

# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Conceitos de Representação de Imagens</b>                 | <b>2</b>  |
| 1.1      | Processamento Digital de Imagens . . . . .                   | 2         |
| 1.2      | Representação da Imagem Digital . . . . .                    | 3         |
| 1.3      | Etapas do Processamento de Imagens . . . . .                 | 4         |
| 1.3.1    | Aquisição de Imagens Digitais . . . . .                      | 4         |
| 1.3.2    | Técnicas de Pré-processamento . . . . .                      | 4         |
| 1.3.3    | Segmentação . . . . .  | 6         |
| 1.3.4    | Pós-processamento . . . . .                                  | 7         |
| 1.3.5    | Extração de Atributos . . . . .                              | 8         |
| 1.3.6    | Classificação e Reconhecimento . . . . .                     | 9         |
| <b>2</b> | <b>Projeto Básico de um Sistema de Visão por Computador</b>  | <b>11</b> |
| 2.1      | Especificação Funcional do Sistema . . . . .                 | 11        |
| 2.2      | O Projeto Base . . . . .                                     | 12        |
| 2.2.1    | Escolha do Tipo de Câmera . . . . .                          | 12        |
| 2.2.2    | Escolha do Ângulo de Visão da Câmera . . . . .               | 13        |
| 2.2.3    | Cálculo do Campo de Visão ( <i>field-of-view</i> ) . . . . . | 13        |
| 2.2.4    | Cálculo da Resolução . . . . .                               | 13        |
| 2.2.5    | Resolução da Imagem . . . . .                                | 14        |
| 2.2.6    | Velocidade de Processamento . . . . .                        | 16        |
| 2.2.7    | Seleção da Câmera . . . . .                                  | 16        |
| 2.2.8    | Seleção da Placa de Aquisição . . . . .                      | 17        |
| 2.2.9    | Seleção de Lentes . . . . .                                  | 18        |
| 2.2.10   | Selecionando a Técnica de Iluminação . . . . .               | 20        |
| <b>3</b> | <b>Conclusão</b>   | <b>23</b> |

# Análise de Imagens e Visão Computacional

Márcio Portes de Albuquerque, Eugenio S. Caner,  
Aline Gesualdi Mello e Marcelo Portes de Albuquerque  
mpa, caner, aline, marcelo @cbpf.br  
*Centro Brasileiro de Pesquisas Físicas*

June 13, 2012

## Abstract

(Resumo da V Escola) Neste curso apresentamos conceitos introdutórios sobre a análise de imagens por computador. Atualmente diversas técnicas são empregadas, desde a aquisição, condicionamento e tratamento dos sinais envolvidos. Em física experimental diversos fenômenos são analisado através de suas imagens, sejam elas óticas, magnéticas, topográficas, etc. O processamento de imagens é uma ferramenta que auxilia a busca pela informação presente na imagem. Em qualquer organização de dados em duas (ou três) dimensões, onde estes estão correlacionados espacialmente, é possível a utilização de técnicas de processamento digital de imagens e visão por computador no auxílio a extração da informação. Este curso está dividido em duas etapas: i) os conceitos para a representação das imagens e algumas técnicas básicas e ii) o projeto básico de um sistema de visão por computador. <sup>1</sup>

## 1 Conceitos de Representação de Imagens

### 1.1 Processamento Digital de Imagens

As imagens são produzidas por uma variedade de dispositivos físicos, tais como câmeras e vídeo câmeras, equipamentos de radiografia, microscópios eletrônicos, magnéticos e de força atômica, radares, equipamento de ultrassom, entre vários outros. A produção e utilização de imagens podem ter diversos objetivos, que vão do puro entretenimento até aplicações militares,

---

<sup>1</sup>As notas de aula deste curso estão disponíveis no site da V Escola do CBPF: <http://mesonpi.cat.cbpf.br/e2004/notas.html>

médicas ou tecnológicas. O objetivo da análise de imagens, seja por um observador humano ou por uma máquina, é extrair informações úteis e relevantes para cada aplicação desejada.

Em geral, a imagem pura, recém adquirida pelo dispositivo de captura, necessita de transformações e realces que a torne mais adequada para que se possa extrair o conteúdo de informação desejada com maior eficiência. O Processamento Digital de Imagens (PDI) é uma área da eletrônica/teoria de sinais em que imagens são convertidas em matrizes de números inteiros, sendo que cada elemento desta matriz é composta por um elemento fundamental: o *pixel* (uma abreviação de *picture element*). A partir desta matriz de pixels que representa a imagem, diversos tipos de processamento digital podem ser implementados por algoritmos computacionais. A aplicação destes algoritmos realiza as transformações necessárias para que se possa, por exemplo, obter uma imagem com os realces pretendidos ou extrair atributos ou informações pertinentes. Assim, o PDI pode ser considerado como a união das áreas de processamento de imagem e visão computacional.

Nesta seção apresentaremos alguns conceitos fundamentais do PDI, procurando introduzir o formalismo e as notações que normalmente são utilizados. Inicialmente será apresentada a forma de representação da imagem digital e em seguida serão descritas as etapas fundamentais de um sistema de PDI.

## 1.2 Representação da Imagem Digital

Uma imagem monocromática é uma função bidimensional  $f(x, y)$  da intensidade luminosa, onde  $x$  e  $y$  denotam coordenadas espaciais, que por convenção:  $x = [1, 2, \dots, M]$  e  $y = [1, 2, \dots, N]$ . O valor de  $f$  no ponto  $(x, y)$  é proporcional ao brilho (ou nível de cinza) da imagem neste ponto, como ilustrado na Figura 1. Esta Figura <sup>2</sup> apresenta uma região em destaque em que se pode observar os pixels e os níveis de cinza ou níveis de luminância de cada um deles.

### Pixel e Conectividade

Um pixel é o elemento básico em uma imagem. A forma mais comum para o pixel é a forma retangular ou quadrada. O pixel é também um elemento de dimensões finitas na representação de uma imagem digital. Freqüentemente, a organização de uma imagem sob a forma de uma matriz de pixels é feita em uma simetria quadrada. Isto se deve a facilidade de implementação eletrônica, seja dos sistemas de aquisição seja dos sistemas de visualização de imagens. 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, isto é, ele é anisotrópico. Esta propriedade faz com que um pixel tenha quatro vizinhos de borda e quatro vizinhos de diagonal, ilustrado na Figura 2. Esta propriedade obriga que seja definido o tipo de conectividade que será utilizada, ou B4 (considerando apenas os vizinhos de borda) ou B8 (considerando os vizinhos de borda e os de diagonal). O segundo problema é conseqüência direta do primeiro, ou seja, as distâncias entre um ponto e seus vizinhos não

---

<sup>2</sup>A imagem “Goldhill” é freqüentemente utilizada para testes e demonstrações em PDI



Figure 1: *Imagem monocromática “Goldhill” com destaque para uma região de  $17 \times 17$  pixels*

é a mesma para qualquer tipo de vizinho. Será igual a 1 para vizinhos de borda e  $\sqrt{2}$  para aqueles na diagonal.

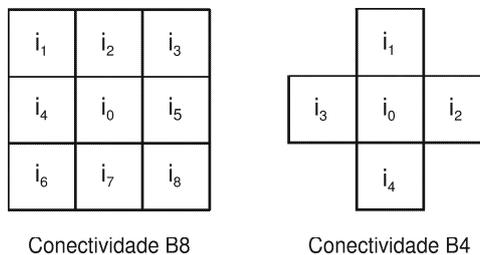


Figure 2: *Ilustração de tipos de conectividade dos pixels vizinhos ao pixel central  $i_0$ . Conectividade B8 apresenta 8 vizinhos, sendo 4 de bordas e 4 diagonais. Conectividade B4 apresenta apenas os pixels de borda.*

### 1.3 Etapas do Processamento de Imagens

Um sistema de processamento de imagens é constituído de diversas etapas, tais como: formação e aquisição da imagem, digitalização, pré-processamento, segmentação, pós-processamento, extração de atributos, classificação e reconhecimento, como ilustra a Figura 3. A seguir, serão apresentadas breves descrições de cada uma destas etapas.

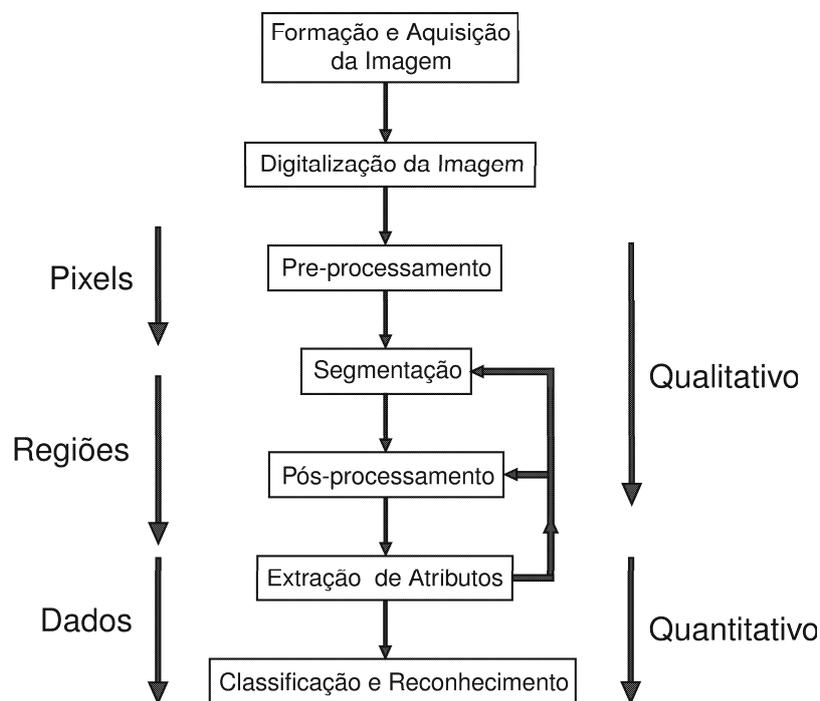


Figure 3: *Etapas de um sistema de PDI.*

### 1.3.1 Aquisição de Imagens Digitais

Dois elementos são necessários para a aquisição digital de imagens. O primeiro é um dispositivo físico que deve ser sensível ao espectro de energia eletromagnético, como por exemplo ao espectro de raio-x, luz ultravioleta, visível, ou infravermelha. Este dispositivo transdutor deve produzir em sua saída um sinal elétrico proporcional ao nível de energia percebido. O segundo, chamado digitalizador, é um dispositivo que converte o sinal elétrico analógico produzido na saída do sensor em um sinal digital.

Em laboratórios de física experimental diversos processos levam a uma organização bi-dimensional de dados, buscando uma representação dos mesmos sob a forma de imagens. A grande vantagem é que esta representação permite que ao observarmos todo o conjunto de dados, nós possamos bucar correlações espaciais entre eles.

### 1.3.2 Técnicas de Pré-processamento

As técnicas de pré-processamento têm a função de melhorar a qualidade da imagem. Estas técnicas envolvem duas categorias principais: métodos que operam no domínio espacial e métodos que operam no domínio da frequência. Técnicas de processamento no domínio espacial baseiam-se em filtros que manipulam o plano da imagem, enquanto que as técnicas de processamento no domínio da frequência se baseiam em filtros que agem sobre o espectro da

imagem. É comum para realçar determinadas características de uma imagem, combinar vários métodos que estejam baseados nestas duas categorias. A Figura 4 ilustra um pré-processamento simples: a aplicação de um filtro mediana, para redução de ruído e em seguida um filtro passa-altas, para realce nos contornos ou bordas dos objetos na imagem.

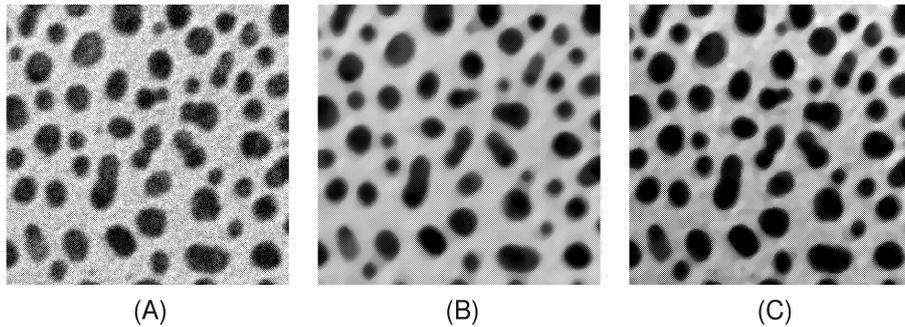


Figure 4: *Exemplo de um pré-processamento simples: (A) Imagem original corrompida com ruído gaussiano, (B) Imagem após a aplicação de um filtro mediana para redução do ruído, e (C) Imagem final, após a aplicação de um filtro passa-altas para realce dos contornos.*

## • O Histograma de Luminância

O histograma de uma imagem digital com  $k$  níveis de cinza é definido por uma função discreta

$$p(k) = \frac{n_k}{n} \quad (1)$$

em que o argumento  $k$  representa os níveis de luminância discretos,  $n_k$  representa o número de pixels na imagem com intensidade  $k$  e  $n$  é o número total de pixels da imagem, ou seja,  $n = M \times N$ . De forma simplificada, podemos afirmar que o histograma de luminância de uma imagem representa a contagem dos níveis de cinza da imagem, podendo informar a distribuição dos pixels dentro dos  $k$  níveis possíveis. O histograma pode ser considerado como uma função distribuição de probabilidades, obedecendo aos axiomas e teoremas da teoria de probabilidades, *i.e.* que  $\sum_k p(k) = 1$ .

O histograma da imagem digital é uma ferramenta bastante útil na etapa de pré-processamento, fornece uma visão estatística sobre a distribuição dos pixels, sobre o contraste da imagem e os níveis de iluminação. Além disso, o histograma é bastante utilizado na etapa de segmentação, principalmente em técnicas que se utilizam da similaridade entre os pixels. O histograma é utilizado com frequência como

sendo uma distribuição estatística dos pixels (“luminância”) na imagem, como por exemplo no caso das técnicas que o utilizam para calcular a entropia da imagem.

Na Figura 5 são apresentadas duas imagens e seus histogramas, sendo que a imagem (A) é uma imagem de baixo contraste, enquanto a imagem (B) possui um maior contraste.

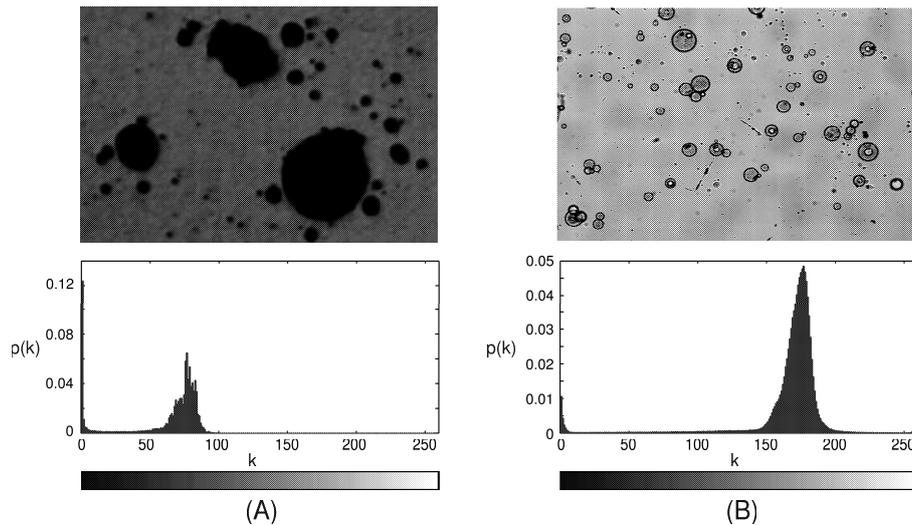


Figure 5: *Exemplo de imagens com histogramas diferenciados. (A) Imagem de baixo contraste e seu histograma de luminância. (B) Imagem de alto contraste e seu histograma de luminância.*

A Figura 6 apresenta outras duas imagens com histogramas de tipos bimodal e multimodal. O histograma bimodal é classificado desta forma devido a presença de dois picos, ou duas regiões de luminância com maior incidência de pixels da imagem. Este é o caso típico de imagens que apresentam objetos e fundo de maneira bem definida, como a imagem (A) da Figura 6. O histograma multimodal apresenta os pixels distribuídos em mais de duas regiões de níveis de cinza, o que se pode perceber através dos picos no histograma em torno destes valores de luminância. Este é o caso da imagem (B) da Figura 6, que apresenta três picos bem definidos.

### 1.3.3 Segmentação

Segmentar uma imagem significa, de modo simplificado, separá-la em suas partes constituintes que se diferenciam entre si. É usual denominar “objetos” da imagem os grupos de pixels de interesse, ou que fornecem alguma informação para o PDI. Da mesma forma, a denominação “fundo” da imagem é utilizada para o grupo de pixels que podem ser desprezados ou que não têm utilidade no PDI. Essas denominações “objeto” e “fundo” possuem

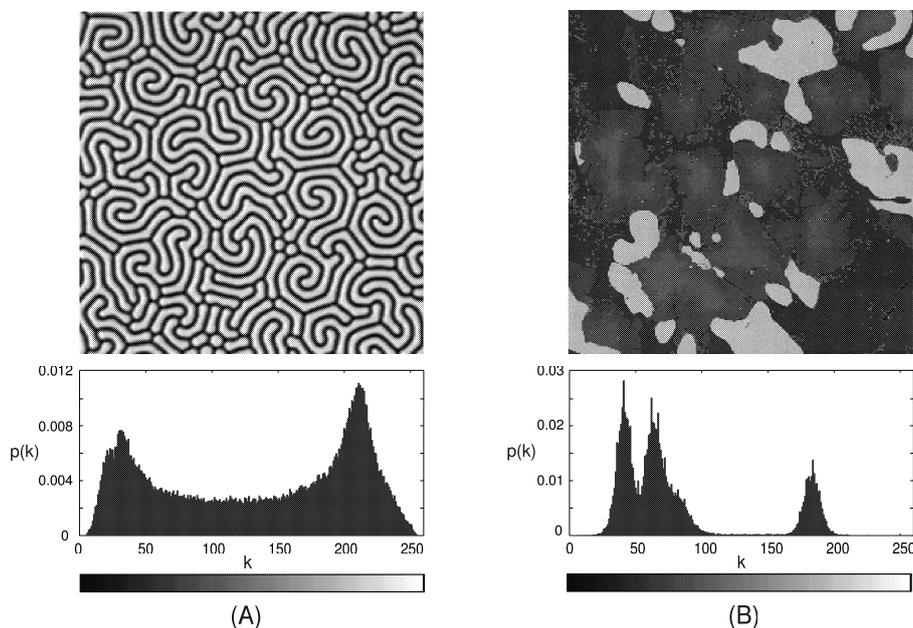


Figure 6: *Exemplo de imagens com histogramas bimodal e multimodal. A imagem (A) possui um histograma bimodal típico, com a presença de dois picos bem definidos. A imagem (B) possui um histograma do tipo multimodal, apresentando mais de duas regiões de concentração dos pixels da imagem.*

uma conotação bastante subjetiva, podendo se referir a grupos de pixels que formam determinadas regiões na imagem sem que representem um objeto, de modo literal, presente na imagem processada.

A segmentação é considerada, dentre todas as etapas do processamento de imagens, a etapa mais crítica do tratamento da informação. É na etapa de segmentação que são definidas as regiões de interesse para processamento e análise posteriores. Como consequência deste fato, quaisquer erros ou distorções presentes nesta etapa se refletem nas demais etapas, de forma a produzir ao final do processo resultados não desejados que podem contribuir de forma negativa para a eficiência de todo o processamento.

Deve ser ressaltado que não existe um modelo formal para a segmentação de imagens. A segmentação é um processo empírico e adaptativo, procurando sempre se adequar às características particulares de cada tipo de imagem e aos objetivos que se pretende alcançar. Apesar de existir uma grande diversidade de técnicas de segmentação de imagens, ainda assim existe atualmente, um grande interesse no estudo e desenvolvimento de novas técnicas.

De um modo geral, as técnicas de segmentação utilizam duas abordagens principais: a similaridade entre os pixels e a descontinuidade entre eles. Sem dúvida, a técnica baseada em similaridade mais utilizada é a chamada *binarização*. A binarização de imagens ou *image thresholding* é uma técnica eficiente e simples do ponto de vista computacional, sendo portanto largamente

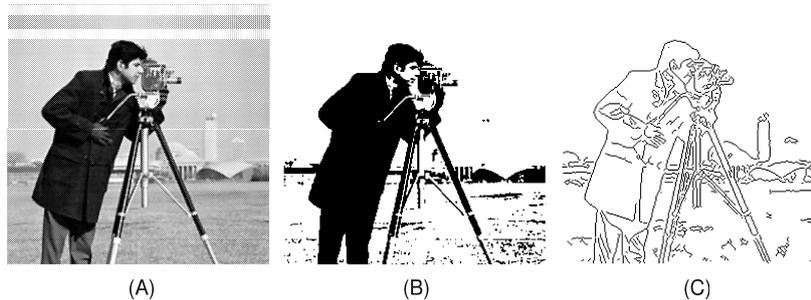


Figure 7: *Dois abordagens para segmentação. (A) Imagem original em níveis de cinza. (B) Imagem segmentada através de uma binarização. (C) Imagem segmentada por detecção de bordas.*

utilizada em sistemas de visão computacional. Este tipo de segmentação é utilizado quando as amplitudes dos níveis de cinza são suficientes para caracterizar os “objetos” presentes na imagem. Na binarização, um nível de cinza é considerado como um limiar de separação entre os pixels que compõem os objetos e o fundo. Nesta técnica, se obtém como saída do sistema uma imagem binária, *i.e.*, uma imagem com apenas dois níveis de luminância: preto e branco. A determinação deste limiar de modo otimizado para segmentação da imagem é o objetivo principal dos diversos métodos de binarização existentes.

As técnicas baseadas em descontinuidade entre os pixels procuram determinar variações abruptas do nível de luminância entre pixels vizinhos. Estas variações, em geral, permitem detectar o grupo de pixels que delimitam os contornos ou bordas dos objetos na imagem. A técnica de segmentação baseada em descontinuidade mais utilizada é a chamada detecção de bordas. A Figura 7 apresenta dois exemplos de segmentação, uma binarização e uma detecção de bordas. A imagem (A) é a imagem original em níveis de cinza, a imagem (B) foi segmentada por binarização e a imagem (C) foi segmentada por detecção de bordas.

### 1.3.4 Pós-processamento

O pós-processamento geralmente é a etapa que sucede a segmentação. É nesta etapa que os principais defeitos ou imperfeições da segmentação são devidamente corrigidos. Normalmente, estes defeitos da segmentação são corrigidos através de técnicas de Morfologia Matemática, com a aplicação em seqüência de filtros morfológicos que realizam uma análise quantitativa dos pixels da imagem.

- **Operações Morfológicas Básicas**

A Morfologia Matemática (MM) é uma das grandes áreas do Processamento Digital de Imagens. Todos os métodos descritos pela MM são fundamentalmente baseados em duas linhas: os operadores booleanos

de conjuntos (união, interseção, complemento etc.) e a noção de forma básica, chamado de “elemento estruturante”. As operações são realizadas sempre entre a imagem e o elemento estruturante. A forma do elemento estruturante é função do tratamento desejado e do tipo de conectividade adotada (B4 ou B8).

Dois operadores básicos são utilizados na maior parte das técnicas de MM: a erosão e a dilatação. Consideremos por exemplo, um objeto  $X$  como um grupo de pixels  $x$  delimitado por uma linha tracejada, como mostrado na Figura 8. A operação de erosão consiste em eliminar do conjunto  $X$  os pixels  $\mathbf{x}$  em função do elemento estruturante  $B$ , tal que:

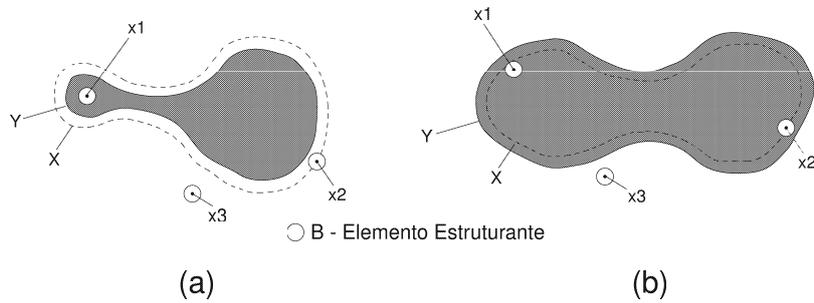


Figure 8: *Operação de erosão (a) e dilatação (b) aplicada a um elemento estruturante  $B$ . Obtemos a forma  $Y$ .  $\mathbf{x}_1$  é um ponto de  $X$  que também pertence a  $Y$ . Em (a)  $\mathbf{x}_2$  pertence a  $X$ , mas não a  $Y$  devido a operação de erosão. Por outro lado, em (b)  $\mathbf{x}_2$  passa a pertencer a  $Y$  devido a dilatação.  $\mathbf{x}_3$  não pertence a ambos  $X$  e  $Y$ .*

$$Y = E^B(X) \rightarrow Y = \{x/B(x) \subset X\} \quad (2)$$

onde  $B(x)$  é o elemento estruturante centrado no pixel  $\mathbf{x}$ . Na prática este procedimento corresponde a construir um novo conjunto de pontos  $Y$ , a partir do conjunto  $X$ , tal que o elemento estruturante esteja inserido totalmente em  $X$ . A Figura 8(a) apresenta o resultado da operação de erosão para 3 pontos distintos da imagem,  $\mathbf{x}_1$ ,  $\mathbf{x}_2$  e  $\mathbf{x}_3$ . O elemento estruturante  $\mathbf{B}$  (representado por um círculo), centrado em  $\mathbf{x}_1$ , está totalmente inserido em  $X$ . O ponto  $\mathbf{x}_1$  pertencerá a forma resultante  $Y$ . O ponto  $\mathbf{x}_2$ , na borda de  $X$ , e  $\mathbf{x}_3$  fora de  $X$  não farão parte de  $Y$ .

A operação dual da Erosão é a Dilatação. Esta operação consiste em dilatar o objeto  $X$  com o elemento estruturante  $B$  tal que:

$$Y = D^B(X) \rightarrow Y = \left(E^B(X^c)\right)^c = \{x/B(x) \cap \subset X \neq \emptyset\} \quad (3)$$

em que  $c$  representa o complemento da operação booleana. A Figura 8(b) apresenta o resultado da operação de dilatação para 3 pontos da imagem.

A operação de erosão permite separar objetos que se tocam. Ao contrário, a operação de dilatação permite preencher furos no interior de um objeto ou mesmo ligá-los. Este resultado dependerá da forma do elemento estruturante. Como as duas operações são iterativas é possível realizar uma seqüência de  $N$  operações de erosão e dilatação sucessivas ou mesmo alternadas. A operação de abertura, uma erosão seguida de uma dilatação, permite eliminar pequenas partículas na imagem (partículas do tamanho do elemento estruturante) e suavizar o contorno dos objetos. Inversamente, a operação de fechamento, uma dilatação seguida de uma erosão, permite fechar canais estreitos que separam objetos, ou suprimir os pequenos furos no seu interior.

### 1.3.5 Extração de Atributos

A etapa final de um sistema de processamento de imagens é aquela em que se extrai as informações úteis da imagem processada. Quando o objetivo do processamento é obter informações numéricas, realiza-se a extração de atributos da imagem.

- **Rotulação ou *Labelização***

A etapa chamada Labelização ou Rotulação é uma etapa intermediária na extração de atributos. Após a etapa de segmentação obtemos uma imagem onde as regiões correspondentes aos “objetos” estão separadas daquelas correspondentes ao “fundo” da imagem. Neste ponto do sistema de processamento, as regiões de interesse estão contiguamente agrupadas por pixels que se tocam. O próximo passo é dar um rótulo (ou *label*) para cada um desses grupos de pixels. Esta identificação permitirá posteriormente parametrizar os objetos segmentados calculando para cada região de pixels contíguos um parâmetro específico, como área ou perímetro por exemplo. A Figura 9 apresenta um exemplo desta técnica para uma imagem constituída de células bem delimitadas entre si. O processo de segmentação separa as regiões pertencentes às células daquelas pertencentes às regiões entre células (fundo), criando um delimitador entre elas. A etapa de “labelização” cria um rótulo que identifica cada uma dessas regiões para que os processos seguintes de tratamento da informação sejam concentrados em cada uma das regiões quer receberam um rótulo.

- **Atributos da Imagem**

Existem basicamente duas classes de medidas: *i*) atributos da imagem como um todo (*field features*), por exemplo número de objetos, área total de objetos, etc. e *ii*) atributos de região (*region features*) que se referem aos objetos independentemente, por exemplo área, perímetro, forma, etc. Os atributos de região podem ser muito sofisticados, permitindo uma nova separação dos objetos em classes de

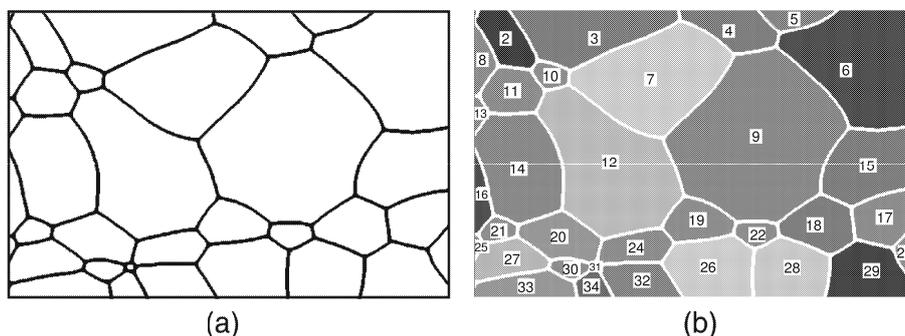


Figure 9: *Imagem ‘Labelizada’*: (a) *Imagem original composta por regiões contíguas de pixels*. (b) *Imagem final após o processo de rotulação. As cores são utilizadas para auxiliar na visualização das regiões*.

similaridades, em função dos parâmetros medidos. A Figura 10 apresenta os principais atributos de região que podem ser extraídos de uma imagem, após as etapas de segmentação e pós-processamento.

### 1.3.6 Classificação e Reconhecimento

O objetivo do reconhecimento é realizar, de forma automática, a “identificação” dos objetos segmentados na imagem. Existem duas etapas no processo de classificação de formas: o aprendizado e o reconhecimento propriamente dito. Na maior parte dos sistemas de reconhecimento de formas, os parâmetros provenientes da etapa de extração de atributos são utilizados para construir um espaço de medida à  $N$  dimensões. Os sistemas de aprendizado irão definir uma função discriminante que separe eficientemente todas as formas representadas neste espaço de medida.

O espaço de medidas é um espaço de dimensão  $N$  onde cada dimensão corresponde a um parâmetro (atributo). A escolha e a qualidade dos parâmetros para construir este espaço é essencial para a realização de um bom processo de reconhecimento de forma. A escolha de um grande número de parâmetros leva a um espaço de grande dimensão e a uma difícil etapa de aprendizado. Um pequeno espaço de medidas pode acarretar em uma baixa caracterização da forma e a muitos erros no processo de reconhecimento. Em alguns casos pode ser interessante normalizar cada eixo para construir um espaço que seja bem distribuído e facilite o processo de classificação.

Podemos dividir o processo de aprendizado em dois tipos: os métodos supervisionados e não supervisionados. No método supervisionado, o classificador, em sua fase de aprendizado, recebe informações de como as classes devem ser identificadas. Por exemplo, em um sistema de reconhecimento de caracteres, existirá classes independentes para todas as letras do alfabeto. Por exemplo a classe das letras “A” será representada pela ocorrência da letra A em suas diferentes variações. O aprendizado consiste então em apresentar ao sistema uma série de objetos “teste” no qual suas classes são previamente conhecidas e definir uma função que separe todas as classes entre si.

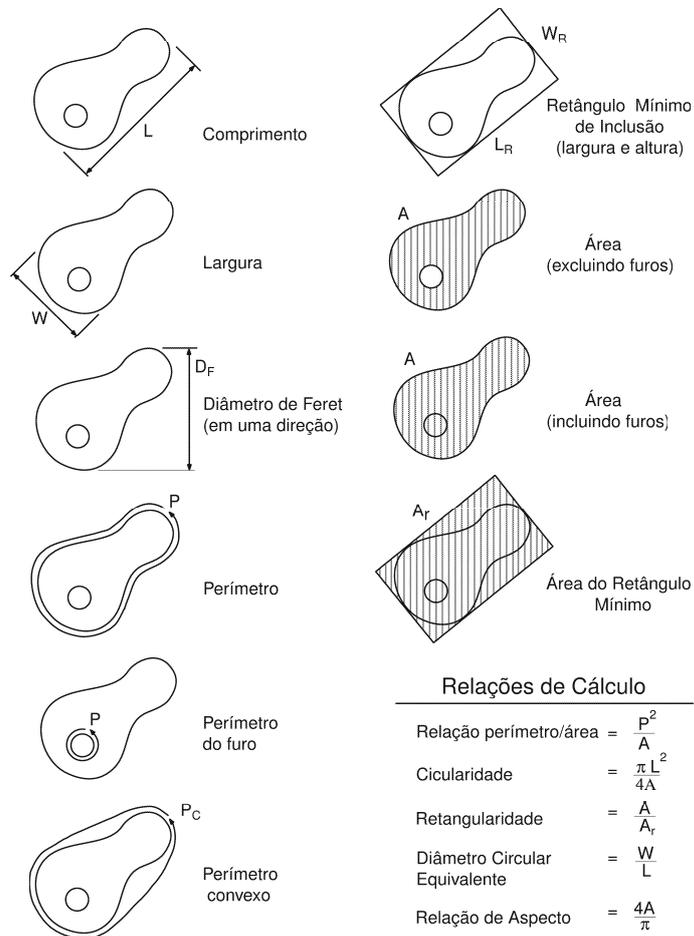


Figure 10: *Principais atributos de região, ou seja, dos objetos independentes presentes na imagem.*

Podemos dizer que o sistema de aprendizado supervisionado age sob a supervisão de um outro sistema de reconhecimento (externo por exemplo) que identificou anteriormente os objetos testes e permitirá a construção correta de seu espaço de medida e sua função discriminante. Durante este processo devemos modificar os parâmetros que compõem o espaço de medida e permitir um melhor ajuste da função discriminante, objetivando sempre que o sistema possa realizar com mais eficiência o processo de classificação. Ao final, é possível determinar a função discriminante responsável pela separação das diversas classes. Este processo pode ser lento e de elevado custo computacional. Somente então os objetos desconhecidos serão fornecidos a este classificador, na fase de reconhecimento.

No caso em que a classificação não é supervisionada, o classificador receberá os objetos desconhecidos e, a partir da medida dos diferentes parâmetros (atributos dos objetos presentes na imagem), ele tentará colocá-los em diferentes classes. A identificação de classes é usualmente realizada a partir da

identificação de agrupamentos em “clusters” de objetos no espaço de medidas.

Existem atualmente vários métodos de reconhecimento de formas. Entre eles podemos citar os métodos baseados em propriedades estatísticas dos objetos (classificador de Bayes), os métodos baseados em técnicas de distância entre os objetos na imagem e suas formas padrões (como as redes neurais artificiais) ou ainda a descrição da forma através de um dicionário ou uma linguagem básica. Neste último caso é definido uma seqüência de elementos básicos que representem as formas dos objetos. Em seguida é construída uma linguagem e formada uma gramática. O reconhecimento é neste caso um reconhecimento sintático baseado nesta gramática [R. C. Gonzalez(1993)]. O reconhecimento por meio de redes neurais artificiais é um método bastante atraente, pois consegue dar um maior grau de liberdade a função discriminante.

## 2 Projeto Básico de um Sistema de Visão por Computador

O desenvolvimento de qualquer sistema requer a determinação dos equipamentos necessários ao seu funcionamento. Podemos aplicar este conceito à construção de um sistema de visão por computador, onde, inicialmente, sabemos a entrada e a saída desejada e desconhecemos seus componentes. A informação de entrada e saída é baseada em especificações funcionais e a lógica do desenvolvimento do sistema pode ser feita como apresentado no diagrama da Figura 11,

O desenvolvimento em seqüência, como sugere a Figura 11, nem sempre pode ser realizado. Por exemplo, não é prático escolher as lentes antes da câmera. Do mesmo modo, há componentes que interagem e devem ser escolhidos em conjunto, como, por exemplo, a câmera e a placa de aquisição devem ser compatíveis.

### 2.1 Especificação Funcional do Sistema

A especificação funcional geralmente é negligenciada no projeto, pois parecem óbvias para aqueles que conhecem as exigências e demandam tempo e esforço para prepará-la. Da mesma forma, algumas pessoas acham as especificações muito rígidas, não fornecendo flexibilidade necessária ao projeto. Entretanto, deixar de produzir uma especificação é um engano, pois esta não precisa ser rígida, podendo ser adaptada a cada problema.

A especificação ideal fornece tanto os requisitos mínimos quanto as melhorias desejadas ao sistema de visão projetado. Desta forma, podemos desenvolver uma proposta que especifica a melhor performance dos equipamentos.

Em geral uma especificação deve fornecer:

- Descrição geral - inclui como a operação será realizada e quais benefícios são esperados do sistema de visão;
- Descrição dos objetos - tamanho, cor, variações, etc.;

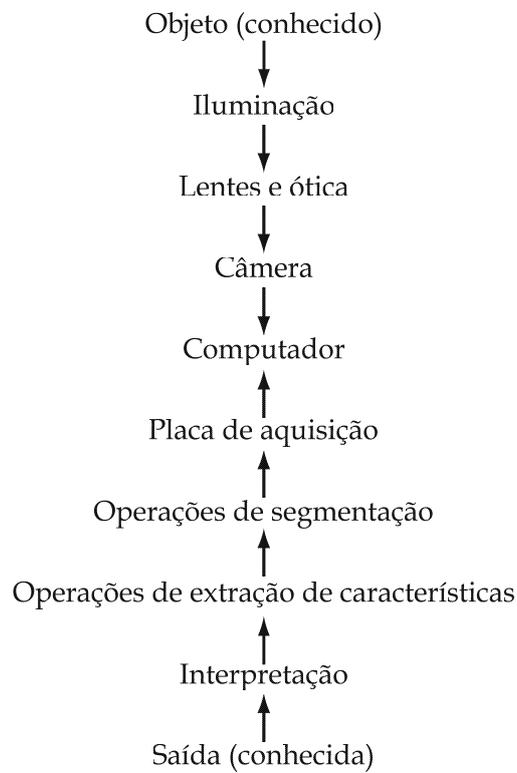


Figure 11: *Diagrama da lógica para o desenvolvimento de um sistema de visão.*

- Exigências de performance - velocidade, acurácia, confiabilidade, etc.;
- Apresentação das partes - o que será incluído no ângulo de visão da câmera;
- Ótica - campo de visão (*field-of-view*), distância do objeto;
- Exigências físicas e mecânicas - limite de tamanho, exigências de montagem;
- Ambiente - iluminação ambiente, temperatura, umidade, sujeira, poeira, névoa;
- Interface do equipamento (I/O) - rede e dispositivos acoplados;
- Interface do operador - controle, tela e privilégios do operador;
- Suporte técnico - treinamento, manutenção e atualizações;

## 2.2 O Projeto Base

Em seguida apresentamos os diversos passos que devem ser executados que contribuem para a obtenção da imagem. São eles:

- Escolha do tipo de câmera;
- Escolha do ângulo de visão da câmera;
- Cálculo do campo de visão (*field-of-view*);
- Cálculo da resolução;
- Velocidade de processamento;
- Seleção do hardware de processamento;
- Seleção da câmera;
- Seleção das lentes;
- Seleção das técnicas de iluminação;
- Seleção da placa de aquisição de imagens;
- Planejar o processamento de imagens;
- Escolher o programa de desenvolvimento.

A ordem apresentada não representa a seqüência de entrada (partes) e saída (imagem) do sistema. Uma possível mudança na ordem de apresentação pode ocorrer sem que prejudique o projeto do sistema de visão. Alguns passos são completados de forma rápida quando comparado com outros, por exemplo, o cálculo da resolução é normalmente mais direto que a seleção de iluminação, e é mais eficiente fazer as coisas fáceis primeiro.

### 2.2.1 Escolha do Tipo de Câmera

Em visão artificial existem câmeras unidimensionais (câmeras de varredura de linhas) e bidimensionais (câmeras de área), Figura 12. A maioria das aplicações de sistema de visão utiliza câmeras de área. Dentro das classes de câmeras existem inúmeras escolhas a serem feitas como, por exemplo, a escolha entre colorido ou monocromático, ângulo de visão, campo de visão, resolução, entre outras.

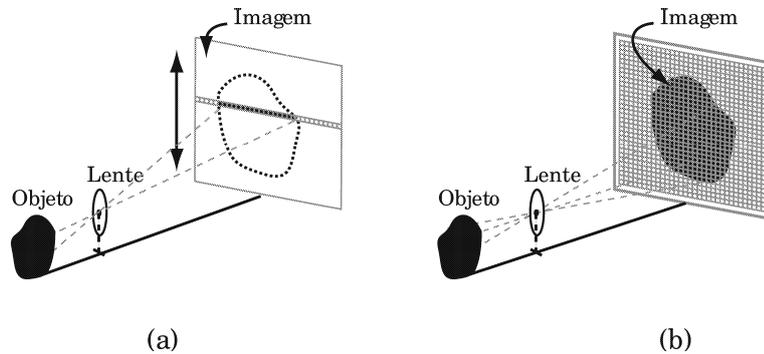


Figure 12: *Tipos de Câmeras.*

### 2.2.2 Escolha do Ângulo de Visão da Câmera

A escolha do posicionamento de câmeras na cena requer o conhecimento de diversos aspectos. Por exemplo, onde os objetos de interesse estão dispostos; quais objetos podem confundir o sistema de visão e onde eles estão localizados; e a disposição (localização e orientação incluindo variações) dos objetos de interesse na cena. Estas informações devem estar disponíveis para a especificação do sistema de visão.

Existem alguns casos especiais que devem ser mencionados. Dentre eles podemos destacar aqueles em que os objetos possuem translação ou rotação em diversas direções. Se o objeto de interesse não pode ser mantido fixo, é necessário determinar um conjunto de vistas que garantam que este estará sempre na cena. Em adição, há o desafio de determinar qual câmera estará vendo as características desejadas durante a operação. Em geral, é melhor ser capaz de restringir parte do objeto e assegurar que as características necessárias são filmadas com um número mínimo de câmeras.

### 2.2.3 Cálculo do Campo de Visão (*field-of-view*)

Uma vez definido o tipo de câmera devemos calcular seu campo de visão. O campo de visão em uma direção pode ser descrito como:

$$FOV = (D_p + L_v)(1 + P_a) \quad (4)$$

onde,  $FOV$  é o campo de visão em uma determinada direção;  $D_p$  é o maior tamanho do objeto na direção do ângulo de visão;  $L_v$  é a variação máxima da localização e orientação do objeto; e  $P_a$  é a fração ou porcentagem que definirá o campo de visão final (incluindo a margem de erro) a partir do campo de visão desejado. O maior tamanho do objeto ( $D_p$ ) e sua variação de localização e orientação ( $L_v$ ) devem estar disponíveis na especificação da aplicação. Veja a Figura 13.

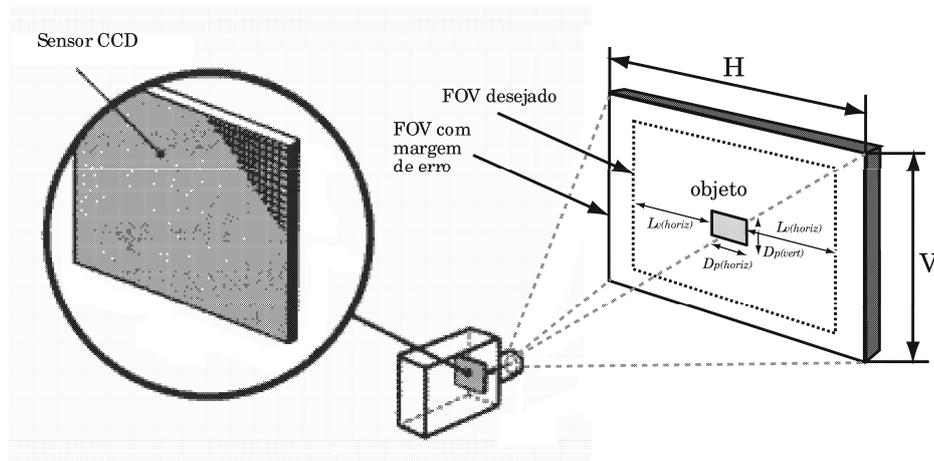


Figure 13: *Campo de Visão (FOV)*.

#### 2.2.4 Cálculo da Resolução

Existem basicamente cinco tipos de resolução utilizados em um sistema de visão computacional, que são:

- Resolução da imagem
- Resolução espacial
- Resolução do objeto de interesse
- Resolução do pixel

#### 2.2.5 Resolução da Imagem

A resolução da imagem é o número de linhas e colunas na imagem, sendo determinada pela câmera, sensor e pela placa de aquisição. Para câmeras de área, a resolução típica de imagens é de  $640 \times 480$  (expressa em colunas e linhas de pixels). Há resoluções mais altas e mais baixas disponíveis no mercado. Para câmeras de varredura de linhas, as resoluções de imagem típicas são 256, 512, 1024 e 2048 (o número de linhas não é exposto porque é sempre igual a 1).

O projetista do sistema ao escolher a câmera e a placa de aquisição determina a resolução da imagem. Quando a câmera e a placa de aquisição possuem o mesmo número de pixels, então existe uma correspondência de um para um entre os pixels adquiridos pela câmera e aqueles fornecidos pela placa de aquisição pelo computador. Quando não há este compartilhamento (como quando utilizamos uma câmera CCTV) então os pixels criados pela placa de aquisição, quando ela periodicamente digitaliza o sinal analógico de vídeo, não serão sincronizados com o detector da imagem dentro do sensor da câmera. Desta forma, a maioria dos projetistas de visão computacional depende da resolução da câmera e da placa de aquisição.

- Resolução espacial: A resolução espacial é a distância entre os centros dos pixels quando estes são mapeados da cena na imagem digital. Ela possui as dimensões do comprimento físico do pixel, por exemplo, 0,1cm/pixel. Ela é determinada pela resolução da imagem e pelo tamanho do campo de visão. Note que para uma dada resolução de imagem, a resolução espacial depende do campo de visão ou da ampliação das lentes. Por exemplo, uma determinada câmera e uma placa de aquisição podem ser configuradas para fotografar um vagão em uma estrada de ferro com resolução espacial de 1 cm/pixel. A mesma câmera e placa de aquisição e, portanto a mesma resolução de imagem, podem ser montadas em um microscópio para fotografar um circuito integrado com resolução espacial de 1 $\mu$ m/pixel. A única diferença entre estas duas situações é a amplitude ótica, Figura 14.

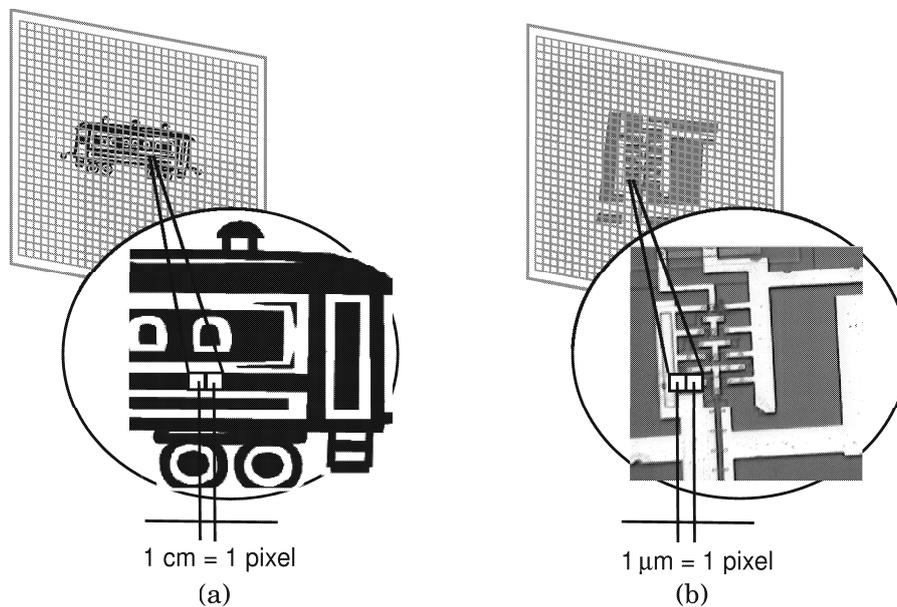


Figure 14: *Resolução Espacial. (a) Situação para 1 cm/pixel - imagem de um vagão de trem; e (b) Situação para 1 $\mu$ m/pixel - imagem de um circuito em microeletrônica.*

Na literatura, a resolução espacial é muitas vezes chamada de resolução do pixel<sup>3</sup>. Nós reservamos o termo resolução do pixel para outro propósito.

- Resolução do objeto de interesse: A resolução do objeto de interesse é a menor resolução possível de um determinado objeto no sistema de imagem. Esta é uma dimensão linear (por exemplo, 0,05mm). Note

<sup>3</sup>É comum encontrarmos o termo *dpi* para expressar a resolução espacial para vários tipos de periféricos. Este termo se refere ao número de pixels por unidade de polegada.

que a definição utiliza a palavra “possível”. Portanto, se um objeto menor do que um pixel aparecer em uma imagem, sua aparência será imperceptível. Isto porque câmeras e placas de aquisição são dispositivos de dados amostrados (um pixel é uma amostra) e, de acordo com o teorema de amostragem de Shannon precisamos de pelo menos dois pixels para a representação de um objeto em uma imagem. Entretanto, na prática, não alcançamos os limites teóricos sendo necessários três ou quatro pixels como tamanho mínimo para a representação de um objeto, Figura 15.

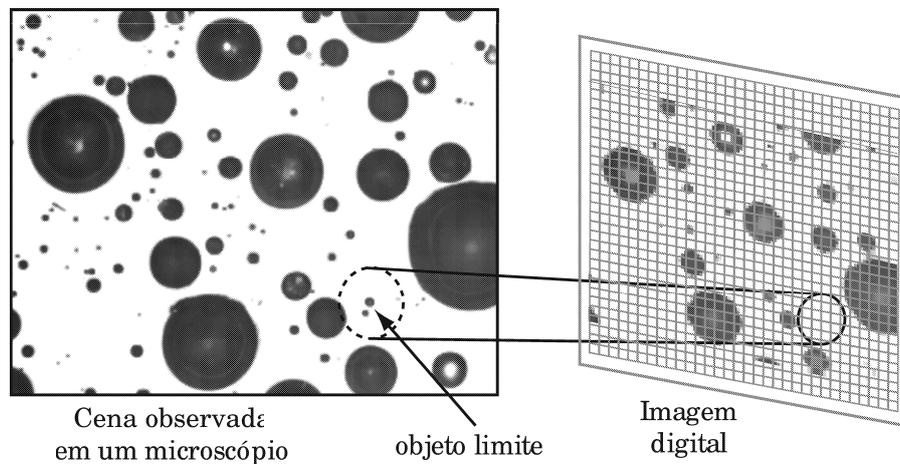


Figure 15: *Limite da Resolução.* Os objetos presentes na cena, cuja as dimensões são inferiores as dimensões do pixel, serão eliminados da imagem digital.

- Resolução do pixel: A resolução do pixel é sua granularidade; isto é, o número de níveis de cinza ou cores que são representadas num pixel. A resolução do pixel é comumente chamada de quantização na literatura científica. A resolução do pixel é função de um conversor analógico/digital localizado tanto na placa de aquisição quanto na câmera. Sistemas de visão monocromáticos normalmente utilizam 8 bits por pixel, fornecendo 256 níveis de cinza. Imagens monocromáticas digitalizadas com 10 ou 12 bits são utilizadas em trabalhos científicos. Sistemas em cores utilizam 8 bits para cada um dos três canais de cores primárias (vermelho, verde e azul) fornecendo 16,777,216 possíveis cores.
- Calculando a resolução: As especificações de uma aplicação nos fornece informação de forma a determinar a resolução do objeto, a resolução da medida, ou as vezes ambas.

Se a especificação requer a detecção de algum objeto pequeno, então este menor objeto determinará a resolução.

$$R_s = FOV/R_i \quad (5)$$

$$R_i = FOV/R_s \quad (6)$$

$$R_m = R_s * M_p \quad (7)$$

$$R_f = R_s * F_p \quad (8)$$

$$R_s = R_f/F_p \quad (9)$$

onde:

- $R_s$  é a resolução espacial (pode ser tanto  $X$  quanto  $Y$ );
- $FOV$  é o campo de visão nas dimensões  $X$  e  $Y$ ;
- $R_i$  é a resolução da imagem; número de pixel em uma linha (dimensão  $X$ ) ou coluna (dimensão  $Y$ );
- $R_m$  é a resolução desejada da medida em unidades físicas (por exemplo, mm);
- $M_p$  é a resolução de medida do sistema de visão em pixels ou fração de pixels;
- $R_f$  é a resolução do objeto de estudo em unidades físicas;
- $F_p$  é o número de pixels que irá determinar o tamanho mínimo de um objeto.

### 2.2.6 Velocidade de Processamento

A estimação da velocidade de processamento permite que evitemos a ocorrência de objetos borrados. A carga total de processamento é igual a soma de todas as cargas de processamento para cada uma das vistas. A melhor estimativa inicial para a carga de processamento é dada medindo-se os dados de entrada em pixels por segundo. Naturalmente, isto não leva em consideração o tipo de processamento necessário. Alguns algoritmos de processamento são intrinsecamente mais rápidos, enquanto outros algoritmos demandam mais recurso de processamento de imagens.

Para estimar a carga de processamento, devemos utilizar a relação:

$$R_p = R_{i(horiz)} \cdot R_{i(vert)} / T_i \quad (10)$$

onde  $R_p$  é a taxa de pixels em pixels/s;  $R_i$  é a resolução da imagem (para uma câmera de varredura em linha, a resolução vertical é 1);  $T_i$  é o menor tempo de aquisição entre duas imagens consecutivas (para câmera de varredura em linha, este valor é igual a  $T_s$ ).

### 2.2.7 Seleção da Câmera

Até este ponto o projetista já decidiu que tipo de câmera usar, varredura em linhas ou área, quantas vistas são necessárias, o campo de visão e qual

resolução da imagem é necessária pela câmera e pela placa de aquisição. A tarefa do agora é fazer a escolha da câmera. O procedimento é primeiramente identificar qualquer característica especial da câmera necessária para a aplicação. Características especiais da câmera devem ser consideradas antes da resolução da imagem, porque esta seleção reduz a seleção disponível para resolução da imagem.

As principais exigências para câmeras de área são:

- Varredura progressiva: o sensor de câmeras que utilizam a tecnologia de varredura progressiva (*progressive scan*) é exposto ao mesmo tempo à cena. As câmeras de varredura entrelaçadas utilizam dois passos, onde a cena é exposta em dois campos separadamente. A varredura progressiva possui vantagens onde há movimento do objeto, pois o quadro é transferido de uma única vez, sem o entrelaçamento.
- Resistência *blooming* (floreescência): câmeras CID (*Charge Injection Device*) possuem as mais elevadas resistências ao *blooming*. Câmeras CID são adequadas às cenas contendo regiões com fortes intensidades luminosas, como apresentada por exemplo na Figura 16. Outras câmeras possuem sensores de imagem com característica *anti-blooming*, e pode ser adequada em algumas aplicações. A alternativa preferida é projetar a iluminação da cena de modo a reduzir o reflexo e o brilho evitando o *blooming*.

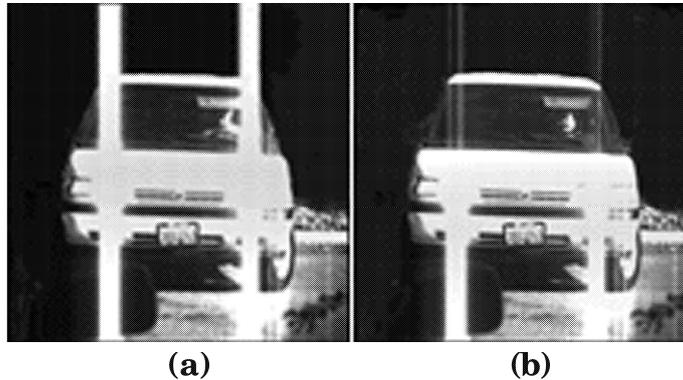


Figure 16: *Efeito de blooming em duas câmeras. (a) câmera normal; (b) câmera CID.*

- Cor: existem dois tipos de câmeras coloridas. Câmeras de cor com único chip são mais baratas, no entanto oferecem uma imagem com resolução de aproximadamente 1/3 da resolução de uma câmera monocromática. Câmeras com três chips de cores possuem maiores resoluções de imagens, no entanto são caras e requerem placa de aquisição mais cara, Figura 18.
- Resolução da imagem: a maioria das câmeras possuem resolução de imagens na faixa utilizada pela CCTV; isto é, aproximadamente  $640 \times 480$

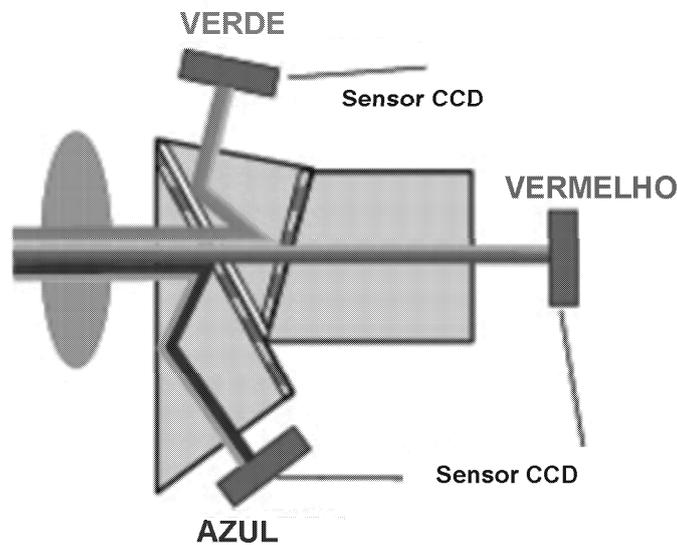


Figure 17: *Camera com três chips de cores.*

pixels. Nesta faixa existem muitos modelos com um conjunto de características que devem ser escolhidas. Para resoluções de imagens significativamente diferentes, a seleção de câmeras é mais restrita. Câmeras de alta resolução têm preços muito elevados.

### 2.2.8 Seleção da Placa de Aquisição

A placa de aquisição (*frame grabber*) é um termo genérico para interface da câmera com o computador, devendo ser determinada por:

- Característica da câmera: a placa de aquisição deve ser compatível com a saída da câmera (analógica ou digital), com a taxa de aquisição e com os tempos da câmera;
- Hardware do computador: a placa de aquisição deve ser compatível mecânica e funcionalmente com o sistema utilizado do computador;
- Atualização da imagem: aplicações que requerem freqüentes atualizações na imagem ou superposição não destrutiva da mesma deve se utilizar de uma placa de vídeo integrada à placa de aquisição;
- Processamento integrado: algumas placas de aquisição têm dispositivos de processamento integrado, podendo ser processadores simples ou DSPs. Dependendo da necessidade de processamento da aplicação esta característica pode ser atrativa;
- Outras entradas e saídas: algumas placas de aquisição têm entradas de sensores de detecção de objetos, fornecendo um sinal de saída para

*trigger* de uma lâmpada estroboscópica. Geralmente, estas entradas e saídas precisam estar sincronizadas com os tempos da câmera.

- Camera Link: é um padrão para comunicação de equipamentos em aplicações de visão computacional. Cientistas e a indústria de vídeo digital definiram este padrão como base para a conexão entre a placa de aquisição e a câmera.

### 2.2.9 Seleção de Lentes

A seleção de lentes envolve principalmente a escolha do tipo da lente e de sua distância focal.

- **Tipos de lentes**

Quase todos os tipos de lentes podem ser utilizados nos sistemas de visão por computador. No entanto, três tipos de lentes são predominantes, que são:

- CCTV (C-mount)
- câmera 35mm
- lentes de aumento

Algumas aplicações utilizam lentes comerciais ou lentes feitas por encomenda. No entanto, o projeto e construção de lentes é caro e demanda tempo. Lentes feitas por encomenda são justificáveis em projetos onde o custo de fabricação é amortizado pela quantidade de sistemas idênticos.

- Lentes CCTV (C-mount): estas lentes são projetadas especificamente para uso de câmeras CCTV, e são amplamente utilizadas em aplicações de sistema de visão. C-mount refere-se a especificação da rosca e da distância focal. A variante de C-mount, a CS-mount, foi desenvolvida para câmeras com sensores de imagens menores (distância focal é 5mm menor).

As lentes são relativamente baratas, pequenas e leves. Elas são projetadas para aplicações de baixo custo e performance como, por exemplo, lentes de sistema de vigilância. Elas não são muito indicadas para aplicações de visão por computador.

- Lentes de câmeras 35mm: lentes para câmeras fotográficas de 35mm têm a melhor relação entre preço e qualidade de imagem. Um grande número de câmeras de área e de varredura de linha vêm equipadas com suporte para lentes de câmera 35mm. Este tipo de lente é fixada à câmera por meio de uma montagem baioneta que foi desenvolvida para facilitar sua troca. Entretanto, ela apresenta um inconveniente na montagem, permitindo que a lente se mova ao redor da montagem no momento de um choque mecânico, vibração ou aceleração. Em alguns casos, modificamos a montagem baioneta incluindo parafusos a fim de acoplar as lentes à câmera.

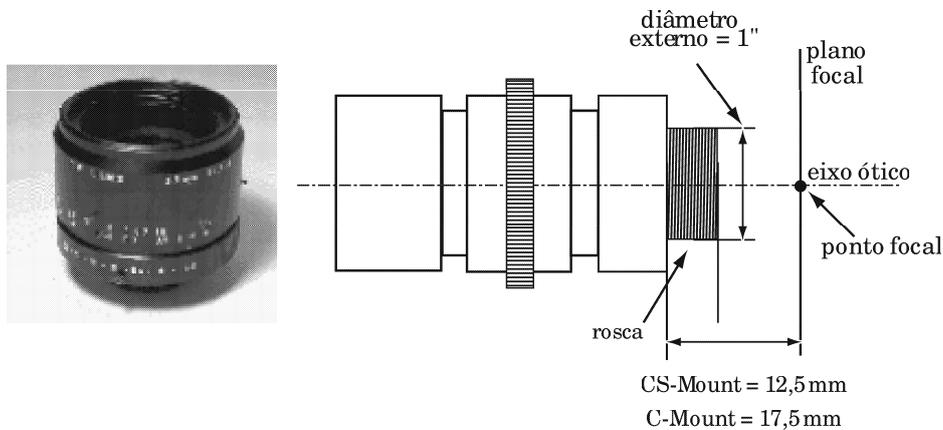


Figure 18: *Exemplo de lente CCTV (C-Mount e CS-Mount).*

- Lentes de aumento: lentes de aumento são projetadas como lentes de campos lisos. Isto é, o plano da imagem e do objeto são lisos. Lentes comuns projetadas para CCTV e câmeras de 35mm permitem campos de curvatura; o plano ótimo do objeto é uma superfície curva onde o plano da imagem é liso. Do mesmo modo, a maioria das lentes comuns são otimizadas para foco no infinito. Lentes de aumento são otimizadas para trabalhar com distâncias próximas.

Lentes de aumento possuem várias limitações: elas são mais difíceis de serem encontradas comercialmente do que as lentes de distância focal fixa, e normalmente não possuem abertura tão grande como as lentes de 35mm e as CCTV, e elas não tem sistema de foco embutido. Existem adaptadores disponíveis que proporcionam foco externos para as lentes de aumento. O custo destas mais o adaptador de foco é normalmente maior que o das lentes para câmeras de 35mm e CCTV.

- **Escolhendo a distância focal**

Até aqui, conhecemos a câmera, o campo de visão e o ângulo de visão da câmera. Desta forma, é possível utilizar estes dados para selecionar a lente.

A Figura 19 ilustra o modelo clássico para lentes finas com e sem abertura do obturador. Este modelo é bastante utilizado para escolha de lentes. Como as lentes atuais são fabricadas com elementos de vidro grosso, devemos esperar alguma discrepância entre os cálculos e os valores reais. A diferença será maior para grandes ampliações ou quando utilizados lentes de grande angular. Para a maioria das situações, os cálculos são satisfatórios e nos permitem determinar o tipo de lente.

As fórmulas para a escolha da lente são:

## Geometria de Lentes Finas

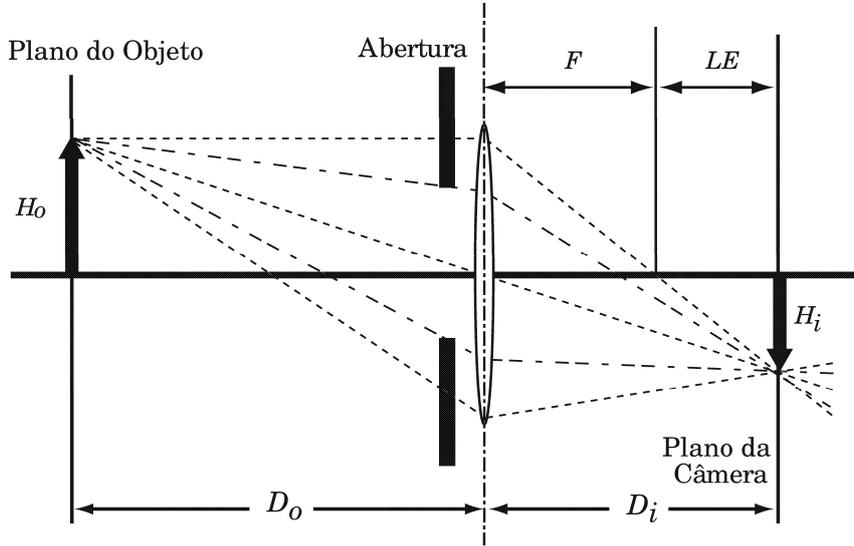


Figure 19: Modelo clássico da geometria de cálculo para lentes finas.

$$M_i = \frac{H_i}{H_o} = \frac{D_i}{D_o} \quad (11)$$

$$F = \frac{D_o * M_i}{1 + M_i} \quad (12)$$

$$D_o = \frac{F(1 + M_i)}{M_i} \quad (13)$$

$$LE = D_i - F = M_i * F \quad (14)$$

$$(15)$$

onde  $M_i$  é a ampliação da imagem;  $H_i$  é a altura da imagem;  $H_o$  é a altura do objeto;  $D_i$  é a distância da imagem;  $D_o$  é a distância do objeto;  $F$  é a distância focal da lente e  $LE$  é a extensão da lente ou a distância que a lente deve estar da imagem de modo a obter foco.

Estes cálculos são simples e devemos estar atento se as unidades são consistentes. O procedimento de escolha das lentes é o seguinte:

1. Escolher a posição a partir das informações especificadas. Se existir uma faixa de distância dos objetos na cena, utilizar o valor intermediário, e siga para o passo 2. Se nenhuma informação for estipulada, utilizar lentes de distância focal mais próximo da maior dimensão do sensor de imagem (diagonal do sensor) e seguir para o passo 4;

2. Calcular a ampliação da imagem utilizando o tamanho do campo de visão previamente determinado e o tamanho do sensor da câmera especificada;
3. Calcular a distância focal necessária utilizando a ampliação e a distância do objeto;
4. Escolher lentes com distância focal perto do valor calculado;
5. Recalcular a distância do objeto para as lentes selecionadas.

### 2.2.10 Selecionando a Técnica de Iluminação

O principal objetivo da iluminação é a otimização do contraste, que pode ser considerado com a qualidade do sinal. O contraste é uma medida da diferença entre duas regiões: o objeto e o fundo da imagem. O primeiro passo na determinação da iluminação é escolher como diferenciar estas duas regiões. A iluminação será definida de forma a enfatizar esta diferença.

O desenvolvimento de projetos de iluminação para imagens de visão computacional envolve cinco parâmetros de iluminação:

- Direção: o controle da direção da luz incidente sobre o aparato é o parâmetro primário que os projetistas de sistemas de visão computacional utilizam para iluminar a cena. Ele é controlado pelo tipo e pelo posicionamento das fontes de iluminação. Geralmente, existem duas abordagens que podem ser adotadas:
  - Direcional, onde a luz incidente vem virtualmente de uma única direção, isto é, ela projeta uma sombra bem definida.
  - Difusa, onde a luz incidente vem de várias direções, isto é, ela projeta sombras e penumbras.
- Espectral: é a cor da luz. É controlada pelo tipo de fonte de luz e por filtros óticos posicionados seja sobre as fontes de iluminação, seja sobre as lentes da câmera.
- Polarização: é a orientação do comprimento de onda. Uma vez que a luz especular refletida (direcional) mantém sua polarização, e que a luz difusa refletida (não direcional) perde sua polarização. Luzes polarizadas podem ser utilizadas em conjunto com polarizadores sobre as lentes da câmera (analisadores) para eliminar a componente especular da luz refletida.
- Intensidade: a intensidade da luz afeta a exposição da câmera. O uso de uma intensidade insuficiente gera baixo contraste ou pode exigir a utilização de amplificadores, que também aumentam o ruído. Por outro lado, o uso excessivo de intensidade resulta em desperdício de energia e no excesso de calor.

- Uniformidade: esta é, possivelmente, uma característica desejável em todos sistema de visão computacional. Uma vez que todas as fontes de luz são atenuadas de acordo com a distância e com o ângulo, é necessário projetar a iluminação de forma a amenizar este efeito.

Alguns parâmetros para se diferenciar duas regiões:

- Refletância: quantidade de luz refletida por um objeto, podendo ser dois tipos distintos de reflexão.
  - Reflexão Especular (Fresnel): o ângulo de reflexão é igual ao ângulo de incidência. A reflexão especular pode ser útil, mas também pode produzir brilhos indesejados.
  - Reflexão Difusa: a luz incidente é espalhada pelo objeto. Em um refletor difuso perfeito, a energia luminosa é espalhada em todas as direções.
- Cor: a distribuição do espectro de energia da luz que é transmitida ou refletida. Existem três maneiras de se visualizar a cor.
  - **Correlação de comprimento de onda direto.** Por exemplo, a luz com cerca de  $550nm$  de comprimento de onda aparenta a cor verde.
  - **Cor Aditiva:** dois ou mais comprimentos de onda são combinados produzindo o efeito de uma cor que espectralmente não está presente no objeto. Por exemplo, que ambas as luzes amarela (cerca de  $620nm$  de comprimento de onda) e azul (cerca de  $480nm$ ) estão presentes, a combinação produz o efeito da cor verde. De fato, não existe nenhuma cor verde presente no espectro de cores.
  - **Cor Subtrativa:** outros comprimentos de onda são removidos. Por exemplo, a diferença em um metal branco como o aço e um metal amarelo como o ouro não se dá pelo fato de que o ouro reflete mais a cor amarela que o aço, mas porque reflete menos cor azul que o aço. A luz branca menos a luz azul produz o efeito da luz amarela.
- Densidade Ótica: a quantidade de luz transmitida por um objeto pode ser desigual em virtude da diferença de material, espessura e propriedades químicas. A densidade pode ser diferente ao longo de todo o espectro, ou mesmo somente em certos pontos do espectro ótico. Geralmente, a luz de fundo é melhor na detecção de diferentes densidades óticas.
- Refração: materiais com diferentes transparências possuem índices diferentes de refração. Assim, eles irão afetar a luz transmitida de formas diferentes. Por exemplo, o ar e o vidro possuem índices de refração diferentes entre si.

- Textura: a textura da superfície pode ser visível no material, ou pode ser muito fina para ser percebida, porém ambas afetam a reflexão. Em alguns casos, a textura é importante, e deve ser enfatizadas pela luz. Em outros casos, a textura não é importante devendo ser minimizada pela iluminação.
- Altura: diferenças de alturas em características específicas dos objetos podem ser enfatizadas com luzes direcionais (às vezes utilizando as sombras) ou minimizadas utilizando uma iluminação bastante difusa.
- Orientação da Superfície: quando a superfície de um objeto varia de acordo com sua orientação relativa, utiliza-se luzes direcionais para permitir a distinção entre as variações; a luz difusa geralmente minimiza as variações na imagem.

Existem duas técnicas principais de iluminação, e diversas variações de cada técnica:

- Luz Frontal
  - **Especular**: a luz é diretamente refletida para fora do objeto em direção à câmera. Isto é útil quando a diferença entre regiões pode ser obtida pela diferença da reflexão especular da luz. Uma vez que cada nível de inclinação da superfície do objeto desloca a luz refletida para ângulos distintos, esta técnica é bastante sensível onde a posição do objeto pode variar.
  - **Fora de eixo**: as luzes são posicionadas de forma a atingir o objeto lateralmente de tal forma que a reflexão ocorra fora das lentes da câmera, porém a reflexão difusa atinge a câmera.
  - **Semi-difusa**: a luz é fornecida de uma grande quantidade de ângulos tal como um grande anel de luzes. Utilizada com cuidado adequado, este tipo de iluminação permite uma boa uniformização em uma pequena região do campo de visão.
  - **Difusa**: luz fornecida virtualmente por todas as direções. Útil quando as características podem ser determinadas através da reflexão difusa, onde é desejada a minimização da reflexão especular ou variação devido à orientação da superfície.
  - **Campo escuro**: luz fornecida por um ângulo próximo a  $90^\circ$  da direção de visão da câmera. Virtualmente, toda a reflexão especular e toda reflexão difusa de uma superfície plana são direcionadas para fora da câmera. Superfícies irregulares, tais como rugosidades, fornecem brilhos especulares bastante intensos que podem ser capturados pela câmera.
- Luz de Fundo
  - **Difusa**: é a forma mais comum de iluminação de fundo, ela consiste de um grande painel translúcido com luzes por de trás dele. É a forma mais simples de montagem e atinge uma uniformidade aceitável.

- **Condensada:** utiliza lentes condensadoras para direcionar o feixe de luz da fonte para as lentes da câmera. Útil para criar luz de fundo com características direcionais.
- **Campo escuro:** luz fornecida por um ângulo próximo de 90° da direção de visão da câmera. Útil quando se procura por artefatos em objetos translúcidos. Sem artefatos presentes, nenhum raio de luz atinge a câmera. Quando existe um anteparo, parte da luz é refletida na direção da câmera.

A técnica de iluminação deve ser escolhida a fim de obter vantagens nas diferenças entre regiões de forma a fornecer o melhor contraste possível. Em geral, a iluminação de fundo é preterida quando a informação necessária pode ser obtida do objeto ou de sua silhueta, e quando é mecanicamente possível realizar a iluminação de fundo.

### 3 Conclusão

Neste curso procuramos abordar aspectos teóricos e práticos para um projeto de um sistema de visão por computador. As técnicas para manipulação de imagens podem parecer complexas e exigem normalmente um bom conhecimento matemático para descreve-las. Recomendamos ao leitor que use como complemento a este texto os slides apresentados em sala de aula, disponíveis na página eletrônica da Escola na Internet <http://mesonpi.cat.cbpf.br/e2004/notas.html>, ou visitem o site do grupo de processamento digital de sinais e imagens (<http://www.cbpf.br/cat/pdsi>).

### References

- [A. Netravali(1999)] A. N. Netravali and B. G. Haskell, 1999. *Digital pictures, representation, compression and standards*. Plenum Press.
- [Pulnix(2001)] PULNIX Imaging Products, 2001. *Oppening the shutters on traffic surveillance*. <http://www.pulnix.com/its/its-op1.html>.
- [Brugge et al.(2001)] M. H. T. Brugge, J. A. G. Nijhuis, L. Spaanenburg and J. H. Sdevens, 2001. *License Plate Recognition, Knowledge-Based Intelligent Techniques in Character Recognition*. Lakhmi C. Jain and Beatrice Lazzerini, <http://www.htsol.com/>.
- [Poltech(2001)] Poltech International Limited Australia, 2001. *Fixed Site Digital Speed Camera and Digital Red Light Camera System*, <http://www.poltech.com.au>.
- [T. Naito(1999)] Naito, T. and Tsukada, T. and Yamada, K. and Kozuka, K., 1999. *License Plate Recognition Method for Inclined Plates Outdoors*. *International Conference on Information Intelligence and Systems*.

- [R. C. Gonzalez(1993)] R. C. Gonzalez and R. E. Woods, 1993. *Digital image processing*, Addison Wesley.
- [O. D. Trier(1995)] O. D. Trier and A. K. Jain, 1995. *Goal-Directed Evaluation of Binarization Methods*, *Transactions on Pattern Analysis and Machine Intelligence*.
- [J. R. Parker(1996)] J. R. Parker, 1996. *Algorithms for Image Processing and Computer Vision*, John Wiley and Sons, Inc.
- [I.E. Abdou(1979)] I.E. Abdou and W.K. Pratt, 1979. *Quantitative design and evaluation of enhancement/thresholding edge detectors*, *Proc. IEEE* 67.
- [A. K. Jain(1989)] A. K. Jain, 1989. *Fundamentals of digital image processing*, Prentice Hall.
- [D. Báez-López(1998)] Báez-López, D. and Ramírez, J. M., 1998. *Pattern Recognition in Automotive Plates*, *Proceedings of the Fifth International Conference on Document Analysis and Recognition*.
- [B. Jahne(1997)] B. Jahne and B. Jahne, 1997. *Digital Image Processing : Concepts, Algorithms, and Scientific Applications*, Springer Verlag.
- [J. Facon(1996)] J. Facon, 1996. *Morfología Matemática: Teoría e Ejemplos*, Jacques Facon.