza da imagem digitalizada, dentre outros. Esta etapa produz à saída uma imagem digitalizada do envelope.

1.4.4 Pré-processamento

A imagem resultante do passo anterior pode apresentar diversas imperfeições, tais como: presença de pixels ruidosos, contraste e/ou brilho inadequado, caracteres (especialmente os dígitos do CEP) interrompidos ou indevidamente conectados etc. A função da etapa de pré-

processamento é aprimorar a qualidade da imagem para as etapas subsequentes. As operações efetuadas nesta etapa são ditas de baixo nível porque trabalham diretamente com os valores de intensidade dos pixels, sem nenhum conhecimento sobre quais deles pertencem aos dígitos do CEP, a outras informações impressas no envelope ou ao fundo. A imagem resultante desta etapa é uma imagem digitalizada de melhor qualidade que a original.

1.4.5 Segmentação

A tarefa básica da etapa de segmentação é a de dividir uma imagem em suas unidades significativas, ou seja, nos objetos de interesse que a compõem. Esta tarefa, apesar de simples de descrever, é das mais difíceis de implementar.

No caso específico do problema do CEP, é possível que o problema seja dividido em duas etapas: em um primeiro momento os algoritmos de segmentação tentarão localizar o CEP do restante das informações para posteriormente, trabalhando sobre esta subimagem, segmentar cada dígito individualmente. Segundo esta linha de raciocínio, este bloco produzirá à saída oito subimagens, cada qual correspondendo a um dígito do CEP.

1.4.6 Extração de Características

Esta etapa procura extrair características das imagens resultantes da segmentação através de descritores que permitam caracterizar com precisão cada dígito e que apresentem bom poder de discriminação entre dígitos parecidos, como o '5' e o '6'. Estes descritores devem ser representados por uma estrutura de dados adequada ao algoritmo de reconhecimento. É importante observar que nesta etapa a entrada ainda é uma imagem, mas a saída é um conjunto de dados correspondentes àquela imagem.

Para maior clareza, suponhamos que os descritores utilizados para descrever um caractere sejam as coordenadas normalizadas x e y de seu centro de gravidade e a razão entre sua altura e largura. Neste caso, um vetor de três elementos é uma estrutura de dados adequada para armazenar estas informações sobre cada dígito processado por esta etapa.

1.4.7 Reconhecimento e Interpretação

Nesta última etapa do sistema, denominamos reconhecimento o processo de atribuição de um rótulo a um objeto baseado em suas características, traduzidas por seus descritores. A tarefa de interpretação, por outro lado, consiste em atribuir significado a um conjunto de objetos reconhecidos. Neste exemplo, uma forma simples de interpretação seria a verificação do CEP em uma base de dados de CEPs válidos, para descobrir se o conjunto dos oito caracteres (cada qual reconhecido individualmente) faz sentido ou não.

1.4.8 Base de Conhecimento

Todas as tarefas das etapas descritas acima pressupõem a existência de um conhecimento sobre o problema a ser resolvido, armazenado em uma base de conhecimento, cujo tamanho e complexidade podem variar enormemente. Idealmente, esta base de conhecimento deveria não somente guiar o funcionamento de cada etapa, mas também permitir a realimentação entre elas. Por exemplo, se a etapa de representação e descrição recebesse 7 caracteres ao invés de 8, ela deveria ser capaz de realimentar a etapa de segmentação (provável responsável pela falha) para que esta procurasse segmentar novamente a subimagem 'suspeita' (aquela de maior largura), buscando dividi-la em duas. Esta integração entre as várias etapas através da base de conhecimento ainda é um objetivo difícil de alcançar e não está presente na maioria dos SVAs existentes atualmente.

Finalmente, cumpre observar que nem todos os SVAs possuem todos estes blocos e que a maioria das técnicas descritas neste livro estão delimitadas até o bloco 'pré-processamento'.

1.5 Estrutura e escopo do livro

Este livro está estruturado em sete capítulos, sendo os dois primeiros dedicados à apresentação de conceitos fundamentais que servirão de base para a compreensão dos capítulos seguintes. O capítulo 2 apresenta os principais aspectos das imagens digitais, caracterizando-as sob o ponto MARQUES FILHO, Ogê; VIEIRA NETO, Hugo. *Processamento Digital de Imagens*, Rio de Janeiro: Brasport, 1999. ISBN 8574520098.

de vista matemático, apresentando suas principais propriedades e ilustrando as principais operações lógicas, aritméticas e transformações geométricas que se pode efetuar sobre elas. O capítulo 3 define histograma de uma imagem e apresenta diversas técnicas de modificação de histograma.

Os capítulos 4, 5 e 6 tratam das técnicas de pré-processamento aplicadas a uma imagem. O capítulo 4 abrange as principais técnicas de filtragem de imagens no domínio espacial ou freqüencial, tanto com o objetivo de realçá-las como de remover ruído eventualmente presente sobre elas. Também são apresentados conceitos de processamento de imagens coloridas e de filtragem adaptativa. Já o capítulo 5 é inteiramente dedicado a técnicas de processamento e análise de imagens baseadas em morfologia matemática. Por sua vez, o capítulo 6 trata de um tema importante e atual que é a compressão de imagens, desde os fundamentos conceituais até os padrões mais utilizados atualmente.

O capítulo 7 serve como referência prática para o leitor interessado em implementar seu próprio sistema de processamento de imagens, abrangendo informações conceituais e práticas sobre o hardware e software disponível para tanto.

O livro contém ainda dois apêndices. O apêndice A apresenta uma visão resumida, porém abrangente, dos principais formatos de arquivos de imagens disponíveis atualmente. O apêndice B traz um conjunto de roteiros de práticas de laboratório utilizando a *toolbox* de Processamento de Imagens do software MATLAB®. Estas práticas são referenciadas nos capítulos correspondentes ao longo do livro.

Exercícios Propostos

1. Seja o diagrama da figura 4, destacando as principais etapas de um sistema de visão artificial. Supondo que se deseje utilizar os conceitos de análise de imagens para reconhecer placas de veículos à entrada e/ou saída de um estacionamento automatizado, comente quais as tarefas que cada bloco desempenhará na solução do problema, indicando os principais aspectos práticos envolvidos e seu impacto no projeto do sistema completo.

2. Em nosso estudo de visão por computador, vimos que as dificuldades em simular o sistema visual humano residem em três frentes:

- Base de dados muito extensa
- Velocidade de processamento muito alta
- Condições de trabalho muito variadas

Explique cada uma delas e comente o grau de dificuldade relativa de cada uma, bem como a relação entre os progressos tecnológicos em software e hardware e seu impacto na minimização destas dificuldades.

3. Qual a diferença entre os termos 'processamento de imagens' e 'reconhecimento de padrões'?

Na Internet⁴

Para maiores informações sobre alguns dos tópicos deste capítulo, sugerimos as páginas a seguir:

"<http://www.jpl.nasa.gov>"

NASA Jet Propulsion Laboratory

Home page do *Jet Propulsion Lab* da NASA.

⁴ Convém lembrar que a referência a informações disponíveis na Internet está sujeita a problemas decorrentes da natureza extremamente dinâmica da rede. É possível que alguns links indicados neste livro sejam removidos de seus servidores, transferidos para outros servidores ou estejam temporariamente fora do ar.

"<http://www.jpl.nasa.gov/missions/ranger/>"

Ranger: Mission to the Moon

Apresenta detalhes técnicos e históricos das missões Ranger.

Para aprofundar ou complementar outros assuntos tratados neste livro, existem inúmeras páginas úteis e interessantes relacionadas às áreas de processamento de imagens e visão por computador na *World Wide Web*. Relacionamos a seguir algumas delas, classificando-as em seis grupos, a saber:

- Top 10: dez páginas selecionadas pelos autores por apresentarem grande quantidade de informação e inúmeros *links* a outros *sites* de interesse.
- Grupos de pesquisa: páginas com *links* para grupos de pesquisa em visão computacional e processamento de imagens no Brasil e no exterior.⁵
- Publicações: neste grupo estão incluídas informações técnicas sobre processamento de imagens e tópicos correlatos, incluindo páginas relativas a periódicos científicos (*journals*), *sites* de revistas técnicas, páginas de perguntas mais frequentes (FAQs), cursos interativos disponíveis na WWW, (trechos de) livros disponíveis *on-line*, notas de aula de professores da área, resenhas de livros etc.
- Eventos: contém *links* para as páginas dos principais simpósios e conferências mundiais na área.
- Imagens: *sites* contendo imagens para teste e ilustração de resultados de algoritmos.

Top 10

"<http://www.cs.cmu.edu/~cil/vision.html>"

Computer Vision Home Page

Excelente ponto de partida para pesquisas na WWW. Desdobra-se em várias páginas específicas, cada qual contendo inúmeros *links* de interesse.

"<http://www.ph.tn.tudelft.nl/PRInfo.html>"

Pattern Recognition Information

Página de referência para assuntos relacionados ao reconhecimento de padrões e temas correlatos.

"<http://www.lpac.ac.uk/SEL-HPC/Articles/VisionArchive.html>"

SEL-HPC Vision and Image Processing Archive

Repositório de artigos técnicos na área de processamento de imagens. Pode ser atualizado dinamicamente.

"<http://iris.usc.edu/Vision-Notes/bibliography/contents.html>"

USC Annotated Computer Vision Bibliography

Extensa e bem estruturada compilação de quase tudo o que já foi publicado na área. Um *bookmark* obrigatório.

"<http://ecvnet.lira.dist.unige.it>"

ECVNet

Relata atividades de um consórcio europeu na área de visão computacional e contém diversos *links* úteis.

⁵ Se o leitor for pesquisador da área, procure verificar se seu grupo de pesquisa está cadastrado nestes *sites*.

"<http://viswiz.gmd.de/MultimediaInfo>"

Multimedia Info & Resources

Excelente *site* para pesquisa de tópicos ligados a sistemas multimídia, processamento de sons, imagens e vídeo e temas correlatos.

"http://reality.sgi.com/employees/rchiang_esd/TI-ImageProc.html"

Technical Information - Image Processing

Outra ótima página repleta de *links* classificados de forma razoavelmente estruturada.

"<http://peipa.essex.ac.uk/>"

The Pilot European Image Processing Archive: Home Page

Repositório de arquivos de interesse nas áreas de processamento de imagens e visão por computador, mantido pela *University of Essex* (Inglaterra).

"<http://www.vision1.com/>"

The Vision and Imaging Technology Resource!

Guia técnico-comercial de produtos, serviços e recursos para desenvolvedores de soluções em visão computacional. Possui uma extensa biblioteca de *links*.

"<http://vision.arc.nasa.gov/VisionScience/VisionScience.html>"

Vision Science: The World-Wide Web Virtual Library

Guia de referência para pesquisa de temas relacionados a visão biológica (principalmente) e computacional na Web.

Grupos de pesquisa

"<http://www.cs.cmu.edu/~cil/v-groups.html>"

Computer Vision: Research Groups

Página com *links* para os principais grupos de pesquisa em processamento de imagens e visão por computador em todo o mundo.

"<http://www.cs.cmu.edu/~cil/txtv-groups.html>"

Computer Vision: Research Groups (text only)

Versão texto da página indicada acima, adequada para reduzir o tempo de carga da página de abertura na tela.

"<http://www.ph.tn.tudelft.nl/PRInfo/groups.html>"

Pattern Recognition Research Groups

Lista de grupos de pesquisa ativos na área, classificados por continente.

Publicações

"<http://www.lpac.ac.uk/SEL-HPC/Articles/VisionArchive.html>"

SEL-HPC Vision and Image Processing Archive

Repositório de artigos técnicos na área de processamento de imagens. Pode ser atualizado dinamicamente. Selecionado pelos autores deste livro como um dos dez *sites* mais relevantes na área.

"<http://www.nr.com/>"

Numerical Recipes in C

Referência obrigatória para programadores na área científica, este livro está agora disponível *on-line*. A versão PostScript de cada capítulo pode ser obtida gratuitamente no endereço acima.

"<http://www.cs.hmc.edu/~fleck/computer-vision-handbook/index.html>"

The Computer Vision Handbook

Ainda em construção, propõe-se a ser um grande livro-texto *on-line* sobre o assunto, com muitos *links* e referências bibliográficas.

"<http://www.khoros.com/dipcourse/dip17sep97/>"

Digital Image Processing (DIP) with Khoros 2

Pioneira e feliz iniciativa do Prof. Roberto de Alencar Lotufo da Unicamp, em parceria com o Prof. Ramiro Jordán da *University of New Mexico* (EUA), é um curso interativo completo de processamento de imagens na rede. Parte do curso exige a plataforma Khoros.

"<http://www.cogs.susx.ac.uk/users/davidy/teachvision/vision0.html>"

Sussex Computer Vision: Introduction to the HTML teach files

Curso *on-line* de visão por computador com exemplos de programas em POP-11 [Barrett et al. 1985].

"<http://www.ime.usp.br/mac/khoros/mmach.old/tutor/mmach.html>"

A Tutorial on Mathematical Morphology

Curso interativo de Morfologia Matemática desenvolvido pela Universidade de São Paulo sobre a plataforma Khoros 2.0. Possui opção para aqueles que desejarem uma visão geral do assunto de forma rápida e independente de hardware ou software.

"<http://www.cs.washington.edu/research/metip/metip.html>"

Mathematics Experiences Through Image Processing (METIP)

Home-page de inovador projeto desenvolvido pela *University of Washington* que se propõe a utilizar operações de processamento de imagens para motivar e facilitar o ensino de matemática. Permite o *download* gratuito dos títulos de software já desenvolvidos no âmbito do projeto.

"http://www.cm.cf.ac.uk/Dave/Vision_index.html"

MSc AI (and Engineering Application) Vision Systems Course Documentation

Curso abrangendo tópicos clássicos em formato eletrônico, porém pouco interativo e com pequeno número de imagens.

"<http://www.inforamp.net/~poynton/Poynton-T-I-Digital-Video.html>"

A Technical Introduction to Digital Video

Contém o índice do livro homônimo, bem como versões *on-line* dos capítulos 1 e 6.

"<http://www.eecs.wsu.edu/IPdb/title.html>"

Digital Image Processing Home Page

Curso de processamento de imagens em formato de hipertexto, ainda em construção, apresentando a teoria básica, exemplos de imagens e trechos de código-fonte em C.

Eventos

"<http://iris.usc.edu/Information/Iris-Conferences.html>"

Computer Vision Conference Listing from USC

Contém bem diagramada agenda de eventos e *links* para as páginas relacionadas a cada evento.

"<http://www.cs.cmu.edu/~cil/v-conf.html>"

Computer Vision: Conferences and Symposia

Lista de conferências e eventos na área, com *links* para as páginas de cada evento e para outras listas de eventos.

"<http://www.ph.tn.tudelft.nl/PRInfo/conferences.html>"

Pattern Recognition Related Conferences

Outra lista de conferências e eventos na área, com *links* para as páginas de cada evento.

"<http://afrodite.lira.dist.unige.it/confs/confs.html>"

Conferences Main Menu

Página de *links* para conferências e eventos, mantida pela ECVNet.

Imagens

"<http://www.ics.forth.gr/ecvnet/imageDB/index.html>"

ECVNet Image DataBases Page

Página contendo *links* para diversos repositórios de imagens na Internet.

"<http://www.cs.cmu.edu/~cil/v-images.html>"

Computer Vision: Test Images

Outra página contendo diversos *links* para diversos repositórios de imagens na Internet.

Bibliografia

- | | |
|-------------------------|--|
| [Barrett et al. 1985] | Barrett, R. et alii, <i>POP-11: a practical language for artificial intelligence</i> . Ellis Horwood Ltd., 1985. |
| [Dougherty 1994] | Dougherty, E.R. (ed.), <i>Digital Image Processing Methods</i> , Marcel Dekker, 1994. |
| [Gonzalez e Woods 1992] | Gonzalez, R.C. e Woods, R.E., <i>Digital Image Processing - Third Edition</i> , Addison-Wesley, 1992. |
| [Lindley 1991] | Lindley, C.A., <i>Practical Image Processing in C</i> , Wiley, 1991. |
| [Marr 1982] | Marr, D., <i>Vision: A Computational Investigation into the Human Representation and Processing of Visual Information</i> , W. H. Freeman and Co., 1982. |
| [Mascarenhas 1990] | Mascarenhas, N.D., "Introdução ao Processamento Digital de Imagens", <i>Anais da I Jornada EPUSP/IEEE de Computação Visual</i> , 1990, 387-420. |
| [Rimmer 1993] | Rimmer, S., <i>Bit-Mapped Graphics</i> , Windcrest Books, 1993. |
| [Schalkoff 1989] | Schalkoff, R.J., <i>Digital Image Processing and Computer Vision</i> , Wiley, 1989. |

Bibliografia Recomendada

Cada capítulo deste livro contém ao seu final indicações bibliográficas referentes ao assunto abordado. Para uma visão mais abrangente das principais fontes de consulta na área, relacionamos a seguir alguns dos mais renomados livros e periódicos disponíveis atualmente.

Alguns dos principais periódicos científicos (*journals*) ligados às áreas de processamento de imagens e visão por computador são (em ordem alfabética):

- *Computer Vision, Graphics and Image Processing*
- *Graphical Models and Image Processing*
- *IEEE Computer Graphics and Applications*
- *IEEE Expert-Intelligent Systems and their Applications*
- *IEEE Multimedia*
- *IEEE Transactions on Image Processing*
- *IEEE Transactions on Information Theory*
- *IEEE Transactions on Medical Imaging*
- *IEEE Transactions on Multimedia*
- *IEEE Transactions on Pattern Analysis and Machine Intelligence*
- *IEEE Transactions on Signal Processing*
- *IEEE Transactions on Systems, Man and Cybernetics*
- *Image and Vision Computing*
- *Image Understanding, International Journal of Computer Vision*
- *Journal of Mathematical Imaging and Vision*
- *Machine Vision and Applications*
- *Pattern Recognition*
- *Proceedings of the IEEE.*
- *Real-Time Imaging*

Dentre os livros-texto consagrados na área, citamos e recomendamos (por ordem alfabética de autor):

- Castleman, K.R., *Digital Image Processing*, Prentice-Hall, 1995.
- Dougherty, E.R. e Giardina, C.R., *Matrix Structured Image Processing*, Prentice-Hall, 1987.
- Gonzalez, R.C. e Woods, R.E., *Digital Image Processing*, Addison-Wesley, 1992.
- Jain, A.K., *Fundamentals of Digital Image Processing*, Prentice-Hall, 1989.
- Jain, R.C., Kasturi, R., e Schunck, B.G., *Machine Vision*, McGraw-Hill, 1995.
- Lim, J.S., *Two-dimensional Signal and Image Processing*, Prentice-Hall, 1990.
- Lindley, C.A., *Practical Image Processing in C*, Wiley, 1991.
- Marr, D., *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman and Co., 1982.
- Myler, H.R. e Weeks, A.R., *Computer Imaging Recipes in C*, Prentice Hall, 1993.
- Pavlidis, T., *Algorithms for Graphics and Image Processing*, Computer Science Press, 1982.
- Pratt, W. K., *Digital Image Processing*, Wiley Interscience, 1991.
- Russ, J. C., *The Image Processing Handbook*, CRC Press, 1995.
- Schalkoff, R.J., *Digital Image Processing and Computer Vision*, Wiley, 1989.
- Sonka, M., Hlavac, V. e Boyle, R., *Image Processing, Analysis and Machine Vision*, Chapman & Hall, 1993.

