

**CBPF - CENTRO BRASILEIRO DE PESQUISAS FÍSICAS**  
**Rio de Janeiro**

**Notas Técnicas**

**CBPF-NT-001/98**

**Janeiro 1998**

**Processamento de Imagens**

**Flávio Luis de Mello & Márcio Portes de Albuquerque**



Flávio Luis de Mello  
fmello@cat.cbpf.br

Márcio Portes de Albuquerque  
mpa@cat.cbpf.br

## *Resumo*

Esta nota técnica tem por objetivo apresentar algumas noções básicas sobre processamento de imagens. A motivação deste trabalho surgiu com o intuito de apresentar as facilidades e dificuldades de se tratar imagens. Pretende-se fornecer informações úteis sobre o pré-processamento, uma fase vital ao processamento de imagens propriamente dito. Espera-se que ao final deste trabalho, o leitor possa ter uma leve compreensão do universo que gira em torno desta disciplina e assim, que ele possa explorar os recursos que as imagens podem oferecer para as suas próprias necessidades.

Inicialmente são discutidos assuntos introdutórios que comentam sobre o hardware envolvido no processo, desde a aquisição da imagem até os dados obtidos através do seu processamento. São apresentadas algumas análises que podem ser realizadas bem como possíveis melhorias que podem ser feitas em uma imagem, e ainda, é feito um paralelo entre o sistema de visão computacional e o humano.

Em seguida, é feito um detalhamento das técnicas de processamento e em especial serão discutidos os elementos principais de um sistema digital de processamento de imagens. São apresentados alguns conceitos interessantes tais como a discretização, a medida de distâncias em imagens digitais, a conectividade entre pixels, as técnicas de filtragem espacial e de segmentação de imagem.

Por fim, é fornecido uma breve coleção de termos amplamente usados nesta disciplina. Eles são descritos de forma clara e sucinta de modo que permitam uma consulta rápida e precisa.

# Índice

<b>RESUMO.....</b>	<b>1</b>
<b>PROCESSAMENTO DE IMAGENS: DIFICULDADES E SOLUÇÕES.....</b>	<b>3</b>
INTRODUÇÃO.....	3
ANÁLISE.....	4
“IMAGE ENHANCEMENT”.....	5
SISTEMA GERAL DE VISÃO.....	5
SISTEMAS EXPERIMENTAIS.....	6
<b>PROCESSAMENTO DE IMAGENS: MÉTODOS E ANÁLISES.....</b>	<b>8</b>
SISTEMA DE PROCESSAMENTO DE IMAGENS.....	8
PIXEL : ORGANIZAÇÃO EM UMA MATRIZ QUADRADA.....	10
DISTÂNCIAS.....	11
FILTRAGEM.....	12
SEGMENTAÇÃO.....	13
RECONHECIMENTO.....	14
CONCLUSÃO.....	15
<b>DICIONÁRIO DE TERMOS.....</b>	<b>16</b>
<b>BIBLIOGRAFIA.....</b>	<b>19</b>

# *Processamento de Imagens: Dificuldades e Soluções*

## *Introdução*

O Processamento de imagens é sem dúvida uma disciplina em crescimento. Nela são abordados diversos temas científicos e em alguns casos de caráter interdisciplinar. Entre eles podemos citar : a análise em multi-resolução e em multi-frequência, a codificação, a transmissão de imagens, a análise estatística, a compreensão de imagens etc. Mas o que faz do processamento de imagens uma disciplina tão particular e tão difícil ? O que faz com que nós ainda não tenhamos sistemas de alta performance de reconhecimento de caracteres ou de formas mais complexas ? E por que esta técnica é solicitada em tantas e diferentes disciplinas ? A resposta a estas questões tem na realidade um caráter filosófico por trás delas. Ao longo desta fase inicial do trabalho vamos discutir alguns problemas e algumas aplicações que demandam técnicas especializadas de processamento de imagens com o intuito de compreender e determinar algumas respostas a estas questões.

A disciplina “Processamento de Imagens” vem na realidade do Processamento de Sinais. Os sinais, como as imagens, são na realidade um suporte físico que carrega no seu interior uma determinada INFORMAÇÃO. Esta INFORMAÇÃO pode estar associada a uma medida (neste caso falamos de um sinal em associação a um fenômeno físico), ou pode estar associada a um nível cognitivo (neste caso falamos de conhecimento). Processar uma imagem consiste em transformá-la sucessivamente com o objetivo de extrair mais facilmente a INFORMAÇÃO nela presente. Cabe neste momento fazer uma breve diferenciação entre o que é Processar uma Imagem e o que é Computação Gráfica, técnica que nós estamos habituados através de seqüências animadas encontradas na televisão ou em filmes de cinema (como “Terminator” ou “Toy Story” da Companhia americana Pixstar/Sun). Quando realizamos Computação Gráfica partimos de uma idéia precisa, ou seja de uma INFORMAÇÃO absoluta do que queremos obter com o filme ou com a imagem animada. De uma forma resumida, podemos dizer que o roteiro do filme já foi idealizado e planejado por antecedência, para só depois chegarmos as imagens. Já o Processamento de Imagens parte da imagem (obtendo uma informação inicial que é geralmente captada por uma camera) ou de uma seqüência de imagens (onde somamos a informação da camera uma informação temporal) para obtermos a “INFORMAÇÃO”. Deste ponto de vista o Processamento de imagens e a computação gráfica são exatamente métodos opostos, mas isto não quer dizer que as técnicas envolvidas em cada caso não possam ser as mesmas ou pelo menos complementares. É evidente que neste sentido processar uma imagem, como é feito pelo sistema visual humano (SVH), é extremamente complexo. Realizar as mesmas tarefas que o SVH, com a ajuda de máquinas, exige por antecedência uma compreensão filosófica do mundo ou dos conhecimentos humanos! Esta característica faz com que o processamento de imagens seja, atualmente, uma disciplina extremamente dependente do sistema no qual ele está associado, não existindo no entanto uma solução única e abrangente para todos os problemas. Daí a NÃO existência, até o momento, de sistemas de análise de imagens complexos e que funcionem para todos os casos.

A análise quantitativa e a interpretação de imagens representa atualmente um ponto de apoio importante em diversas disciplinas científicas. Tal é o caso por exemplo na ciência dos materiais, na biofísica, na medicina, na física da matéria condensada, etc. Na realidade a diversidade de aplicações do processamento de imagens, está associada diretamente a análise da INFORMAÇÃO que falamos acima. Pois em todas estas disciplinas estamos na realidade em busca de informações quantitativas que representem um fenômeno estudado. Quando observamos do ponto de vista da ótica, uma imagem é um conjunto de pontos que convergem para formar um todo, mas podemos dizer de uma maneira mais ampla que uma imagem é o suporte para efetuarmos troca de INFORMAÇÕES. Por exemplo a imagem é a que você lê este artigo ou aquela que você identifica os objetos a sua volta. Mesmo se inicialmente o termo imagem estava associado somente ao domínio da luz visível, atualmente é muito freqüente ouvirmos falar de imagens quando uma grande quantidade de dados estão representados sob uma forma bidimensional (por exemplo: as imagens acústicas, sísmicas, de satélites, infravermelhas, magnéticas). Os métodos recentes de exploração automática desta informação permitiu o desenvolvimento de técnicas complexas, que podem ser globalmente classificadas em dois grandes ramos. O primeiro está associado a uma ANÁLISE da informação e o segundo representa as técnicas que permitam obter uma MELHORIA (do termo em inglês "ENHANCEMENT") significativa da imagem. Nos parágrafos posteriores vamos discutir cada uma destas linhas.

## *Análise*

O termo análise está relacionada a parte do tratamento onde existe uma descrição da INFORMAÇÃO presente na imagem. Esta parte é chamada de parametrização e é nela que várias medidas quantitativas (parâmetros) são utilizadas para descrever diferentes informações dentro de uma imagem. Algumas aplicações típicas são: a determinação do número de células presentes em um tecido biológico, o cálculo das formas dos contornos de uma célula ou ainda a determinação da distribuição de uma população específica de um conjunto de células. As técnicas dedicadas a análise de imagens podem variar enormemente segundo a sua complexidade e a necessidade em tempo de processamento. É nesta área que vamos encontrar um nível de complexidade mais elevada. Um exemplo prático e quando da classificação automática de células doentes dentro de um conjunto de células observadas em microscopia. Esta análise específica demanda soluções dadas em técnicas de "CLASSIFICAÇÃO E RECONHECIMENTO DE FORMAS", que discutiremos brevemente em uma seção posterior. Neste caso devemos rapidamente medir vários parâmetros, pertinentes ao problema, na imagem, como por exemplo: a superfície, a forma de cada célula, sua quantidade, o número de células vizinhas a uma dada célula e a densidade de células em uma dada região. Em seguida comparamos estas medidas com várias classes de células organizadas em uma base de dados, catalogadas anteriormente. Obteremos então uma classificação das células com uma dada probabilidade de serem células doentes ou normais.

## “Image Enhancement”

O termo “enhancement” esta associado a melhoria da qualidade de uma imagem, com o objetivo posterior de ser julgado por um observador humano. De uma forma geral nós vamos trabalhar nos níveis de cinza da imagem, transformando-os para aumentar o contraste ou para colocar em evidência alguma região de interesse particular. Alguns exemplos deste tipo de técnica é a subtração da imagem por uma imagem referência, a utilização de cores-falsas, a utilização de filtros espaciais (conhecidos como “convolution kernels”), a correção de deformações espaciais devido a ótica ou devido a uma variação de inomogeneidade da iluminação de fundo. Os sistemas dedicados a melhorar a qualidade da imagem trabalham geralmente muito rápido, pois são construídos em “hardware” ou “firmware”, permitindo rapidamente ao usuário um julgamento sobre várias imagens processadas, segundo o tipo de tratamento. Esta técnica é encontrada na maioria dos programas de tratamento de imagens ou fotografias que estão atualmente no mercado, mas com algoritmos implementados em software, para computadores do tipo PC, como por exemplo: o Adobe Photoshop (da Adobe Sistemas), o Aldus PhotoStyler (da Kodak), o Corel PhotoPaint (da “Corel Corporation”) etc.

## Sistema Geral de Visão

Um sistema geral de visão, isto é, um sistema que deve responder a todas as situações, como é o caso do Sistema Visual Humano é extremamente desenvolvido e complexo. O interesse em se compreender um tal sistema é fundamentalmente de orientar a pesquisa de uma máquina de reconhecimento genérica com as mesmas características do SVH. Atualmente algumas correntes científicas consideram o cérebro humano uma máquina com grande capacidade de explorar as informações sensoriais. Filosoficamente não se pretende, com estas linhas de pesquisas, retirar as qualidades humanas associadas ao cérebro ou aos homens. Mas devemos atentar para alguns pontos fundamentais: os computadores atuais são extremamente eficientes para trabalharem com base de dados, cálculos numéricos e formais, mas eles não conseguem realizar eficientemente algumas tarefas “simples” realizadas por animais e seres-humanos. Fica então uma questão, que é parte integrante das linhas de pesquisas modernas em processamento de imagens: Seria realmente o computador a máquina mais adaptada para resolver estes problemas ? Um exemplo prático é o processo de reconhecimento de uma pessoa. O trabalho realizado pelo cérebro nos parece muito simples e bastante elementar ; quando *observado de perto* o cérebro humano é capaz, a partir de uma grande quantidade de informações de luminância e crominância (captada pelos olhos) de realizar esta tarefa de reconhecimento em apenas 150ms. Os computadores atuais são na pratica sistemas experimentais que nos ajudam a melhor compreender estas tarefas e conseqüentemente modelizá-las e reproduzi-las.

## Sistemas Experimentais

Quando o sistema se restringe a sistemas experimentais, isto é, ele é orientado para a solução de um problema específico, a tarefa prática é de uma certa forma simplificado, pois vamos poder controlar alguns parâmetros que um sistema geral de visão não tem acesso. A dificuldade está no fato de que devemos dar uma descrição objetiva ao problema e para aquelas análises que são científicas é ainda importante que esta descrição seja quantitativa. O objetivo neste caso é desenvolver um sistema de tratamento da informação onde cada fase vai exigir métodos para retirar da imagem os parâmetros que exprimam esta informação. Fica assim evidente que o processamento de imagens, quando associado a sistemas experimentais, é bastante dependente do problema a ser resolvido. Neste instante as fases de pré-tratamento (seja da imagem ou fora da imagem, como o controle da iluminação) podem ter um peso importante quando queremos aumentar a performance do sistema. Um sistema de processamento de imagens pode ser de uma maneira geral dividido nas seguintes etapas:

- 1 - Fases de pré-tratamento (correção de iluminação, uso de colorantes...).
- 2 - Aquisição (possível compactação)
- 3 - Melhoramento ("image enhancement")
- 4\* - Segmentação da informação
- 5\* - Determinação de grandezas (área, perímetro, forma, descrição estrutural, topologia,...)
- 6\* - Reconhecimento (classificação)
- 7\* - Associação das grandezas ao problema : determinação de funções de correlação espacial ou temporal, análise de seqüência de imagens, etc.

\* - Fases com extrema dependência ao problema.

## Conclusão

O processamento de imagens está baseado em métodos matemáticos que permitem descrever quantitativamente imagens das mais diversas origens. Uma imagem pode, de alguma forma, ser descrita independentemente do que ela representa e, *a priori*, todos os parâmetros que tem uma característica bidimensional ou topológico são convenientes. Em cada objeto definido em um espaço 2D nós podemos efetuar medidas de superfície, perímetros, comprimentos, espessura, posição, etc, e em seguida deduzir grandezas estatísticas de uma forma automática. É importante ressaltar que a análise automática é imprescindível quando queremos efetuar transformações sucessivas na imagem.

Neste trecho do trabalho discutimos questões importantes e pertinentes, essenciais para a compreensão de um sistema de processamento de imagens aplicado. Na seção seguinte vamos entrar mais em detalhes nas técnicas de processamento e em especial discutiremos os elementos principais de um sistema digital de processamento de imagens. Discutiremos também as técnicas e

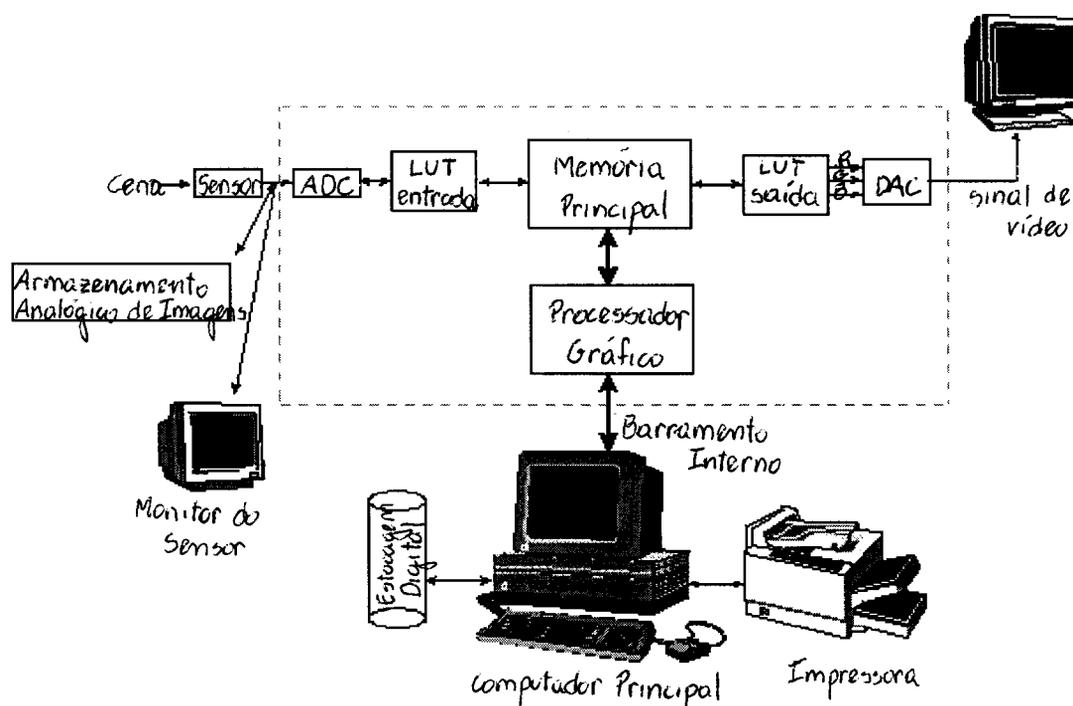
os problemas que podemos encontrar na implementação de sistemas de processamento de imagens: como a discretização, a medida de distâncias em imagens digitais, a conectividade entre pixels etc. Veremos também técnicas de filtragem espacial e de segmentação de imagem.

## Processamento de Imagens: Métodos e Análises

### Sistema de Processamento de Imagens

Na seção anterior descrevemos conceitos fundamentais que estão atrás do processamento de qualquer imagem. Nesta parte do trabalho vamos discutir técnicas e métodos aplicados em processamento de imagem. Vamos primeiro discutir os elementos principais num sistema de processamento digital de imagens, em seguida vamos falar dos efeitos de uma imagem discreta e as conseqüências de medidas de distância em tais imagens. Vamos descrever conceitos básicos de filtragem espacial utilizado em imagens e falar um pouco de reconhecimento e análise de formas.

Um sistema de processamento de imagens pode ser dividido em diversos componentes, como mostrados na figura 1.



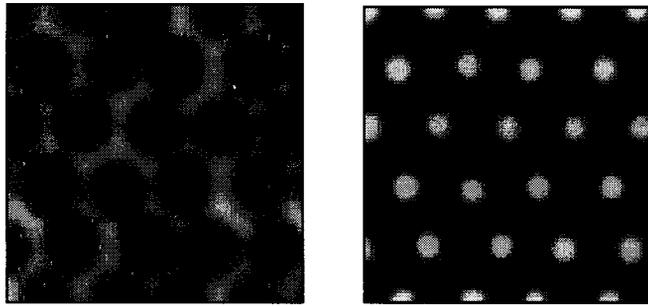
**Figura 1** - Diagrama em blocos de um sistema genérico de processamento de imagens. O retângulo tracejado indica as partes encontradas atualmente nos sistemas de processamento de imagens comercializados. O poder do processamento está associado a potencialidade do processador gráfico e a capacidade de armazenamento de imagens na memória principal. A organização desta memória pode variar segundo a resolução ou a necessidade de estocagem de grande quantidade de imagens.

O módulo de entrada é constituído pelos sensores que captam a informação em duas dimensões e em seguida, enviam através de sinais padronizados de vídeo, à um conversor analógico digital. Devemos lembrar que uma imagem digital é discretizada espacialmente (ou seja em  $x$  e  $y$ ) e em luminância (níveis de cinza). Uma imagem típica de  $640 \times 480$  por 256 níveis de cinza tem 307.200 bytes sem compressão. Alguns exemplos de módulos de aquisição são: as câmeras CCD ou

Vidicom, os scanners, algumas sondas dedicadas (como sistema de infravermelho, mapeamento magnético de uma superfície, sensoriamento de radares, etc). A câmera é o elemento de aquisição de imagem mais freqüente, e podemos nos interessar principalmente na suas características de sensibilidade, resolução e função de transferência entre a imagem captada da cena e a intensidade do sinal de vídeo fornecido. Algumas câmeras mais sofisticadas permitem ainda controlar o offset ou ganho do sinal de vídeo, possibilitando uma correção do contraste da imagem. Em alguns casos podemos ainda ajustar a iluminação de fundo ou realçar os contornos dos objetos através da execução de funções específicas como derivadas espaciais.

Após a conversão da imagem pelo modulo de conversão analógica digital, cada intensidade indexa uma tabela de conversão. Esta tabela é conhecida como "Look-up table" ou simplesmente LUT. As LUTs são na realidade memórias de acesso rápido que conseqüentemente possibilitam a modificação de toda a imagem em curtos intervalos de tempos. Podemos por exemplo configurar uma LUT com uma função específica e corrigir o contraste de uma imagem que esta sendo adquirida pelo ADC. A imagem digital é em seguida armazenada em uma memória principal de acesso exclusivo de um processador gráfico. Este processador é bastante dependente do sistema de processamento de imagens utilizado. Em alguns casos ele pode mesmo não existir, deixando todo o processamento para o processador principal no computador de controle.

No caso de imagens coloridas cada componente da imagem ("red", "green" e "blue") é armazenada separadamente na memória principal. Neste trabalho vamos considerar somente imagens em níveis de cinza, pois grande parte da análise de imagens coloridas são feitas recombinação das componentes de cores e trabalhando com a informação de luminância somente. A imagem de saída é geralmente reconstruída através de conversores digitais - analógicos (DAC), onde são somados pulsos de sincronização para geração do sinal de vídeo. Antes de ser reconvertido o sinal pode passar por outra LUT, desta vez a LUT de saída, que nos permite ajustarmos regiões da imagem em função da intensidade dos pixels. Podemos assim associar falsas cores e observarmos "on-line" diferentes regiões, separadas por diferentes níveis de cinza (técnica conhecida como "*density slice*"). A figura 2 apresenta um exemplo da técnica de "falsas cores" usada em uma imagem de aglomerações de líquidos magnéticos. Resumidamente as células observadas na figura correspondem a uma maior densidade do líquido magnético, observada através de uma técnica conhecida como *Densimetria*. A fonte luminosa envia um feixe de luz através do líquido. Cada pixel na imagem digital representa na realidade uma medida da absorção da luz pelo líquido. Nas regiões mais escuras a absorção da luz é maior, nas mais claras esta é menor. As falsas cores permitem rapidamente relacionar as regiões equivalentes (em intensidade luminosa) em diferentes partes da imagem.



*Figura 2* - Uso de falsas cores : A imagem da esquerda corresponde a uma estrutura hexagonal de um líquido magnético semi-transparente observada em detalhes. Cada célula apresenta aproximadamente 0.5cm de diâmetro. Cada ponto da imagem corresponde a uma medida de absorção da luz. A absorção é maior no centro das células que nas suas bordas. Numa imagem em níveis de cinza é praticamente impossível identificar as mesmas intensidades (absorção) em diferentes objetos. A imagem da direita corresponde a mesma da esquerda, mas com outra LUT de cores. Cada nível, ou regiões contíguas são associadas a diferentes cores, indo do preto ao amarelo, passando pelo vermelho.

Um conceito importante em processamento de imagens é a Região de Interesse. Por Região de Interesse (ROI - “Region Of Interest” ou AOI - “Area Of Interest”) entende-se como sendo a região definida pelo operador (ou automaticamente a partir de parâmetros obtidos na própria imagem) onde o processamento estará totalmente concentrado. Podemos, por exemplo, definir uma região de interesse onde sabemos por antecedência que a iluminação de fundo é constante ou foi corrigida. Em geral em processamento de imagens iremos sempre encontrar problemas nas bordas da imagem. A criação de regiões de interesse faz com que apareçam também problemas nas bordas das regiões de interesse. A solução deste problema é completamente dependente do problema ao qual o processamento de imagens está associado, i.e., ao que conhecemos e podemos levar em consideração anteriormente ao processamento.

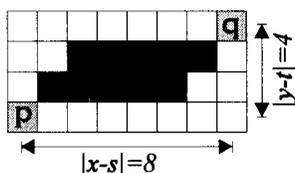
### *Pixel : organização em uma matriz quadrada.*

Um pixel (“picture element” ou ainda “pel”) é a unidade de amostragem (retangular ou quadrada) de uma imagem e também um elemento de dimensões FINITAS na representação de uma imagem digital. Frequentemente, a organização de uma imagem sob a forma de uma matriz de pixels é feita em uma simetria quadrada, i.e., na forma de um tabuleiro de xadrez. Isto se deve a facilidade de implementação eletrônica seja dos sistemas de aquisição ou de visualização de imagens. É importante lembrar que este tipo de organização provoca o aparecimento de dois problemas importantes nas técnicas de processamento. Em primeiro lugar um pixel não apresenta as mesmas propriedades em todas as direções, i.e., ele é anisotrópico. Esta propriedade faz com que um pixel tenha 4 vizinhos de *borda* e 4 vizinhos de *diagonal*. Esta propriedade nos força a definir o tipo de conectividade que vamos trabalhar, ou D4 (onde levamos em consideração apenas os vizinhos de borda) ou em D8 (onde levamos em consideração os vizinhos de borda e os de diagonal). O segundo problema é consequência direta do primeiro, ou seja as distâncias entre um ponto e seus vizinhos não é a mesma segundo o tipo de vizinho (ela é igual a 1 para vizinhos de borda e  $\sqrt{2}$  para aqueles na diagonal). Alguns tipos de algoritmos que são sensíveis a este problema são : as operações morfológicas que usam uma matriz de análise do tipo 3x3, as operações de

esqueletização em análise de formas e principalmente as transformações de Distâncias (na transformação de distância cada ponto da imagem ao invés de representar uma intensidade luminosa, representa uma distância de um dado ponto referência ao ponto calculado). A solução para este tipo de problema é a correção dos valores calculados através de mascaras (pequenas matrizes) que ajustam ou ponderam estas distâncias em função da direção. O leitor deve atentar para as possíveis complicações em medidas de distâncias quando trabalhamos com um pixel que não seja de forma quadrada.

## Distâncias

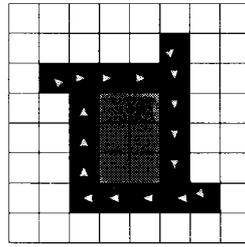
Quando trabalhamos com imagens é muito freqüente medirmos distâncias. Em muitos casos usamos a definição clássica de distância euclidiana entre dois pontos. Porém devido aos problemas apresentados no parágrafo anterior se torna evidente que a simetria quadrada interfere na execução de alguns algoritmos. Existem diferentes medidas de distância em função da conectividade utilizada. Por exemplo se considerarmos dois pontos  $p(x,y)$ ,  $q(s,t)$ , da figura 3, teremos as seguintes definições de distância.



**Figura 3** - Conexão entre dois pontos  $p(x,y)$  e  $q(s,t)$ . A distância entre estes dois pontos pode ser determinada de diversas maneiras. Os números 1 e 2 representam os pixels que seriam percorridos somente uma vez nas duas representações possíveis quando saímos de  $p$  para  $q$ . O número 3 indica os pixels percorridos duas vezes.

- Distância Euclidiana :  $De(p, q) = \sqrt{(x - s)^2 + (y - t)^2}$
- Distância “Manhattan” (em analogia a organização em quarteirões regulares de cidade de “Manhattan” - Nova York) :  $D4(p, q) = |x - s| + |y - t|$ . É importante lembrar que em D4 só podemos nos movimentar para a horizontal e para vertical. Os caminhos mostrados pela figura 3 não servem para D4.
- Distância Tabuleiro de Xadrez (em analogia a um tabuleiro de Xadrez) :  
 $D8(p, q) = \text{Max}(|x - s|, |y - t|)$ . Veja os dois caminhos representados na figura 3.

Por exemplo, considere a medida do perímetro de uma determinada forma, como aquela mostrada na figura 4. É importante lembrar que antes de começarmos a medir distâncias devemos antes isolar a forma do resto da image (discutiremos este tópico no capítulo dedicado a *Segmentação*) e depois devemos conectar todos os pixels que pertençam ao seu contorno. Na forma representada na figura 4 estamos supondo já conhecermos todos os pontos que pertençam a sua borda, para em seguida calcularmos o seu perímetro.



*Figura 4* - Determinação do perímetro de uma forma usando D8. Após isolar o objeto do fundo da imagem, devemos identificar os pontos que pertençam ao seu contorno. Escolhemos um ponto inicial ao acaso e em seguida estes pontos devem ser conectados para efetuarmos o cálculo do seu comprimento. O leitor deve atentar para o fato de que segundo o tipo de conectividade adotada teremos valores diferentes para o perímetro.

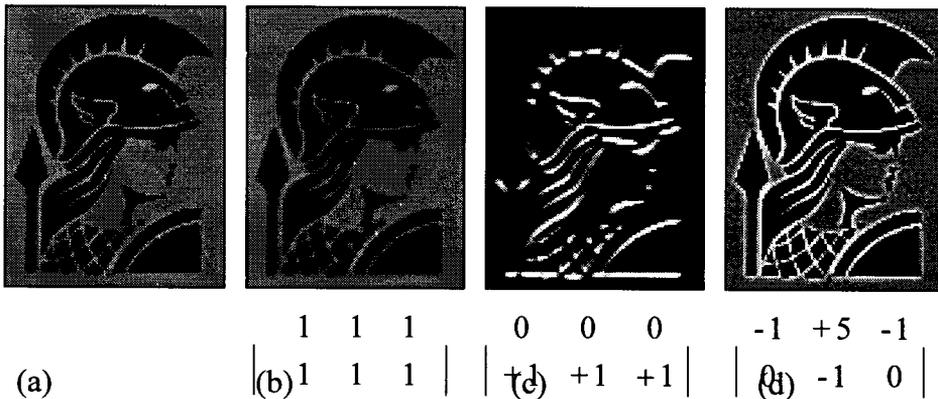
A seqüência composta pelas flechas indicadas na figura correspondem as conexões que devemos fazer para calcularmos o perímetro da forma. O valor determinado para o perímetro será diferente, segundo as diferentes conectividades adotadas. O valor obtido poderá ser usado em seguida para reconhecermos esta forma entre outras formas. Desta maneira é muito importante uma compatibilidade entre a conectividade adotada e os dados armazenados numa Base de Dados.

## Filtragem

Em processamento de imagens, os filtros lineares são geralmente descritos através de matrizes de “convolução”. Não iremos aqui entrar em detalhes de filtros lineares mas apenas descrever basicamente alguns pontos importantes. Um filtro numérico vai influenciar a variação da frequência espacial em uma imagem. Na frequência temporal a escala usada é geralmente o Hertz ( $s^{-1}$ ), já em uma imagem usamos o 1/metro ( $m^{-1}$ ) ou 1/pels ( $pix^{-1}$ ). O termo frequência espacial é análogo ao termo frequência temporal e ela descreve a velocidade de modificação de uma luminosidade em uma imagem. Na prática, para realizarmos uma operação de filtragem espacial, devemos escolher uma matriz de dimensão  $n \times n$  com valores que dependem do filtro que queremos usar, seja ele passa baixa (filtrando as altas frequências), passa faixa (filtrando uma região específica de frequências espaciais) ou passa alta (filtrando as baixas frequências). Em uma imagem as altas frequências correspondem as modificações abruptas dos níveis de cinza, i.e., as bordas dos objetos. As baixas frequências correspondem as variações suaves dos níveis de cinza. Logo quando queremos evidenciar os contornos de um determinado objeto podemos usar filtros do tipo passa-alta. Em outros casos podemos estar interessado na forma da iluminação de fundo, onde devemos usar filtros passa-baixa para eliminarmos todas as altas frequências correspondendo a borda dos objetos.

Resumidamente na figura 5 apresentamos 3 filtros clássicos representados em uma pequena matriz  $3 \times 3$  e seu efeito numa imagem exemplo (Fig. 5a). Toda as operações realizadas consistem em gerar uma nova imagem onde cada ponto desta imagem corresponde a uma soma dos produtos termo a termo da matriz  $3 \times 3$  com imagem original. Por exemplo a aplicação do filtro Médio (todos os elementos igual a 1 - Fig. 5b) em uma imagem, equívale a um imagem final onde cada pixel

corresponde a media local, na imagem original, dele com os 8 pixels vizinhos. Veja o seu efeito na figura 5b. Considere agora o efeito do filtro passa-alta da figura 5c. Todas as variações fortes horizontais, que correspondem as bordas horizontais do objeto, ficaram na imagem final filtrada. Não detalhamos aqui os problemas de representação em números inteiro dos níveis de cinza e de operações em “ponto flutuante”, como média ou médias ponderadas. Neste caso a imagem deve ser renormalizada em função de máximos e mínimos.



**Figura 5** - Matrizes de Filtragem de uma imagem. (a) - Imagem original. (b) - A primeira equivale à um filtro Médio (passa-baixas). Os contornos perdem ligeiramente o “foco”. (c) - Filtro de detecção de variações bruscas no sentido horizontal (Passa-Alta). (d) - Filtro também Passa-Alta que melhora o contraste da imagem realçando os contornos da Minerva - UFRJ.

## Segmentação

A segmentação consiste na primeira etapa de processamento da imagem quando consideramos uma análise do ponto de vista da INFORMAÇÃO nela presente. O termo segmentação vem do termo em inglês “*image segmentation*”, criado durante os anos 80. Esta área representa até hoje uma linha de pesquisa importante do processamento de imagens, principalmente por ela estar na base de todo o processamento da informação. Segmentar consiste na realidade em dividir a imagem em diferentes regiões, que serão posteriormente analisadas por algoritmos especializados em busca de informações ditas de “alto-nível”. Por exemplo cada pixel na imagem da figura 2 poderia ser segmentadas em duas regiões: aqueles pertencentes as células e aqueles pertencentes ao fundo da imagem. A imagem obtida neste caso é composta por apenas duas regiões, por exemplo uma região branca (fundo) e outra preta (células/objeto) e esta é conhecida como *Imagem Binária*. Devido as grandes facilidades na manipulação deste tipo de imagens, principalmente porque reduzimos consideravelmente a quantidade de dados, elas são enormemente utilizadas.

Existem diversas técnicas de segmentação de imagens, mas não existe nenhum método único que seja capaz de segmentar todos os tipos de imagem. Globalmente, uma imagem em níveis de cinza pode ser segmentada de duas maneiras: ou consideramos a semelhança entre os níveis de cinza ou consideramos as suas diferenças, pois é na borda dos objetos que estão as diferenças entre os níveis de cinza. A detecção de um contorno de um objeto, através de matrizes do tipo Passa-Alta, é um exemplo de técnicas baseado nas diferenças. Neste caso estamos segmentando as imagens em regiões que pertencem a borda do objeto e aquelas que pertencem ao interior ou fundo da imagem.

## *Reconhecimento*

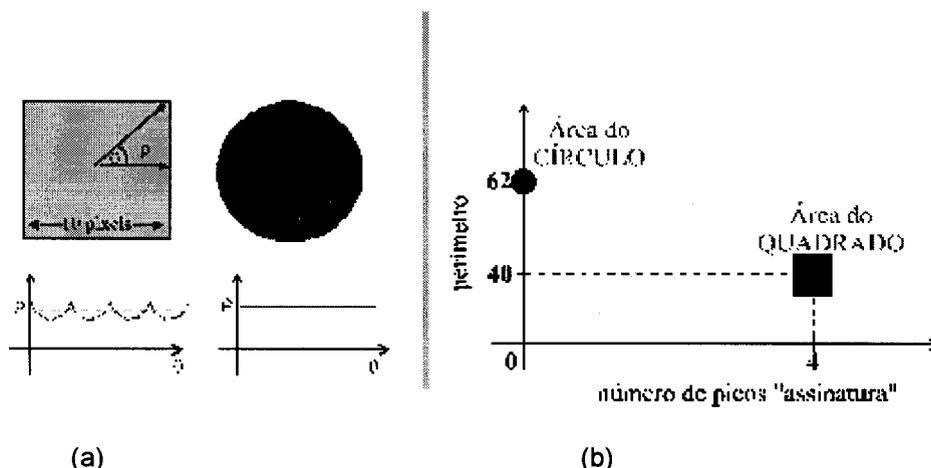
Reconhecimento é a parte do processamento que vai classificar os objetos à partir de informações encontradas na imagem, geralmente em acordo com um banco de dados. Esta fase é geralmente aplicada após uma fase de segmentação da imagem e de parametrização. A fase de parametrização identifica e calcula alguns parâmetros (pré-determinados) nos objetos segmentados (como exemplo de parâmetro podemos citar o perímetro de uma determinada forma).

A título de exemplo considere a “assinatura” de uma forma. A representação de uma forma pela sua assinatura é simples e permite facilmente compreender a análise envolvida para no reconhecimento. Por “assinatura de uma forma” se compreende como a função representativa da distância de todos os pontos da borda do objeto em relação ao seu centro de gravidade. Veja na figura 6 a representação de duas formas clássicas (o QUADRADO e o CÍRCULO) e suas respectivas assinaturas. A assinatura de um CÍRCULO é uma única linha reta, paralela ao eixo das abcissas; já a assinatura de um QUADRADO é uma forma particular contento 4 picos, associados aos seus 4 vértices. Repare que a representação de uma forma através da sua assinatura faz com que a função obtida seja independente da orientação da forma em relação a imagem e se renormalizada podemos ainda eliminar a dependência com a escala (zoom).

Em seguida ao cálculo da assinatura devemos montar um espaço de parâmetros, como mostrado na figura 6d. Suponhamos que o nosso espaço seja de 2-dimensões, onde representamos no eixo X o número de picos encontrado na assinatura da forma e no eixo Y o valor do seu perímetro. No gráfico da figura 6d o ponto em vermelho representa a forma CÍRCULO e o azul a forma QUADRADA. A linha pontilhada que divide o espaço em duas regiões é considerada o limite entre a forma CÍRCULO e a forma QUADRADA para o espaço 2D adotado. O QUADRADO tem um perímetro igual à 40 pixels e o CÍRCULO igual à  $2\pi r = 62.8$  pixels. O CÍRCULO não apresenta nenhum pico na sua assinatura; já o QUADRADO apresenta 4. Repare que uma eventual imprecisão na medida do perímetro ou do número de picos na assinatura não atrapalhará a identificação das formas.

Para terminar é importante lembrar que para diferenciarmos CÍRCULOS de QUADRADOS poderíamos utilizar apenas um dos dois parâmetros para separar as duas formas, porém o objetivo

aqui é introduzir as noções de “espaço de medida” e reconhecimento que a do cálculo preciso entre CÍRCULO e QUADRADOS. Vale a pena lembrar que estes espaços são muito importantes e extremamente utilizados em reconhecimento de formas, seja por métodos clássicos ou aqueles que usam Redes Neurais.



**Figura 6** - Uso de “espaço de medidas” no reconhecimento de duas Formas. A esquerda as formas Quadrado e Círculo são representadas com as suas respectivas “assinaturas” (vide texto). A direita o espaço de medida é representado pelo perímetro e pelos picos encontrados na função assinatura.

## Conclusão

Nesta seção discutimos técnicas básicas e introdutórias de processamento de imagens e quando foi o caso exemplificamo-as com exemplos práticos diretamente em uma imagem. Como citado na primeira parte deste trabalho devemos lembrar que o processamento de imagens é uma técnica extremamente dependente do problema que queremos resolver. Muitos dos procedimentos usados podem ser caracterizados de rudimentares ou mesmo artesanais pois serão específicos para cada caso.

Devemos comentar ainda que o investimento em equipamento modernos de processamento de imagens são ainda muito caros. Uma boa placa de processamento de imagens com um bom software de análise pode custar mais de 100 mil reais (e são dificilmente encontrados no Brasil). Alguns fabricantes importantes de sistemas de processamento de imagens são: Data Translation, Matrox, Hamamatsu, e outros. Do ponto de vista do grande público podemos encontrar sistemas de processamento de imagens mais simples, mas em geral muito dependente do processador do computador onde esta associado a placa de processamento. Um sistema barato pode custar na faixa de 1.500Us\$, permitindo apenas a aquisição e alguns tratamentos que visam melhorar apenas a qualidade visual da imagem. É importante lembrar que existem, para o mundo Macintosh e Unix, sistemas de boa qualidade de domínio público que foram desenvolvidos por organizações governamentais americanas. Para o Macintosh o software é o NIH-Image do Instituto Americano de Saúde (“National Institute of Health”) ver : [ftp://anonymous : zippy.nih.gov \(pub/nih-image\)](ftp://anonymous.zippy.nih.gov/pub/nih-image). Para o Unix existe o sistema Khoros, desenvolvido pela Universidade do Novo México, ver <http://www.khoros.unm.edu/>.

## Dicionário de Termos

**Área de Interesse.** Entende-se como sendo a região definida pelo operador (ou automaticamente a partir de parâmetros obtidos na própria imagem) onde o processamento estará totalmente concentrado. Podemos, por exemplo, definir uma região de interesse onde sabemos por antecedência que a iluminação de fundo é constante ou foi corrigida por ter esta característica.

**Arquitetura RISC (*Reduced Instruction Set Computer*).** Arquitetura moderna implementada nos microprocessadores (e.g., Motorollas Power PC). Técnica usada com o objetivo de acelerar o processamento de poucas e eficientes instruções. Esta arquitetura implementa técnicas de "pipeline" nas microoperações internas do processador, permitindo a realização de uma instrução a cada ciclo de relógio.

**Binário.** Um sistema a dois estados simbólicos. Geralmente são representados em 0 e 1.

**Bitmap x Vetorial.** Considere por exemplo uma imagem como sendo uma Pintura ou um Modelo bi/tri-dimensional de elementos básicos.

No primeiro caso a imagem é armazenada como uma imagem "scaneada", i.e., uma série de pontos armazenados sob a forma de linhas e colunas, onde cada ponto representa a intensidade luminosa recebida pelo "scanner". Cada um destes pontos é chamado de "pixel" (picture element) quando lidamos com um monitor de vídeo, ou de "pel" para o caso de uma impressora. Associado a cada pixel podemos também guardar informações de cor em uma forma numérica (ou digital). A imagem inteira é conhecida como imagem bitmap (mapa de bits).

No segundo caso podemos construir a imagem através de formas elementares e geométricas. Estas formas podem ser pontos,

linhas, círculos ou polígonos (por exemplo estes podem ter o seu interior pintados de alguma cor) etc. Neste caso a imagem pode ser considerada como sendo uma descrição bi ou tridimensional e armazenada sob a forma matemática de vetores, descrevendo cada componente básico.

### VANTAGENS E DESVANTAGENS:

Com uma descrição em modo BITMAP, os dados são armazenados diretamente numa memória de vídeo (tipo vídeo RAM). O armazenamento e a leitura deste tipo de imagens é relativamente rápido e fácil. Entretanto o tamanho dos arquivos finais (estacados em disco) de imagens bitmaps são diretamente associados ao tamanho e a resolução da imagem. Se for feito um "zoom" na imagem você terá a sensação de ver pixels grandes e quadrados pois foi atingida a resolução máxima da imagem. O trabalho com imagens do tipo bitmap podem ser extremamente complexos e muito dependente de softwares profissionais (ex. Adobe Photoshop) que disponibilizam ferramentas poderosas para a seleção de áreas precisas da imagem.

Uma imagem descrita na forma vetorial precisará de uma técnica de "RENDERIZAÇÃO", i.e., de uma conversão dos elementos básicos (pontos, círculos, ...) na forma de uma imagem, porque a maior parte das placas de vídeo modernas possibilitam apenas o trabalho com memórias de vídeo organizadas sob forma bitmap. Por outro lado, uma imagem vetorizada não perderá em definição (resolução) ao ser reduzida ou ampliada, pois cada elemento pode ser recalculado em função da nova resolução. Estas imagens são bastante utilizadas pois permitem se adaptar função da qualidade de impressão que temos disponível. Podemos trabalhar com uma imagem na tela do computador (96dpi) e enviar para impressora objetos com maiores resolução (ex.600dpi). Outra grande vantagem de imagens vetoriais é poder selecionar

facilmente diferentes objetos no seu interior e consequentemente isola-los para trabalhar.

**CCD (*Charge Coupled Device*).** Um dispositivo "charge coupled device", dito de carga-acoplada é um dispositivo semicondutor onde um conjunto isolado de cargas são transportados de uma posição do semicondutor para uma posição adjacente, através de pulsos de "clock" por portas digitais seqüenciais. As câmeras que usam tais dispositivos transferem suas cargas de uma região sensível a intensidade luminosa a uma região protegida da intensidade luminosa. O intervalo de tempo define o tempo de exposição da camera à luz.

**Computação Gráfica.** Método no qual os dados de entrada são processados e o produto final é uma imagem ou uma seqüência de imagens (filme). Esta pode ser vista através de um dispositivo de saída gráfica.

**Corrente de Fundo.** A corrente gerada em um fotosensor colocado em um ambiente totalmente escuro. O ruído de fundo ("dark noise") é expresso sob a forma de uma corrente.

**Espaço de Medidas.** Espaço a N-dimensões onde cada dimensão (ou eixo) representa uma "feature" (característica) do objeto. Nos bons espaço para reconhecimentos de formas as nuvens de pontos devem ser isoladas uma das outras.

**Imagem Binária (*Binary Image*).** Uma imagem constituída de somente dois níveis de cinza, e.g, um deles representando o objeto o outro representando o fundo. São bastante usadas em processamentos e análise de imagens devido a simplicidade e velocidade de tratamento.

**Imagem RGB (*Full Color*).** A imagem é representada através das três cores básicas R(Vermelho), G(Verde) e B(Azul). Cada uma das componentes ocupa 8 bits (ver Representação em 8 bits). Temos uma representação real em 24 bits e podemos

representar 16.7 milhões de cores. Os padrões atuais (JPG e TIFF adotam este tipo de representação).

**Look-up Table.** Para se explicar melhor a filosofia de uma LUT será feita uma analogia de uma LUT com o trabalho com cores usando lápis coloridos. Quando se escreve em papel, podemos escolher diversas cores: preto, branco, azul, vermelho, etc. Porem em computação as cores tem significado prático e temos que representá-las através de números. Devido a evolução histórica dos sistemas eletrônicos encontramos hoje dois tipos de representação de números para termos cores.

*Primeiro:* As placas de vídeo tem o que chamamos habitualmente de palette de cores. Se voltarmos aos nossos lápis de cores, podemos chamar o lápis preto de 1 o azul de 2, o vermelho de 3 etc. Em sistemas de vídeo temos a presença da "cor de fundo ou do papel" (background color), numerada normalmente de cor 0(zero). Com um sistema monocromático temos apenas a possibilidade de representar duas cores (a cor 0 e a cor 1). Em sistemas de 4-cores, podemos representar as cores, 0, 1, 2 e 3. Como pode-se ver o pixel colorido é referenciado na memória de vídeo com a sua respectiva cor. O array de cores é chamado analogamente de palette e todos os pixels na imagem indexam um elemento desta palette. Devido a esta referenciação chamamos este tipo de codificação de "look-up table encoding". Com este tipo de indexação podemos expandir o sistema para indexar três componentes de cores, a R (red-vermelha), G (green-verde) e a B (blue-azul).

Uma das grandes vantagens deste tipo de representação é a redução de trabalho de um algoritmo de troca de uma cor na imagem e principalmente no belo efeito de troca de cores que podemos conseguir mexendo sucessivamente a palette de cores. A maior desvantagem fica associado ao tamanho da palette. Atualmente estamos restritos a palettes de 2, 4, 16 ou 256 cores.

*Segundo:* Para evitar este problema de tamanho da palette podemos usar a representação de cores direta. Com esta

especificação vamos armazenar cada componente de cor (R,G e B) diretamente na memória do computador em vez de associá-lo a uma tabela. Naturalmente fica impossível fazer efeitos de animações com cores com este tipo de representação, porém em contra partida temos agora muito mais cores para representar um pixel. Quando temos um sistema de representação direta das cores RGB em uma representação de 8 bits para cada um dos canais de cores, chamamos o sistema de "true color", pois trabalhamos com 16 milhões de possibilidades de cores.

**Níveis de Cinza.** Variação dos valores de branco até preto passando por níveis intermediários de cinza. Tipicamente a cor preta é associada ao nível 0 e a branca ao nível 255 (ver representação em 8bits).

**Número f ou f/Stop.** Relação entre a distância focal de uma lente e seu diâmetro de entrada.

**Padrão RS-170.** Padrão definido pela EIA (Electronics Industries Association) para os sistemas de televisão monocromático. Especifica um uma amplitude máxima (pico-a-pico) de 1.4 volts, incluindo os pulsos de sincronismo vertical e horizontal.

**Processamento de Dados.** Método no qual o computador recebe uma quantidade de dados e após processamento devolve, na saída, dados da mesma natureza; como por exemplo na movimentação diária de uma conta bancária, onde débitos e créditos são processados para fornecer o saldo no final de cada dia.

**Processamento Digital de Imagens.** Método pelo qual o sistema recebe como entrada uma imagem que, após processada, produz outra imagem como saída. Essa imagem pode ser observada em um dispositivo gráfico. Um exemplo clássico desta área é o processamento de imagens

provenientes de satélite, onde o objetivo é colorir ou realçar detalhes de regiões. Veja um sistema padrão aqui.

**Reconhecimento de Formas.** Classificação de objetos ou imagens em categorias pré-determinadas, usualmente através de métodos estatísticos. Existem os métodos supervisionados (no qual o sistema exige a presença de um supervisor para aceitar criar as classes) e não supervisionados (no qual o sistema-algoritmo tenta sozinho separar o espaço de medidas diferentes categorias ou classes)

**Representação em 8 bits.** Uso de um byte (conjunto de 8 bits) para representar valores entre 0 e 255 = (2 elevado a 8 potência)-1.

**Transformada de Fourier.** Técnica utilizada para converter dados de um domínio X (espaço) para o domínio 1/X (frequência espacial). A transformada de Fourier representa a soma de uma série de formas de onda senoidais com diferentes amplitudes, fase e frequência. Pode ser utilizada em processamento digital de imagens quando queremos conhecer frequências espaciais de um determinado padrão.

**Vetorial x Bitmap.** ver Bitmap x Vetorial .

**Visão Computacional.** A área de visão computacional tem por finalidade obter, a partir de uma imagem (entrada), informações geométricas, topológicas ou físicas sobre o cenário que deu origem a esta imagem. Nesta área podemos lançar mão de técnicas de reconhecimento de formas com o objetivo de classificar padrões ou objetos dentro da imagem. Este domínio tem, por exemplo, grandes aplicações na robótica onde o objetivo é prover o sentido da visão aos robôs.

*Bibliografia*

GONZALEZ, Rafael C. & WINTZ, Paul. "Digital Image Processing". 2ed. Addison-Wesley Publishing Company, Reading, Massachusetts, 1987.

SCHALKOFF, Robert J. "Digital Image Processing and Computer Vision". John Wiley & Sons, New York, 1989.